The Expression, Purification & Characterization of a Recombinant Cellulase from *Cryptococcus gattii*

By:

Dylan Moodley

(1438301)

Dissertation

Submitted in fulfilment of the requirements for the degree

Master of Science

in

Molecular and Cell Biology

in the Faculty of Science, University of Witwatersrand, Johannesburg, South Africa

Supervisor: Dr Angela Botes

Co-supervisor: Dr Ikechukwu A. Achilonu

May 2022

Declaration

I, Dylan Moodley (1438301), am a student registered for the degree of

Master of Science (by dissertation) in the academic year 2020.

I hereby declare the following:

- ☐ I am aware that plagiarism (the use of someone else's work without their permission and/or without acknowledging the original source) is wrong.
- ☐ I confirm that the work submitted for assessment for the above degree is my own unaided work except where explicitly indicated otherwise.
- ☐ I have not submitted this work before for any other degree or examination at this or any other University.
- ☐ I have followed the required conventions in referencing the thoughts and ideas of others.
- ☐ I understand that the University of the Witwatersrand may take disciplinary action against me if there is a belief that this is not my own unaided work or that I have failed to acknowledge the source of the ideas or words in my writing.

Signed: 

Date: 26th of May 2022

Abstract

Cryptococcus gattii (Sanfelice) Vuillemin is an opportunistic pathogen and a primary aetiological agent of cryptococcosis. There are approximately 1 000 000 new cases of cryptococcosis reported per annum, resulting in nearly 625 000 deaths. *Cryptococcus gattii* has been isolated from a number of different environmental sources such as avian guano, soil and most commonly woody debris which is rich in cellulose. An open reading frame within the genome of *C. gattii* (WM276) has been annotated as a putative cellulase, however the structure and function of this gene-product remain unresolved. This is particularly glaring given that infection by this yeast is linked to environmental sources. As such, the overall aim of this project was to recombinantly overexpress, purify, and characterize the putative cryptococcal cellulase in terms of its structure and function. The construction of an automated homology model revealed that despite the presence of the MBP-tag, the cellulase adopted a typical $(\alpha/\beta)_8$ TIM barrel fold, which is indicative of the glycosyl hydrolase family 5 (GH5). The fusion enzyme, MBP-cellulase, was successfully overexpressed in *Escherichia coli* as a soluble protein and affinity purified to homogeneity using amylose affinity chromatography. Investigation of the structural character of the enzyme revealed that MBP-cellulase has a secondary structure composed primarily of α-helices with the presence of some β-sheets. Tertiary structure analysis highlighted that MBP-cellulase undergoes a conformational change in order to accommodate its ligand, sodium carboxymethyl cellulose, suggesting that the MBP does not block the active site. Quaternary structure analysis revealed that MBP-cellulase exists as a 90 kDa monomer. Finally, insights into MBP-cellulase's functionality delineated the fusion enzyme's optimal cellulolytic parameters, which were a pH of 6, a temperature of 50 ˚C and the presence of 500 mM sodium chloride. The presence of divalent metal ions had a negligible to deleterious effect on MBP-cellulase's cellulolytic ability, which potentially rules out the enzyme's requirement for a metallic co-factor. To the best of our knowledge, this work represents the first overexpression, purification, and characterization of a cellulase from *C. gattii* WM276. Moreover, this is the first illustration of the structural character of a cellulase from any cryptococcal pathogen. The thermotolerant and halotolerant nature of this particular cellulase, makes it useful for industrial applications, and adds to our understanding of the pathogen's environmental physiology.

Dedications

I dedicate this work to my mother and father, for their unconditional love, support, and sacrifices made, to give me opportunities that they never had.

I would also like to dedicate this work to the incomparable Marlene Denise Pillay.

"There's Always a Rainbow at The End of Every Rain"

- Prince

Acknowledgements

I shall be eternally thankful to my supervisor, Dr Angela Botes, for motivating me to pursue and complete this project. Your unfailing support and assistance throughout this effort motivated me to persevere when things were difficult.

Additionally, I would like to thank my co-supervisor, Dr Ikechukwu Achilonu, and the Wits Protein Structure-Function Research unit for enabling me to utilise their facilities.

I am eternally grateful to Ms Sasha Richardson, Mrs Ateret Ben-David, Ms Nikita Nankoo and Dr Sanchia Moodley for believing in me and for being constant sources of compassion, comfort, and support.

I would like to express my gratitude to the National Research Foundation for their financial assistance during the course of this degree.

To all my colleagues at the Botes Lab, I would like to express my gratitude for your constant guidance and assistance.

Finally, I would like to thank my mother and father for always encouraging me to achieve my ambitions regardless of the circumstances.

Table of Contents

List of Figures

List of Tables

List of Symbols and Abbreviations

%: percent

~: approximately

˚C: degrees Celsius

<: less than

±: plus minus

α: alpha

µg: microgram

µL: microliter

µM: micromolar

A.U: absorbance units

AIDS: acquired immune deficiency syndrome

AMPs: antimicrobial peptides

C. bacillisporus: *Cryptococcus bacillisporus*

C. decagattii: Cryptococcus decagattii

C. deuterogattii: Cryptococcus deutoerogattii

C. gattii: *Cryptococcus gattii*

C. neoformans: *Cryptococcus neoformans*

C. tetragattii: Cryptococcus tetragattii

c: protein concentration (g/mL)

CATΔ9: chloramphenicol acetyl transferaseΔ9

CBM: carbohydrate binding module

CD: catalytic domain

CMC-Na^+: carboxymethyl cellulase sodium salt

CNS: central nervous system

DNS: 3,5-dinitrosalicylic acid

E. coli: *Escherichia coli*

E6: early protein 6

et al: and others

Far-UV CD: far ultraviolet circular dichroism

g: gravitational force

GFP: green fluorescent protein

GH: glycosyl hydrolase

GH: glycosyl hydrolase family

GRAVY: grand average of hydropathicity

GST: glutathione-s-transferase

h: hours

HIV: human immunodeficiency virus

IPTG: isopropyl β-D-1-thiogalactopyranoside

kDa: kilodalton

l: path length

L-DOPA: l-3,4-dihydroxypheny-alalanine

M: molar

MBP: maltose binding protein

MBP-cellulase: maltose binding protein-cellulase

min: minutes

mL: millilitre

mM: millimolar

n: number of amino acid residues

Na_3PO_4: sodium phosphate

NaCl: sodium chloride

nm: nanometre

P. dulcis: Prunus dulcis

PAGE: polyacrylamide gel electrophoresis

PAMPs: pathogen-associated molecular patterns

pI: isoelectric point

PRRs: pattern recognition receptors

ROS: reactive oxygen species

SDS: sodium dodecyl sulphate

SEC: size exclusion chromatography

T. reesei: Trichoderma reesei

T. viridae: *Trichoderma viridae*

TEV: tobacco etch virus

THF: tetrahydrofuran

TIM: triose phosphate isomerase

TIMP: tissue inhibitor of metalloprotease

TNFs: tumour necrosis factors

V. carnescens: *Vishniacozyma carnescens*

β: beta

Θ: ellipticity (mdeg)

Chapter One: Literature Review

1.1 Introduction

Cryptococcus gattii (Sanfelice) Vuillemin (genotype VGI) is an opportunistic yeast pathogen that has been repeatedly isolated from various environmental sources including guano, soil, and particularly, woody material. Along with its sister species, *Cryptococcus neoformans*, *C. gattii* is associated with lethal infections in both animals and humans.

Cryptococcosis, an invasive fungal disease, accounts for up to 1 000 000 new infections annually as well as 625 000 deaths per annum (Mada *et al*., 2019; Park *et al*., 2009). While *C. neoformans* accounts for up to 95 % of infections, *C. gattii* accounts for the remaining 5 % of reported infections (Zavala and Baddley, 2020). Interestingly, unlike *C. neoformans*, *C. gattii* has shown a predilection for healthy (non-immuno-compromised) individuals (La Hoz and Pappas, 2013). These patients exhibit a higher incidence of both pulmonary and neurological lesions (cryptococcomas), a higher incidence of neurological symptoms and have a weaker response to antifungal therapies (Chen *et* al., 2014; La Hoz and Pappas, 2013; Saag *et al*., 2000). Similarly, cryptococcal meningitis caused by *C. gattii* can result in mortality rates as high as 44 % (Severo *et al*., 2009), whilst meningitis caused by *C. neoformans* has a mortality rate of around 25 % (Saag *et al*., 2000). As such *C. gattii* is largely considered to be more virulent that its sister species.

In a local context, nearly 20 000 strains of *C. gattii* and *C. neoformans* have been isolated from 25 African countries as of 2013 (Cogliati, 2013). The vast majority of these isolates (80 %) originate from South Africa (Cogliati, 2013). Moreover, of the nearly 230 000 global cryptococcal meningitis cases being reported per annum, approximately 75 % of these cases are in Southern African countries (Rajasingham *et al*., 2017). Given the opportunistic nature of both *C. gattii* and *C. neoformans*, the high incidence of cryptococcosis in sub-Saharan Africa is attributed to the high incidence of immune supressing infections, such as human immunodeficiency virus (HIV) and acquired immunodeficiency syndrome (AIDS), which is prevalent amongst populations from those particular countries (Alimi and Keku, 2021).

1.2 Pathogenesis and Treatment

Pathogenesis is the process by which a particular organism induces a disease state in its host (Rauwane *et al*., 2020). In the case of *C. gattii*, the initial site of infection is usually the lungs due to the inhalation of infectious basidiospores, or inadequately encapsulated *C. gattii* cells, that are deposited in the alveoli (Shao *et al*., 2005). Typical host responses rely on tumour necrosis factors (TNFs), interferon-γ and interleukin-2, which culminates as a granulomatous inflammatory response (Perfect, 2005). If the host's immune response is adequate, the phagosome will develop, acidify upon fusion with a lysosome (giving rise to a phagolysosome), and result in the clearance *C. gattii* (Kwon-Chung *et al*., 2014). In some cases, *C. gattii* will remain as a latent infection within phagolysosomes, a fusion between a phagocyte, *C. gattii* cells and lysosomes (Qin *et al*., 2011). The phagolysosomes can remain dormant in an asymptomatic host for many years, until such a time where host immune function is suppressed, and the yeast is permitted to disseminate throughout the body (Sun *et al*., 2010).

Cryptococcosis is most lethal after *C. gattii* has spread throughout the central nervous system (CNS), causing swelling and lesions (Hurtado *et al*., 2019). Central nervous system symptoms include swelling of the brain's surrounding membrane (meningitis), the brain itself (encephalitis), or both (meningoencephalitis), all of which are fatal due to the sudden rise in intracranial pressure (Harris *et al*., 2011).

Although both *C. neoformans* and *C. gattii* can cause disease in immunocompetent hosts, that are otherwise healthy without impaired immune function, infection rates by *C. gattii* are significantly higher in this regard (Lui *et al*., 2006; Maziarz and Perfect, 2016). Furthermore, *C. gattii* is known to be more resistant to antifungal therapies, such as amphotericin B, fluconazole and flucytosine, when compared to *C. neoformans* (Illnait-Zaragozí *et al.*, 2013). Moreover, *C. gattii* infections of the brain are known to react more slowly to conventional treatment and often require more diagnostic follow-ups when compared to *C. neoformans* infections (Tsunemi *et al*., 2001).

Typical treatment strategies include the use of amphotericin B, a compound derived from *Streptomyces nodosus*, in conjunction with flucytosine (Bennett *et al*., 1979). This first line of treatment is then followed by long term maintenance therapy with fluconazole (Kwon-Chung *et al*., 2014). Common dosage strategies for the treatment of cryptococcosis include 1 mg/kg/day of amphotericin B and 100 mg/kg/ day of flucytosine, followed by eight weeks of consolidation treatment, in the form of fluconazole at a dose of 800 mg/day (World Health

Organization, 2018). As a result of extensive antifungal usage in both the medical and agricultural sectors, a variety of fungal pathogens, including *C. gattii*, have developed resistance against traditional therapeutic strategies (Fisher *et al*., 2018). For example, research has shown that the use of tebuconazole as an agricultural pesticide may have caused *C. gattii* to develop hetero-resistance to azoles used in clinical practise, such as fluconazole (Bastos *et al*., 2018).

1.3 Virulence Factors

Cryptococcus gattii, in addition to its ability to persist as a dormant infection, has developed a variety of survival responses and virulence factors that enable it to persist in the phagosome or avoid the host's phagocytic response entirely (Kronstad *et al*., 2011).

Virulence is the severity of the disease condition, and is potentiated by virulence factors (Rauwane *et al*., 2020). Virulence can be seen as the pathogen's capacity to cause illness and the extent to which it is infectious to its host (Shapiro-Ilan *et al*., 2005). Virulence factors such as polysaccharide encapsulation, melanin biosynthesis and the ability to grow at physiological temperature (37 ˚C) aid *C. gattii* in colonising host niches, ultimately resulting in tissue damage and local and systemic inflammation (Fraser *et al*., 2005; de Nies *et al*., 2021). These properties belonging to virulence factors are necessary for pathogens to form an infection and span a wide spectrum, contributing both directly and indirectly to disease processes (Wu *et L*., 2008).

The cryptococcal polysaccharide capsule is a sophisticated, complicated structure made of a combination of glucuronoxylomannan, galactoxylomannan, and select mannoproteins (Leopold-Wager *et al*., 2016). The capsule prevents dehydration of the yeast in an environmental setting, but also facilitates the evasion of phagocytosis, initiated by the host's immune response (Aksenov *et al*., 1973; Kozel *et al*., 1988). During host infection the capsule is able to thicken, which masks pathogen-associated molecular patterns (PAMPs) targeted to *C. gattii*, which are subsequently recognised by pattern recognition receptors (PRRs) on the host's cells (Doering, 2009). This allows *C. gattii* to escape phagocytosis while also aiding in its protection against reactive oxygen species (ROS) and antimicrobial peptides (AMPs).

Cryptococcus gattii utilizes a laccase enzyme as a means of surviving within host macrophages, promoting proliferation within pulmonary sites during initial lung infection and impeding the host's Th17 cytokine response; as well as averting neutrophil recruitment and immune function (Hansakon *et al*., 2019). The laccase enzyme also assists in the breaking down of host-

generated compounds and converting them into highly reactive intermediates which cause harm to the host (Zhu and Williamson, 2003). For example, laccase's inherent iron oxidase activity, facilitates the breakdown of harmful hydrogen peroxide within the phagolysosome into water and oxygen, shielding *C. gattii* from the fungicidal action of alveolar macrophages (Liu *et al*., 1999). The laccase enzyme also produces a melanin intermediate known as dopachrome (Zhu and Williamson, 2004).

The production of melanin by *C. gattii* provides protection against temperature variations, ultraviolet (UV) radiation, enzymatic degradation, and from amoebic predation, to name a few benefits (Doering *et* al., 2008; Wang and Casadevall, 1994; Rosas, 1997). Melanin is a universally present pigment, which is typically negatively charged, hydrophobic and has a high molecular weight (Gómez and Nosanchuk, 2003). Melanin synthesis is brought about by the catalysis of catecholamine substrates, such as l-3,4-dihydroxyphenyl-alanine (L-DOPA), by copper-dependent laccases (Nosanchuk and Casadevall, 2006). Heat energy is collected by melanin, and slowly dispersed, which defends *C. gattii* from heat-induced damage (Rosas and Casadevall, 2006). Likewise, melanin strengthens the cell wall, protecting it against mechanical shearing induced by ice crystal formation when *C. gattii* is exposed to colder temperatures (Rosas and Casadevall, 2006). Melanin protects *C. gattii* against photodamage owing to melanin's innate ability to absorb a wide range of electromagnetic radiation (Hill, 1992). Melanised *C. gattii* cells are able to evade enzymatic degradation due to the steric hindrance encountered, thereby preventing the destruction of *C. gattii* (Jacobson, 2000). Furthermore, melanisation of *C. gattii* aids its evasion from host phagocytosis and has also been implicated in the emerging antifungal drug resistance of this yeast pathogen (Van Duin *et al*., 2002; Wang *et al*., 1995).

Lastly, researchers have identified a suite of secreted enzymes that play a role in cryptococcal virulence. Urease is capable of hydrolysing the substrate urea into two compounds, ammonia and carbamate, which have been implicated in promoting passage across the blood-brain barrier (Almeida *et al*., 2015). Similarly, a recently identified metalloprotease was found to be required for invasion of the CNS in murine infection (Vu *et al*., 2014). Phospholipases, hydrolyse phospholipids found in the host's pulmonary surfactant and plasma membrane, allowing colonisation of host tissue (Chen *et al*., 1997). Furthermore, it promotes fungal survival by maintaining cell wall integrity (Siafakas *et al*., 2007) and delivering nutrients that *C. gattii* may use as the sole source of carbon during the infection (Noverr *et al*., 2003).

1.4 Environmental Distribution and Habitats

As evidenced above, the medical implications surrounding *C. gattii* have been extensively scrutinized over the decades (Watkins *et al*., 2017; Chen *et al*., 2014). However, its environmental capabilities and physiology remain poorly understood.

Like other Tremellales fungi, *C. gattii* is often isolated from soil and decaying woody material, particularly woody material that has a waxy texture, such as *Pseudotsuga menziesii*, *Eucalyptus camaldulensis* and *Prunus dulcis* trees (Litvintseva *et al*., 2011; Springer *et al*., 2014; Chowdhary *et al*., 2011). In Australia, *C. gattii* is an endemic pathogen and is often found in association with *Eucalyptus* trees (Pfeiffer and Ellis, 1992). Previously, *C. gattii* was thought to be environmentally limited to tropical or sub-tropical locations such as Australia, however an outbreak in temperate British Columbia, Canada in 1999 was linked to *C. gattii* (Kidd *et al*., 2004). From Canada, it is suggested that *C. gattii* has progressed to locations within the Pacific Northwest of the United States, such as Oregon and Washington (Datta *et al*., 2009; MacDougall *et al*., 2007). Its recent isolation in North America and Northern Europe indicates that members of the *C. gattii* species complex may have a broader geographical distribution than previously thought (Figure 1.1) (Francisco *et al*., 2021; May *et al*., 2015). *Cryptococcus gattii* (VGI) and *C. deuterogattii* (VGII) are widespread throughout the world, whilst *C. bacillisporus* (VGIII), *C. tetragattii* (VGIV) and *C. decagattii* (VGIV/VGIIIc) are less common. *Cryptococcus bacillisprous* is most prevalent in South America but has also been identified in North and Central America, Australia, New Zealand, and South Asia (Springer and Chaturvedi, 2010). *Cryptococcus tetragattii* has been frequently found in Central America, South America and some Southern African countries as well as India (Springer and Chaturvedi, 2010). The presence of *C. decagattii* is limited to South America, specifically Argentina (Berejnoi *et al*., 2019).

In terms of climate preferences, *C. gattii* (VGI) and *C. deuterogattii* (VGII) are most often isolated from locations which have temperate climates, such as the Pacific Northwest, the Vancouver Islands and Oregon (Byrnes and Marr, 2011). On the other hand, *C. bacillisporus* (VGIII), *C. tetragattii* (VGIV) and *C. decagattii* (VGIV/VGIIIc) are most often isolated from regions which have tropical climates, such as South America, South Asia and parts of Africa such as Zambia (Edwards *et al*., 2021). The global spread of *C. gattii* is potentiated through the pathogen's ability to persist in air, water currents and through biological vehicles, such as animals, birds, insects, and trees (Alimi and Keku, 2021). Furthermore, the ability of *C. gattii* to form associations with the vesicular elements of woody material in trees, in addition to its

ability to persist in the vasculature of *P. dulcis* (almond trees) and its proliferation into soil (Escandon *et al*., 2002; Ren *et al*., 2006), suggest that *C. gattii* can be propagated to different locations through the export of economically valuable wood and trees (Ellis and Pfeiffer, 1990).

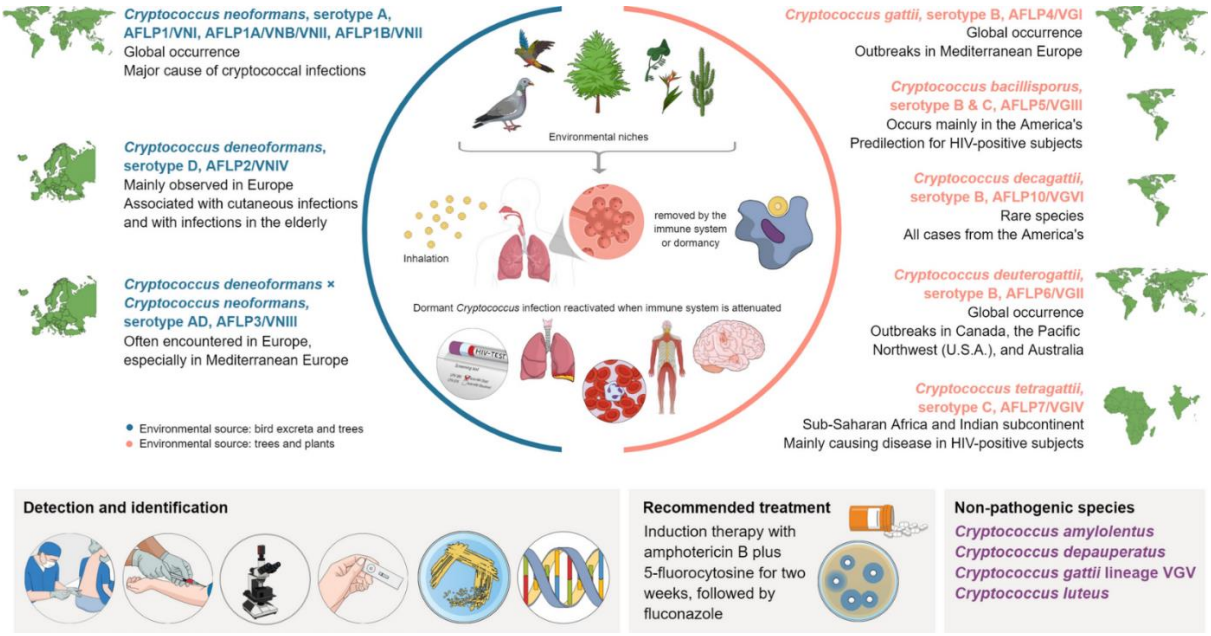

Figure 1.1: The prevalence and distribution of both medically relevant and environmental members of the *Cryptococcus gattii* and *Cryptococcus neoformans* species complex. Additionally, infographics representing the pathogens' proposed environmental niches, detection and identification techniques, life cycle and pathogenesis, as well as the recommended treatment strategies are illustrated (Francisco *et al*., 2017).

1.5 Lignocellulose

When considering that *C. gattii* has been routinely isolated from decaying woody material (Chowdhary *et al*., 2011; Pfeiffer and Ellis, 1992; Springer *et al*., 2014; Vreulink *et al*., 2020), it would not be unlikely that *C. gattii* is capable of utilizing lignocellulosic material as a potential carbon source.

Lignocellulosic biomass is considered the most abundantly available biopolymer and source of potential renewable energy on Earth (Ghaemi *et* al., 2019; Zoghlami and Paës, 2019). Numerous sources of lignocellulosic material exist, including sugarcane bagasse, paper waste from the paper industry, cereal straw, maize stover, and agricultural wastes such as empty fruit and vegetable tree branches (Xu and Li, 2017).

Lignocellulose is a complex polymer composed of the polysaccharides cellulose and hemicellulose; and the aromatic polymer lignin (Zoghlami and Paës, 2019). Cellulose is a linear homopolymeric polysaccharide which consists of repeating units of dextrorotatory glucose (D-glucose) and accounts for 40-60 % of lignocellulose (Harris *et al*., 2010). Longer cellulose chains are thought to have a higher degree of hydrogen bonding and hence are more difficult to hydrolyse, whilst shorter cellulose chains have an overall weaker hydrogen bonding system and thus easier enzyme accessibility (Liu *et al*., 2010). Enzymatic hydrolysis is achieved through the use of cellulases (Siu-Rodas *et al*., 2018). Cellulases' hydrolytic mechanism varies depending on the site of catalysis, with endocellulases acting on amorphous regions of cellulose, by cleaving the 1,4-β-D-glycosidic bonds from within crystalline cellulose (Langston *et al*., 2011). Exocellulases are sub-divided into two types, with type I cellulases sequentially cleaving linkages from the reducing end of the cellulosic chain, and type II cellulases sequentially cleaving linkages from the non-reducing end of the cellulosic chain (Siu-Rodas *et al*., 2018).

Hemicellulose represents a non-uniform group of biopolymers which constitute between 20 - 35 % of lignocellulose (Chandel *et al*., 2018). Hemicellulose polymers include glucomannans, mannans, xylans and xyloglucans which are made up of monosaccharide subunits such as arabinose, galactose, glucuronic acid, mannose and xylose (McKendry, 2002). Hemicellulose has a degree of polymerisation which falls in the range of 60-200, which is significantly lower than cellulose, which has a degree of polymerisation which can be in the range of 100 to well over 10 000, making hemicellulose pre-treatment less harsh and laborious (Chen, 2014; Zoghlami and Paës, 2019). The pre-treatment (chemical or enzymatic) of hemicellulose is an important step in achieving complete hydrolysis of lignocellulose, since hemicellulose has a high degree of intrinsic acetylation, which often impedes cellulose recognition by cellulolytic enzymes (Pan *et al*., 2006).

Lignin is the second most prevalent biopolymer and comprises up to 40 % of lignocellulose (Ahmad and Pant, 2018). It is a very intricate crystalline polymer consisting of several different phenylpropane subunits, including coniferyl, coumaryl and sinapyl alcohols (Chen, 2014). Lignin is the primary contributor to the mechanical rigidity and overall hydrophobicity of the plant's cell wall (Zoghlami and Paës, 2019). It is widely established that lignin has a detrimental effect on the bioconversion of cellulose into fermentable sugars (Santos *et al*., 2012). Generally, the amount of lignin is negatively correlated with an enzyme's cellulolytic activity, as lignin physically restricts polysaccharide accessibility, as it acts as a physical barrier

that prevents enzymes from accessing and successfully binding to cellulose (Sorek *et al*., 2014). Additionally, owing to its hydrophobicity, lignin can permanently adsorb enzymes during enzymatic hydrolysis, which decreases cellulolytic productivity (Kumar and Wyman, 2009). Pre-treatment of lignin requires the use of tetrahydrofuran (THF) in combination with water and dilute sulfuric acid, which strips lignin from cellulose, thereby permitting access to cellulase enzymes for enhanced hydrolysis (Patri *et al*., 2021).

1.6 Cellulases and Cellulose hydrolysis

Cellulases are enzymes produced primarily by bacteria, fungi, protozoans and were recently discovered in green microalgae (Guerriero *et al*., 2018). In general, cellulases are responsible for cellulolysis (Gupta *et al*., 2012), a process which sees the degradation of cellulose into either monosaccharide glucose units, disaccharides such as cellobiose, or oligosaccharides of varying lengths (Barkalow and Whistler, 2014).

Cellulases are classified as glycosyl hydrolases (GH) (Xia *et al*., 2008). Glycosyl hydrolases are a group of enzymes which are able to cleave glycosidic bonds within glycans, glycoconjugates and glycosides (Vuong and Wilson, 2010). Glycosyl hydrolases are grouped on the basis of their amino acid sequence and may be further categorised as retaining or inverting enzymes according to the stereochemical configuration of the enzyme's product(s) (Sinnott, 2010). Retaining GHs typically need an acid or base and a nucleophilic residue, which catalyses the reaction by a double-displacement mechanism, while inverting GHs require just a single catalytic acid and catalytic base residue, and the reaction occurs via a single-displacement mechanism (Thuan and Sohng, 2013).

The complete hydrolysis of cellulose is achieved via the action of three separate cellulases; 1) endoglucanase (endocellulases), 2) exoglucanase (exocellulases) and 3) β-glucosidase (Lynd *et al*., 2002). Endoglucanases cleave in a disordered fashion at amorphous loci, generating oligosaccharide chains of varying lengths. Exoglucanases act on the generated oligosaccharide chains in a sequential manner at either the reducing or non-reducing end. There are two major types of exoglucanases which are distinguishable based on their reaction products, namely: 1) glucanohydrolases, which yield glucose and 2) cellobiohydrolases, which yield the disaccharide cellobiose. Lastly, β-glucosidases hydrolyse cellobiose into glucose monomers (Teeri, 1997).

There are six key steps involved in the hydrolysis of cellulose into its end-products namely, 1) adsorption, 2) lateral diffusion, 3) initial catalysis, 4) processive catalysis, 5) decomplexation and 6) desorption or re-complexation (Fox *et al*., 2011). Initially, the exocellulase carbohydrate binding module (CBM) adsorbs (adheres) to the non-amorphous (crystalline) segments of the cellulose chain, it then moves along the cellulose chain, thereafter, forming a complex with its catalytic domain (CD). The initial catalytic event sees the formation of either glucose, cellobiose or cellotriose. Subsequent processive catalysis permits the formation of cellobiose only. The CD then de-complexes from the cellulose chain where it can either re-complex with an adjacent cellulosic chain or desorb entirely (Fox *et al*., 2011).

Whilst cellulase enzymes are produced by a number of different organisms, cellulases from bacteria and fungi in particular, are of industrial value due to their shorter generation time, low-cost cultivation and their capacity for survival under adverse environments (Jayasekara and Ratnayake, 2019).

1.6.1 Bacterial Cellulases

There are a vast number of bacterial species that are capable of cellulolytic activity, with one of the most commonly occurring bacterial cellulase producers being *Acetivibrio cellulolyticus* (Dassa *et al*., 2012). Members of the genera *Bacillus, Cellulomonas* and *Clostridium* are also well-known cellulase producers (Sadhu, 2013). Additionally, cellulases from bacteria are classified into 14 distinct GH families (Ejaz *et al*., 2021).

Cellulases produced by aerobic bacteria are structurally distinct from those produced by anaerobic bacteria (Doi and Kosugi, 2004). Aerobic bacteria typically produce cellulase enzymes (both endo- and exo-acting) for extracellular secretion to act synergistically on an external substrate (e.g., a plant's cell wall) (Lynd *et al*., 2002; Mansfield *et al*., 1999). All endocellulases and exocellulases from aerobic organisms have a signal peptide, which is abundant in proline, threonine and serine residues, that direct the enzyme for secretion to the external environment (Ejaz *et al*., 2021). On the other hand, anaerobic bacteria make use of large extracellular complexes known as cellulosomes, which are classically composed of a single scaffolding protein (scaffoldins) in conjunction with multiple attached enzymes (Bayer *et al*., 1985). The arrangement of cellulosomes sees up to 11 catalytic enzymes attached to a structural scaffolding protein, which is attached to both the bacterial cell's envelope and the substrate (Schwarz, 2001). When the scaffoldin successfully binds to cellulose, it induces an

increase in the hydrolytic ability of each enzymatic component towards the attached substrate (Schwarz, 2001).

An important consideration is that not all anaerobic bacterial species capable of cellulosome biosynthesis are capable of cellulolytic activity (Doi and Kosugi, 2004). Although *C. acetobutylicum* encodes a 650 kDa cellulosome (Sabatha *et al*., 2002), it exhibits only marginal specific activity against soluble cellulosic substrates such as carboxymethyl cellulose (CMC), is unable to grow on cellulosic material nor does it have the ability to hydrolyse insoluble and crystalline cellulose (Lee *et al*., 1985).

1.6.2 Fungal Cellulases

Basidiomycota are one of the major phyla that constitute the fungal kingdom (Coelho *et al*., 2017). Species within this phylum vary from edible mushrooms such as *Agaricus bisporus* (button mushrooms) to lethal pathogens such as *C. gattii* (De Mattos-Shipley *et al*., 2016; Coelho *et al*., 2017). Basidiomycetes are recognized as one of the most abundant sources of cellulolytic enzymes, which is unsurprising, given that many basidiomycetes are saprophytic, growing on decaying woody material (Lynd *et al*., 2002). The optimal pH and optimal temperature of numerous cellulases from different basidiomycetes fall within distinct ranges of 4.0-5.0 and 37-60 ˚C, respectively (Rouau & Odier, 1986; Hamada *et al.*, 1999). Interestingly, a study undertaken by De Garcia and co-workers (2012) highlighted the cold-tolerant nature of cellulases produced by several Tremellales species. It was found that cellulases produced by *Bulleribasidium foliicola*, *Fonsecazyma tronadorensis*, *Papiliotrema frias*, *Vishniacozyma carnescens* and *V. victoriae* retained their activity at temperatures as low as 5 °C.

Structurally, fungal cellulases have two conserved domains, a large CD and at least one or more smaller CBM. These domains are separated by a flexible arm/linker region whose length varies from 6-109 residues in length (Sonan *et al*., 2007). The CBM acts to bring the catalytic domain near the cellulosic substrate, whilst the CD facilitates hydrolysis. Some endocellulases appear to function without an existing CBM, however all exocellulases require the presence and consequent binding of a CBM to properly enable hydrolysis (Ståhlberg *et al*., 1991).

Endo and exocellulases, belonging to glycosyl hydrolase family 5 (GH5), have conserved active site residues of two aspartates and/or two glutamates, which facilitate the acid-catalysed

hydrolysis of cellulose (Ardèvol and Rovira, 2015). One carboxylate residue acts as a nucleophile, whilst the other carboxylate acts as a proton-donor (White and Rose, 1997). The active site of endocellulases is generally found in a binding site cleft which is composed of four (±1) amino acid residues, whilst the active site of exocellulases houses a similarly structured cleft, which can bind up to ten glycosyl units (Pereira *et al*., 2010; Rouvinen *et al*., 1990).

The presence of a cellulase within the *C. gattii* genome would not be unexpected. Considering that *C. gattii* has a woody ecological niche, it is unsurprising that these pathogens would make use of both endo- and exoglucanases to act synergistically on woody material as a means of colonization and nutrient acquisition (Springer *et al*., 2017; Wang *et al*., 2020). A vast number of microorganisms, particularly Basidiomycota have exhibited the ability to utilise cellulose as a primary carbon source (Cohen *et al*., 2005; Martinez *et al*., 2004). Similarly, a putative cellulase was identified in the genome of *C. neoformans* only, which correlates with woody material being *C. neoformans*' true ecological niche (Loftus, 2005). Subsequently, numerous *C. neoformans* strains have been shown to possess limited cellulase activity (Botes *et al*., 2008).

1.7 Cellulases in Agriculture, Industry, Medicine and Molecular Biology

Enzymes are increasingly being identified as biocatalytic replacements for reactions that traditionally require chemical catalysts, particularly due to the 'green' nature of biologically derived catalysts (Sher *et al*., 2021). In 2019, the global market for industrial enzymes was valued at US$10 billion. This market is expected to continue growing and is anticipated to reach a value of approximately US$15 billion in 2022 (Sher *et al*., 2021).

Cellulases have found wide application in agriculture, industry, and medicine (Kuhad *et al*., 2011). This is attributed to cellulose, the enzyme's substrate, being the most abundant polysaccharide on Earth (George and Sabapathi, 2015). This is also reinforced by the non-specificity of certain cellulases from the GH5, GH7 and GH8 families, that can also hydrolyse non-cellulosic polymers such as chitosan, hemicellulose, and lichenin (Xia *et al*., 2008; Yan and Wu, 2013).

1.7.1 Application in Biofuel Production

The use of biocatalysts as a manner through which crude cellulosic materials, such as sugarcane bagasse, rice straw, sawdust, and other agricultural wastes, are saccharified is seen as the most popular and increasingly demanding application of cellulase enzymes (Kuhad *et al*., 2011).

Cellulases play a key role in the production of biofuels by releasing sugars, which are subsequently fermented into bioethanol (Naik *et al*., 2010). The pre-treatment of cellulosic material is associated with high temperatures and acid or alkaline treatment, and the subsequent neutralization thereof, results in an increased salt concentration. It is therefore recommended that a halotolerant enzyme is used for downstream processing in order to mitigate biocatalytic hindrance. Interestingly, Sepcic and Zalar (2010) found that certain *Cryptococcus* spp maintained their enzymatic activity in an environment supplemented with 10 % (w/v) sodium chloride.

1.7.2 Application in The Alcohol Industry

Cellulases are used to improve the fermentation process, overall quality, and the yield of alcoholic beverages such as beer and wine (Bamforth, 2009). Cellulases are often added at two phases of the brewing process, namely mashing and preliminary fermentation, to enhance glucan hydrolysis, improve wort uniformity, and permit more effective filtering (Bamforth, 2009). Similarly, Galanate and co-workers (1998) highlighted that cellulases were found to increase the extraction of the first wine must by up to 35 %, increase the wine must's filtration rate by an average of 75 %, decrease filtering time by one to two hours, decrease wine must viscosity by an average of 50 %, increase energy savings of up to 40 % during bioreactor cooling and yield a marked improvement in end-product stability.

1.7.3 Application in The Paper and Pulp Industry

Since the 1980's, more focus has been paid to enzymatic recycling of pulp in the paper and pulp industry, particularly via the use of cellulases, as opposed to more conventional chemical options that are not considered ecologically friendly (Mai *et al*., 2004). Notably is the fact that cellulases are capable of de-inking and removing particulate contaminants without impairing the tensile strength or brightness of the paper (Sher *et al*., 2021). Additionally, cellulases may aid in the bioremediation of paper wastes and paper bleaching (Bhat, 2000). The overall benefit of cellulase treatment in the paper and pulp sector is its capacity to improve the final product's aesthetic, textural, and tensile properties without necessitating the use of chlorine, which is ecologically detrimental (Hossein *et al*., 2015).

1.7.4 Role in Agriculture

Cellulases play an important role in agriculture, in that they act in conjunction with pectinases to increase the growth of crops and can negatively affect certain plant pathogens such as *Botrytis cinerea*, *Rhizoctonia solani* and a number of species within the *Pythium* genus (Bhat, 2000). Certain fungal cellulases, such as those secreted by *Trichoderma harzianum*, can hydrolyse the cell wall of plant pathogens, thereby protecting crops and preventing yield losses (Bailey and Lumsden, 1998). Cellulase producing species from the genera *Trichoderma* and *Penicillium* also play beneficial roles in agriculture by enhancing seed germination, accelerating plant development, and blooming, strengthening the root system, and increasing overall yields (Bailey and Lumsden, 1998).

1.7.5 Application in the Textile Industry

Cellulases are industrially valuable due to their texture-refining properties (Kuhad *et al*., 2011). For instance, their favoured use in cotton garment production is attributed to their ability to enzymatically eliminate rough cotton fibrils and improve the dulling of colour, which happens when clothing is washed frequently (Hebeish and Ibrahim, 2007). Endo-acting cellulases tend to be more effective in the texture refining process, which improves the visual aspect, texture, and colour of garments without the need for chemical coating (Galante *et al*., 1998). Endocellulases can restore the texture and colour of these clothes by enzymatically hydrolysing the loose microfibrils (Ibrahim *et al*., 2011). Similarly, cellulases are often included as an active component in detergents due to their colour-enhancing and dirt-removing abilities (Juturu and Wu, 2014).

1.7.6 Food Biotechnology

In conjunction with pectinases and xylanases, cellulases are involved in the preparation of fruit and vegetable juices where they play key roles in extracting and clarifying juices by decreasing the viscosity of the initial puree (Bhat, 2000). Due to the maceration afforded by enzymes, there is a marked increase in yield and production line processing, without incurring additional costs (Kuhad *et al*., 2011). Additionally, cellulase enzymes may be utilised to extract carotenoids and other plant metabolites that can be used to colour food goods and beverages artificially (Sher *et al*., 2021). Chen and co-workers (2011) highlighted the ability of cellulases

to assist in the modification process of certain flavonoids through transglycosylation, as well as the overall extraction of flavonoids through plant cell wall hydrolysis. This is vital given that flavonoids show promise as potential treatments for cardiac disease and certain types of cancers (García-Lafuente *et al*., 2009; Stalikas, 2007).

1.7.7 Role in Animal Feed

Cellulases are used in fortifying animal feed (Godfrey and West, 1996). Animal feed contains a high proportion of constituents known as anti-nutritional factors, which includes cellulose, complex oligosaccharides, inulin and pectin. Cellulases, in conjunction with other hydrolytic enzymes, can hydrolyse these anti-nutritional factors into easily digestible compounds, such as glucose, thereby enhancing animal feed nutrition and subsequent digestion (Ali *et al*., 1995). Animal feed typically needs a heat pre-treatment step, which aims to inactivate any harmful microbial pathogens. Thus, the employment of a thermophilic cellulase, would allow both enzymatic transformation and heat pre-treatment to be carried out concurrently (Bhat, 2000). Furthermore, the synergistic action of cellulase consumption with the fermentation occurring inside the animals' digestive system creates increased levels of propionic acid, which is bacteriostatic, consequently inhibiting the growth of hazardous bacteria without the demand for supplemental antibiotics (Pazarlioğlu *et al*., 2005).

1.7.8 Role in Medicine

Medical implants are an invaluable tool in modern medicine (Khatoon *et al*., 2018). Their utilization varies from artificial heart valves, pacemakers, and stents to orthopaedic implants such as screws, pins, and plates (Bronzino *et al*., 2013; Mclatchie *et al*., 2013). Unfortunately, up to 50 % of all medical implants are colonized by biofilm forming pathogens (Finch *et al*., 1995) such as *Candida albicans*, *Klebsiella pneumoniae*, *Pseudomonas aeruginosa*, *Staphylococcus aureus* and *S. epidermis* (Chen *et al*., 2013; Nobile and Johnson, 2015). Cellulases from the filamentous fungi *P. funiculosum* and *T. reseei* were investigated for their efficacy to degrade biofilms formed by *P. aeruginosa* on glass slides (Loiselle and Anderson, 2003). Both cellulases worked optimally in a moderately acidic environment (pH 5) and were shown to decrease the biofilm areal density by as much as 88 % (Loiselle and Anderson, 2003). Similarly, cellulases have been investigated for their ability to induce the dissociation of

biofilms (Trivedi *et al*., 2016). Recombinant cellulase from the fungus *T. viride* outperformed other recombinant enzymes (amylase, DNAse, lipase and proteinase K) in its ability to degrade and inhibit biofilm formation by *Mycobacterium tuberculosis* (Trivedi *et al*., 2016). This is of particular importance, especially when considering that biofilms can protect the pathogen from changes in environmental conditions, host immune cells and the action of antibiotics (Maya-Hoyos et al., 2015; Stewart and William-Costerton, 2001).

1.7.9 Uses in Molecular Biology

Lastly, individual genetic elements comprising of cellulase encoding genes have been investigated for their inclusion as molecular tools in biochemical and analytic studies (Bayer *et al*., 1994). The CBM domain has been shown to perform well as an affinity partner for heterologous proteins and their subsequent immobilization on a cellulose matrix (Poole *et al*., 1991). A CD and the associated N-terminus from a *Bacillus* derived cellulase was shown to act as both an effective solubility enhancer and secretion signal for the export of heterologous overexpressed proteins in *Escherichia coli* (Gao *et al*., 2015). Similarly, a cellulase promotor derived from *T. reesei* was shown to elicit high degrees of recombinant overexpression in heterologous hosts, allowing for efficient accumulation of heterologous proteins (Saloheimo and Niku-Paavola, 1991).

As seen above, the characterization of novel cellulases is an important endeavour to a number of industries given their broad applications. The characterization of this particular cellulase will also provide insight into the environmental metabolism and physiology of *C. gattii*.

1.8 Aim and Objectives

The aim of this research project was to express, purify and characterize a recombinant cellulase derived from *Cryptococcus gattii* (WM276).

The following objectives are highlighted:

1. Use protein bioinformatics to construct an automated homology model and assess the predicted fold of the fusion enzyme.
2. Overexpress the recombinant maltose binding protein-cellulase (MBP-cellulase) in *E. coli* T7 and purify the enzyme using amylose affinity purification.

3. Characterize the enzyme's structure using far-ultraviolet circular dichroism (far UV-CD), intrinsic tryptophan fluorescence and native polyacrylamide gel electrophoresis.
4. Characterize the enzyme's functionality in terms of optimal pH, temperature, salt concentration and the effect of different divalent metal ions, via the 3,5-dinitrosalicylic acid (DNS) activity assay.

Chapter Two: Materials and Methods

2.1. Protein Bioinformatics

2.1.1 Sequence Information

The gene sequence encoding a putative cellulase, CGB_E4160W, was procured from GenBank (accession number: XM_003194284). This nucleotide sequence originates from the genome of *Cryptococcus gattii* (strain: WM276; serotype: B) on chromosome E (NC_014942.1).

2.1.2 Assessment of Theoretical Parameters

The use of protein bioinformatics is an invaluable tool in predicting numerous theoretical traits regarding proteins that are to be overexpressed (Rodríguez-Ruiz *et al*., 2019). The amino acid sequence (primary structure) of an enzyme can be used to determine the enzyme's isoelectric point (pI), theoretical molecular weight, instability index (which is often correlated to the *in vivo* half-life of a protein), grand average of hydropathicity (GRAVY) (a determinant of solubility in a recombinant host) and molar extinction coefficient (Rodríguez-Ruiz *et al*., 2019). The primary structure of a protein can also be used to probe the presence of certain conserved domains, secondary structures, secretion signal sequences and possible transmembrane topologies (Bakheet and Doig, 2009). A membrane protein (high instability index, positive GRAVY value and predicted transmembrane topologies), for example, will need lower temperature induction and inducer concentrations to slow protein synthesis, resulting in larger quantities of soluble protein when overexpressed in *Escherichia coli* (Wagner *et al*., 2006). As such, it is vital to ascertain these traits to optimise expression conditions to enhance recombinant protein solubility in the host, specifically the maltose binding protein-cellulase (MBP-cellulase), which is the focus of this study.

ProtParam (ExPASy - ProtParam tool, 2020) was used to determine MBP-cellulase's pI, molecular weight, instability index and the GRAVY value, by using MBP-cellulase's amino acid sequence as the input (Rodríguez-Ruiz *et al*., 2019). To determine the presence or absence of a signal peptide, SignalP (version 5; https://services.healthtech.dtu.dk/service.php?SignalP-5.0), in conjunction with Phobius (version 1; https://phobius.sbc.su.se/) and Predisi (version 1; http://www.predisi.de/), were used with MBP-cellulase's amino acid sequence as the input (Hiller *et al*., 2004). The native amino acid sequence, including the signal peptide, was used to determine these parameters (Appendix 1; Figure A1).

2.1.3 Construction of an Automated Homology Model

An automated homology model, using MBP-cellulase's amino acid sequence as the input, was constructed using The Phyre2 web portal (http://www.sbg.bio.ic.ac.uk/phyre2) for protein modelling, prediction and analysis using 'intensive' mode (Kelley *et al*., 2015). The intensive mode of Phyre2 aims to construct a comprehensive full-length homology model of a given amino acid sequence (Appendix 1; Figure A2) by combining multiple template models with simplified *ab initio* folding simulations (Kelley *et al*., 2015). The homology model was constructed using multiple MBPs and cellulase templates, whose structures have been resolved using X-ray crystallography.

2.2 Strains and Vectors

Bacterial strains and vectors used during this study are listed in Table 2.1.

Table 2.1: *Escherichia coli* strains and expression vector used for the expression of recombinant maltose binding protein-cellulase (MBP-cellulase)

Strain/Vector	Genotype/Characteristic	Source
E. coli T7	*lacI*^{q}, *lon*$^{-}$, *OmpT*$^{-}$, *McrA*-, *McrBC*-, *EcoBr*-, *Mrr*-, *hsdSB*$^{-}$	New England Biolabs (MA, USA)
E. coli T7 + empty plasmid	*lacI*^{q}, *lon*$^{-}$, *OmpT*$^{-}$, *McrA*-, *McrBC*-, *EcoBr*-, *Mrr*-, *hsdSB*$^{-}$, *Amp*^{R}, *MalE*	This study
E. coli T7 + pMAL-CGB_E4160W	*lacI*^{q}, *lon*$^{-}$, *OmpT*$^{-}$, *McrA*-, *McrBC*-, *EcoBr*-, *Mrr*-, *hsdSB*$^{-}$, *Amp*^{R}, *MalE*+CGB_E4160W	This study
pMAL-CGB_E4160W	Maltose binding protein (*MalE*), Amp^{R}, 6xHis, Thrombin cleavage site, *Ptac* promotor, cellulase (CGB_E4160W)	This study (constructed by GenScript)

2.3 Construction, Transformation, Expression and Purification

2.3.1 Construction of the Expression Vector

The cryptococcal cellulase gene sequencing (accession number XM_003194284) was codon optimised (using the GenSmart™ Codon Optimization software; www.genscript.com/gensmart-free-gene-codon-optimization.html) for *E. coli* and cloned into the expression vector, pMAL-c5x by GenScript Corporation (NJ, USA) (Figure 2.1). A putative signal peptide was identified using SignalP, Phobius and Predisi (Hiller et al., 2004) between the 17 and 18th amino acid position and was omitted from the synthesis of pMAL-CGB_E4160W, as *E. coli* will not be able to recognise a eukaryotic signal peptide (Ma *et al*., 2020).

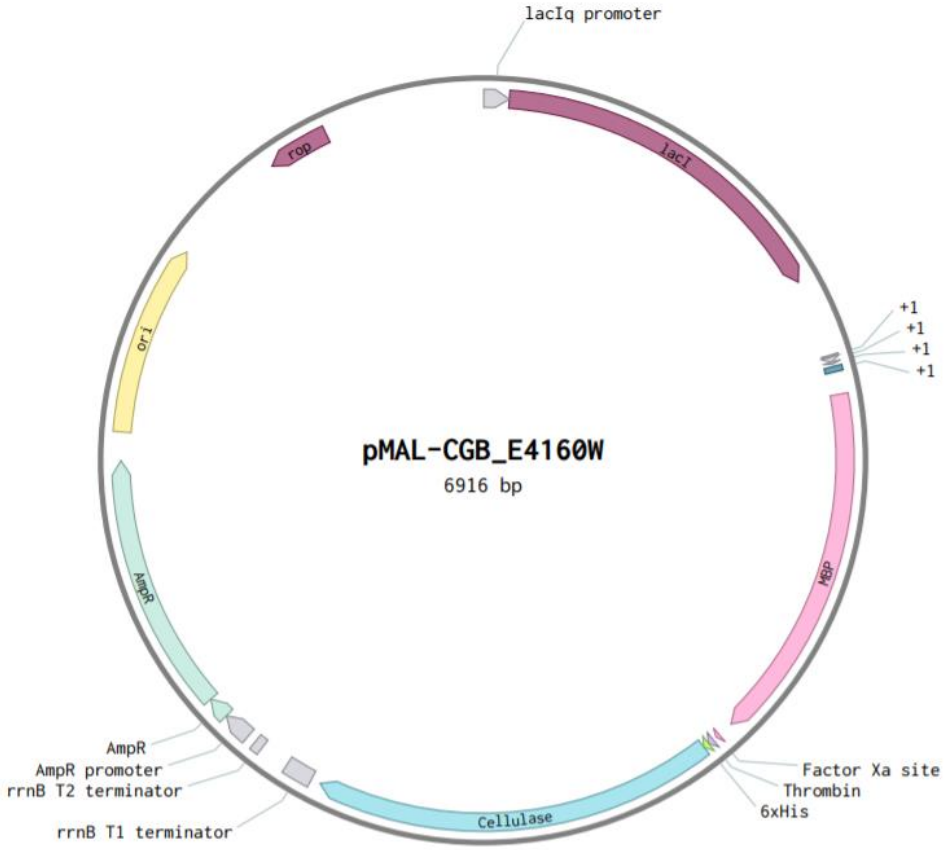

Figure 2.1: Vector map of pMAL-CGB_E4160W, obtained from Benchling (https://www.benchling.com/). The expression vector is 6916 base pairs (bp) in size and encodes the following: maltose binding protein (MBP) fused to a putative cellulase gene derived from *Cryptococcus gattii* (WM276), under the control of the inducible tac promoter (Ptac), with a factor Xa site, a thrombin cleavage site and a hexahistidine tag on the cellulase's N-terminus.

The pMAL-c5X vector was chosen due to an N-terminal MBP, which has a potent solubilising effect on its passenger protein (Rosano, 2014). Rosano and Ceccarelli (2009) highlighted that the presence of 'rare codons' in plant genes expressed in *E. coli* sufficiently slowed down the rate of translation which facilitated the accumulation of soluble protein, which suggests that codon-optimisation might increase translation to such an extent that the majority of the recombinant protein is misfolded. The expression vector, pMAL-CGB_E4160W, was re-engineered to encode a hexahistidine tag and a thrombin cleavage site (Figure 2.1) for downstream purification processes, which was done by adding the nucleotides ctggtgccgcgcggcagc and caccaccaccaccatcac to encode a thrombin cleavage site (LVPRGS) and hexahistidine tag (HHHHHH) respectively. Lastly, the inducible Ptac promotor was selected for controlled expression in *E. coli* (De Boer *et al*., 1983). The promotor is repressed by the *lac* repressor and is inducible (derepressed) by the addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) (De Boer *et al*., 1983).

2.3.2 Transformation of Competent E. coli T7 Cells

Prior to transformation, 4 µg of lyophilised vector DNA (pMAL-CGB_E4160W) was reconstituted in 100 µL of autoclave-sterilised Milli-Q water to a final concentration of 40 ng/mL. Transformation of *E. coli* T7 cells was achieved through the utilisation of an already established heat-shock method (Green and Sambrook, 2018). Briefly, 80 ng of pMAL-CGB_E4160W was added to chemically ($MgCl_2$-$CaCl_2$) competent *E. coli* T7 cells and allowed to incubate on ice for 45 min. The transformation mix was then subjected to heat-shock in a water bath set at 42 ˚C for exactly 1 min, after which the cells were placed back on ice for an additional 2 min. The transformation mix was then added to 1 mL of Luria-Bertani (LB, pH 7.0) broth (Sigma Aldrich; Missouri, USA) and incubated for 1 h at 37 ˚C with shaking at 200 rotations per min (RPM). Three different volumes (25, 50 and 100 µL) of the transformation mix were plated on separate LB agar plates (pH 7.0) supplemented with 100 µg/mL ampicillin (LB^{Amp}) (VWR Life Science; Pennsylvania, United States) and incubated overnight at 37 ˚C. Four large and individually localised colonies were randomly selected and sub-cultured on LB^{Amp} agar and incubated at 37 ˚C overnight. In order to confirm successful bacterial transformation, pMAL-CGB_E4160W was extracted using the ZymoPURE™ Plasmid Miniprep Kit (Zymo Research; California (CA), USA) following the manufacturer's recommended protocol. Briefly, transformed *E. coli* cells were resuspended in a lysis solution

and passed through a silica column binding the DNA. The silica column was then washed to remove non-specifically bound materials, such as cellular debris and genomic DNA, after which the vector DNA was eluted. Extracted pMAL-CGB_E4160W was subjected to either a single restriction enzyme (RE) digest with FastDigest-*Nde*I (Thermo Fisher Scientific; Massachusetts (MA), USA) or a double RE digest with FastDigest-*Nde*I and FastDigest-*EcoR*I (Thermo Fisher Scientific; MA, USA). In the case of the single RE digest, 400 ng of extracted pMAL-CGB_E4160W was digested with 2000 units of FastDigest-*Nde*I, whereas 80 ng of extracted pMAL-CGB_E4160Wwas used in the case of the double digest in the presence of 400 units of FastDigest-*Nde*I and FastDigest-*EcoR*I. Both single and double RE digestions were allowed to proceed for 30 min at 37 ˚C, after which the reaction was terminated at 80 ˚C for 5 min. Undigested and digested pMAL-CGB_E4160W was visualized on a 0.8 % (w/v) agarose gel supplemented with 1 µg/mL ethidium bromide (EtBr). The agarose gel was run for 45 minutes at 120 V and then visualised using ultraviolet (UV) light, in a ChemiDoc™ MP Imaging System (Bio-Rad; CA, USA).

2.3.3 Heterologous Expression Trials

Escherichia coli transformants were inoculated into a 5 mL pre-induction culture composed of freshly prepared LB broth (pH 7.0) supplemented with 100 µg/mL ampicillin and incubated overnight at 37 ˚C with rotation (200 RPM). A 1:100 dilution of the overnight culture into 200 mL freshly prepared rich broth (LB supplemented with 0.2 % (w/v) glucose and 100 µg/mL ampicillin) for subsequent amylose affinity purifications was prepared. The 200 mL induction culture was then incubated at 37 ˚C with rotation (200 RPM) until an optical density (OD) of 0.35-0.55 (OD_{600}) was reached. The culture was immediately cooled on ice for 10 min prior to supplementation with 200 µM IPTG to induce gene expression via Ptac. To investigate the effect of induction temperature on recombinant protein solubility, induction was carried out under varying conditions, specifically; 30 ˚C for 6 h; 20 ˚C for 20 h; 18 ˚C for 22 h and 16 ˚C for 24 h. Induction temperature was gradually decreased in order to slow down protein translation as a consequence of an increase in protein folding time (Galluccio *et al*., 2020). This strategy has been shown to increase the percentage of soluble and stable recombinant protein produced in *E. coli* (Sandomenico *et al*., 2020). Subsequent to induction, cells were harvested via centrifugation at 4000 × *g* for 20 min at 4 ˚C. The cell pellets were stored at -80 ˚C until further processing. Frozen cell pellets were thawed on ice for 15 min and resuspended in ice-

cold amylose affinity binding buffer (20 mM Tris-HCl; 200 mM NaCl; pH 7.4). Resuspended cells were sonicated on ice intermittently for 7 x 30-seccycles at 40 % amplitude of 20 kHz frequency to disrupt the bacterial cell wall (Kim *et al*., 2019). The soluble extracts were separated from the insoluble cell debris via centrifugation at 20 000 x *g* for 15 min at 4 ˚C. The soluble extracts were assayed using sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE) to verify optimal expression conditions for soluble MBP-cellulase (~90 kDa). Untransformed *E. coli* T7 cells and transformed but uninduced *E. coli* transformants served as negative controls and were treated as described above.

2.3.4 Amylose Affinity Purification

Amylose affinity purification utilises the inherent affinity of the MBP to bind to α-(1→4)-maltodextrin (amylose) beads (Pattenden and Thomas, 2008). Once immobilised, the MBP fusion protein can be eluted using maltose, which binds to the MBP and dislodges it from the amylose matrix. For the preparation of the column, 4 mL of Dextrin Sepharose High-Performance resin (Cytiva Life Sciences; MA, USA) in a 50 % (w/v) slurry with 20 % (v/v) ethanol was packed into a 2.5 x 10 cm Econo-Column® chromatography column (BioRad; CA, USA) which corresponds to a bed volume of 2 mL. The ethanol was drained from the column and the resin was rinsed with 10 column volumes (i.e., 20 mL) of Milli-Q water. In order to prime the resin for optimal binding of the MBP-cellulase, the resin was-pre-equilibrated with 10 column volumes of amylose affinity binding buffer. For preparation of the soluble cell extracts, frozen cell pellets derived from induced cultures (0.2 mM IPTG; 16 ˚C; 24 h) of transformed *E. coli* T7 cells were thawed on ice for approximately 15 min. The cell pellets were resuspended in amylose affinity binding buffer and processed as described in section 3.2.2. The soluble extracts were passed through a 0.45 µm cellulose acetate syringe filter to remove any particulate matter that would have otherwise clogged the column and disrupted the purification process. The column was pre-equilibrated with 15 column volumes of amylose affinity binding buffer in order to enhance binding of affinity tagged protein. The filtered soluble extracts were diluted in excess amylose affinity binding buffer in a 1:5 ratio to increase the contact time between the resin and maximise binding of the MBP-cellulase within the soluble extracts. The diluted soluble extracts were then added to the column, after which the column was then washed with 25 column volumes of amylose affinity binding buffer to remove any non-specifically bound material. The immobilised MBP-cellulase fusion protein

was eluted using amylose affinity elution buffer (20 mM tris-HCl; 200 mM NaCl; 10 mM maltose; pH 7.4) All chromatographic fractions were assayed using SDS-PAGE to qualitatively determine the purity of the MBP-cellulase fusion protein.

2.3.5 Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

Sodium dodecyl sulphate-polyacrylamide gel electrophoresis is a technique that has become commonplace in determining a protein's purity, relative size, and solubility (Bergmann-Leitner *et al*., 2008). All expressed protein samples and fractions from the amylose affinity purification were subjected to vertical discontinuous SDS-PAGE comprised of a 4 % (w/w) stacking gel and a 10 % (w/v) resolving gel (Laemmli, 1970). The 10 % (w/v) resolving gel (10 % [w/v] acrylamide; 1.35 % [w/v] bisacrylamide; 0.1 % SDS; 0.375 M Tris-HCl pH 8.8; 4.4 mM ammonium persulfate; 0.2 % [v/v] Tetramethylethylenediamine [TEMED]) was cast and immediately overlayed with 1 mL isopropanol to ensure a smooth interface between the resolving and stacking gel (Tewfick *et al*., 2018). The 10 % (w/v) resolving gel was permitted to polymerise at room temperature for 30 min. Once polymerized, the isopropanol was removed and the resolving gel was overlayed with a 4 % (w/v) stacking gel (4 % [w/v] acrylamide; 0.35 % [w/v] bis-acrylamide; 0.125 M Tris-HCl pH 6.8; 0.01 % [w/v] SDS, 4.4 mM ammonium persulfate and 0.01 % [v/v] TEMED). Protein fractions to be loaded onto the SDS-PAGE gels were prepared by mixing the sample in a 1:1 ratio with 2 x Laemmli buffer (62.5 mM Tris-HCl; 2 % [w/v] SDS; 25 % [v/v] glycerol; 0.02 % [w/v] bromophenol blue; pH 6.8) to partially denature and impart an overall negative charge to the proteins (Pitt-Rivers and Impiombato, 1968). Beta-mercaptoethanol to a final concentration of 5 % (v/v) was added, before boiling the sample for 5 min at 95 ˚C, in order to completely denature the proteins and reduce all intermolecular di-sulfide bonds (Emerson and Ghiorse, 1993). The denatured protein samples were loaded and allowed to stack in the stacking gel for 40 min at 80 V. Once samples entered the resolving gel, the voltage was increased to 120 V, and the samples were run for an additional 100 min. The gels were subsequently stained with Coomassie stain (0.1 % [w/v] Coomassie R250; 10 % [v/v] glacial acetic acid; 40 % [v/v] methanol) overnight. Gels were then transferred to a de-stain solution (20 % [v/v] methanol, 10 % [v/v] glacial acetic acid) for 3 to 4 h until blue protein bands were visible against a clear background.

2.3.6 Protein Quantification

Prior to quantification all elution fractions were dialysed against 1000 volumes of 10 mM sodium phosphate (Na_3PO_4) buffer (pH 7), with two buffer changes at 2-h intervals and a third and final buffer change proceeded overnight at 4 ˚C. Low concentration sodium phosphate buffer was chosen as the dialysate as it removes excess salts, which would have otherwise introduced noise to subsequent structural studies such as far-ultraviolet range (< 200 nm) (Greenfield, 2006; Kumagai *et al*., 2017). Purified MBP-cellulase was easily and readily quantified using the NanoDrop Spectrophotometer ND-1000 (Thermo Fisher Scientific; MA, USA). Using a NanoDrop spectrophotometer is advantageous over traditional spectrophotometry as the surface tension of liquids is exploited. Only small volumes (0.5 – 2 µL) of samples are required, thus limited protein samples are sacrificed (Desjardins *et al*., 2009). The NanoDrop works based on the absorbance of a protein at wavelengths of 280 nm, or 205 nm if the protein lacks aromatic residues, with respect to the Beer-Lambert law (Anthis and Clore, 2013). The Beer-Lambert law stipulates that the absorbance (A) of a protein at the aforementioned wavelengths, in this case 280 nm, is dependent on the extinction coefficient of the protein when all cysteine residues are reduced (ε), the concentration of the protein (c) and the incident light's pathlength (l), such that $A = \varepsilon \times c \times l$ (Grimsley and Pace, 2003). Therefore, the molar concentration (M) of the purified MBP-cellulase was calculated by substituting A_{280} for c as the subject of the formula, such that $c = \frac{A}{\varepsilon \times l}$. The answer was then multiplied by 1000 000 to obtain the micromolar concentration (µM) of purified MBP-cellulase. A blank reading consisting of 10 mM sodium phosphate buffer was taken to ensure that only the concentration of MBP-cellulase was determined. All readings were conducted in triplicate.

2.4 Structural Characterisation

2.4.1 Far-Ultraviolet Circular Dichroism (far-UV CD)

The use of far-ultraviolet circular dichroism (far-UV CD) enables the characterisation of the recombinant MBP-cellulase's secondary structure, namely, the fractions of the MBP-cellulase's structure derived from α-helices, β-sheets, and random disordered coils (Reed and Kinzel, 1984). This is due to the chirality of amino acids. More specifically, the peptide bonds, inherently optically active in the far-UV range (180-250 nm), differentially absorb left- and right-handed polarised light in an unequal manner (Kelly and Price, 2000). For instance,

α-helices have a positive CD ($A_{left} > A_{right}$) at 190 nm and two negative values ($A_{left} < A_{right}$) at 222 and 208 nm. Conversely, β-sheets have a positive CD at 196 nm and a negative CD at 218 nm. Lastly, random coils have a positive CD at 212 nm and a negative CD at 195 nm. A total of 0.2 mg/mL dialysed MBP-cellulase suspended in 20 mM sodium phosphate buffer (pH 7.4) was subjected to far-UV CD analysis. In general, a desirable protein concentration for CD analysis is in the range of 0.05-0.2 mg/mL (Kelly and Price, 2000). The following parameters were set for CD spectra collection on a Jasco J-810 Circular Dichroism Spectropolarimeter: sensitivity (100 mdeg), start wavelength (260 nm), end wavelength (190 nm), data pitch (0.5nm), scanning mode (continuous), scanning speed (1000 nm/sec), response (1 second) and bandwidth (5 nm) (Kelly and Price, 2000). A blank containing only sodium phosphate buffer (20 mM; pH 7.4) was also assessed to obtain the contribution of the buffer to the CD spectrum. The CD spectrofluorometer's output data is in ellipticity units ([θ]) and a total of five readings were recorded. The data was then converted into mean residual ellipticity [θ]MWR with the equation:

$$[\theta] = \frac{(100 \times \theta)}{C \times n \times l}$$

θ = ellipticity (mdeg)

n = number of amino acid residues

l = path length (1 mm)

c = protein concentration (g/mL)

2.4.2 Intrinsic Tryptophan Fluorescence

Intrinsic fluorescence is an important characteristic belonging to proteins that contain an appreciable number of aromatic residues (Akbar *et al*., 2016). The fluorescence results from any aromatic amino acids (phenylalanine, tryptophan, and tyrosine) containing pi-bonds in their aromatic moiety (Antosiewicz and Shugar, 2016). Tryptophan absorbs UV-light at 295 nm and has an emission maximum at 350 nm, depending on the degree of quenching. Quenching is a phenomenon wherein the fluorescence intensity is decreased due to the tryptophan's close proximity to protonated groups such as histidine and lysine. Quenching of tryptophan occurs due to its energy transfer to other aromatic residues through Förster resonance energy transfer (FRET) (Ghisaidoobe and Chung, 2014). Tryptophan residues are

also very sensitive to their environmental conditions, which allows the study of conformational changes under different environmental conditions, such as changes in temperature, pH, ligand-binding, and the concentration of denaturing agents such as urea. A total of 10 µM MBP-cellulase in 10 mM sodium phosphate buffer (pH 7) was subjected to intrinsic tryptophan fluorescence analysis. Generally, 0.1-0.5 µM of purified protein is sufficient to obtain accurate readings (Royer, 2006). The resultant difference in tryptophan spectra and intensity under three varying conditions was assessed, specifically, the native, denatured and ligand-bound state of MBP-cellulase. The observation of a blue- or red-shift was observed, depending on the initial degree of quenching, in other words, in relation to the degree of fluorescence in the native conformation (Royer, 2006). Native MBP-cellulase was assayed in the presence of 10 mM sodium phosphate buffer (pH 7) only, denatured MBP-cellulase was assayed in the presence of 10 mM sodium phosphate buffer (pH 7) supplemented with 8 M urea, whilst ligand-bound MBP-cellulase was assayed in the presence of carboxymethylcellulose sodium salt (CMC-Na^+), at a concentration of 0.5 % (w/v). The following parameters were set on the Jasco FP-8200 Spectrofluorometer: excitation bandwidth (2.5 nm), emission bandwidth (5 nm), the response time (1 sec), sensitivity (high), excitation wavelength (295 nm), measurement range (280-500 nm) and data interval (200 nm/min) with three accumulations (Ghisaidoobe and Chung, 2014). A blank reading containing only 10 mM sodium phosphate buffer (pH 7), was taken before the experimental readings of the MBP-cellulase fusion and subtracted from the experimental readings. All readings were conducted in triplicate.

2.4.3 Native Polyacrylamide Gel Electrophoresis

Analysis of the quaternary structure of the MBP-cellulase was performed using native polyacrylamide gel electrophoresis (PAGE), which can accurately determine the oligomeric state of a particular protein (Moutaoufik *et al*., 2017; Zhang *et al*., 2017). Native PAGE permits the preservation of the protein's oligomeric state due to the absence of denaturants, such as SDS (Padula *et al*., 2017). Purified MBP-cellulase was subjected to discontinuous native PAGE comprised of a 4 % (w/v) stacking gel and a 10 % (w/v) resolving gel (Laemmli, 1970). The 10 % (w/v) resolving gel and the 4 % stacking gel were prepared as described in section 2.3.5 however, SDS was excluded. Purified MBP-cellulase to be loaded onto native the PAGE gel was prepared by mixing the fusion protein in a 1:1 ratio with 2 x loading buffer (62.5 mM Tris-HCl; 25 % [v/v] glycerol; 0.02 % [w/v] bromophenol blue; pH 6.8) (Pitt-Rivers and

Impiombato, 1968). Protein samples were loaded and allowed to stack in the stacking gel for 50 min at 50 V. Once samples entered the resolving gel, the voltage was increased to 100 V, and the samples were run for an additional 120 min. The gels were subsequently stained with Coomassie stain (0.1 % [w/v] Coomassie R250; 10 % [v/v] glacial acetic acid; 40 % [v/v] methanol) overnight. Gels were then transferred to a de-stain solution (20 % [v/v] methanol, 10 % [v/v] glacial acetic) acid for 3 to 4 h until blue protein bands were visible against a clear background.

2.5 Functional Characterisation

2.5.1 Congo Red Assay

In order to qualitatively assay the ability of the purified MBP-cellulase to hydrolyse CMC-Na^+, a 48-well plate Congo Red assay was performed (Kim *et al*., 2008). This assay is adapted from the zymogram, which is a used to qualitatively assay an enzyme's ability to degrade its respective substrate that is co-polymerised in a gel based medium, such as a PAGE gel or agar slab (Cano-Ramírez *et al*., 2016). The enzymatic hydrolysis of the substrate can be visualised through selective staining (Congo Red; 0.1 % w/v) and de-staining (sodium chloride; 1 M). This technique is most often employed to investigate cellulolytic and xylanolytic activities (Eriksson and Pettersson, 1973). A substrate solution containing 1.5 % (w/v) agar and 1 % (w/v) CMC-Na^+ (pH 6) was aliquoted in volumes of 500 µL into a 48-well plate and allowed to set for 20 min at room temperature. Once set, 100 µL of each protein fraction obtained from the amylose affinity chromatography (section 2.3.4) was aliquoted onto the set agar and allowed to incubate at 50 ˚C for 2-h. The wells were subsequently flooded with 0.1 % (w/v) Congo Red solution and incubated at room temperature for 30 min. The excess Congo Red was poured off, and the wells were de-stained with 1 M sodium chloride until decolourised. Decolourization denotes CMC-Na^+ hydrolysis, since it represents areas where the substrate, CMC-Na^+, has been hydrolysed. (Ni *et al*., 2014; Schwarz *et al*., 1987). All reactions were performed in triplicate.

2.5.2 3,5-Dinitrosalicylic Acid (DNS) Activity Assay

The 3,5-dinitrosalicylic acid (DNS) activity assay determines the presence of a free carbonyl group, which is detected as reducing sugars. When subjected to high alkalinity, the aldehyde

functional group is oxidised to the corresponding acid, whilst DNS is concurrently reduced to 3-amino-5-nitrosalicylic acid (Miller, 1959).

2.5.2.1 Preparation of DNS Reagent and a Glucose Standard Curve

The reagent, DNS, was prepared by dissolving 2.18 g of DNS in 80 mL of 0.5 M sodium hydroxide (NaOH), which was heated and continually stirred at 70 ˚C. Once dissolved, 25 g of Rochelle salt (Potassium sodium tartrate tetrahydrate), 0.2 g of phenol and 0.05 g of sodium metabisulphite were added and dissolved. Once the solution reached room temperature, distilled water was added to a final volume of 100 mL. The solution was stored in the dark to avoid photo-oxidation of the sodium metabisulphite present within the reagent (Jain *et al*., 2020; Miller 1959). To identify the cellulolytic capabilities of MBP-cellulase, a standard curve representing absorbance at an optical density of 575 nm (OD_{575}) as a function of known glucose concentrations was constructed. Known concentrations of glucose (0.2; 0.4; 0.6; 0.8 and 1 mg/mL) were added to 3 mL of DNS reagent and the reaction mixture was incubated in rapidly boiling water for 10 min and cooled on ice until the colour stabilised. The glucose standards were then read spectrophotometrically at 575 nm (OD_{575}) after diluting the samples 1:10 in distilled water (Adney and Baker, 2008). Once constructed, an equation for the curve was derived, and the R^2 value was calculated using Microsoft Excel (Excel version 2016). All reactions were performed in triplicate. Although cellulase's enzyme activity is typically defined as the amount of enzyme that liberates 1 µmol glucose per minute, MBP-cellulase could not be defined as such due to the nature of the cellulase's fusion to the MBP tag. Specific activity could also not be obtained, due to inefficient cleavage of the MBP tag.

2.5.2.2 Assaying Soluble Protein Extracts for Cellulase Activity

In order to determine the cellulolytic capabilities of the recombinant MBP-cellulase, 250 µg of soluble protein extracts (which contain *E. coli*'s cytoplasmic proteins) from induced cultures of transformed *E. coli* cells were incubated in the presence of 0.5 % (w/v) CMC-Na^+ in 50 mM sodium phosphate buffer (pH 6) for 1 h at 50 ˚C. Glucose was omitted from the cultivation medium for the preparation of soluble protein extracts (section 2.3.3) that were to be assayed for reducing sugar content using the DNS activity assay. The addition of DNS reagent terminated the reaction, and the reactions were subsequently incubated for 10 min in rapidly boiling water. Once cooled on ice, the absorbance at an optical density of 575 nm

(OD_{575}) was obtained. The glucose amount (µg/mL) released from CMC-Na^+ hydrolysis was extrapolated from the constructed glucose standard curve (Luo *et al*., 2019). Soluble extracts from uninduced and untransformed cultures and resuspended pellets from all three conditions (induced empty vector, uninduced, untransformed) served as negative controls. Cellulase from *Trichoderma reesei* (30 units) (Sigma: C8546-2.5KU) served as a positive control.

2.5.2.3 Optimal pH, Temperature, NaCl and Metal-ion Determination

In order to be industrially applicable, cellulases need to have the intrinsic ability to withstand industrial conditions, which include higher temperatures, high or low pHs and comparatively high salt concentrations, as well as the presence of metal ions (Lin *et al*., 2016). Optimal parameters for CMC-Na^+ hydrolysis by the MBP-cellulase was determined by assaying 250 µg of soluble extracts in the presence of 0.5 % CMC-Na^+ at different temperatures, pH levels, sodium chloride concentrations and in the presence of varying divalent metal ions, such as calcium and manganese, for 1 h. For optimal pH determination, soluble protein extracts from induced cultures of *E. coli* T7 transformed with pMAL-CGB_E4160W were incubated in different buffers over a pH range of 3-8. All buffers had a 50 mM concentration of either sodium citrate (for pH 3-5) or sodium phosphate (for pH 6-8) and were incubated at 50 ˚C. All assays were carried out as previously described (section 2.5.2.2). The spectrophotometric absorbance value of the pH at which the cellulolytic activity was the highest was taken as 100 %. The remaining absorbance values from sub-optimal pHs were divided by the highest value and multiplied by 100 to compute percentage relative activity (Herlet *et al*., 2017). For optimal temperature determination, soluble protein extracts from induced cultures of *E. coli* T7 transformed with pMAL-CGB_E4160W were incubated at different temperatures over the range of 20-70˚C at a pH of 6. All reactions were carried out as previously described (section 2.5.2.2). The spectrophotometric absorbance value of the temperature at which the cellulolytic activity was the highest was taken as 100 %. The remaining absorbance values from sub-optimal reaction temperatures were divided by the highest value and multiplied by 100 to compute percentage relative activity. In order to determine the optimal NaCl concentration for CMC-Na^+ hydrolysis, soluble protein extracts from induced cultures of *E. coli* T7 transformed with pMAL-CGB_E4160W were incubated in the presence of increasing sodium chloride concentrations in the range of 0-1000 mM. All reactions were carried out as previously described (section 2.5.2.2). The cellulolytic absorbance value of the sodium chloride concentration at which the highest hydrolytic activity was taken as 100 %. The remaining

absorbance values from sub-optimal sodium chloride concentrations were divided by the highest value and multiplied by 100 to compute percentage relative activity. The effects of different metal-ions on cellulolytic activity were investigated by adding 5 mM of calcium (Ca^{2+}), cobalt (Co^{2+}), iron (Fe^{2+}), lead (Pb^{2+}), manganese (Mn^{2+}) and nickel (Ni^{2+}) to the assay, which had a pH of 6 and proceeded at 50 ˚C. All reactions were carried out as previously described (section 2.5.2.2). The spectrophotometric absorbance value of the metal-ion reaction at which the cellulolytic activity was the highest was taken as 100 %. The remaining absorbance values from sub-optimal metal-ions were divided by the highest value and multiplied by 100 to compute percentage relative activity.

Chapter Three: Results and Discussion

The aim of the study was to express, purify and characterize a recombinant cellulase derived from *Cryptococcus gattii* (WM276; VGI). The predicted effect of the maltose binding protein (MBP) tag on the cellulase's solubility, structure and physicochemical properties were also investigated.

The MBP-cellulase fusion was overexpressed and purified to homogeneity using amylose affinity chromatography. The purified MBP-cellulase's structural character was delineated using far-ultraviolet circular dichroism, tryptophan fluorescence and native polyacrylamide gel electrophoresis.

The cellulolytic ability of MBP-cellulase was confirmed using a Congo Red assay. The optimal parameters, temperature, pH, sodium chloride concentration and the effect of divalent metal ions, for MBP-cellulase's functionality were then investigated using the 3,5-dinitrosalicylic acid (DNS) assay.

3.1. Protein Bioinformatics

3.1.1 Identification of a Signal Peptide

A signal peptide which targets the cellulase for secretion from the cell, with a length of 17 amino acids (position 2-18) was identified (Table 3.1). Proteins with signal peptides are either targeted to the secretory pathway, targeted for membrane insertion, or targeted to a number of different organelles, such as the endoplasmic reticulum or Golgi apparatus (Nielsen, 2017). This particular signal peptide was subsequently omitted from the design and synthesis of the expression vector pMAL-CGB E4160W. This omission is due to the fact that signal peptides, particularly those derived from cellulases, have been shown to inhibit the synthesis of recombinant enzymes and impair the enzyme's function when overexpressed in *E. coli* (Hu *et al*., 2020). Wang and colleagues (2019) also overexpressed a basidiomycetous GH5 endoglucanase in *E. coli* in the absence of its 23 amino acid signal peptides in order to produce more soluble enzyme. It is a well understood practise that the elimination of the hydrophobic signal peptide of GH5 endoglucanases, results in a higher degree of soluble expression, without compromising enzyme structure and function (Gopal and Kumar, 2013). All subsequent bioinformatic analyses were performed with the signal peptide omitted.

Table 3.1: Prediction of a signal peptide for extracellular secretion of the cellulase using Predisi (predisi.de/), Phobius (phobius.sbc.su.se/) and signalP (cbs.dtu.dk/services/SignalP/), using the primary structure of the cellulase from *Cryptococcus gattii* (WM276) (Hiller *et al*., 2004; Kall *et al*., 2007).

Software	Signal Peptide Length (Amino Acid Residues)	Cleavage position	Probability
PrediSi	17	17-18	0.5476
Phobius	17	17-18	0.9999
SignalP	17	17-18	0.9934

3.1.2 Construction of an Automated Homology Model

An automated homology model of the MBP-cellulase was constructed (Figure 3.1) to computationally determine the impact of the MBP tag on the cellulase's structure, which could ultimately affect enzymatic function (Guo and Zhu, 2012). The MBP-cellulase's primary structure (Appendix 1, Figure A2) was used as the template, from which a three-dimensional structure was built using Phyre2's intensive mode (Kelley *et* al., 2015). The intensive mode makes use of *ab initio* modelling that is capable of using multiple templates to construct the homology model.

In total, 773 residues out of 814 (95 %) were modelled at > 90 % confidence (Kelley *et al*., 2015). The residues covering the MBP, 2-371 (45 % coverage of the MBP-cellulase) were modelled using the template, c6vlsD (Jiang *et al*., 2020). The template c6vlsD corresponds to a maltose/maltodextrin-binding periplasmic protein derived from *E. coli* K-12 and was modelled with 99 % identity to the template with a confidence score of 100 (Jiang *et al*., 2020). The residues covering the cellulase, 416-814 (48 % coverage of MBP-cellulase) were modelled using the template d1h4pa (Taylor *et al*., 2004). The template d1h4pa corresponds to an exo-1,3-beta glucanase overexpressed in *Saccharomyces cerevisiae* and was modelled with 43 % identity to the template with a confidence score of 100 (Taylor *et al*., 2004). The homology model was visualized using PyMOL (Rigsby and Parker, 2016; Figure 3.1).

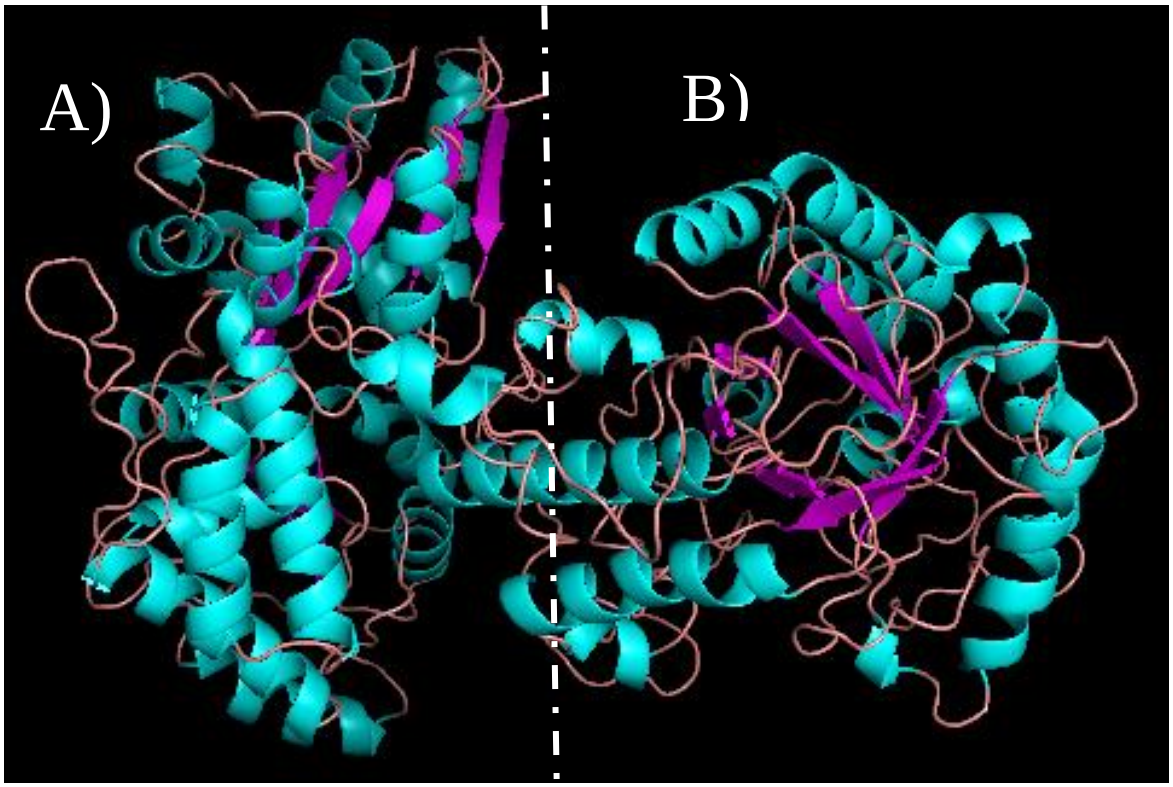

Figure 3.1: Homology model representing the maltose binding protein-cellulase (MBP-cellulase) fusion protein constructed using Phyre2 (Kelley *et al*., 2015) (http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index). A) represents the MBP tag modelled against c6vlsD with a confidence score of 100 (Jiang *et al*., 2020). B) represents the cryptococcal cellulase (WM276) modelled against d1h4pa with a confidence score of 100 (Taylor *et al*., 2004). The fused cellulase (B) still retains a typical $(\alpha/\beta)_8$ triose phosphate isomerase (TIM) barrel fold associated with glycosyl hydrolase family 5 (GH5) enzymes. The automated homology model of MBP-cellulase was visualized in PyMOL (Rigsby and Parker, 2016).

From Figure 3.1, it is evident that despite the presence of a large 43 kDa MBP fusion tag (Figure 1A), the passenger cellulase retains a $(\alpha/\beta)_8$ triose phosphate isomerase (TIM) barrel fold, which is typical of enzymes belonging the glycosyl hydrolase family 5 (GH5) (Wierenga, 2001). The TIM barrel fold is the most prevalent structure among enzymes, accounting for roughly 10 % of all proteins with deduced structures. (Orengo *et al*., 1994). Glycosidase enzymes are the most abundant type of enzymes with a TIM barrel fold, accounting for 50 % of all documented TIM barrels. (Sterner and Höcker, 2005). Members of GH5 have conserved catalytic residues, two glutamates, with one residue acting as the proton donor and the other residue acting as the nucleophile (Tseng *et al*., 2011). To carry out hydrolysis, all GH5 enzymes

use the classic Koshland double-displacement mechanism utilising the two catalytic glutamates (Yuan *et al*., 2020).

There have been several studies identifying endo-acting cellulases in cryptococcal species (Midiri *et al*., 2020; Thongekkaew *et al*., 2008). Botes and co-workers (2008) reported that 36 strains of *C. neoformans* tested positive for endoglucanase activity when cultured on carboxymethyl cellulose (CMC) agar and subjected to a Congo Red assay. Moreover, the gene encoding an endo-acting cellulase derived from, *C. neoformans* was overexpressed, purified, and functionally characterized by Midiri and colleagues (2020). They found that the lysate from transformed and induced cultures of *E. coli* BL21 (DE3) were readily able to hydrolyze CMC. Similarly, Thongekkaew and co-workers (2008) overexpressed, purified, and functionally characterized an enzyme originating from *Cryptococcus* sp. S-2 that was capable of hydrolysing CMC-sodium (CMC-Na^+) salt and CMC. However, no hydrolysis was evident when an insoluble crystalline substrate such as Avicel or cellulose were used. Such an observation, whereby only soluble substrates are hydrolyzed, indicates that the cellulase is likely endo-acting (Wang *et al*., 2019). Cellulases which exhibit high hydrolytic ability against insoluble substrates such as Avicel often denote that that the cellulase is exo-acting (Ma *et al*., 2019; Soares Júnior *et al*., 2013; Ventorino *et al*., 2016).

3.1.3 Theoretical Parameter Determination

The use of protein bioinformatics is an invaluable tool in terms of predicting numerous theoretical traits regarding proteins that are to be overexpressed (Rodríguez-Ruiz *et al*., 2019. The primary structure of the MBP-cellulase (Appendix 1, Figure A2) was used to determine the enzyme's isoelectric point (pI), theoretical molecular weight (in kDa), and molar extinction coefficient (Roy, 2007; Table 3.2). The amino acid sequence of a protein can also be used to probe the presence of certain conserved domains, secondary structures, secretion signal sequences and possible transmembrane topologies (Bakheet and Doig, 2009). As such, it is vital to ascertain these traits to optimize expression conditions as a means of enhancing MBP-cellulase's solubility.

The estimated molecular weight of the MBP-cellulase, a fusion of the MBP tag (43 kDa) and a cellulase from *C. gattii* (47 kDa), is approximately 90 kDa (Table 3.2). The outstanding 0.8 kDa is attributed to the remaining structural properties encoded by the pMAL-CGB_E4160W expression vector (Figure 2.1), such as a thrombin cleavage recognition site, a

factor Xa cleavage recognition site, a hexahistidine tag and a string of ten asparagine residues. The ten asparagine repeats act as a spacer, which functions to shield the MBP from its fusion partner, thereby increasing the probability of successful binding of the MBP fusion construct to the amylose resin during the purification process (New England Biolabs, 2018; www.neb.com).

Table 1.2: Predicted theoretical parameters of maltose binding protein-cellulase (MBP-cellulase) derived using Expasy ProtParam (Garg *et al*., 2016) (web.expasy.org/protparam/). The theoretical parameters were predicted based on the primary structure of MBP-cellulase (Appendix 1, Figure A2).

Theoretical Parameter	**Maltose binding protein-cellulase**
Number of amino acid residues	814
Estimated molecular weight	90.8 kDa
Isoelectric point	4.97
Extinction coefficient (reduced cysteines)	215090 M^{-1} cm^{-1}
Aliphatic Index	75.20
Instability index	25.24
Grand average of hydropathicity (GRAVY)	-.0477

The pI of a protein is the pH at which the protein carries no net charge (Kaur *et al*., 2020). The pI of the MBP-cellulase was calculated to be 4.97 (Table 3.2), which suggests that it is likely able to maintain its hydrolytic ability at lower pH values due to the presence of 103 negatively charged glutamate and aspartate residues. The MBP tag and cellulase have 51 negatively charged residues. The remaining glutamate residue forms part of the factor Xa cleavage site in the spacer region (Ganghadhar *et al*., 2016). The high prevalence of negatively charged residues have been suggested to contribute to some GH5 enzymes having a lower optimal pH (Dey and Roy, 2018). This trait of cryptococcal cellulases being acid-tolerant is not uncommon when considering that a recombinant cellulase derived from *C. neoformans* was found to have a pI of 4.56 (Midiri *et al*., 2020) and an optimal catalytic pH of 4. The presence of an acid-tolerant cellulase secreted by *C. gattii* is supported by the findings of Kidd and co-workers

(2007), who found that *C. gattii* was tolerant of acidic conditions, since *C. gattii* samples were mainly isolated from acidic soils (Deng and Tabatabai, 1994). This is further supported by the fact that previous studies have observed that several cryptococcal species preferentially colonize soil with an acidic pH (Gadanho *et al*., 2006; Vishniac, 2006).

The extinction coefficient of the MBP-cellulase, was determined to be 215090 M^{-1} cm^{-1} (Table 3.2) at a wavelength of 280 nm (Gasteiger, 2003). This value is attributed to the high proportion of aromatic residues (cysteine, tryptophan, and tyrosine) present within the amino acid sequence (Behbahani *et al*., 2016; Appendix 1, Figure A2). In the MBP-cellulase's primary structure, there are five cysteine residues (0.6%), 28 tryptophan residues (3.4%) and 41 tyrosine residues (5.0%). The MBP has no cysteine, eight tryptophan (28.6 %) and 15 tyrosine (36.9 %) residues, whilst the passenger cellulase has five cysteine (100 %), 20 tryptophan (71.4 %) and 26 tyrosine (63.4 %) residues. The computationally determined extinction coefficient enables the investigation of protein-protein and protein-ligand interactions in solution (Gill and Von Hippel, 1989; Mohan, 2012). Such interactions can be monitored through the utilization of intrinsic tryptophan fluorescence (Akbar *et al*., 2016; Vivian and Callis, 2001). Since the number of tryptophan residues is directly proportional to the fluorescence intensity, then a higher amount of tryptophan residues is directly associated with an increase in the intensity of the emission maximum (Sindrewicz *et al*., 2019).

The aliphatic index describes the proportion of the protein sequence, wherein specific non-polar residues contain aliphatic sidechains such as alanine, isoleucine, leucine, and valine (Enany, 2014). Ikai (1980) observed that the aliphatic index of a particular protein is often directly proportional to its predicted thermostability. High aliphatic indices are not an uncommon trait amongst fungal cellulases. *Aspergillus fumigatus* produces a cellulase which had an almost identical aliphatic index (75.05) compared to that of MBP-cellulase (75.20), however its thermostability was not investigated experimentally (Ahmed *et al*., 2013). In addition, a thermotolerant and halotolerant GH5 endoglucanase derived from the metagenome of a hot spring, had its thermostability partially attributed its high aliphatic index (86.64) (Joshi *et al*., 2021).

The theoretical instability index of a particular protein is a parameter by which one can predict both *in vitro* and *in vivo* stability of a given protein (Kaur *et al*., 2020). The predicted instability index of the MBP-cellulase is 25.24 (Table 3.2) and is, therefore, considered to be experimentally stable, as its value is < 40 (Käll *et al*., 2004). Proteins whose instability index

exceeds 40 are considered experimentally unstable. A study undertaken by Sanjaya and co-workers (2021) aimed to characterize 27 GH5 family microbial cellulases from uncultured microorganisms using *in silico* methods. Of the 27 cellulases computationally analysed, only six had instability indices above 40. Furthermore, only 26 out of 27 cellulases had a pI below 7 and all 27 cellulases had an aliphatic index in the range of 62-84, emphasizing the thermotolerant nature of GH5 cellulases.

Lastly, the grand average of hydropathicity (GRAVY) score is a predictor of hydrophilicity versus hydrophobicity (Kaur, 2020). Its value lies in the range of -2 to 2, with negative values denoting hydrophilicity and positive values denoting hydrophobicity (Kyte and Doolittle, 1982). Increased hydrophilicity is often associated with globular proteins and enhanced solubility, whilst increased hydrophobicity is associated with membrane proteins that are insoluble (Wilkins *et al*., 1998) The MBP-cellulase's GRAVY score is negative, -.0477 (Table 3.2), which indicates that it is relatively hydrophilic. Interestingly, the GRAVY value of the unfused cellulase without the signal peptide is -0.501, whilst inclusion of the signal peptide results in a more positive value of -0.417, which suggests that the signal peptide makes the enzyme less soluble upon overexpression in *E. coli*. Grand average of hydropathicity scores amongst other GH5 family cellulases are typically negative, which suggests that the cellulases from this family are hydrophilic and likely soluble upon overexpression (Sanjaya *et al*., 2021).

3.1.4 Prediction of MBP Fusion on Cellulase Solubility

The effect of the MBP fusion tag on cellulase solubility upon overexpression in *E. coli* was computationally determined using SoluProt (Hon *et al*., 2021). SoluProt has been shown to achieve higher rates of accuracy (58.5 %) when predicting *in vivo* solubility when compared to other, well-known protein solubility predictive software, such as PROSO II (58.0%) and CamSol (54.1%) (Almagro Armenteros *et* al., 2019; Smialowski *et al*., 2012; Sormanni *et al*., 2015).

The primary structure of the putative cryptococcal cellulase (without the signal peptide) was used as the first input, whilst the primary structure of the cellulase with the N-terminal MBP was used as the second input (Table 3.3).

Soluprot revealed a marked increase in the predicted solubility of the cellulase when fused to the MBP (Table 3.3) with values closer to 1.0 indicating increased solubility (Kapust and

Waugh, 1999). This is in line with previous studies that have shown that the MBP fusion tag is associated with increased solubility of its passenger protein when overexpressed in *E. coli* (Sluchanko *et al*., 2016).

Table 3.3: Predicted solubility, upon overexpression in *Escherichia coli*, of the putative cellulase derived from *Cryptococcus gattii* (WM276). Theoretical solubilities were obtained via SoluProt (Hon *et al*., 2021) (loschmidt.chemi.muni.cz/soluprot/) using the primary structure of the maltose binding protein-cellulase (MBP- cellulase) fusion protein or the native unfused cellulase.

Fusion State	**Solubility Score as determined via SoluProt**
Cellulase fused to N-terminal MBP	0.949
Unfused Cellulase	0.669

Oftentimes, the consequence of heterologous expression of eukaryotic proteins in *E. coli* is the insolubility of the target protein and the formation of inclusion bodies, which are non-functional aggregates of misfolded proteins (Singh *et al*., 2015). This phenomenon is most often as a result of codon bias, which differs greatly across organisms (McNulty *et al*., 2003). Codon bias arises as a consequence of the genetic code's redundancy, which includes 64 distinct codons that encode 20 amino acids in addition to three stop codons (Xu *et al*., 2021). The fact that distinct codons are preferentially used in particular coding sequences provides the foundation for codon bias, which is ubiquitous across all species (Sharp and Li, 1986). Codon bias is another method of influencing protein production since highly expressed genes have more frequently utilised codons, which translates to a greater abundance of their corresponding transfer RNA (tRNA) (Ikemura, 1981; Krüger and Sørensen, 1998). Similarly, low-expression genes have a greater proportion of comparably uncommon codons, which delays protein synthesis since there are considerably fewer matching tRNAs available (Krüger and Sørensen, 1998; Tuller *et al*., 2010). The decrease in soluble protein production and the formation of insoluble aggregates is often attributed to the difference in synonymous codon bias between the expression system (in this case, *E. coli*) and the natural host from which the gene is derived (in this case, *C. gattii*) (Angov *et al*., 2008).

3.2 Heterologous Protein Expression

The 90 kDa MBP-cellulase, encoded on the pMAL-CGB_E4160W vector, was successfully overexpressed in *E. coli* T7 at varying temperatures using a consistent concentration (0.2 mM) of the inducer, isopropyl β-D-1-thiogalactopyranoside (IPTG). Optimal expression of soluble MBP-cellulase was achieved at lower temperatures (16 – 20 ˚C) and longer induction times (20 – 24 h) (Figure 3.2).

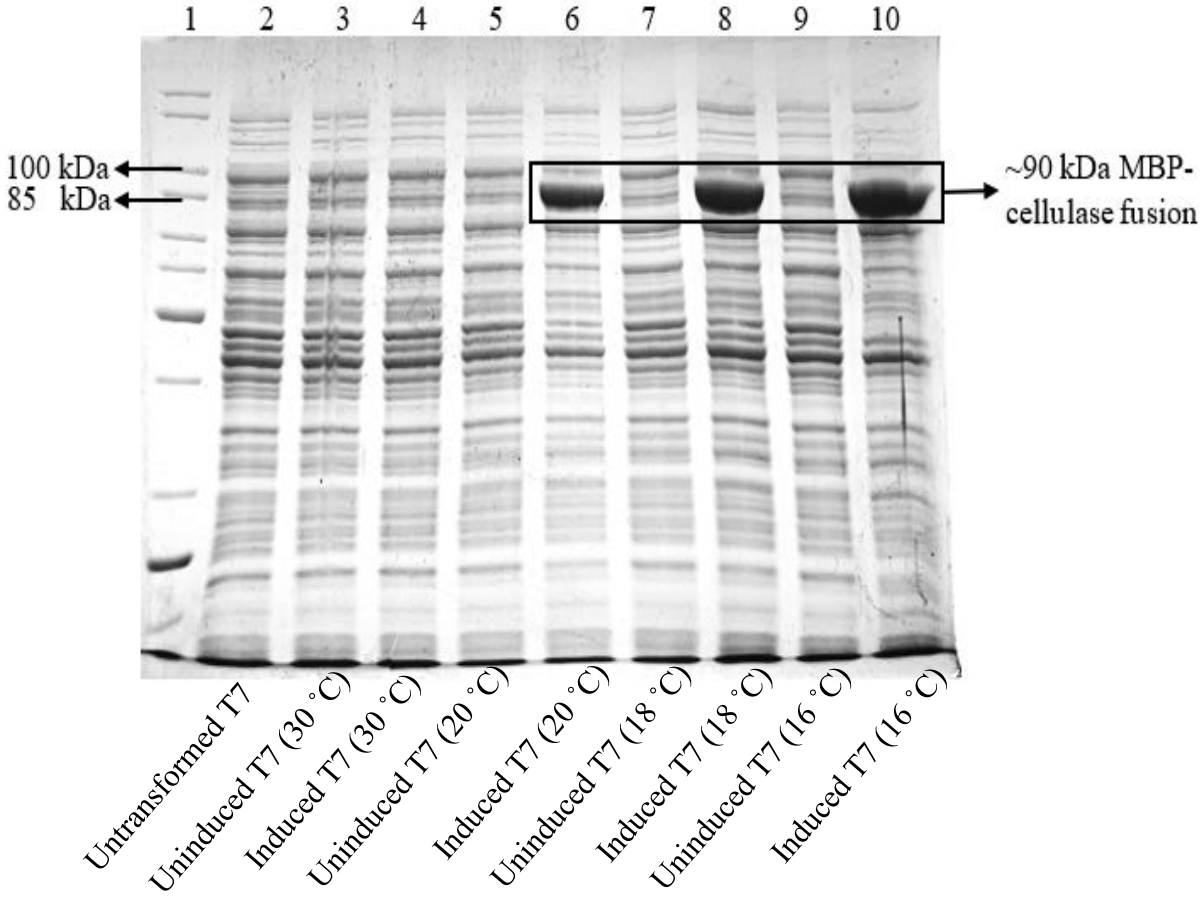

Figure 3.2: Coomassie stainedsodium dodecyl sulphate-polyacrylamide (SDS-PAGE) gel illustrating heterologous overexpression trials of maltose binding protein-cellulase (MBP-cellulase) in *Escherichia coli T7* cells transformed with pMAL-CGB_E4160W, as a function of decreasing temperature. All *E. coli T7* cultures were induced with 0.2 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG). Lane 1: Unstained Protein Standard (New England Biolabs); lanes 2-10: expression conditions at different temperatures (16 – 30 ˚C).

A common consequence with respect to heterologous expression of eukaryotic proteins in *E. coli* is protein misfolding, which in conjunction with insufficient processing by molecular chaperones, leads to the formation of insoluble inclusion bodies (Sørensen and Mortensen,

2005). Although solubilization and subsequent refolding of inclusion bodies is possible, the consequent loss of secondary structure often results in a significant loss of protein bioactivity (Datar *et al*., 1993). A study undertaken by Finley and co-workers (2004) found that only 228 out of 2078 genes from *Caenorhabditis elegans* were soluble when expressed in *E. coli*. As discussed previously, such an observation is most likely attributed to the discrepancy in codon usage bias between the expression host and the host from which the gene was derived (Gustafsson *et al*., 2004).

There are a number of strategies that are routinely employed as a means of circumventing inclusion body formation. One such strategy, which is often a relatively simple parameter to manipulate, is induction temperature (Baldwin, 1986; Vera *et al*., 2007). A decrease in induction temperature directly correlates to a decrease in the rate of protein synthesis, which often gives the recombinant protein time to correctly fold, thereby preventing aggregation (Schellman, 1997). Higher induction temperatures were also found to decrease recombinant protein expression due to the increased mRNA degradation associated with higher temperatures (Kennel, 1986). These findings support the observation that the MBP-cellulase was only expressed in the soluble fraction when induction temperatures were relatively low (≤ 20 ˚C).

Another factor which is relatively easy to manipulate is the concentration of the inducing agent, which in this case is IPTG (Margawati *et al*., 2017). Increasing concentrations of IPTG have a similar effect to that of increasing temperature, resulting in an overall increase in the rate of protein synthesis (Silaban *et al.*, 2019). Rizkia and colleagues (2015) determined that, in their case, increasing the IPTG concentrations beyond 0.01 mM resulted in a higher proportion of the target recombinant protein localizing in the insoluble fraction, as opposed to the soluble fraction of the supernatant. Additionally, greater IPTG concentrations were observed to inhibit *E. coli* growth owing to an increased metabolic load, which would have a detrimental effect on total protein synthesis (Nazarali *et al*., 2015). The choice of 0.2 mM IPTG for the expression of MBP-cellulase is relatively low enough to not inflict too significant of a metabolic load on *E. coli*, yet at the same time permit the correct folding of MBP-cellulase.

The use of a codon-optimized gene sequence may have also aided the correct folding of the MBP-cellulase. Due to the discrepancy in specific codon usage frequencies between the native host (in this case, *C. gattii*) and the expression host (*E. coli*), the presence of too many rare codons results in translational pausing, which is often proceeded by dissociation of the transcribed mRNA from the ribosomal machinery (Yang *et al*., 2019). Curran (1995)

determined that the rare codon, 'CGA', encoding an arginine residue, leads to ribosomal stalling in *E. coli* and results in premature translational termination, yielding truncated polypeptides. According to GenScript, the native cellulase DNA sequence from *C. gattii* had a codon adaption index (CAI) of 0.42 (without the signal peptide), but the optimised cellulase DNA sequence had a CAI of 0.96. While a CAI of 1.0 is regarded as desirable for expression, values greater than 0.8 are deemed adequate for high expression in *E. coli*. There was a total of 20 low-frequency (< 30%) codons in the original cellulase sequence, which was projected to decrease translational efficiency and perhaps disengage translation owing to ribosomal stalling (GenScript, 2020, www.genescript.com).

3.3 Amylose Affinity Purification

A final tactic for the circumvention of inclusion body formation, and enhancement of protein solubility, lies in the use of fusion tags such as glutathione-S-transferase (GST), an MBP or thioredoxin tag. When Kapust and Waugh (1999) interrogated the ability of the aforementioned tags to enhance the solubility of six frequently insoluble passenger proteins, namely, chloramphenicol acetyl transferaseΔ9 (CATΔ9), early protein 6 (E6), green fluorescent protein (GFP), p16, tissue inhibitor of metalloprotease (TIMP) and tobacco etch virus (TEV) protease, the MBP tag outperformed the GST and thioredoxin tags in both solubility enhancement and final yield.

Although large in molecular weight, the MBP does not seem to impede or deter its passenger protein's enzymatic activity (Momin *et al*., 2019). An endo-acting xylanase belonging to GH5 was expressed as a translational fusion with an N-terminal MBP and was found to possess significant catalytic activity against birch wood xylan (Mitreva-Dautova *et al*., 2006). Braithwaite and co-workers (1995) fused an MBP to an endo-β-1,4-mannanase originating from *Pseudomonas fluorescens*, and highlighted, that irrespective of MBP's large molecular weight, the MBP-mannanase fusion protein was nonetheless able to hydrolyse its substrate. Jindou and colleagues (2006) overexpressed a GH9 cellulase originating from *Acetivibrio cellulolyticus* as a fusion with an N-terminal MBP. The researchers found that this fusion protein was able to effectively hydrolyse CMC, despite the presence of the MBP tag.

A single-step purification of MBP-cellulase utilising amylose resin in conjunction with gravity-flow chromatography was utilized to purify the fusion protein (Figure 3.3).

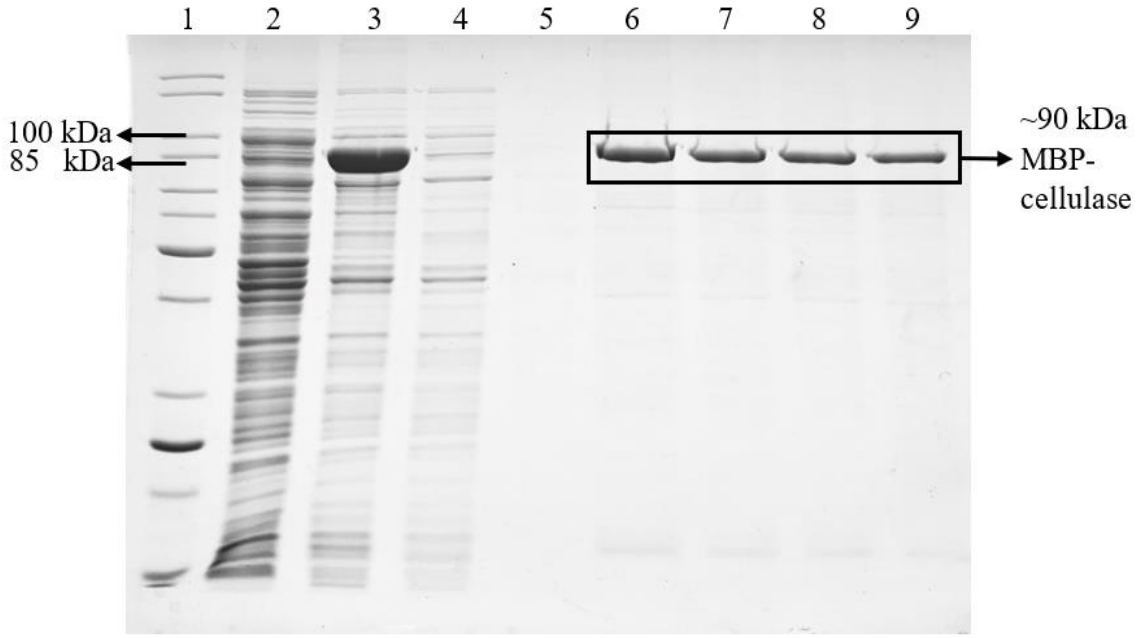

Figure 3.3: Coomassie stained sodium dodecyl sulphate-polyacrylamide (SDS-PAGE) gel illustrating single-step purification of maltose binding protein-cellulase (MBP-cellulase) fusion protein from soluble extracts of induced *Escherichia coli* T7 cells using amylose affinity chromatography. Lane 1: unstained Protein Standard (New England Biolabs); lane 2: soluble extracts from untransformed *E. coli* T7 cells; lane 3: soluble extracts from induced *E. coli* T7 cells; lane 4: amylose column flow-through; lane 5: amylose column wash fraction; lanes 6-9: isocratic elutions of purified MBP-cellulase fusion protein.

Amylose affinity purification relies on the interactions between MBP and α-(1-4)- glycosidic linkages, such as those found in amylose and starch polymers (Dippel and Boos, 2005). Due to high affinity binding of maltose to MBP, maltose can be used to displace MBP fusion enzymes from the amylose matrix (Szmelcman *et al*., 1976).

Despite the MBP's potent solubilizing effect on its fusion partner, there is an almost 30 % failure rate with regards to the proper binding of MBP to the amylose matrix (Pattenden and Thomas, 2008). These failures are most often attributed to C-terminal fusion of the MBP tag as opposed to N-terminal fusion, as C-terminal MBP tags tends to block the ligand binding site of MBP (Sachdev and Chirgwin, 1998). Another high attributor to binding inefficiency, is the use of non-ionic detergents such as Triton X-100 and Tween 20, which are traditionally used during protein purification as they aid cell lysis and prevent non-specific binding of contaminating proteins (Tan and Ting, 2000). These non-ionic detergents tend to have a negative impact on hydrophobic interactions which necessitate proper binding of the MBP onto

the amylose resin, which is why they were excluded from the binding buffer (Pattenden and Thomas, 2008).

The pMAL-c5X vector, like other next generation pMAL vectors, contains an asparagine spacer region between the MBP and protease cleavage sites. This spacer region is thought to result in a tighter binding of the MBP to the amylose matrix, which may contribute to the success of MBP-cellulase purification using amylose affinity chromatography (NEB, 2018).

Cleavage of the cryptococcal cellulase from the MBP tag using thrombin was not experimentally effective (Appendix 1; Figure A3), since there were still significant amounts of MBP-cellulase, which remained even when a buffer for optimal thrombin cleavage (20 mM Tris–HCl; pH 8.4; 150 mM NaCl; 25 mM $CaCl_2$) was used (Liu *et al*., 2008). The existence of residual fusion enzyme is a frequent outcome of protease cleavage (Frey and Görlich, 2014). In addition, the increased temperatures, and times necessary for optimal cleavage (24 h at room temperature) could accelerate the breakdown of MBP-cellulase or irreversibly inactivate it (Frey and Görlich, 2014). Protease-mediated cleavage of fusion proteins is very often inefficient, expensive and can degrade the protein of interest due to non-specific cleavage (Jenny *et al*., 2003). Furthermore, cleavage of larger affinity tags, such as MBP, is known to destabilise the passenger protein and induce a permanent loss of enzymatic function (Momin *et al*., 2019; Zhang *et al*., 2020). As such, the MBP tag was not cleaved from the passenger cellulase due to both inefficient cleavage, potential loss of cellulolytic activity after the cleavage reaction proceeded and the decrease in heterologous protein solubility (Raran-Kurussi and Waugh, 2012).

3.4 Far-Ultraviolet Circular Dichroism

Far-ultraviolet circular dichroism (far-UV CD) was employed as a means of characterizing the MBP-cellulase's secondary structure (Figure 3.4). The use of far-UV CD in the elucidation of a protein's secondary structure gives vital information regarding whether or not the protein is folded correctly and in the correct conformation (Greenfield, 2006). Proteins which occur as aggregates tend to exhibit light scattering and a flattened CD spectra when compared to correctly folded proteins (Venyaminov and Yang, 1996).

As shown in Figure 3.4, MBP-cellulase exhibited a CD maximum at 190 nm and two CD minima at 208 nm and, to a lesser extent 218 nm. This suggests that the MBP-cellulase fusion enzyme is primarily α-helical, with the presence of some β-sheets. The MBP component of the

fusion protein is dominantly α-helical (Gilardi *et al.*, 1997). This, in conjunction with the eight conserved α-helices of the canonical TIM barrel fold (Figure 3.1B), may have contributed to the observed maximum at 190 nm and minimum at 208 nm. The observed peaks are also consistent with that of Pimentel and colleagues (2017), who discovered that a GH5 endoglucanase generated from a metagenome library, had a CD maximum at 190 nm and two CD minima at 209 and 219 nm. Additionally, a second GH5 endoglucanase isolated from a sugarcane soil metagenome had a very comparable CD spectrum in its native state with a maximum at approximately 195 nm and two minima at 208 and 217nm, (Alvarez *et al.*, 2013). Additionally, a GH5 cellulase from the fungus *Volvariella volvacea* displayed two distinct CD minima at 210 and 222 nm, which is a hallmark of proteins that exhibit both α-helix and β-sheet properties (Zheng and Ding, 2013).

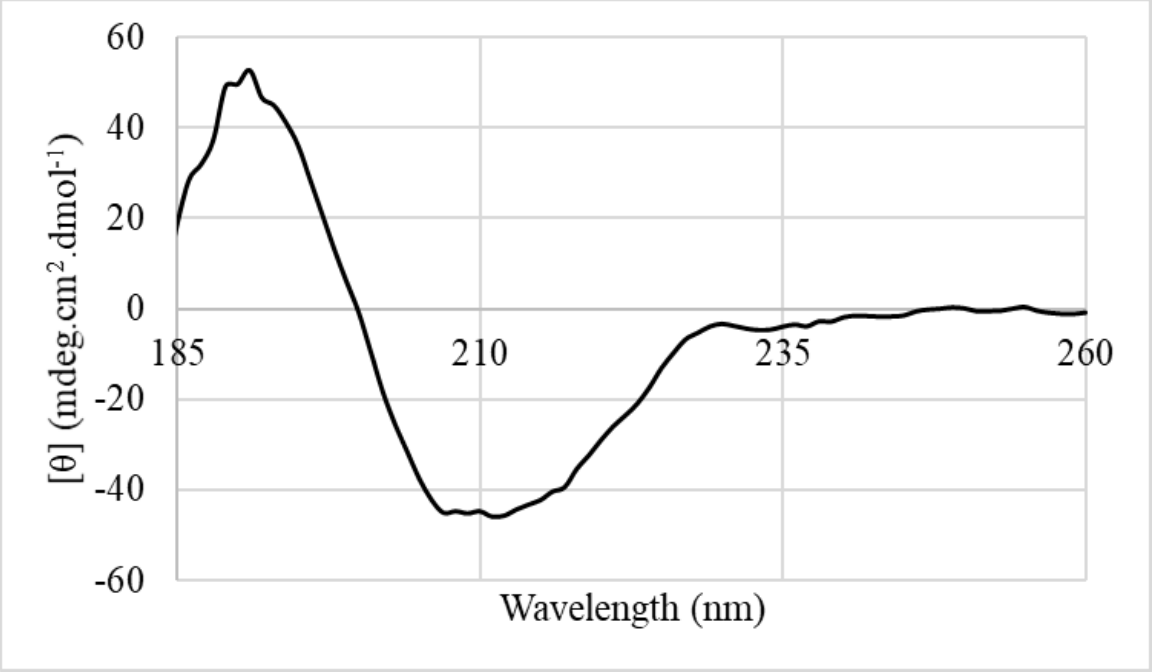

Figure 3.4: Maltose binding protein-cellulase (MBP-cellulase; 10 μM) in 10 mM sodium phosphate buffer (pH 7) exhibits a dominant α-helical secondary structure due to the presence of a distinct peak at 190 nm and decline at 208 nm. A slight decline at 218 nm is indicative of β-sheets (Corrêa and Ramos, 2009).

3.5 Intrinsic Tryptophan Fluorescence

The tertiary structure of MBP-cellulase, was evaluated through the use of intrinsic tryptophan fluorescence (Figure 3.5). The analysis of a protein's tertiary structure using intrinsic tryptophan fluorescence provides valuable insights into a protein's conformation when its local environment is altered, for example in the presence of a ligand whose binding causes a perturbation in the protein's structure (Ghisaidoobe and Chung, 2014; Souza *et al*., 2016).

The majority of proteins exhibit a maximum emission at ~350 nm which is due to the interaction of tryptophan's benzene ring with the aqueous environment (Teale and Weber, 1957). Excitation at 295 nm is used in order to prevent background noise, generated from other aromatic amino acids, such as tyrosine (Möller and Denicola, 2002), and to a lesser extent phenylalanine (Protein Fluorescence, 2006). Phenylalanine has a lower quantum yield compared to both tyrosine and tryptophan of approximately 0.02 in water, thus emission from this particular residue is largely avoided (Chen, 1967). On the other hand, the quantum yield of tyrosine and tryptophan are 0.2 and 0.14 in water, respectively (Schmid, 2001; Teale and Weber, 1957).

Based on its primary structure (Appendix 1; Figure A2), the MBP-cellulase contains 28 tryptophan residues which comprises 3.4 % of the total amino acid sequence (814 amino acids). Eight of the tryptophan residues are contributed by the MBP tag, whilst the passenger cellulase contains the remaining 20 tryptophan residues. Conformational changes were induced in the presence of MBP-cellulase's ligand, CMC-Na^+, and in the presence of a high concentration of a denaturant, in this instance urea, to induce unfolding (Ghisaidoobe and Chung, 2014; Povarova *et al*., 2010).

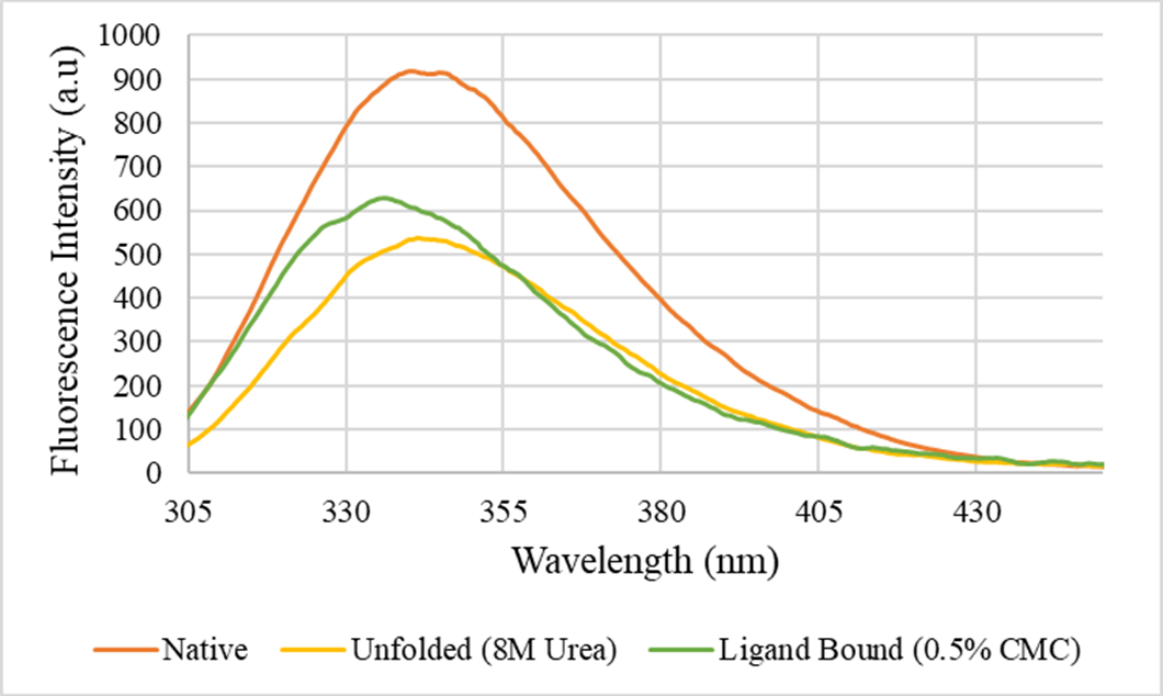

Figure 3.5: Investigation of the tertiary structure of maltose binding protein-cellulase (MBP-cellulase) with intrinsic tryptophan fluorescence. Native MBP-cellulase (10 µM; orange line) in sodium phosphate buffer (10 mM; pH 7) exhibited a high tryptophan fluorescent intensity emission at 340 nm. Denatured MBP-cellulase (10 µM; yellow line) in sodium phosphate buffer (10 mM, pH 7) with 8 M urea exhibited a low tryptophan fluorescent intensity emission and a red shift to 342 nm. Ligand binding with 0.5 % (w/v) carboxymethylcellulose (CMC; green line) resulted in a blue shift (emission at 335 nm) and lower fluorescent intensity which indicates increased hydrophobicity when compared to the native state (orange line).

Like other GH5 endoglucanase, in its native state the MBP-cellulase exhibited an emission maximum at 340 nm after initial excitation at 295 nm (Figure 3.5, orange line) (Sharma and Guptasarma 2017; Souza and colleagues 2016). It should be mentioned that the eight tryptophan residues from the MBP tag may have caused fluorescent energy transfer between other tryptophan residues via a histidine residue (Ghisaidoobe and Chung, 2014). This has been routinely observed in cases where multiple tryptophan residues are present in a particular protein, however this would only have an impact on the fluorescence intensity and not the emission maximum value (Ghisaidoobe and Chung, 2014).

When denatured, MBP-cellulase produced a red-shift (increase in emission maximum) from 340 to 342 nm, which although minor, still represents a change in MBP-cellulase's conformation, and a drop in the florescence intensity from 918 to 535 absorbance units (A.U) (Figure 3.5, yellow line). A red-shift in the presence of 8 M urea indicates that the tryptophan

residues are substantially exposed to the buffer environment, implying that MBP-cellulase has been entirely unfolded (Balsera *et al*., 2004).

Finally, a blue-shift (decrease in emission maximum) and a decrease in florescence intensity occurred as a consequence of the MBP-cellulase's binding to a ligand, CMC-Na^+, (Figure 3.5, green line). Such an observation is indicative of an increase in hydrophobicity, which suggests that the tryptophan residues are less exposed to the buffer environment and become more buried within MBP-cellulase's core (Dusa *et al*., 2006). This observation is unsurprising, when considering that a change in the emission spectra of a particular protein is often linked to a conformational change due to pH shifts, changes in temperature and ligand binding (Möller and Denicola, 2002).

3.6 Native Polyacrylamide Gel Electrophoresis

The utilisation of native polyacrylamide gel electrophoresis (PAGE) is a well understood technique which is employed in the determination of a protein's quaternary structure (Penchala *et al*., 2013; Tagoe *et al*., 2007). However, this technique is more suitable for cytosolic or secreted proteins, such as MBP-cellulase, as opposed to membrane proteins (Pollock *et al*., 2019). This is primarily attributed to the requirement for detergents that membrane proteins need in order to aide their solubilisation, which may unintentionally disrupt non-covalent interactions between protein subunits and interfere with accurate quaternary structure determination (Crichton *et al*., 2013). Overall, native PAGE offers an attractive and non-disruptive manner of describing the quaternary conformation of limited amounts of a particular protein while also producing results comparable to more costing techniques such as size-exclusion chromatography (SEC) (Zhang et al; Pollock *et al*., 2019).

From Figure 3.6, it appears that MBP-cellulase exists in a monomeric configuration, with a molecular weight of approximately 90 kDa. While the MBP fusion tag is a known monomer (Reuten *et al*., 2016) research has shown that the tag can conform to its passenger protein's native oligomeric state (Coscia *et al*., 2016). Coscia and co-workers (2016) observed that MBP existed as a dodecameric fusion when fused to glutamine synthase, which is dodecameric in nature. Therefore, it can be ruled out that the MBP tag would have influenced the native oligomeric state of the passenger cellulase.

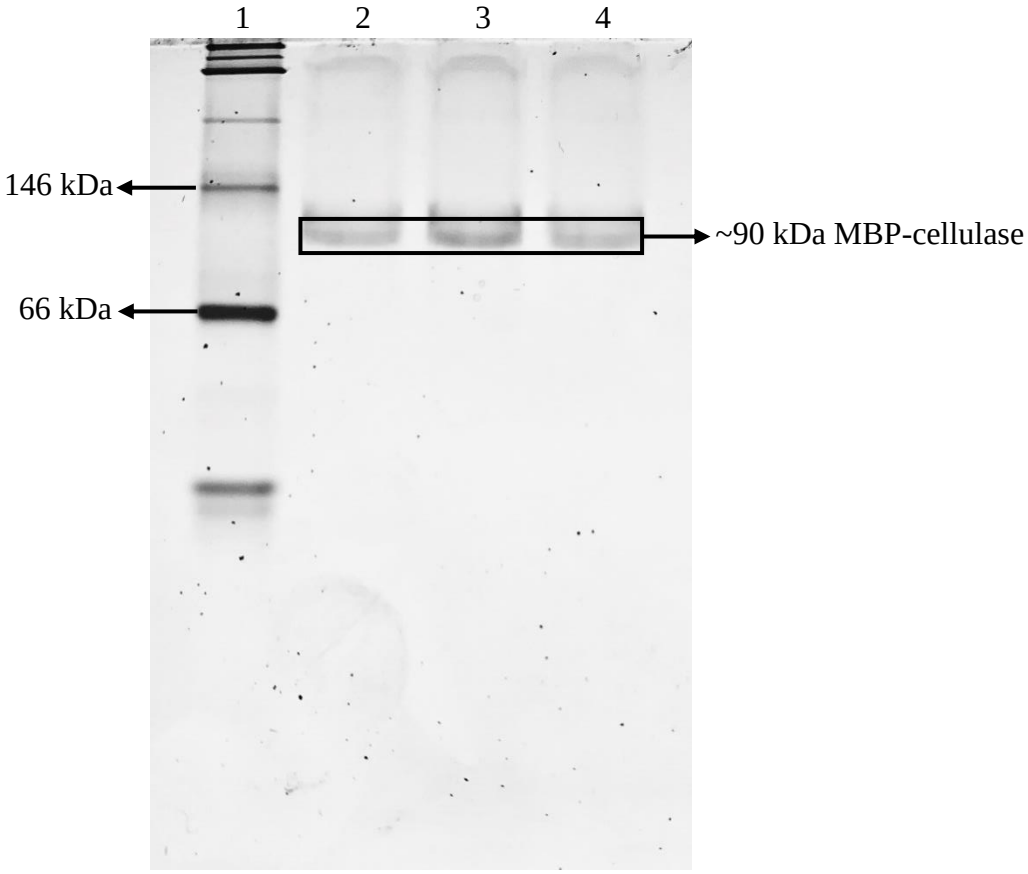

Figure 3.6: Native polyacrylamide gel electrophoresis (PAGE) of maltose binding protein-cellulase (MBP-cellulase) fusion protein purified using amylose affinity chromatography. The MBP-cellulase fusion protein appears to exist primarily as a monomer, since the native molecular weight corresponds to the denatured molecular weight, which denotes the absence of oligomerization. Lane 1: NativeMark™ Unstained Protein Standard (Thermo Fisher Scientific; MA, USA); lanes 2-4: isocratic elutions of purified MBP-cellulase fusion protein.

Cellulase enzymes, particularly those belonging to GH5, have been shown to be functional in a monomeric state (Badieyan *et al*., 2011) and are thus regarded as monomeric proteins (Zarafeta and Zheng *et al*., 2018. Alvarez *et al*., 2013; Lafond *et al*., 2016). There are, however, rare cases where certain cellulases from GH5 have displayed activity as oligomers. A GH5 cellulase, originating from the gut of an earthworm (*Aporrectodea caliginosa*), existed as an 85 kDa homodimer with a temperature optimum of 55 ˚C and a pH optimum of 9.5.

3.7 Congo Red Assay

In order to qualitatively assay and confirm the ability of MBP-cellulase to hydrolyse CMC-Na^+, all of the fractions acquired during amylose affinity chromatography were subjected to a

Congo Red assay (Carder, 1986). Classically, Congo Red is used in plating assays on agar plates as a method of screening for microorganisms which possess cellulolytic activity (Carder, 1986). The assay is based on the chemical interaction between the dye, Congo Red, and intact β-(1-4)-bound-D-glucopyranosyl units within CMC-Na^+ (unhydrolyzed substrate) (Teather and Wood, 1982). Thus, one may see the degradation of CMC-Na^+ by observing the emergence of zones of clearing, or more precisely, by observing the dye being washed away as CMC is hydrolysed (Sharma and Guptasarma, 2017).

Congo Red is not confined to whole cell assays that rely on the secretion and subsequent diffusion of active cellulase into the surrounding environment (Carder, 1986; Liang *et al*., 2014). Azizi and co-workers (2015) utilized CMC-containing plates to qualitatively assay cell-free extracts (supernatant) for cellulolytic activity. In addition, Valencia Jiménez and colleagues (2014) used culture supernatant from a strain of *Pichia pastoris* overexpressing an endoglucanase as a method of detecting CMC digestion using Congo Red. As such, based on a liquid Congo Red assay (Carder, 1986) the fractions used to assay for cellulolytic ability were the soluble extracts from an untransformed *E. coli* T7 negative control, soluble extracts from a transformed and induced *E. coli* T7 culture harbouring pMAL-CGB_E4160W, column flow-through, column wash, and four isocratic elutions containing purified MBP-cellulase (Figure 3.7).

The results, suggest that MBP-cellulase is capable of CMC-Na^+ digestion (Figure 3.7, lanes 5-8), and that endogenous intracellular proteins from *E. coli* T7 cells (Figure 3.7, lanes 1 and 3) are not capable of utilizing CMC-Na^+ as a substrate. Similarly, the ability of the cryptococcal cellulase to digest CMC-Na+ suggests that it is endo-acting (Farrow and Arnold, 2011).

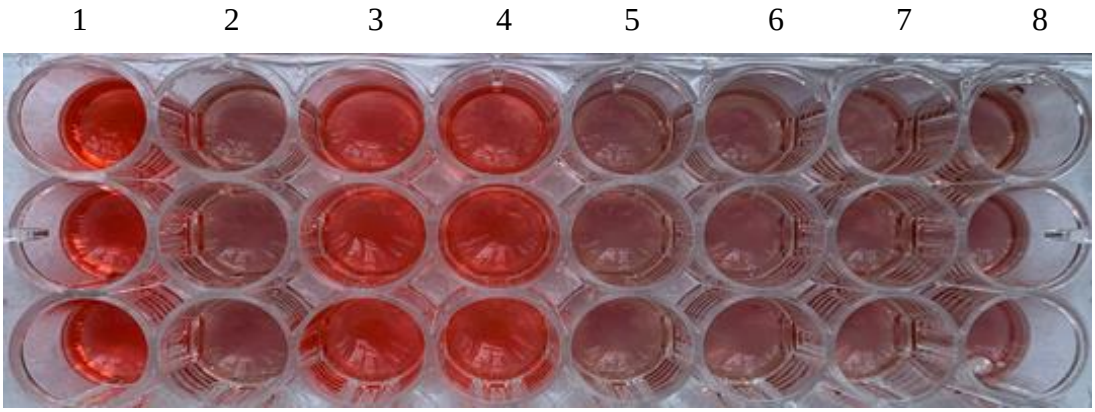

Figure 3.7: Congo Red assay to qualitatively determine the ability of maltose binding protein-cellulase (MBP-cellulase) to hydrolyse the substrate, carboxymethylcellulose sodium salt (CMC-Na^+; 0.5% w/v). Cellulolytic ability is based on decolourisation from red to grey-brown. Column 1: soluble extracts from untransformed *Escherichia coli* T7 cells; column 2: soluble extracts from induced *E. coli* T7 cells; column 3: amylose column flow-through; column 4: amylose column wash fraction; columns 5-8: isocratic elutions of purified MBP-cellulase fusion protein. The assay was performed in triplicate.

3.8 3,5-Dinitrosalicylic Acid (DNS) Assay

3.8.1 Assaying Soluble Extracts for Cellulase Activity

The 3,5-ninitrosalicylic acid (DNS) assay is a well-known technique with regards to detecting reducing sugars (Gusakov et al., 2011; Luo et al., 2019). Developed by Sumner and Graham (1921) and later refined by Miller (1959), the DNS assay is based on the reduction of the aromatic compound, 3,5-dinitrosalicylic acid, into 3-amino-5-nitrosalicylic acid by the reducing sugars generated by CMC-Na^+ hydrolysis (Marsden et al., 2007). With the help of a glucose standard curve representing the absorbance of known concentrations of glucose at 575 nm (Figure 3.8) one can estimate the amount of reducing sugars resulting from the hydrolysis of the substrate.

In order to confirm MBP-cellulase's cellulolytic activity and eliminate the possibility of MBP or *E. coli*'s endogenous proteins contributing to CMC-Na^+ hydrolysis, the lysates from four different *E. coli* T7 strains were subjected to the 3,5-dinitrosalicylic acid (DNS) activity assay (Figure 3.9).

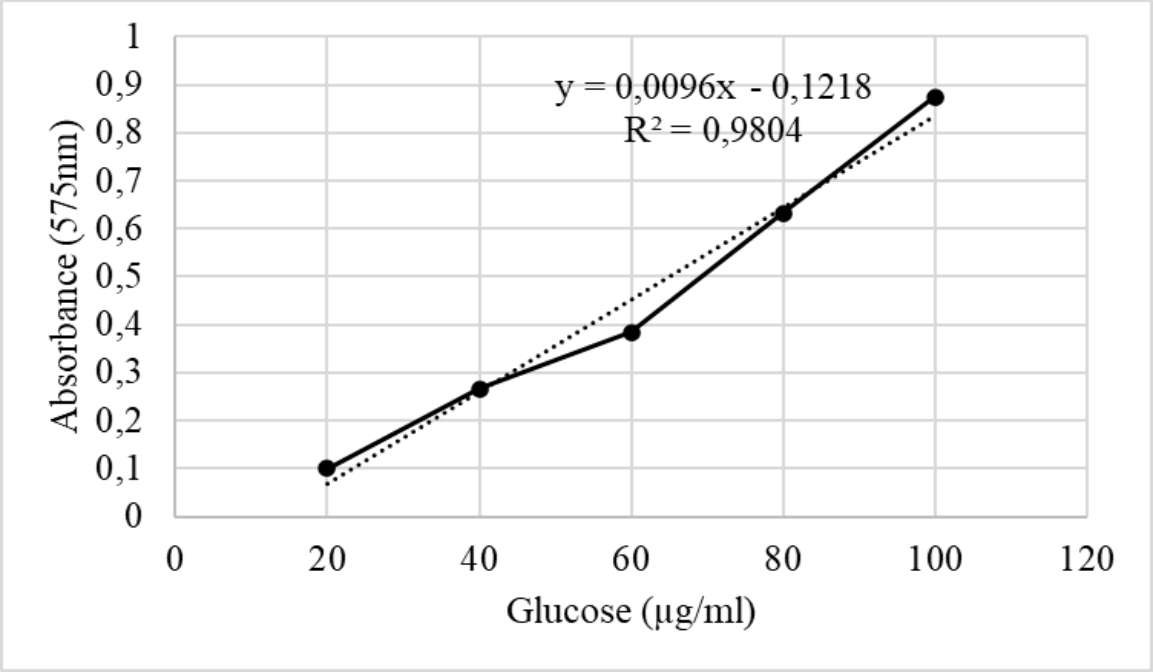

Figure 3.8: Glucose standard curve representing the absorbance of known concentrations of glucose at 575 nm. The derived equation, y = 0.0096(x) – 0.1218, was utilized as a means of calculating the amount of glucose released, based on the absorbance of enzymatic reactions after the addition of 3,5-dinitrosalicylic acid reagent and rapid boiling for 10 minutes. An R^2 value of 0.9804 indicates a direct positive correlation between the concentration of glucose and the absorbance at 575 nm.

Figure 3.9 underscores the fact that the soluble extracts from a culture of *E. coli* T7 harbouring pMAL-CGB_E4160W and induced with 0.2 mM IPTG were able to generate more glucose as a consequence of CMC-Na^+ hydrolysis, when compared to supernatants originating from cultures of untransformed and uninduced *E. coli* T7 cultures, as well as an induced culture harbouring an empty pMAL vector. When using an ANOVA test, there was a significant difference ($p = 1.4 \times 10^{-8}$) when the soluble extracts from induced cells harbouring pMAL-CGB_E4160W were incubated with CMC-Na^+, indicating that there was significantly more glucose released as a result of the cellulolytic ability of MBP-cellulase.

Although not statistically significant (p = 0.8), the soluble extracts from uninduced cells harbouring pMAL-CGB_E4160W resulted in slightly more CMC-Na^+ hydrolysis compared to the untransformed and empty vector controls. This observation is likely attributed to leaky expression (absence of the inducer; IPTG) of MBP-cellulase, which can occur when a gene is under the control of the Ptac promoter (Royo, 2005).

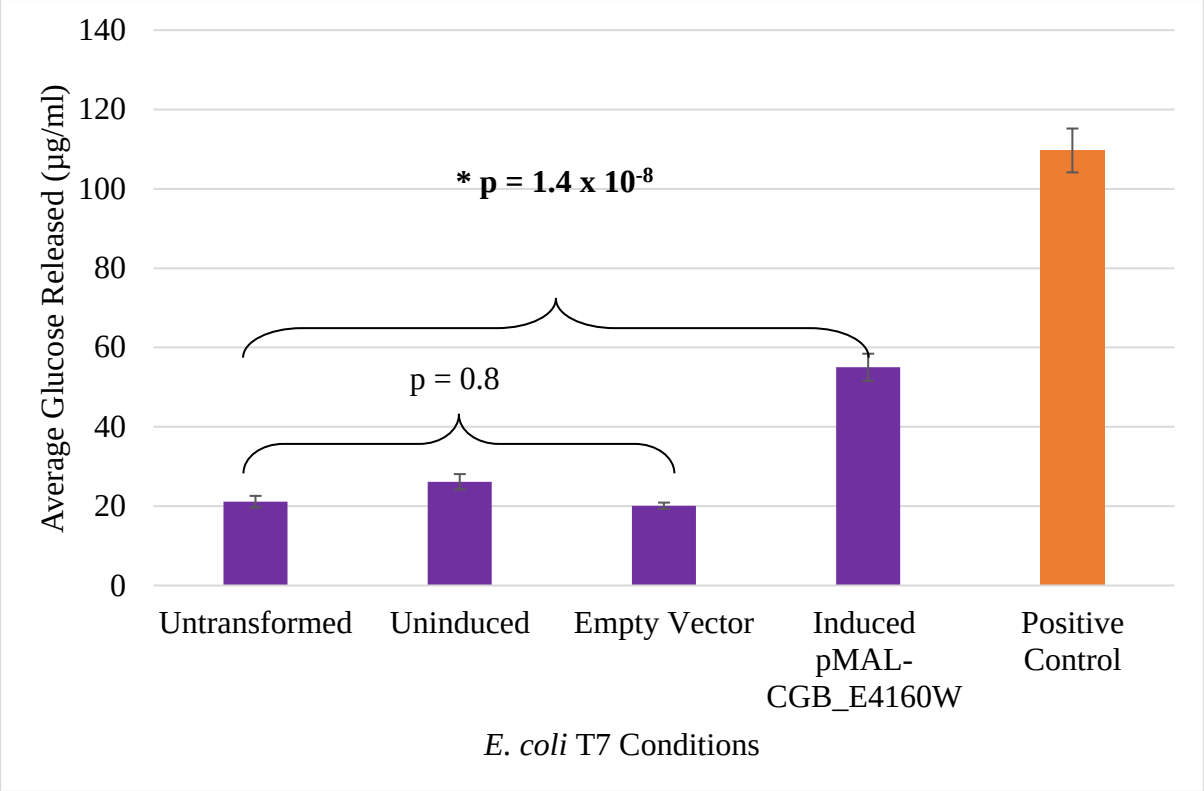

Figure 3.9: CMCase activity of four different *Escherichia coli* T7 strains (untransformed T7 cells, transformed uninduced cells, transformed induced T7 cells with an empty plasmid and transformed induced cells habouring pMAL-CGB_E4160W) with respect to CMC-Na^+ hydrolysis. Error bars represent the standard deviation of three replicates. No significant difference was observed between the three negative controls since the p-value > 0.05. There is a significant difference when comparing the three control conditions to the experimental, as evidenced by a p-value < 0.05. Cellulase from *Trichoderma reesei* served as the positive control.

The use of crude cell lysates or cell-free supernatants in the determination of enzyme activity is comparable to the use of purified protein, however using cell lysates facilitates more rapid determination of an enzyme's hydrolytic ability, without the need for time consuming purification processes (Garcia *et al*., 2018).

Midiri and co-workers (2020) were similarly able to readily detect CMC hydrolysis from lysates of *E. coli* transformed with a pET-21b vector, harbouring a gene encoding a cellulase from *C. neoformans*. This approach enables rapid determination of enzyme activity, and also mitigates the unpredictable loss of enzyme function and solubility, as a result of extended processing times during the purification process (Dako *et al*., 2012).

3.8.2 Optimal pH, Temperature, Sodium chloride and Metal-Ion Determination

In an industrial context, enzymes must operate in adverse circumstances such as high temperatures, extreme pHs, high salinity, and high concentrations of detergents, organic solvents and in the presence of metal ions, all of which frequently result in enzyme inactivation or denaturation (Zarafeta *et al*., 2016).

Once it was confirmed that the soluble extracts from transformed and induced *E. coli* T7 strains displayed cellulolytic activity, the optimum temperature, pH, sodium chloride concertation and the effect of metal ions on the enzyme using lysate was investigated.

As illustrated in Figure 3.10, MBP-cellulase displays maximum activity at 50 ˚C and retained over 50 % of its relative maximal activity at 60 ˚C. A cellulase from *Cryptococcus* sp. S-2 was maximally catalytic between 40-50 ˚C (Thongekkaew *et al*., 2008), whilst a cellulase from *C. neoformans* exhibited maximum activity at physiological temperature, 37 ˚C (Midiri *et al*., 2020). The cellulase from *C. neoformans* did, however, retain over 80 % of its maximum activity at 50 ˚C.

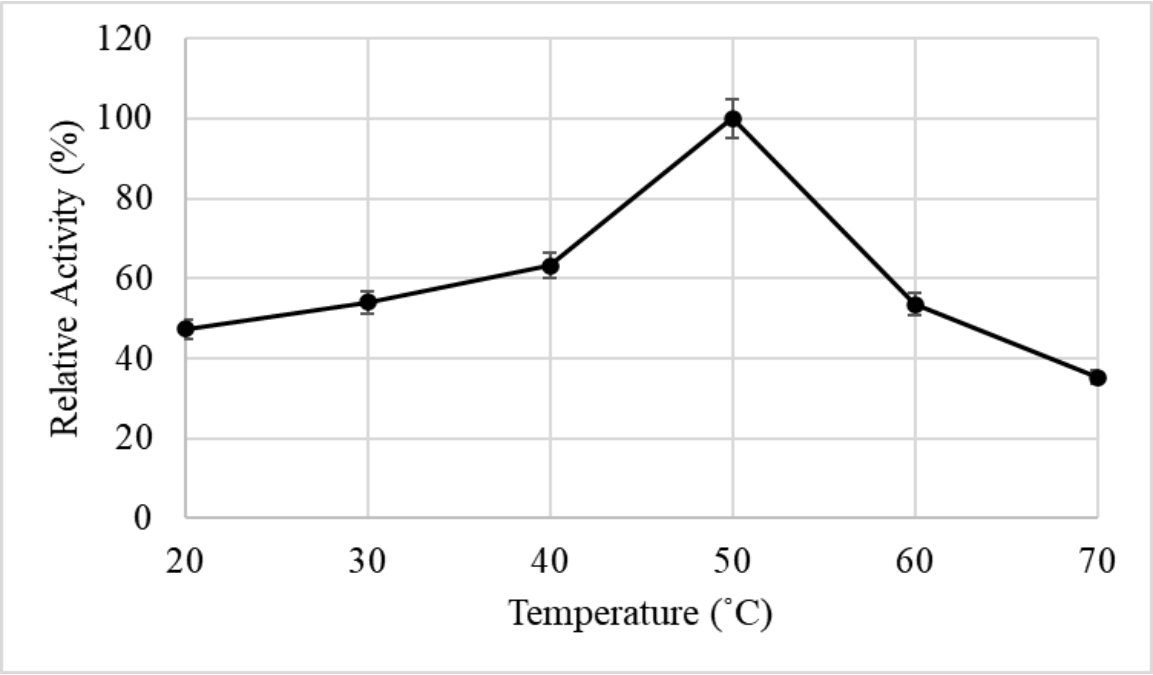

Figure 3.10: Optimal temperature determination of maltose binding protein-cellulase using the 3,5-dinitrosalicylic acid (DNS) assay (Miller, 1959). The optimal temperature for MBP-cellulase was found to be 50˚C. Error bars represent the standard deviation of three replicates.

The higher optimal temperature of the MBP-cellulase compared to the cellulase from *C. neoformans* may be attributed to the fact that members of the *C. gattii* complex are able to persist in temperatures higher than that of *C. neoformans* (Fernandes *et al*., 2016). This is especially apparent since *C. gattii*'s isolation was initially limited to tropical and sub-tropical climates, which tend to be warmer than temperate climates (Kwon-Chung and Bennet, 1984).

Another variable which may have facilitated the *C. gattii* cellulase's higher optimal temperature, may have been its fusion to the N-terminal MBP. As discussed previously, the MBP is an endogenous *E. coli* protein which has been extensively studied and is well understood to be a potent fusion partner that enhances the thermostability and solubility of its passenger protein (Kapust and Waugh, 1999; LaVallie *et al*., 1993). The use of MBP as a fusion partner for industrially important enzymes may confer beneficial qualities on the enzyme's capacity to tolerate the severe conditions encountered in the industrial setting. Such an observation was made apparent by Kuang and co-workers (2006), who observed that a heparinase, when fused to MBP, retained more of its activity after prolonged incubation at 4 ˚C. Furthermore, it was found that by substituting a cysteine residue at position 297 with a serine, in tandem with MBP fusion, the MBP-heparinase fusion's activity increased by over 30% due to a thermostability enhancement (Chen *et al*., 2012).

The optimum pH of the MBP-cellulase was determined to be 6 (Figure 3.11). This is slightly higher than that reported in previous studies looking at cellulases from other cryptococcal species (Midiri *et al*., 2020; Thongekkaew *et al*., 2008). Thongekkaew and co-workers (2008) reported a pH optimum of 3.5 for a recombinant cellulase from *Cryptococcus sp.* S-2, whilst Midiri and co-workers (2020) reported a pH optimum of 4 for a recombinant cellulase from *C. neoformans*. However, several other GH5 cellulases derived from other fungi, such as *Aspergillus japonicus, Neosartorya fischeri, Thermoascus aurantiacus* and *T. reesei* have a pH optimum which falls within the range of 3-6 (Maleki et al., 2020).

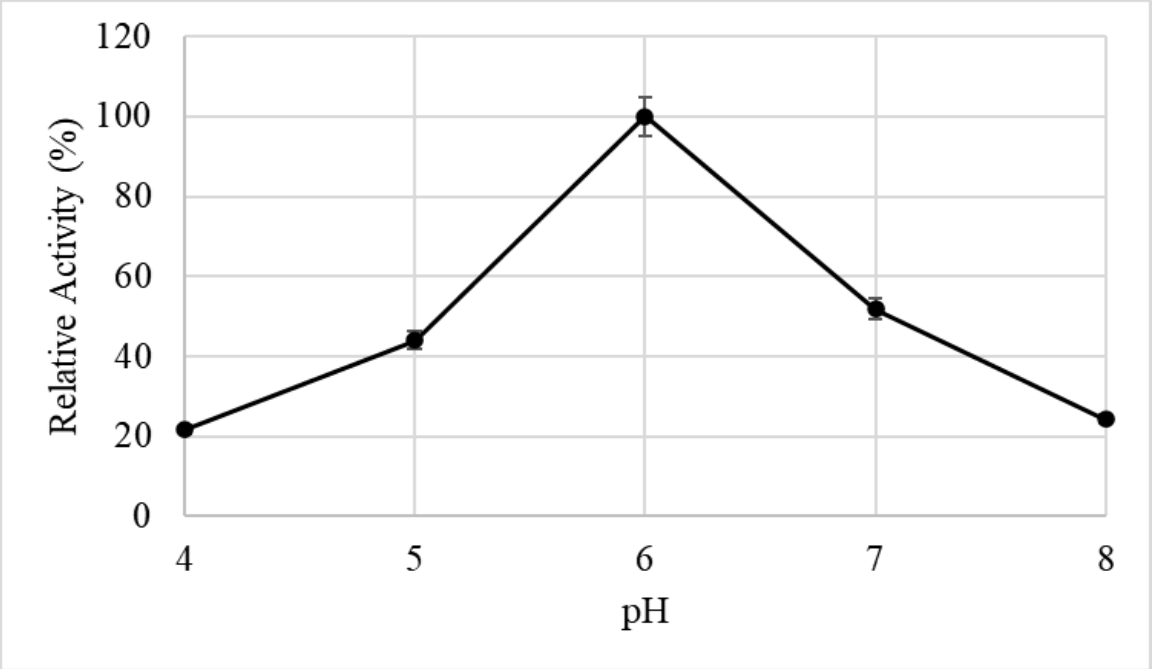

Figure 3.11: Optimal pH determination of maltose binding protein-cellulase using the 3,5-dinitrosalicylic acid (DNS) assay (Miller, 1959). Citrate buffer 50 mM was used for pH 4-5 while 50 mM sodium phosphate buffer was used for pH 6-8. The optimal pH from enzymatic activity of maltose binding protein-cellulase (MBP-cellulase) was found to be 6. Error bars represents the standard deviation of three replicates.

From Figure 3.11, it is evident that MBP-cellulase is active within a relatively wide pH range, with its optimal activity being achieved at a pH of 6, which makes this cellulase suitable for industrial applications (Zarafeta *et al*., 2016). However, there is a decline in activity to only 40 % of its maximum at a pH of 5, and a slightly lesser decline in activity to only 46 % of its maximum at a pH of 7. This finding is not unusual given that Yang and Dang (2011) characterized a GH5 cellulase that was also active over a wide range but experienced a significant decline in activity outside of its pH optima. That particular cellulase exhibited maximal activity at a pH of 4.5, which then declined to 60 % and 80% of its maximum when subjected to a pH of 3.5 and 5.5, respectively.

The MBP tag is not known to significantly alter its passenger enzyme's optimal pH. Ushasree and co-workers (2012) highlighted that a periplasmic phytase enzyme from *E. coli* exhibited an optimal pH at 4, as both a free enzyme as well as a fusion with an N-terminal MBP tag. Zhang and colleagues (2020) also underscored that an L-Amino Acid Deaminase derived from *Proteus mirabilis* had an optimal pH of 8, both in its native state, and when fused to an MBP

tag. It is therefore unlikely that the MBP tag has altered the optimal pH of the cryptococcal cellulase, however, further experiments will need to be conducted.

Considering that pre-treatment of cellulosic biomass needs strong alkaline agents followed by acid neutralisation, the consequence of excessive salinity is a common consequence (Klinke *et al*., 2004). Figure 3.12 illustrates the effect that increasing sodium chloride concentration had on the cellulolytic ability of MBP-cellulase.

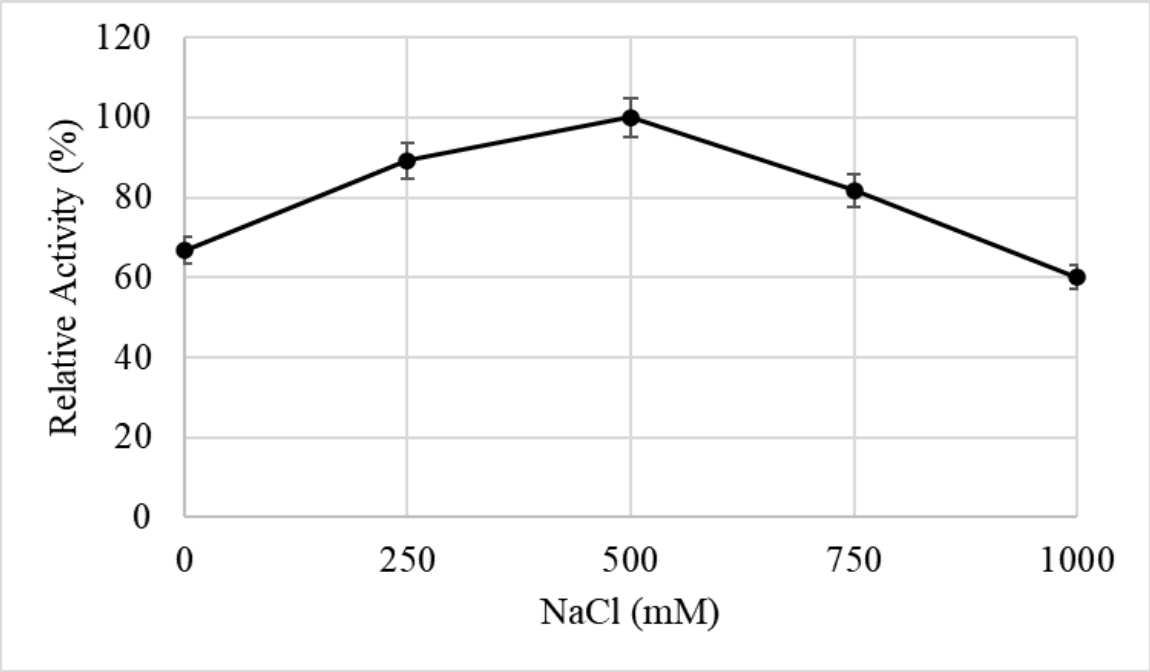

Figure 3.12: Optimal sodium chloride concentration determination of maltose binding protein-cellulase using the 3,5-dinitrosalicylic acid (DNS) assay (Miller, 1959). The optimal sodium chloride concentration was found to be 500 mM. Error bars represents the standard deviation of three replicates.

The MBP-cellulase exhibited optimal enzymatic action at a sodium chloride concentration of 500 mM, and still retained 80 and 60 % of its maximum activity at 750 mM and 1000 mM sodium chloride, respectively. Previously studied cellulases from *Cryptococcus* sp. S-2 and *C. neoformans* were not assayed for their ability to hydrolyse CMC in the presence of increasing salt concentrations (Midiri *et al*., 2020; Thongekkaew *et al*., 2008). However, it stands to reason that a cellulase from *C. gattii* would possess a halotolerant cellulase, when considering that *C. gattii* has been shown to persist in water with a sodium chloride concentration of 20 % (w/v) for over 350 days (Kidd *et al*., 2007).

Lastly, as can be seen in Figure 3.13, the addition of different divalent metal ions did not positively impact the enzymatic activity of MBP-cellulase, ruling out the possibility that the tested metals could act as potential co-factors (Ma *et al*., 2020).

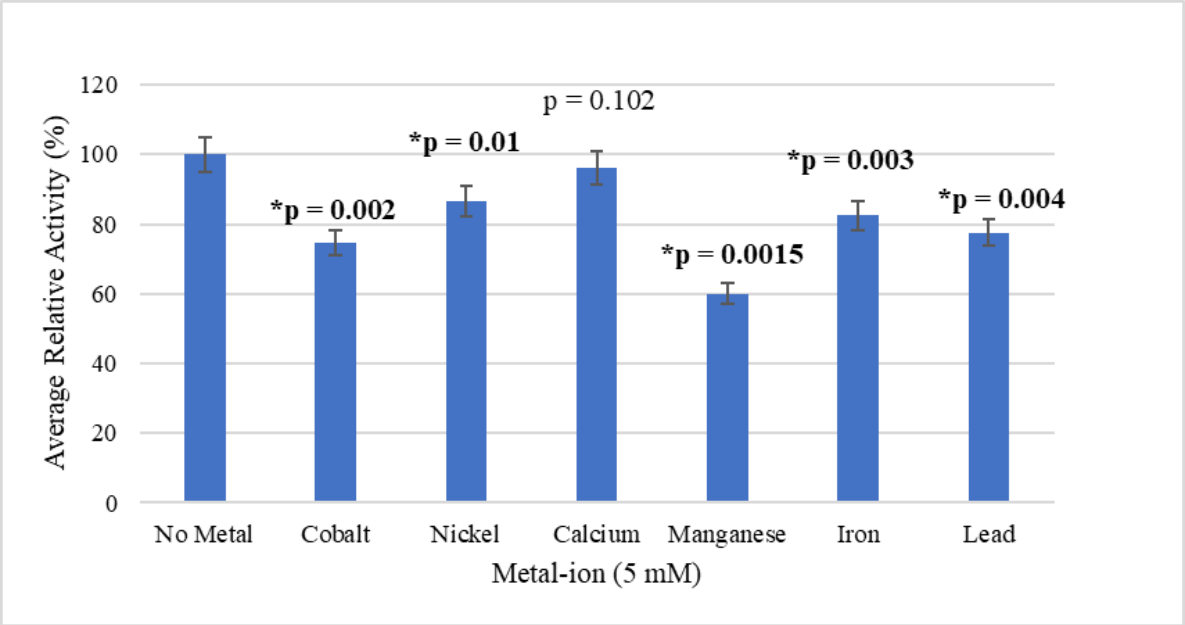

Figure 3.13: The effect of different divalent metal ions on maltose binding protein-cellulase using the 3,5-dinitrosalicylic acid (DNS) assay (Miller, 1959). All reactions proceeded in the presence of either no metal (negative control) or 5 mM of calcium, cobalt, iron, lead, manganese or nickel, The addition of the metal ions did not have a meaningful impact on the hydrolytic ability of maltose binding protein-cellulase. Error bars represents the standard deviation of three replicates. At an α-value of 0.05, there was a significant difference between the no metal control and all the tested metal ions, with the exception of calcium, characterized by a p-value < 0.05 when using an ANOVA test.

Gu and colleagues (2021) observed that a GH5 cellulase, originating from *A. fumigatus*, was marginally inhibited by the presence of 5 mM of different metal ions. Patel and Shah (2021) similarly found that the addition of calcium, cobalt and manganese had detrimental effects on the activity of a GH5 cellulase from the brown-rot fungus, *Fomitopsis meliae*. Of all the monovalent and divalent metal ions tested in this study, only calcium did not have a significantly negative impact on MBP-cellulase's cellulolytic ability, as evidenced by the very high p-value (0.102). In particular, 5 mM manganese had the most deleterious impact on enzymatic activity, which decreased from 100 % (no metals added) to only 60 % when the metal was present.

It is important to consider that there are certain limitations which prevent the accurate determination of released sugars using the DNS activity assay in the presence of certain metal ions (Akbar *et al*., 2006). Akbar and colleagues (2006) determined that the presence of aluminium, barium, calcium, cobalt, iron, lead, manganese, silver, and zinc lead to an overestimation and exaggerated the amount of reducing sugars released, whilst magnesium, copper, cadmium, and mercury lead to an underestimation of the reducing sugars released, which is indicative of the metals' deleterious effects on the reducing power of glucose.

3.9 Conclusion

In conclusion, the results obtained from the computational studies indicated that, despite the cellulase's fusion to the MBP tag, the cellulase retained its TIM barrel fold which is associated with GH5 enzymes. Structural characterization revealed that the MBP-cellulase existed as a functional monomer and had a similar secondary structure to other GH5 cellulases and undergoes a conformational change to accommodate its ligand, CMC-Na^+. Activity data revealed that the MBP-cellulase had a pH and temperature optima which was comparable to that of other cryptococcal cellulases, as well as a tendency to perform better in the presence of 500 mM sodium chloride and in the absence of divalent metal ions.

Chapter Four: Summary

The medically relevant environmental yeast, *Cryptococcus gattii*, is one of the primary aetiological agents of cryptococcosis. The emerging drug resistance, pathogenesis and virulence factors of this yeast have been extensively studied and are well understood. However, the environmental physiology of *C. gattii* and its true ecological niche are points of contention and remain unclear.

A novel gene derived from the genome of *C. gattii* (WM276) was sequenced and annotated as a putative cellulase. To gain insights into the environmental metabolism of *C. gattii*, and elucidate the functionality of the cellulase protein, the gene, CGB_E4160W, was codon-optimized and cloned into a pMAL-c5X vector for recombinant overexpression in *Escherichia coli* T7. The recombinant cellulase, overexpressed as a fusion protein with an N-terminal maltose binding protein (MBP), was predominantly present in the soluble fraction and was subsequently purified in a single step, using amylose affinity chromatography. The MBP was not cleaved from the cellulase due to protease inefficiency, however protein bioinformatics revealed that inclusion of the MBP tag had a negligible effect on the cellulase's structure. Despite the presence of an N-terminal MBP tag, a homology model of the fusion enzyme displayed the classical $(\alpha/\beta)_8$ TIM barrel fold, which is characteristic of cellulases from glycosyl hydrolase family 5 (GH5).

The structural character of the MBP-cellulase was investigated using far-ultraviolet circular dichroism (far-UV CD) for secondary structure, and intrinsic tryptophan fluorescence as a means of monitoring tertiary structure under the influence of different local environments. Far-UV CD revealed a primarily α-helical secondary structure with β-sheets to a lesser extent, which is jointly attributed to the primarily α-helical secondary structure of MBP, as well as the conserved canonical triose phosphate isomerase barrel fold associated with members of the GH5 group of enzymes. Intrinsic tryptophan fluorescence revealed that the MBP-cellulase, in its native state, has its tryptophan residues partially exposed to the local environment. |In the presence of a ligand, carboxymethylcellulose sodium salt ($CMC\text{-}Na^+$), the MBP-cellulase becomes more hydrophobic, possibly as a consequence of ligand accommodation, which preludes hydrolysis.

The impure soluble extracts from transformed and induced *E. coli* cultures were subjected to a 3,5-dinitrosalicylic acid (DNS) assay and were found to exhibit endo-acting cellulase activity. The optimum cellulolytic parameters for the MBP-cellulase were determined to be 50 ˚C, pH 6 and in the presence of 500 mM sodium chloride. This fusion enzyme has potential

biotechnological applications, given its thermotolerance, halotolerance and its ability to perform optimally at an acidic pH, however further studies using the purified cellulase should be undertaken.

To the best of our knowledge, this work represents the first overexpression, purification, and characterization of a cellulase from *C. gattii* WM276. Moreover, this is the first illustration of the structural character of a cellulase from any cryptococcal pathogen. These findings have formed the foundation of understanding the environmental physiology of *C. gattii*, and further supports the notion that woody material, is in fact, the true ecological niche of *C. gattii*.

Chapter Five: References

Ahmad, E. and Pant, K. (2018). Lignin Conversion: A Key to the Concept of Lignocellulosic Biomass-Based Integrated Biorefinery. *Waste Biorefinery*, pp.409-444.

Akbar, S., Sreeramulu, K. and Sharma, H. (2016). Tryptophan fluorescence quenching as a binding assay to monitor protein conformation changes in the membrane of intact mitochondria. *Journal of Bioenergetics and Biomembranes*, 48(3), pp.241-247.

Aksenov, S. I. Babyeva, I. P. Golubev, V. (1973). On the mechanism of adaptation of micro-organisms to conditions of extreme low humidity. *Life Science Space Research*, (11) pp. 55-61.

Ali, S., Hall, J. and Soole, K. (1995). Targeted expression of microbial cellulases in transgenic animals, *Carbohydrate Bioengineering, Progress in Biotechnology*, (10) pp. 279–293.

Alimi, I. and Keku, E. (2021). Epidemiological Trends for Cryptococcosis in Swaziland (Eswatini), Southern Africa. *Research Ideas and Outcomes*, 19(7), pp.38-47

Almagro Armenteros, J., Tsirigos, K., Sønderby, C., Petersen, T., Winther, O., Brunak, S., von Heijne, G. and Nielsen, H. (2019). SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nature Biotechnology*, 37(4), pp.420-423.

Almeida, F., Wolf, J. and Casadevall, A. (2015). Virulence-Associated Enzymes of *Cryptococcus neoformans*. *Eukaryotic Cell*, 14(12), pp.1173-1185.

Alvarez, T., Paiva, J., Ruiz, D., Cairo, J., Pereira, I., Paixão, D., de Almeida, R., Tonoli, C., Ruller, R., Santos, C., Squina, F. and Murakami, M. (2013). Structure and Function of a Novel Cellulase 5 from Sugarcane Soil Metagenome. *PLoS ONE*, 8(12), p.e83635.

Angov, E., Hillier, C., Kincaid, R. and Lyon, J. (2008). Heterologous Protein Expression Is Enhanced by Harmonizing the Codon Usage Frequencies of the Target Gene with those of the Expression Host. *PLoS ONE*, 3(5), p.e2189.

Anthis, N. and Clore, G. (2013). Sequence-specific determination of protein and peptide concentrations by absorbance at 205 nm. *Protein Science*, 22(6), pp.851-858.

Antosiewicz, J. and Shugar, D. (2016). UV–Vis spectroscopy of tyrosine side-groups in studies of protein structure. Part 2: selected applications. *Biophysical Reviews*, 8(2), pp.163-177.

Ardèvol, A. and Rovira, C. (2015). Reaction Mechanisms in Carbohydrate-Active Enzymes: Glycoside Hydrolases and Glycosyltransferases. Insights from ab Initio Quantum Mechanics/Molecular Mechanics Dynamic Simulations. *Journal of the American Chemical Society*, 137(24), pp.7528-7547.

Aspeborg, H., Coutinho, P., Wang, Y., Brumer, H. and Henrissat, B. (2012). Evolution, substrate specificity and subfamily classification of glycoside hydrolase family 5 (GH5). *BMC Evolutionary Biology*, 12(1).

Badieyan, S., Bevan, D. and Zhang, C. (2011). Study and design of stability in GH5 cellulases. *Biotechnology and Bioengineering*, 109(1), pp.31-44.

Bailey, B. A. and Lumsden R. D, (1998). Direct effects of *Trichoderma* and *Gliocladium* on plant growth and resistance to pathogens, In *Trichoderma* & *Gliocladium*—Enzymes,

Bakheet, T. and Doig, A. (2009). Properties and identification of human protein drug targets. *Bioinformatics*, 25(4), pp.451-457.

Baldwin, R. (1986). Temperature dependence of the hydrophobic interaction in protein folding. *Proceedings of the National Academy of Sciences*, 83(21), pp.8069-8072.

Balsera, M., Menéndez, M., Sáiz, J., de las Rivas, J., Andreu, J. and Arellano, J. (2004). Structural Stability of the PsbQ Protein of Higher Plant Photosystem II. *Biochemistry*, 43(44), pp.14171-14179.

Bamforth, C. (2009). Current perspectives on the role of enzymes in brewing. *Journal of Cereal Science*, 50(3), pp.353-357.

Barcellos, V., Martins, L., Fontes, A., Reuwsaat, J., Squizani, E., de Sousa Araújo, G., Frases, S., Staats, C., Schrank, A., Kmetzsch, L. and Vainstein, M. (2018). Genotypic and Phenotypic Diversity of *Cryptococcus gattii* VGII Clinical Isolates and Its Impact on Virulence. *Frontiers in Microbiology*, 9(4), pp.12-19.

Barkalow, G. D. and Whistler, R. L. (2014). Cellulose. *AccessScience*, 10(2), pp.22-28.

Bastos, R., Carneiro, H., Oliveira, L., Rocha, K., Freitas, G., Costa, M., Magalhães, T., Carvalho, V., Rocha, C., Ferreira, G., Paixão, T., Moyrand, F., Janbon, G. and Santos, D. (2018). Environmental Triazole Induces Cross-Resistance to Clinical Drugs and Affects Morphophysiology and Virulence of *Cryptococcus gattii* and *C. neoformans*. *Antimicrobial Agents and Chemotherapy*, 62(1).

Bayer, E., Morag, E. and Lamed, R. (1994). The cellulosome — A treasure-trove for biotechnology. *Trends in Biotechnology*, 12(9), pp.379-386.

Bayer, E., Setter, E. and Lamed, R. (1985). Organization and distribution of the cellulosome in *Clostridium thermocellum*. *Journal of Bacteriology*, 163(2), pp.552-559.

Behbahani, M., Mohabatkar, H. and Nosrati, M. (2016). Analysis and comparison of lignin peroxidases between fungi and bacteria using three different modes of Chou's general pseudo amino acid composition. *Journal of Theoretical Biology*, 411, pp.1-5.

Bennett, J., Dismukes, W., Duma, R., Medoff, G., Sande, M., Gallis, H., Leonard, J., Fields, B., Bradshaw, M., Haywood, H., McGee, Z., Cate, T., Cobbs, C., Warner, J. and Alling, D. (1979). A Comparison of Amphotericin B Alone and Combined with Flucytosine in the Treatment of Cryptoccal Meningitis. *New England Journal of Medicine*, 301(3), pp.126-131.

Berejnoi, A., Taverna, C., Mazza, M., Vivot, M., Isla, G., Córdoba, S. and Davel, G. (2019). First case report of cryptococcosis due to *Cryptococcus decagattii* in a pediatric patient in Argentina. *Revista da Sociedade Brasileira de Medicina Tropical*, 52.

Bergmann-Leitner, E., Mease, R., Duncan, E., Khan, F., Waitumbi, J. and Angov, E. (2008). Evaluation of immunoglobulin purification methods and their impact on quality and yield of antigen-specific antibodies. *Malaria Journal*, 7(1).

Bhat, M. (2000). Cellulases and related enzymes in biotechnology. *Biotechnology Advances*, 18(5), pp.355-383.

Botes, A., Boekhout, T., Hagen, F., Vismer, H., Swart, J. and Botha, A. (2008). Growth and Mating of *Cryptococcus neoformans var. grubii* on Woody Debris. *Microbial Ecology*, 57(4), pp.757-765.

Braithwaite, K., Black, G., Hazlewood, G., Ali, B. and Gilbert, H. (1995). A non-modular endo-β-1,4-mannanase from *Pseudomonas fluorescens* subspecies *cellulosa*. *Biochemical Journal*, 305(3), pp.1005-1010.

Byrnes, E. and Marr, K. (2011). The Outbreak of *Cryptococcus gattii* in Western North America: Epidemiology and Clinical Issues. *Current Infectious Disease Reports*, 13(3), pp.256-261.

Cano-Ramírez, C., Santiago-Hernández, A., Rivera-Orduña, F., Pineda-Mendoza, R., Zúñiga, G. and Hidalgo-Lara, M. (2016). One-step zymogram method for the simultaneous detection of cellulase/xylanase activity and molecular weight estimation of the enzyme. *Electrophoresis*, 38(3-4), pp.447-451.

Carder, J. (1986). Detection and quantitation of cellulase by Congo red staining of substrates in a cup-plate diffusion assay. *Analytical Biochemistry*, 153(1), pp.75-79.

Casadevall, A., Steenbergen, J. and Nosanchuk, J. (2003). 'Ready made' virulence and 'dual use' virulence factors in pathogenic environmental fungi — the *Cryptococcus neoformans* paradigm. *Current Opinion in Microbiology*, 6(4), pp.332-337.

Chandel, A., Garlapati, V., Singh, A., Antunes, F. and da Silva, S. (2018). The path forward for lignocellulose biorefineries: Bottlenecks, solutions, and perspective on commercialization. *Bioresource Technology*, 264, pp.370-381.

Chen, L. (2004). VFDB: a reference database for bacterial virulence factors. *Nucleic Acids Research*, 33(Database issue), pp.D325-D328.

Chen, M., Yu, Q. and Sun, H. (2013). Novel Strategies for the Prevention and Treatment of Biofilm Related Infections. *International Journal of Molecular Sciences*, 14(9), pp.18488-18501.

Chen, R. (1967). Fluorescence Quantum Yields of Tryptophan and Tyrosine. *Analytical Letters*, 1(1), pp.35-42.

Chen, S., Huang, Z., Wu, J., Chen, Y., Ye, F., Zhang, C., Yatsunami, R., Nakamura, S. and Xing, X. (2012). Combination of site-directed mutagenesis and calcium ion addition for enhanced production of thermostable MBP-fused heparinase I in recombinant *Escherichia coli*. Applied Microbiology and Biotechnology, 97(7), pp.2907-2916.

Chen, S., Meyer, W. and Sorrell, T. (2014). *Cryptococcus gattii* Infections. Clinical *Microbiology Reviews*, 27(4), pp.980-1024.

Chen, S., Slavin, M., Heath, C., Playford, E., Byth, K., Marriott, D., Kidd, S., Bak, N., Currie, B., Hajkowicz, K., Korman, T., McBride, W., Meyer, W., Murray, R. and Sorrell, T. (2012). Clinical Manifestations of *Cryptococcus gattii* Infection: Determinants of Neurological Sequelae and Death. *Clinical Infectious Diseases*, 55(6), pp.789-798.

Chen, S., Xing, X., Huang, J. and Xu, M. (2011). Enzyme-assisted extraction of flavonoids from Ginkgo biloba leaves: Improvement effect of flavonol transglycosylation catalyzed by *Penicillium decumbens* cellulase. *Enzyme and Microbial Technology*, 48(1), pp.100-105.

Chowdhary, A., Randhawa, H., Prakash, A. and Meis, J. (2011). Environmental prevalence of *Cryptococcus neoformans* and *Cryptococcus gattii* in India: An update. *Critical Reviews in Microbiology*, 38(1), pp.1-16.

Coelho, M., Bakkeren, G., Sun, S., Hood, M. and Giraud, T. (2017). Fungal Sex: The Basidiomycota. *Microbiology Spectrum*, 5(3).

Cogliati, M. (2013). Global Molecular Epidemiology of *Cryptococcus neoformans* and *Cryptococcus gattii*: An Atlas of the Molecular Types. *Scientifica*, 2013, pp.1-23.

Cohen, R., Suzuki, M. and Hammel, K. (2005). Processive Endoglucanase Active in Crystalline Cellulose Hydrolysis by the Brown Rot Basidiomycete *Gloeophyllum trabeum*. *Applied and Environmental Microbiology*, 71(5), pp.2412-2417.

Corrêa, D. and Ramos, C. (2009). The use of circular dichroism spectroscopy to study protein folding, form and function. *African Journal of Biochemistry Research*, 3(5), pp.164-173.

Coscia, F., Estrozi, L., Hans, F., Malet, H., Noirclerc-Savoye, M., Schoehn, G. and Petosa, C. (2016). Fusion to a homo-oligomeric scaffold allows cryo-EM analysis of a small protein. *Scientific Reports*, 6(1), pp.232-239

Crichton, P., Harding, M., Ruprecht, J., Lee, Y. and Kunji, E. (2013). Lipid, Detergent, and Coomassie Blue G-250 Affect the Migration of Small Membrane Proteins in Blue Native Gels. *Journal of Biological Chemistry*, 288(30), pp.22163-22173.

Curran, J. (1995). Decoding with the A:I wobble pair is inefficient. *Nucleic Acids Research*, 23(4), pp.683-688.

Datta, K., Bartlett, K., Baer, R., Byrnes, E., Galanis, E., Heitman, J., Hoang, L., Leslie, M., MacDougall, L., Magill, S., Morshed, M. and Marr, K. (2009). Spread of *Cryptococcus gattii* into Pacific Northwest Region of the United States. *Emerging Infectious Diseases*, 15(8), pp.1185-1191.

de Boer, H., Comstock, L. and Vasser, M. (1983). The tac promoter: a functional hybrid derived from the trp and lac promoters. *Proceedings of the National Academy of Sciences*, 80(1), pp.21-25.

De Mattos-Shipley, K., Ford, K., Alberti, F., Banks, A., Bailey, A. and Foster, G. (2016). The good, the bad and the tasty: The many roles of mushrooms. *Studies in Mycology*, 85, pp.125-157.

de Nies, L., Lopes, S., Busi, S., Galata, V., Heintz-Buschart, A., Laczny, C., May, P. and Wilmes, P. (2021). PathoFact: a pipeline for the prediction of virulence factors and antimicrobial resistance genes in metagenomic data. *Microbiome*, 9(1).

Deng, S. and Tabatabai, M. (1994). Cellulase activity of soils. *Soil Biology and Biochemistry*, 26(10), pp.1347-1354.

Desjardins, P., Hansen, J. and Allen, M. (2009). Microvolume Protein Concentration Determination Using the NanoDrop 2000c Spectrophotometer. *Journal of Visualized Experiments*, (33).

Dey, P. and Roy, A. (2018). Molecular structure and catalytic mechanism of fungal family G acidophilic xylanases. *3 Biotech*, 8(2).

Dippel, R. and Boos, W. 2005. The Maltodextrin System of *Escherichia coli*: Metabolism and Transport. *Journal of Bacteriology*, 187(24), pp.8322-8331.

Doering, T. (2009). How Sweet it is! Cell Wall Biogenesis and Polysaccharide Capsule Formation in *Cryptococcus neoformans*. *Annual Review of Microbiology*, 63(1), pp.223-247.

Doering, T., Nosanchuk, J., Roberts, W. and Casadevall, A. (2008). Melanin as a potential cryptococcal defence against microbicidal proteins. *Medical Mycology*, 37(3), pp.175-181.

Doi, R. and Kosugi, A. (2004). Cellulosomes: plant-cell-wall-degrading enzyme complexes. *Nature Reviews Microbiology*, 2(7), pp.541-551.

Dusa, A., Kaylor, J., Edridge, S., Bodner, N., Hong, D. and Fink, A. (2006). Characterization of Oligomers during α-Synuclein Aggregation Using Intrinsic Tryptophan Fluorescence. *Biochemistry*, 45(8), pp.2752-2760.

Edwards, H., Cogliati, M., Kwenda, G. and Fisher, M. (2021). The need for environmental surveillance to understand the ecology, epidemiology and impact of *Cryptococcus* infection in Africa. *FEMS Microbiology Ecology*, 97(7).

Ejaz, U., Sohail, M. and Ghanemi, A. (2021). Cellulases: From Bioactivity to a Variety of Industrial Applications. Biomimetics, 6(3), p.44.

Ellis, D. and Pfeiffer, T. (1990). Ecology, life cycle, and infectious propagule of *Cryptococcus neoformans*. *The Lancet*, 336(8720), pp.923-925.

Enany, S. (2014). Structural and functional analysis of hypothetical and conserved proteins of *Clostridium tetani*. *Journal of Infection and Public Health*, 7(4), pp.296-307.

Eriksson, K. and Pettersson, B. (1973). A zymogram technique for the detection of carbohydrases. *Analytical Biochemistry*, 56(2), pp.618-620.

Escandón, P., Huérfano, S. and Castañeda, E. (2002). Inoculación experimental de Terminalia catappa con un aislamiento ambiental de *Cryptococcus neoformans* var. *gattii* serotipo C. *Biomédica*, 22(4), p.524.

Eubel, H. and Millar, A. (2009). Systematic Monitoring of Protein Complex Composition and Abundance by Blue-Native PAGE. *Cold Spring Harbor Protocols*, 2009(5), p.5221.

Farrow, M. and Arnold, F., 2011. High Throughput Screening of Fungal Endoglucanase Activity in Escherichia coli. *Journal of Visualized Experiments*, (54).

Fernandes, K., Dwyer, C., Campbell, L. and Carter, D. (2016). Species in the *Cryptococcus gattii* Complex Differ in Capsule and Cell Size following Growth under Capsule-Inducing Conditions. *mSphere*, 1(6).

Finch R G, Hill P, and Williams P. (1995) *Staphlococci*—the emerging

Fisher, M., Hawkins, N., Sanglard, D. and Gurr, S. (2018). Worldwide emergence of resistance to antifungal drugs challenges human health and food security. *Science*, 360(6390), pp.739-742.

Fox, J., Levine, S., Clark, D. and Blanch, H. (2011). Initial- and Processive-Cut Products Reveal Cellobiohydrolase Rate Limitations and the Role of Companion Enzymes. *Biochemistry*, 51(1), pp.442-452.

Fraser, J., Giles, S., Wenink, E., Geunes-Boyer, S., Wright, J., Diezmann, S., Allen, A., Stajich, J., Dietrich, F., Perfect, J. and Heitman, J. (2005). Same-sex mating and the origin of the Vancouver Island *Cryptococcus gattii* outbreak. *Nature*, 437(7063), pp.1360-1364.

Frey, S. and Görlich, D. (2014). A new set of highly efficient, tag-cleaving proteases for purifying recombinant proteins. *Journal of Chromatography A*, 1337, pp.95-105.

Gadanho, M., Libkind, D. and Sampaio, J. (2006). Yeast Diversity in the Extreme Acidic Environments of the Iberian Pyrite Belt. *Microbial Ecology*, 52(3), pp.552-563.

Galante Y, M,. DeConti A,. Monteverdi R. (1998). Application of *Trichoderma* enzymes in food and feed industries. In: Harman GF, Kubicek CP, editors. *Trichoderma and Gliocladium—Enzymes.* Vol. 2. London, UK: Taylor & Francis. pp. 311–326.

Galluccio, M., Pantanella, M., Giudice, D., Brescia, S. and Indiveri, C. (2020). Low temperature bacterial expression of the neutral amino acid transporters SLC1A5 (ASCT2), and SLC6A19 (B0AT1). *Molecular Biology Reports*, 47(9), pp.7283-7289.

Ganghadhar, C., Rohit, B. and Basappa, B. (2016). In silico characterization of beta-galactosidase using computational tools. *Journal of Bioinformatics and Sequence Analysis*, 8(1), pp.1-11.

Gao, D., Wang, S., Li, H., Yu, H. and Qi, Q. (2015). Identification of a heterologous cellulase and its N-terminus that can guide recombinant proteins out of *Escherichia coli*. *Microbial Cell Factories*, 14(1).

Garcia, D., Mohr, B., Dovgan, J., Hurst, G., Standaert, R. and Doktycz, M. (2018). Elucidating the potential of crude cell extracts for producing pyruvate from glucose. *Synthetic Biology*, 3(1).

García-Lafuente, A., Guillamón, E., Villares, A., Rostagno, M. and Martínez, J. (2009). Flavonoids as anti-inflammatory agents: implications in cancer and cardiovascular disease. *Inflammation Research*, 58(9), pp.537-552.

Gasteiger, E., 2003. ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Research*, 31(13), pp.3784-3788.

George, J. and Sabapathi, S. (2015). Cellulose nanocrystals: synthesis, functional properties, and applications. Nanotechnology, Science and Applications, p.45.

Ghaemi, F., Abdullah, L. and Ariffin, H. (2019). Lignocellulose Structure and the Effect on Nanocellulose Production. *Lignocellulose for Future Bioeconomy*, pp.17-30.

Ghisaidoobe, A. and Chung, S. (2014). Intrinsic Tryptophan Fluorescence in the Detection and Analysis of Proteins: A Focus on Förster Resonance Energy Transfer Techniques. *International Journal of Molecular Sciences*, 15(12), pp.22518-22538.

Gilardi, G., Mei, G., Rosato, N., Agro, A. and Cass, A. (1997). Spectroscopic properties of an engineered maltose binding protein. *Protein Engineering Design and Selection*, 10(5), pp.479-486.

Gill, S. and von Hippel, P. (1989). Calculation of protein extinction coefficients from amino acid sequence data. *Analytical Biochemistry*, 182(2), pp.319-326.

Godfrey, T. and West, S. (1996) Textiles, in *Industrial Enzymology*, pp. 360–371, Macmillan Press, London, UK, 2nd edition.

Gómez, B. and Nosanchuk, J. (2003). Melanin and fungi. *Current Opinion in Infectious Diseases*, 16(2), pp.91-96.

Gopal, G. and Kumar, A., 2013. Strategies for the Production of Recombinant Protein in *Escherichia coli*. *The Protein Journal*, 32(6), pp.419-425.

Greenfield, N. (2006). Using circular dichroism spectra to estimate protein secondary structure. *Nature Protocols*, 1(6), pp.2876-2890.

Grimsley, G. and Pace, C. (2003). Spectrophotometric Determination of Protein Concentration. *Current Protocols in Protein Science*, 33(1).

Gu, X., Lu, H., Zhang, L. and Meng, X. (2021). A Thermophilic GH5 Endoglucanase from *Aspergillus fumigatus* and Its Synergistic Hydrolysis of Mannan-Containing Polysaccharides. *Catalysts*, 11(7), p.862.

Guerriero, G., Sergeant, K., Legay, S., Hausman, J., Cauchie, H., Ahmad, I. and Siddiqui, K. (2018). Novel Insights from Comparative In Silico Analysis of Green Microalgal Cellulases. *International Journal of Molecular Sciences*, 19(6), p.1782.

Guo, F. and Zhu, G., 2012. Presence and removal of a contaminating NADH oxidation activity in recombinant maltose binding protein fusion proteins expressed in Escherichia coli. *BioTechniques*, 52(4), pp.247-253.

Gupta, P., Samant, K. and Sahu, A. (2012). Isolation of Cellulose-Degrading Bacteria and Determination of Their Cellulolytic Potential. *International Journal of Microbiology*, 2012, pp.1-5.

Gusakov, A., Kondratyeva, E. and Sinitsyn, A. (2011). Comparison of Two Methods for Assaying Reducing Sugars in the Determination of Carbohydrase Activities. *International Journal of Analytical Chemistry*, 2011, pp.1-4.

Gustafsson, C., Govindarajan, S. and Minshull, J. (2004). Codon bias and heterologous protein expression. *Trends in Biotechnology*, 22(7), pp.346-353.

Hamada, N., Ishikawa, K., Fuse, N., Kodaira, R., Shimosaka, M., Amano, Y., Kanda, T. and Okazaki, M. (1999). Purification, characterization and gene analysis of exo-cellulase II (Ex-2) from the white rot basidiomycete *Irpex lacteus*. *Journal of Bioscience and Bioengineering*, 87(4), pp.442-451.

Hansakon, A., Ngamskulrungroj, P. and Angkasekwinai, P. (2019). Contribution of laccase expression to immune response against *Cryptococcus gattii* infection. *Infection and Immunity*.

Harman, G. F., and Kubicek,C. P. Eds., vol. 2 of *Biological Control and Commercial Applications*, pp. 327–342, Taylor & Francis, London, UK.

Harris, D., Bulone, V., Ding, S. and DeBolt, S. (2010). Tools for Cellulose Analysis in Plant Cell Walls. *Plant Physiology*, 153(2), pp.420-426.

Harris, J., Lockhart, S., Debess, E., Marsden-Haug, N., Goldoft, M., Wohrle, R., Lee, S., Smelser, C., Park, B. and Chiller, T. (2011). *Cryptococcus gattii* in the United States: Clinical Aspects of Infection with an Emerging Pathogen. *Clinical Infectious Diseases*, 53(12), pp.1188-1195.

Hebeish, A., Ibrahim N, A. (2007). The impact of frontier sciences on textile industry. *Colourage*. 54, pp.41–55.

Herlet, J., Kornberger, P., Roessler, B., Glanz, J., Schwarz, W., Liebl, W. and Zverlov, V. (2017). A new method to evaluate temperature vs. pH activity profiles for biotechnological relevant enzymes. *Biotechnology for Biofuels*, 10(1).

Hill, H. (1992). The function of melanin or six blind people examine an elephant. *BioEssays*, 14(1), pp.49-56.

Hiller, K., Grote, A., Scheer, M., Munch, R. and Jahn, D. (2004). PrediSi: prediction of signal peptides and their cleavage positions. *Nucleic Acids Research*, 32(Web Server), pp.W375-W379.

Hossein, M., Talaeipour, M., Hemmasi, A., Bazyar, B. and Mahdavi, S. (2015). Effects of Sequencing Enzyme Application and Refining on DIP Properties Produced From Mixed Office Waste Paper.10(3).

Hu, X., Fan, G., Liao, H., Fu, Z., Ma, C., Ni, H. and Li, X. (2020). Optimized soluble expression of a novel endoglucanase from *Burkholderia pyrrocinia* in *Escherichia coli*. *3 Biotech*, 10(9).

Hurtado, J., Castillo, P., Fernandes, F., Navarro, M., Lovane, L., Casas, I., Quintó, L., Marco, F., Jordao, D., Ismail, M., Lorenzoni, C., Martinez-Palhares, A., Ferreira, L., Lacerda, M., Monteiro, W., Sanz, A., Letang, E., Marimon, L., Jesri, S., Cossa, A., Mandomando, I., Vila, J., Bassat, Q., Ordi, J., Menéndez, C., Carrilho, C. and Martínez, M. (2019). Mortality due to *Cryptococcus neoformans* and *Cryptococcus gattii* in low-income settings: an autopsy study. *Scientific Reports*, 9(1).

Ibrahim, N., EL-Badry, K., Eid, B. and Hassan, T. (2011). A new approach for biofinishing of cellulose-containing fabrics using acid cellulases. *Carbohydrate Polymers*, 83(1), pp.116-121.

Ikai, A. (1980). Thermostability and Aliphatic Index of Globular Proteins. *The Journal of Biochemistry*, 88(6), pp.1985-1988.

Ikemura, T. (1981). Correlation between the abundance of Escherichia coli transfer RNAs and the occurrence of the respective codons in its protein genes: A proposal for a synonymous codon choice that is optimal for the *E. coli* translational system. *Journal of Molecular Biology*, 151(3), pp.389-409.

Illnait-Zaragozí, M., Ortega-Gonzalez, L., Hagen, F., Martínez-Machin, G. and Meis, J. (2013). Fatal *Cryptococcus gattii* genotype AFLP5 infection in an immunocompetent Cuban patient. *Medical Mycology Case Reports*, 2, pp.48-51

Jacobson, E. (2000). Pathogenic Roles for Fungal Melanins. *Clinical Microbiology Reviews*, 13(4), pp.708-717.

Jain, A., Jain, R., & Jain, S. (2020). Basic Techniques in Biochemistry, Microbiology and Molecular Biology. *Springer Protocols Handbooks.*

Jayasekara, S. and Ratnayake, R. (2019). Microbial Cellulases: An Overview and Applications. *Cellulose*.

Jenny, R., Mann, K. and Lundblad, R. (2003). A critical review of the methods for cleavage of fusion proteins with thrombin and factor Xa. *Protein Expression and Purification*, 31(1), pp.1-11.

Jiang, K., Zhang, Y., Chen, Z., Wu, D., Cai, J. and Gao, X. (2020). Structural and Functional Insights into the C-terminal Fragment of Insecticidal Vip3A Toxin of *Bacillus thuringiensis*. *Toxins*, 12(7), p.438.

Jindou, S., Xu, Q., Kenig, R., Shulman, M., Shoham, Y., Bayer, E. and Lamed, R. (2006). Novel architecture of family-9 glycoside hydrolases identified in cellulosomal enzymes of *Acetivibrio cellulolyticus* and *Clostridium thermocellum*. *FEMS Microbiology Letters*, 254(2), pp.308-316.

Joshi, N., Kaushal, G. and Singh, S. (2021). nik. *Biotechnology and Bioengineering*, 118(4), pp.1531-1544.

Juturu, V. and Wu, J. (2014). Microbial cellulases: Engineering, production and applications. *Renewable and Sustainable Energy Reviews*, 33, pp.188-203.

Käll, L., Krogh, A. and Sonnhammer, E. (2004). A Combined Transmembrane Topology and Signal Peptide Prediction Method. *Journal of Molecular Biology*, 338(5), pp.1027-1036.

Kall, L., Krogh, A. and Sonnhammer, E. (2007). Advantages of combined transmembrane topology and signal peptide prediction--the Phobius web server. *Nucleic Acids Research*, 35, pp.W429-W432.

Kapust, R. and Waugh, D. (1999). *Escherichia coli* maltose binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Science*, 8(8), pp.1668-1674.

Kaur, A., Pati, P., Pati, A. and Nagpal, A. (2020). Physico-chemical characterization and topological analysis of pathogenesis-related proteins from *Arabidopsis thaliana* and *Oryza sativa* using in-silico approaches. *PLOS ONE*, 15(9), p.e0239836.

Kelley, L., Mezulis, S., Yates, C., Wass, M. and Sternberg, M. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols*, 10(6), pp.845-858.

Khatoon, Z., McTiernan, C., Suuronen, E., Mah, T. and Alarcon, E. (2018). Bacterial biofilm formation on implantable devices and approaches to its treatment and prevention. *Heliyon*, 4(12), p.e01067.

Kidd, S., Chow, Y., Mak, S., Bach, P., Chen, H., Hingston, A., Kronstad, J. and Bartlett, K. (2007). Characterization of Environmental Sources of the Human and Animal Pathogen *Cryptococcus gattii* in British Columbia, Canada, and the Pacific Northwest of the United States. *Applied and Environmental Microbiology*, 73(5), pp.1433-1443.

Kidd, S., Hagen, F., Tscharke, R., Huynh, M., Bartlett, K., Fyfe, M., MacDougall, L., Boekhout, T., Kwon-Chung, K. and Meyer, W. (2004). A rare genotype of *Cryptococcus gattii* caused the cryptococcosis outbreak on Vancouver Island (British Columbia, Canada). *Proceedings of the National Academy of Sciences*, 101(49), pp.17258-17263.

Kim, J., Copeland, C., Padumane, S. and Kwon, Y. (2019). A Crude Extract Preparation and Optimization from a Genomically Engineered *Escherichia coli* for the Cell-Free Protein Synthesis System: Practical Laboratory Guideline. *Methods and Protocols*, 2(3), p.68.

Kim, S., Lee, C., Han, B., Kim, M., Yeo, Y., Yoon, S., Koo, B. and Jun, H. (2008). Characterization of a gene encoding cellulase from uncultured soil bacteria. *FEMS Microbiology Letters*, 282(1), pp.44-51.

Klinke, H., Thomsen, A. and Ahring, B. (2004). Inhibition of ethanol-producing yeast and bacteria by degradation products produced during pre-treatment of biomass. *Applied Microbiology and Biotechnology*, 66(1), pp.10-26.

Kozel, T., Pfrommer, G., Guerlain, A., Highison, B. and Highison, G. (1988). Role of the Capsule in Phagocytosis of *Cryptococcus neoformans*. *Clinical Infectious Diseases*, 10(Supplement 2), pp.S436-S439.

Kronstad, J., Attarian, R., Cadieux, B., Choi, J., D'Souza, C., Griffiths, E., Geddes, J., Hu, G., Jung, W., Kretschmer, M., Saikia, S. and Wang, J. (2011). Expanding fungal pathogenesis: *Cryptococcus* breaks out of the opportunistic box. *Nature Reviews* Microbiology, 9(3), pp.193-203.

Krüger, M. and Sørensen, M. (1998). Aminoacylation of hypomodified tRNAGlu *in vivo*. *Journal of Molecular Biology*, 284(3), pp.609-620.

Kuang, Y., Xing, X., Chen, Y., Ye, F., Chen, Y., Yan, Y., Liu, Z. and Bi, R. (2006). Production of heparin oligosaccharides by fusion protein of MBP–heparinase I and the enzyme thermostability. *Journal of Molecular Catalysis B: Enzymatic*, 43(1-4), pp.90-95.

Kuhad, R., Gupta, R. and Singh, A. (2011). Microbial Cellulases and Their Industrial Applications. *Enzyme Research*, 2011, pp.1-10.

Kumagai, P., Araujo, A. and Lopes, J. (2017). Going deep into protein secondary structure with synchrotron radiation circular dichroism spectroscopy. *Biophysical Reviews*, 9(5), pp.517-527.

Kumar, R. and Wyman, C. (2009). Cellulase adsorption and relationship to features of corn stover solids produced by leading pretreatments. *Biotechnology and Bioengineering*, 103(2), pp.252-267.

Kwon-Chung, K. and Bennet, J. (1984). Epidemiological Differences Between the Two Varieties of *Cryptococcus neoformans*. *American Journal of Epidemiology*, 120(1), pp.123-130.

Kyte, J. and Doolittle, R. (1982). A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology*, 157(1), pp.105-132.

La Hoz, R. and Pappas, P. (2013). Cryptococcal Infections: Changing Epidemiology and Implications for Therapy. *Drugs*, 73(6), pp.495-504.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head

Lafond, M., Sulzenbacher, G., Freyd, T., Henrissat, B., Berrin, J. and Garron, M. (2016). The Quaternary Structure of a Glycoside Hydrolase Dictates Specificity toward β-Glucans. *Journal of Biological Chemistry*, 291(13), pp.7183-7194.

Langston, J., Shaghasi, T., Abbate, E., Xu, F., Vlasenko, E. and Sweeney, M. (2011). Oxidoreductive Cellulose Depolymerization by the Enzymes Cellobiose Dehydrogenase and Glycoside Hydrolase 61. *Applied and Environmental Microbiology*, 77(19), pp.7007-7015.

LaVallie, E., DiBlasio, E., Kovacic, S., Grant, K., Schendel, P. and McCoy, J. (1993). A Thioredoxin Gene Fusion Expression System That Circumvents Inclusion Body Formation in the *E. coli* Cytoplasm. *Nature Biotechnology*, 11(2), pp.187-193.

Lee, S., Forsberg, C. and Gibbins, L. (1985). Cellulolytic Activity of *Clostridium acetobutylicum*. *Applied and Environmental Microbiology*, 50(2), pp.220-228.

Leopold-Wager, C., Hole, C., Wozniak, K. and Wormley, F. (2016). *Cryptococcus* and Phagocytes: Complex Interactions that Influence Disease Outcome. *Frontiers in Microbiology*, 7.

Lin, L., Liu, X., Zhou, Y., Guan, L., He, J. and Huang, W. (2016). A novel pH-stable, endoglucanase (JqCel5A) isolated from a salt-lake microorganism, *Jonesia quinghaiensis*. *Electronic Journal of Biotechnology*, 24, pp.56-62.

Lin, X., Hull, C. and Heitman, J. (2005). Sexual reproduction between partners of the same mating type in Cryptococcus neoformans. *Nature*, 434(7036), pp.1017-1021.

Litvintseva, A., Carbone, I., Rossouw, J., Thakur, R., Govender, N. and Mitchell, T. (2011). Evidence that the Human Pathogenic Fungus *Cryptococcus neoformans* var. *grubii* May Have Evolved in Africa. *PLoS ONE*, 6(5), p.e19688.

Liu, D., Xu, R., Dutta, K. and Cowburn, D., 2008. N-terminal cysteinyl proteins can be prepared using thrombin cleavage. *FEBS Letters*, 582(7), pp.1163-1167.

Liu, H., Sale, K., Holmes, B., Simmons, B. and Singh, S. (2010). Understanding the Interactions of Cellulose with Ionic Liquids: A Molecular Dynamics Study. *The Journal of Physical Chemistry B*, 114(12), pp.4293-4301.

Liu, I, Liu, M. and Shergill, K. (2006). The Effect of Spheroplast Formation on the Transformation Efficiency in Escherichia coli DH5α. *Journal of Experimental Microbiology and Immunology*. 9, pp.81-85.

Liu, L., Tewari, R. and Williamson, P. (1999). Laccase Protects *Cryptococcus neoformans* from Antifungal Activity of Alveolar Macrophages. *Infection and Immunity*, 67(11), pp.6034-6039.

Loftus, B. (2005). The Genome of the Basidiomycetous Yeast and Human Pathogen *Cryptococcus neoformans*. *Science*, 307(5713), pp.1321-1324.

Loiselle, M. and Anderson, K. (2003). The Use of Cellulase in Inhibiting Biofilm Formation from Organisms Commonly Found on Medical Implants. *Biofouling*, 19(2), pp.77-85.

Luo, K., Kim, N., You, S. and Kim, Y. (2019). Colorimetric Determination of the Activity of Starch-Debranching Enzyme via Modified Tollens' Reaction. *Nanomaterials*, 9(9), p.1291.

Lynd, L., Weimer, P., van Zyl, W. and Pretorius, I. (2002). Microbial Cellulose Utilization: Fundamentals and Biotechnology. *Microbiology and Molecular Biology Reviews*, 66(3), pp.506-577.

MacDougall, L., Kidd, S., Galanis, E., Mak, S., Leslie, M., Cieslak, P., Kronstad, J., Morshed, M. and Bartlett, K. (2007). Spread of *Cryptococcus gattii* in British Columbia, Canada, and Detection in the Pacific Northwest, USA. *Emerging Infectious Diseases*, 13(1), pp.42-50.

Mada, P., Jamil, R. and Alam, M. (2019). *Cryptococcus (Cryptococcosis)*. Treasure Island: StatPearls Publishing.

Mai, C., Kues, U. and Militz, H. (2004). Biotechnology in the wood industry. *Applied Microbiology and Biotechnology*, 63(5), pp.477-494.

Maleki, M., Shahraki, M., Kavousi, K., Ariaeenejad, S. and Hosseini Salekdeh, G. (2020). A novel thermostable cellulase cocktail enhances lignocellulosic bioconversion and biorefining in a broad range of pH. *International Journal of Biological Macromolecules*, 154, pp.349-360.

Mansfield, S., Mooney, C. and Saddler, J. (1999). Substrate and Enzyme Characteristics that Limit Cellulose Hydrolysis. *Biotechnology Progress*, 15(5), pp.804-816.

Margawati, E., Fuad, A., Indriawati, Ridwan, M. and Volkandari, S. (2017). Optimization of expression JTAT protein with emphasis on transformation efficiency and IPTG concentration. *Journal of Genetic Engineering and Biotechnology*, 15(2), pp.515-519.

Marsden, W., Gray, P., Nippard, G. and Quinlan, M. (2007). Evaluation of the DNS method for analysing lignocellulosic hydrolysates. *Journal of Chemical Technology and Biotechnology*, 32(7-12), pp.1016-1022.

Martinez, D., Larrondo, L., Putnam, N., Gelpke, M., Huang, K., Chapman, J., Helfenbein, K., Ramaiya, P., Detter, J., Larimer, F., Coutinho, P., Henrissat, B., Berka, R., Cullen, D. and Rokhsar, D. (2004). Genome sequence of the lignocellulose degrading fungus *Phanerochaete chrysosporium* strain RP78. *Nature Biotechnology*, 22(6), pp.695-700.

May, R., Stone, N., Wiesner, D., Bicanic, T. and Nielsen, K. (2015). *Cryptococcus*: from environmental saprophyte to global pathogen. *Nature Reviews Microbiology*, 14(2), pp.106-117.

Maya-Hoyos, M., Leguizamón, J., Mariño-Ramírez, L. and Soto, C. (2015). Sliding Motility, Biofilm Formation, and Glycopeptidolipid Production in *Mycobacterium colombiense* Strains. *BioMed Research International*, 2015, pp.1-11.

Maziarz, E. and Perfect, J. (2016). Cryptococcosis. *Infectious Disease Clinics of North America*, 30(1), pp.179-206.

McKendry, P. (2002). Energy production from biomass (part 1): overview of biomass. *Bioresource Technology*, 83(1), pp.37-46.

McLatchie, G.; Borley, N.; Chikwe, J. (2013). *Oxford Handbook of Clinical Surgery. Oxford*, UK: OUP Oxford. p. 794.

McNulty, D., Claffee, B., Huddleston, M. and Kane, J. (2003). Mistranslational errors associated with the rare arginine codon CGG in *Escherichia coli*. *Protein Expression and Purification*, 27(2), pp.365-374.

Midiri, A., Mancuso, G., Lentini, G., Famà, A., Galbo, R., Zummo, S., Giardina, M., De Gaetano, G., Teti, G., Beninati, C. and Biondo, C. (2020). Characterization of an immunogenic cellulase secreted by *Cryptococcus* pathogens. *Medical Mycology*, 58(8), pp.1138-1148.

Miller, G. (1959). Use of Dinitrosalicylic Acid Reagent for Determination of Reducing Sugar. *Analytical Chemistry*, 31(3), pp.426-428.

Mitreva-Dautova, M., Roze, E., Overmars, H., de Graaff, L., Schots, A., Helder, J., Goverse, A., Bakker, J. and Smant, G. (2006). A Symbiont-Independent Endo-1,4-β-Xylanase from the Plant-Parasitic Nematode *Meloidogyne incognita*. *Molecular Plant-Microbe Interactions®*, 19(5), pp.521-529.

Mohan, R. (2012). Computational structural and functional analysis of hypothetical proteins of *Staphylococcus aureus*. *Bioinformation*, 8(15), pp.722-728.

Momin, A., Hameed, U. and Arold, S. (2019). Passenger sequences can promote interlaced dimers in a common variant of the maltose binding protein. *Scientific Reports*, 9(1).

Moutaoufik, M., Morrow, G., Finet, S. and Tanguay, R. (2017). Effect of N-terminal region of nuclear *Drosophila melanogaster* small heat shock protein DmHsp27 on function and quaternary structure. *PLOS ONE*, 12(5), p.e0177821.

Naik, S., Goud, V., Rout, P. and Dalai, A. (2010). Production of first- and second-generation biofuels: A comprehensive review. *Renewable and Sustainable Energy Reviews*, 14(2), pp.578-597.

Nazarali, A., Ji, S., Hu, Y., An, Y., Fang, N., Li, Y. and Jin, H. (2015). The Optimization of Soluble PTEN Expression in *Escherichia coli*. *The Open Biochemistry Journal*, 9(1), pp.42-48.

New England Biolabs. (2018). pMAL™ Protein Fusion & Purification System. *PROTEIN EXPRESSION & ANALYSIS*.

New England Biolabs. T7 Express Competent *E. coli* (High Efficiency) | NEB. Available at: <https://international.neb.com/products/c2566-t7-express-competent-e-coli-high-efficiency#Product%20Information> [Accessed 18 December 2021].

Ni, J., Wu, Y., Yun, C., Yu, M. and Shen, Y. (2014). cDNA Cloning and Heterologous Expression of an Endo-β-1,4-glucanase from The Fungus-Growing Termite *Macrotermes barneyi*. *Archives of Insect Biochemistry and Physiology*, 86(3), pp.151-164.

Nielsen, H. (2017). Predicting Secretory Proteins with SignalP. *Methods in Molecular Biology*, pp.59-73.

Nobile, C. and Johnson, A., 2015. *Candida albicans* Biofilms and Human Disease. *Annual Review of Microbiology*, 69(1), pp.71-92.

Nosanchuk, J. and Casadevall, A. (2006). Impact of Melanin on Microbial Virulence and Clinical Resistance to Antimicrobial Compounds. *Antimicrobial Agents and Chemotherapy*, 50(11), pp.3519-3528.

Noverr, M., Cox, G., Perfect, J. and Huffnagle, G. (2003). Role of PLB1 in Pulmonary Inflammation and Cryptococcal Eicosanoid Production. *Infection and Immunity*, 71(3), pp.1538-1547.

Orengo, C., Jones, D. and Thornton, J. (1994). Protein superfamilles and domain superfolds. *Nature*, 372(6507), pp.631-634.

Padula, M., Berry, I., O′Rourke, M., Raymond, B., Santos, J. and Djordjevic, S. (2017). A Comprehensive Guide for Performing Sample Preparation and Top-Down Protein Analysis. *Proteomes*, 5(4), p.11.

Pan, X., Gilkes, N. and Saddler, J. (2006). Effect of acetyl groups on enzymatic hydrolysis of cellulosic substrates. *Holzforschung*, 60(4), pp.398-401.

Park, B., Wannemuehler, K., Marston, B., Govender, N., Pappas, P. and Chiller, T. (2009). Estimation of the current global burden of cryptococcal meningitis among persons living with HIV/AIDS. *AIDS*, 23(4), pp.525-530.

Patel, A. and Shah, A. (2021). Purification and characterization of novel, thermostable and non-processive GH5 family endoglucanase from *Fomitopsis meliae* CFA 2. *International Journal of Biological Macromolecules*, 182, pp.1161-1169.

Patri, A., Mohan, R., Pu, Y., Yoo, C., Ragauskas, A., Kumar, R., Kisailus, D., Cai, C. and Wyman, C. (2021). THF co-solvent pretreatment prevents lignin redeposition from interfering with enzymes yielding prolonged cellulase activity. *Biotechnology for Biofuels*, 14(1).

Pattenden, L. and Thomas, W. (2008). Amylose Affinity Chromatography of Maltose binding Protein. *Affinity Chromatography*, pp.169-190.

Pazarlioğlu, N., Sariişik, M. and Telefoncu, A. (2005). Treating denim fabrics with immobilized commercial cellulases. *Process Biochemistry*, 40(2), pp.767-771.

Penchala, S., Connelly, S., Wang, Y., Park, M., Zhao, L., Baranczak, A., Rappley, I., Vogel, H., Liedtke, M., Witteles, R., Powers, E., Reixach, N., Chan, W., Wilson, I., Kelly, J., Graef, I. and Alhamadsheh, M. (2013). AG10 inhibits amyloidogenesis and cellular toxicity of the familial amyloid cardiomyopathy-associated V122I transthyretin. *Proceedings of the National Academy of Sciences*, 110(24), pp.9992-9997.

Pereira, J., Chen, Z., McAndrew, R., Sapra, R., Chhabra, S., Sale, K., Simmons, B. and Adams, P. (2010). Biochemical characterization and crystal structure of endoglucanase Cel5A from the hyperthermophilic *Thermotoga maritima*. *Journal of Structural Biology*, 172(3), pp.372-379.

Perfect, J. (2005). *Cryptococcus neoformans*: A sugar-coated killer with designer genes. *FEMS Immunology & Medical Microbiology*, 45(3), pp.395-404.

Perfect, J., Dismukes, W., Dromer, F., Goldman, D., Graybill, J., Hamill, R., Harrison, T., Larsen, R., Lortholary, O., Nguyen, M., Pappas, P., Powderly, W., Singh, N., Sobel, J. and Sorrell, T. (2010). Clinical Practice Guidelines for the Management of Cryptococcal Disease: 2010 Update by the Infectious Diseases Society of America. *Clinical Infectious Diseases*, 50(3), pp.291-322.

Pfeiffer, T. and Ellis, D. (1992). Environmental isolation of *Cryptococcus neoformans var. gattii* from *Eucalyptus tereticornis* . , 30(5), pp.407-408.

Pimentel, A., Ematsu, G., Liberato, M., Paixão, D., Franco Cairo, J., Mandelli, F., Tramontina, R., Gandin, C., de Oliveira Neto, M., Squina, F. and Alvarez, T. (2017). Biochemical and biophysical properties of a metagenome-derived GH5 endoglucanase displaying an unconventional domain architecture. *International Journal of Biological Macromolecules*, 99, pp.384-393.

Pitt-Rivers, R. and Impiombato, F. (1968). The binding of sodium dodecyl sulphate to various proteins, *Biochemical Journal* 109(5), pp.825-830.

Pollock, N., Rai, M., Simon, K., Hesketh, S., Teo, A., Parmar, M., Sridhar, P., Collins, R., Lee, S., Stroud, Z., Bakker, S., Muench, S., Barton, C., Hurlbut, G., Roper, D., Smith, C., Knowles, T., Spickett, C., East, J., Postis, V. and Dafforn, T. (2019). SMA-PAGE: A new method to examine complexes of membrane proteins using SMALP nano-encapsulation and native gel electrophoresis. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1861(8), pp.1437-1445.

Poole, D., Durrant, A., Hazlewood, G. and Gilbert, H. (1991). Characterization of hybrid proteins consisting of the catalytic domains of *Clostridium* and *Ruminococcus* endoglucanases, fused to *Pseudomonas* non-catalytic cellulose-binding domains. *Biochemical Journal*, 279(3), pp.787-792.

Povarova, O., Kuznetsova, I. and Turoverov, K. (2010). Differences in the Pathways of Proteins Unfolding Induced by Urea and Guanidine Hydrochloride: Molten Globule State and Aggregates. *PLoS ONE*, 5(11), p.e15035.

Primhak, R. and Tanner, M. (2001). Alpha-1 antitrypsin deficiency. *Archives of Disease in Childhood*, 85(1), pp.2-5.

Qin, Q., Luo, J., Lin, X., Pei, J., Li, L., Ficht, T. and de Figueiredo, P. (2011). Functional Analysis of Host Factors that Mediate the Intracellular Lifestyle of *Cryptococcus neoformans*. *PLoS Pathogens*, 7(6), p.e1002078.

Rajasingham, R., Smith, R., Park, B., Jarvis, J., Govender, N., Chiller, T., Denning, D., Loyse, A. and Boulware, D. (2017). Global burden of disease of HIV-associated cryptococcal meningitis: an updated analysis. *The Lancet Infectious Diseases*, 17(8), pp.873-881.

Raran-Kurussi, S. and Waugh, D. (2012). The Ability to Enhance the Solubility of Its Fusion Partners Is an Intrinsic Property of Maltose binding Protein but Their Folding Is Either Spontaneous or Chaperone-Mediated. *PLoS ONE*, 7(11), p.e49589.

Rauwane, M., Ogugua, U., Kalu, C., Ledwaba, L., Woldesemayat, A. and Ntushelo, K. (2020). Pathogenicity and Virulence Factors of *Fusarium graminearum* Including Factors Discovered Using Next Generation Sequencing Technologies and Proteomics. *Microorganisms*, 8(2), p.305.

Reed, J. and Kinzel, V. (1984). Near- and far-ultraviolet circular dichroism of the catalytic subunit of adenosine cyclic 5'-monophosphate dependent protein kinase. *Biochemistry*, 23(7), pp.1357-1362.

Ren, P., Springer, D., Behr, M., Samsonoff, W., Chaturvedi, S. and Chaturvedi, V. (2006). Transcription Factor STE12α Has Distinct Roles in Morphogenesis, Virulence, and Ecological Fitness of the Primary Pathogenic Yeast *Cryptococcus gattii*. *Eukaryotic Cell*, 5(7), pp.1065-1080.

Rigsby, R. and Parker, A. (2016). Using the PyMOL application to reinforce visual understanding of protein structure. *Biochemistry and Molecular Biology Education*, 44(5), pp.433-437.

Rodríguez-Ruiz, H., Garibay-Cerdenares, O., Illades-Aguiar, B., Montaño, S., Jiang, X. and Leyva-Vázquez, M. (2019). In silico prediction of structural changes in human papillomavirus type 16 (HPV16) E6 oncoprotein and its variants. *BMC Molecular and Cell Biology*, 20(1).

Rosano, G. and Ceccarelli, E. (2009). Rare codon content affects the solubility of recombinant proteins in a codon bias-adjusted *Escherichia coli* strain. *Microbial Cell Factories*, 8(1), p.41.

Rosano, G. and Ceccarelli, E. (2014). Recombinant protein expression in *Escherichia coli*: advances and challenges. *Frontiers in Microbiology*, 5.

Rosas, A. (1997). Melanization affects susceptibility of *Cryptococcus neoformans* to heat and cold. *FEMS Microbiology Letters*, 153(2), pp.265-272.

Rouau, X. and Odier, E. (1986). Purification and properties of two enzymes from *Dichomitus squalens* which exhibit both cellobiohydrolase and xylanase activity. *Carbohydrate Research*, 145(2), pp.279-292.

Rouvinen, J., Bergfors, T., Teeri, T., Knowles, J. and Jones, T. (1990). Three-dimensional structure of cellobiohydrolase II from *Trichoderma reesei*. *Science*, 249(4967), pp.380-386.

Roy, A. (2017). In Silico Analysis, Structure Modeling and Phosphorylation Site Prediction of Vitellogenin Protein from *Gibelion Catla*. *Journal of Applied Biotechnology & Bioengineering*, 3(1).

Royer, C. A. (2006). Probing Protein Folding and Conformational Transitions with Fluorescence. *Chemical Reviews*, 106(5), 1769–1784.

Royo, J. (2005). A new generation of vectors with increased induction ratios by overimposing a second regulatory level by attenuation. *Nucleic Acids Research*, 33(19), pp.e169-e169.

Saag, M., Graybill, R., Larsen, R., Pappas, P., Perfect, J., Powderly, W., Sobel, J. and Dismukes, W. (2000). Practice Guidelines for the Management of Cryptococcal Disease. *Clinical Infectious Diseases*, 30(4), pp.710-718.

Sabatha, F., Belaich, A. and Soucaille, P. (2002). Characterization of the cellulolytic complex (cellulosome) of *Clostridium acetobutylicum*. *FEMS Microbiology Letters*, 217(1), pp.15-22.

Sachdev, D. and Chirgwin, J. (1998). Order of Fusions between Bacterial and Mammalian Proteins Can Determine Solubility in *Escherichia coli*. *Biochemical and Biophysical Research Communications*, 244(3), pp.933-937.

Sadhu, S. (2013). Cellulase Production by Bacteria: A Review. *British Microbiology Research Journal*, 3(3), pp.235-258.

Saloheimo, M. and Niku-Paavola, M. (1991). Heterologous Production of a Ligninolytic Enzyme: Expression of the *Phlebia Radiata* Laccase Gene in *Trichoderma reesei*. *Bio/Technology*, 9(10), pp.987-990.

Sandomenico, A., Sivaccumar, J. and Ruvo, M. (2020). Evolution of *Escherichia coli* Expression System in Producing Antibody Recombinant Fragments. *International Journal of Molecular Sciences*, 21(17), p.6324.

Sanjaya, R., Putri, K., Kurniati, A., Rohman, A. and Puspaningsih, N. (2021). *Journal of Genetic Engineering and Biotechnology*, 19(1).

Santos, R., Lee, J., Jameel, H., Chang, H. and Lucia, L. (2012). Effects of hardwood structural and chemical characteristics on enzymatic hydrolysis for biofuel production. Bioresource Technology, 110, pp.232-238.

Schellman, J. (1997). Temperature, stability, and the hydrophobic interaction. *Biophysical Journal*, 73(6), pp.2960-2964.

Schmid, F. (2001). *Biological macromolecules: UV-visible spectrophotometry*. 1st ed. Hoboken, NJ, USA: John Wiley & Sons.

Schwarz, W. (2001). The cellulosome and cellulose degradation by anaerobic bacteria. *Applied Microbiology and Biotechnology*, 56(5-6), pp.634-649.

Schwarz, W., Bronnenmeier, K., Gräbnitz, F. and Staudenbauer, W. (1987). Activity staining of cellulases in polyacrylamide gels containing mixed linkage β-glucans. *Analytical Biochemistry*, 164(1), pp.72-77.

Sepcic, K. and Zalar, P. (2010). Low Water Activity Induces the Production of Bioactive Metabolites in Halophilic and Halotolerant Fungi. *Marine Drugs*, 9(1), pp.43-58.

Severo, C., Xavier, M., Gazzoni, A. and Severo, L. (2009). Cryptococcosis in children. *Paediatric Respiratory Reviews*, 10(4), pp.166-171.

Shao, X., Mednick, A., Alvarez, M., van Rooijen, N., Casadevall, A. and Goldman, D. (2005). An Innate Immune System Cell Is a Major Determinant of Species-Related Susceptibility Differences to Fungal Pneumonia. *The Journal of Immunology*, 175(5), pp.3244-3251.

Shapiro-Ilan, D., Fuxa, J., Lacey, L., Onstad, D. and Kaya, H. (2005). Definitions of pathogenicity and virulence in invertebrate pathology. *Journal of Invertebrate Pathology*, 88(1), pp.1-7.

Sharma, P. and Guptasarma, P. (2017). Endoglucanase activity at a second site in *Pyrococcus furiosus* triosephosphate isomerase – Promiscuity or compensation for a metabolic handicap? *FEBS Open Bio*, 7(8), pp.1126-1143.

Sharp, P. and Li, W. (1986). Codon usage in regulatory genes in *Escherichia coli* does not reflect selection for 'rare' codons. *Nucleic Acids Research*, 14(19), pp.7737-7749.

Sher, H., Zeb, N., Zeb, S., Ali, A., Aleem, B., Iftikhar, F. (2021). Microbial Cellulases: A Review on Strain Development, Purification, Characterization and their Industrial Applications. *Journal of Bacteriology and Mycology*. 8(5), pp1180

Siafakas, A., Sorrell, T., Wright, L., Wilson, C., Larsen, M., Boadle, R., Williamson, P. and Djordjevic, J. (2007). Cell Wall-linked Cryptococcal Phospholipase B1 Is a Source of Secreted Enzyme and a Determinant of Cell Wall Integrity. *Journal of Biological Chemistry*, 282(52), pp.37508-37514.

Silaban, S., Gaffar, S., Simorangkir, M., Maksum, I. and Subroto, T. (2019). Effect of IPTG Concentration on Recombinant Human Prethrombin-2 Expression in *Escherichia coli* BL21(DE3) ArcticExpress. *IOP Conference Series: Earth and Environmental Science*, 217, p.012039.

Sindrewicz, P., Li, X., Yates, E., Turnbull, J., Lian, L. and Yu, L. (2019). Intrinsic tryptophan fluorescence spectroscopy reliably determines galectin-ligand interactions. *Scientific Reports*, 9(1).

Singh, A., Upadhyay, V., Upadhyay, A., Singh, S. and Panda, A. (2015). Protein recovery from inclusion bodies of *Escherichia coli* using mild solubilization process. *Microbial Cell Factories*, 14(1).

Siu-Rodas, Y., Calixto-Romo, M., Guillén-Navarro, K., Sánchez, J., Zamora-Briseño, J. and Amaya-Delgado, L. (2018). *Bacillus subtilis* with endocellulase and exocellulase activities isolated in the thermophilic phase from composting with coffee residues. *Revista Argentina de Microbiología*, 50(3), pp.234-243.

Sluchanko, N., Tugaeva, K., Faletrov, Y. and Levitsky, D. (2016). High-yield soluble expression, purification and characterization of human steroidogenic acute regulatory protein (StAR) fused to a cleavable Maltose binding Protein (MBP). *Protein Expression and Purification*, 119, pp.27-35.

Smialowski, P., Doose, G., Torkler, P., Kaufmann, S. and Frishman, D. (2012). PROSO II - a new method for protein solubility prediction. *FEBS Journal*, 279(12), pp.2192-2200.

Soares Júnior, F., Dias, A., Fasanella, C., Taketani, R., Lima, A., Melo, I. and Andreote, F. (2013). Endo-and exoglucanase activities in bacteria from mangrove sediment. *Brazilian Journal of Microbiology*, 44(3), pp.969-976.

Sonan, G. K., Receveur-Brechot, V., Duez, C., Aghajari, N., Czjzek, M., Haser, R., and Gerday, C. (2007) The linker region plays a key role in the adaptation to cold of the cellulase from an Antarctic bacterium. *Biochemistry Journal*. 407, 293– 302

Sorek, N., Yeats, T., Szemenyei, H., Youngs, H. and Somerville, C. (2014). The Implications of Lignocellulosic Biomass Chemical Composition for the Production of Advanced Biofuels. *BioScience*, 64(3), pp.192-201.

Sørensen, H. and Mortensen, K. (2005). Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. *Journal of Biotechnology*, 115(2), pp.113-128.

Sormanni, P., Aprile, F. and Vendruscolo, M. (2015). The CamSol Method of Rational Design of Protein Mutants with Enhanced Solubility. *Journal of Molecular Biology*, 427(2), pp.478-490.

Souza, T., Araujo, J., da Silva, V., Liberato, M., Pimentel, A., Alvarez, T., Squina, F. and Garcia, W. (2016). Chemical stability of a cold-active cellulase with high tolerance toward surfactants and chaotropic agent. *Biotechnology Reports*, 9, pp.1-8.

Springer, D. and Chaturvedi, V. (2010). Projecting Global Occurrence of *Cryptococcus gattii*. *Emerging Infectious Diseases*, 16(1), pp.14-20.

Springer, D., Billmyre, R., Filler, E., Voelz, K., Pursall, R., Mieczkowski, P., Larsen, R., Dietrich, F., May, R., Filler, S. and Heitman, J. (2014). *Cryptococcus gattii* VGIII Isolates Causing Infections in HIV/AIDS Patients in Southern California: Identification of the Local Environmental Source as Arboreal. *PLoS Pathogens*, 10(8), p.e1004285.

Springer, D., Mohan, R. and Heitman, J. (2017). Plants promote mating and dispersal of the human pathogenic fungus Cryptococcus. *PLOS ONE*, 12(2), p.e0171695.

Ståhlberg, J., Johansson, G. and Pettersson, G. (1991). A New Model for Enzymatic Hydrolysis of Cellulose Based on the Two-Domain Structure of Cellobiohydrolase I. *Nature Biotechnology*, 9(3), pp.286-290.

Stalikas, C. (2007). Extraction, separation, and detection methods for phenolic acids and flavonoids. *Journal of Separation Science*, 30(18), pp.3268-3295.

Sterner, R. and Höcker, B. (2005). Catalytic Versatility, Stability, and Evolution of the (βα)8-Barrel Enzyme Fold. *Chemical Reviews*, 105(11), pp.4038-4055.

Stewart, P. and William-Costerton, J. (2001). Antibiotic resistance of bacteria in biofilms. *The Lancet*, 358(9276), pp.135-138.

Sumner, J. and Graham, V. (1921). Dinitrosalycilic acid: A Reagent For The Estimation of Sugar in Normal and Diabetic Urine. *Journal of Biological Chemistry*, 47(1), pp.5-9.

Sun H., Alexander, B., Lortholary, O., Dromer, F., Forrest, G., Lyon, G., Somani, J., Gupta, K., del Busto, R., Pruett, T., Sifri, C., Limaye, A., John, G., Klintmalm, G., Pursell, K., Stosor, V., Morris, M., Dowdy, L., Munoz, P., Kalil, A., Garcia-Diaz, J., Orloff, S., House, A., Houston, S., Wray, D., Huprikar, S., Johnson, L., Humar, A., Razonable, R., Fisher, R., Husain, S., Wagener, M. and Singh, N. (2010) . Unrecognized Pretransplant and Donor-Derived Cryptococcal Disease in Organ Transplant Recipients. *Clinical Infectious Diseases*, 51(9), pp.1062-1069.

Szmelcman, S., Schwarts, M., Silhavy, T. and Boos, W. (1976). Maltose Transport in *Escherichia coli* K12. A Comparison of Transport Kinetics in Wild-Type and lamba-Resistant Mutants with the Dissociation Constants of the Maltose binding Protein as

Measured by Fluorescence Quenching. *European Journal of Biochemistry*, 65(1), pp.13-19.

Tagoe, C., Tagoe, C., Reixach, N., Tagoe, C., Reixach, N., Friske, L., Mustra, D., French, D., Gallo, G. and Buxbaum, J. (2007). In vivostabilization of mutant human transthyretin in transgenic mice. *Amyloid*, 14(3), pp.227-236.

Tan, Y. and Ting, A. (2000). Non-ionic detergent affects the conformation of a functionally active mutant of Bcl-XL. *Protein Engineering, Design and Selection*, 13(12), pp.887-892.

Taylor, S., Ferguson, A., Bergeron, J. and Thomas, D. (2004). The ER protein folding sensor UDP-glucose glycoprotein–glucosyltransferase modifies substrates distant to local changes in glycoprotein conformation. *Nature Structural & Molecular Biology*, 11(2), pp.128-134.

Teather, R. and Wood, P. (1982). Use of Congo red-polysaccharide interactions in enumeration and characterization of cellulolytic bacteria from the bovine rumen. *Applied and Environmental Microbiology*, 43(4), pp.777-780.

Teeri, T. (1997). Crystalline cellulose degradation: new insight into the function of cellobiohydrolases. *Trends in Biotechnology*, 15(5), pp.160-167.

Tewfick, M., Serag, W. and Soliman, B. (2018). Detection of antibacterial protein in *Bacillus sphaericus*-treated *Culex pipiens* (Diptera: Culicidae). *Egyptian Journal of Biological Pest Control*, 28(1).

Thongekkaew, J., Ikeda, H., Masaki, K. and Iefuji, H. (2008). An acidic and thermostable carboxymethyl cellulase from the yeast *Cryptococcus sp*. S-2: Purification, characterization and improvement of its recombinant enzyme production by high cell-density fermentation of Pichia pastoris. *Protein Expression and Purification*, 60(2), pp.140-146.

Trivedi, A., Mavi, P., Bhatt, D. and Kumar, A. (2016). Thiol reductive stress induces cellulose-anchored biofilm formation in *Mycobacterium tuberculosis*. *Nature Communications*, 7(1).

Tseng, C., Ko, T., Guo, R., Huang, J., Wang, H., Huang, C., Cheng, Y., Wang, A. and Liu, J. (2011). Substrate binding of a GH5 endoglucanase from the ruminal fungus *Piromyces rhizinflata*. *Acta Crystallographica Section F Structural Biology and Crystallization Communications*, 67(10), pp.1189-1194.

Tsunemi, T., Kamata, T., Fumimura, Y., Watanabe, M., Yamawaki, M., Saito, Y., Kanda, T., Ohashi, K., Suegara, N., Murayama, S., Makimura, K., Yamaguchi, H. and Mizusawa, H. (2001). Immunohistochemical Diagnosis of *Cryptococcus neoformans* var. *gattii* Infection in Chronic Meningoencephalitis: The First Case in Japan. *Internal Medicine*, 40(12), pp.1241-1244.

Tuller, T., Carmi, A., Vestsigian, K., Navon, S., Dorfan, Y., Zaborske, J., Pan, T., Dahan, O., Furman, I. and Pilpel, Y. (2010). An Evolutionarily Conserved Mechanism for Controlling the Efficiency of Protein Translation. *Cell*, 141(2), pp.344-354.

Ushasree, M., Gunasekaran, P. and Pandey, A. (2012). Single-step Purification and Immobilization of MBP–phytase Fusion on Starch Agar Beads: Application in Dephytination of Soy Milk. *Applied Biochemistry and Biotechnology*, 167(5), pp.981-990.

Valencia Jiménez, A., Wang, H. and Siegfried, B. (2014). Expression and Characterization of a Recombinant Endoglucanase from Western Corn Rootworm, in *Pichia pastoris*. *Journal of Insect Science*, 14(1).

Van Duin, D., Casadevall, A. and Nosanchuk, J. (2002). Melanization of *Cryptococcus neoformans* and *Histoplasma capsulatum* Reduces Their Susceptibilities to Amphotericin B and Caspofungin. *Antimicrobial Agents and Chemotherapy*, 46(11), pp.3394-3400.

Ventorino, V., Ionata, E., Birolo, L., Montella, S., Marcolongo, L., de Chiaro, A., Espresso, F., Faraco, V. and Pepe, O. (2016). Lignocellulose-Adapted Endo-Cellulase Producing *Streptomyces* Strains for Bioconversion of Cellulose-Based Materials. *Frontiers in Microbiology*, 7.

Venyaminov, S. and Yang, J. (1996). Determination of Protein Secondary Structure. *Circular Dichroism and the Conformational Analysis of Biomolecules*, pp.69-107.

Vera, A., González-Montalbán, N., Arís, A. and Villaverde, A. (2007). The conformational quality of insoluble recombinant proteins is enhanced at low growth temperatures. *Biotechnology and Bioengineering*, 96(6), pp.1101-1106.

Vishniac, H. (2006). A Multivariate Analysis of Soil Yeasts Isolated from a Latitudinal Gradient. *Microbial Ecology*, 52(1), pp.90-103.

Vivian, J. and Callis, P. (2001). Mechanisms of Tryptophan Fluorescence Shifts in Proteins. *Biophysical Journal*, 80(5), pp.2093-2109.

Vreulink, J., Boekhout, T., Vismer, H. and Botha, A. (2020). The growth of *Cryptococcus gattii* MATα and MATa strains is affected by the chemical composition of their woody debris substrate. *Fungal Ecology*, 47, p.100943.

Vu, K., Tham, R., Uhrig, J., Thompson, G., Na Pombejra, S., Jamklang, M., Bautos, J. and Gelli, A. (2014). Invasion of the Central Nervous System by *Cryptococcus neoformans* Requires a Secreted Fungal Metalloprotease. *mBio*, 5(3).

Vuong, T. and Wilson, D. (2010). Glycoside hydrolases: Catalytic base/nucleophile diversity. *Biotechnology and Bioengineering*, 107(2), pp.195-205.

Wagner, S., Bader, M., Drew, D. and de Gier, J. (2006). Rationalizing membrane protein overexpression. *Trends in Biotechnology*, 24(8), pp.364-371.

Wang, B., Hu, S., Yu, X., Jin, L., Zhu, Y. and Jin, F. (2020). Studies of Cellulose and Starch Utilization and the Regulatory Mechanisms of Related Enzymes in Fungi. *Polymers*, 12(3), p.530.

Wang, W., Archbold, T., Lam, J., Kimber, M. and Fan, M. (2019). A processive endoglucanase with multi-substrate specificity is characterized from porcine gut microbiota. *Scientific Reports*, 9(1).

Wang, Y. and Casadevall, A. (1994). Decreased susceptibility of melanized *Cryptococcus neoformans* to UV light. *Applied and Environmental Microbiology*, 60(10), pp.3864-3866.

Wang, Y., Aisen, P. and Casadevall, A. (1995). *Cryptococcus neoformans* melanin and virulence: mechanism of action. *Infection and Immunity*, 63(8), pp.3131-3136.

Watkins, R., King, J. and Johnston, S. (2017). Nutritional Requirements and Their Importance for Virulence of Pathogenic *Cryptococcus* Species. *Microorganisms*, 5(4), p.65.

White, A. and Rose, D. (1997). Mechanism of catalysis by retaining β-glycosyl hydrolases. *Current Opinion in Structural Biology*, 7(5), pp.645-651.

Wierenga, R. (2001). The TIM-barrel fold: a versatile framework for efficient enzymes. *FEBS Letters*, 492(3), pp.193-198.

Wilkins, M., Gasteiger, E., Sanchez, J., Bairoch, A. and Hochstrasser, D. (1998). Two-dimensional gel electrophoresis for proteome projects: The effects of protein hydrophobicity and copy number. *Electrophoresis*, 19(8-9), pp.1501-1505.

Wu, H., Wang, A. and Jennings, M. (2008). Discovery of virulence factors of pathogenic bacteria. *Current Opinion in Chemical Biology*, 12(1), pp.93-101.

Xia, W., Liu, P. and Liu, J. (2008). Advance in chitosan hydrolysis by non-specific cellulases. *Bioresource Technology*, 99(15), pp.6751-6762.

Xu, F. and Li, Y. (2017). Biomass Digestion. *Encyclopedia of Sustainable Technologies*, pp.197-204.

Xu, Y., Liu, K., Han, Y., Xing, Y., Zhang, Y., Yang, Q. and Zhou, M. (2021). Codon usage bias regulates gene expression and protein conformation in yeast expression system *P. pastoris*. *Microbial Cell Factories*, 20(1).

Yan, S. and Wu, G. (2013). Secretory pathway of cellulase: a mini-review. *Biotechnology for Biofuels*, 6(1), p.177.

Yang, J. and Dang, H. (2011). Cloning and characterization of a novel cold-active endoglucanase establishing a new subfamily of glycosyl hydrolase family 5 from a psychrophilic deep-sea bacterium. *FEMS Microbiology Letters*, 325(1), pp.71-76.

Yang, Q., Yu, C., Zhao, F., Dang, Y., Wu, C., Xie, P., Sachs, M. and Liu, Y. (2019). eRF1 mediates codon usage effects on mRNA translation efficiency through premature termination at rare codons. *Nucleic Acids Research*, 47(17), pp.9243-9258.

Yuan, Y., Zhang, X., Zhang, H., Wang, W., Zhao, X., Gao, J. and Zhou, Y. (2020). Degradative GH5 β-1,3-1,4-glucanase PpBglu5A for glucan in *Paenibacillus polymyxa* KF-1. *Process Biochemistry*, 98, pp.183-192.

Zarafeta, D., Kissas, D., Sayer, C., Gudbergsdottir, S., Ladoukakis, E., Isupov, M., Chatziioannou, A., Peng, X., Littlechild, J., Skretas, G. and Kolisis, F. (2016). Discovery and Characterization of a Thermostable and Highly Halotolerant GH5 Cellulase from an Icelandic Hot Spring Isolate. *PLOS ONE*, 11(1), p.e0146454.

Zavala, S. and Baddley, J. (2020). Cryptococcosis. *Seminars in Respiratory and Critical Care Medicine*, 41(01), pp.069-079.

Zhang, D., Jing, X., Fan, A., Liu, H., Nie, Y. and Xu, Y. (2020). Active Expression of Membrane-Bound L-Amino Acid Deaminase from Proteus mirabilis in Recombinant Escherichia coli by Fusion with Maltose binding Protein for Enhanced Catalytic Performance. *Catalysts*, 10(2), pp.215-226.

Zhang, Y., Zhou, J., Ardejani, M., Li, X., Wang, F. and Orner, B. (2017). Designability of Aromatic Interaction Networks at *E. coli* Bacterioferritin B-Type Channels. *Molecules*, 22(12), pp.2184-2195.

Zheng, F. and Ding, S. (2013). Processivity and Enzymatic Mode of a Glycoside Hydrolase Family 5 Endoglucanase from *Volvariella volvacea*. *Applied and Environmental Microbiology*, 79(3), pp.989-996.

Zheng, F., Tu, T., Wang, X., Wang, Y., Ma, R., Su, X., Xie, X., Yao, B. and Luo, H. (2018). Enhancing the catalytic activity of a novel GH5 cellulase GtCel5 from *Gloeophyllum trabeum* CBS 900.73 by site-directed mutagenesis on loop 6. *Biotechnology for Biofuels*, 11(1). pp.1-13

Zhu, X. and Williamson, P. (2004). Role of laccase in the biology and virulence of. *FEMS Yeast Research*, 5(1), pp.1-10.

Zhu, X. and Williamson, P. (2003). A CLC-type chloride channel gene is required for laccase activity and virulence in *Cryptococcus neoformans*. *Molecular Microbiology*, 50(4), pp.1271-1281.

Zoghlami, A. and Paës, G. (2019). Lignocellulosic Biomass: Understanding Recalcitrance and Predicting Hydrolysis. *Frontiers in Chemistry*, 7. pp.1-11

Appendix 1

MRLLSPFFIVPLAVVSALPALQTVDDQKRSVNVGWPYGTNKIRGVNIGGWLVTEP
FITPSLFEATGNNDIVDEWTFCQYQDYNTAQAALKNHWDTWFTEDDFARIAAAG
LNHVRIPIGFWAYDVQGGEPYIQGQADYLDRAIGWARNHNLAVIIDLHGAPGSQ
NGYDNSGRRGNADWATDNTNVERTKNVIAQLSQKYSDPQYYGVVTALALLNEP
ATYLNDQLLQTARQYWYDAYGAARYPFGNSDKSGLALVIHDGFQPLSTFDSYM
VEPEFEDVLLDTHNYQVFNDEYVAWNWDQHISSVCNLASTYSSSPLWLVVGEW
SLASTDCAKYLNGRGINARYDGSYPGSSYIGSCEDKSNDVSKFSNEYKDFMHKF
WNVQTQLYEQNGQGWIHWTWKTESAADWSYEAGLDGGWIPWDAGSHDVSLSS
LCG

Figure A1: Amino acid sequence of the native cellulase from *Cryptococcus gattii* (WM276). The pink segment represents the cellulase's signal peptide, which targets the enzyme for extracellular secretion. The cellulase's amino acid sequence was extracted from GenBank (accession number: XM_003194284).

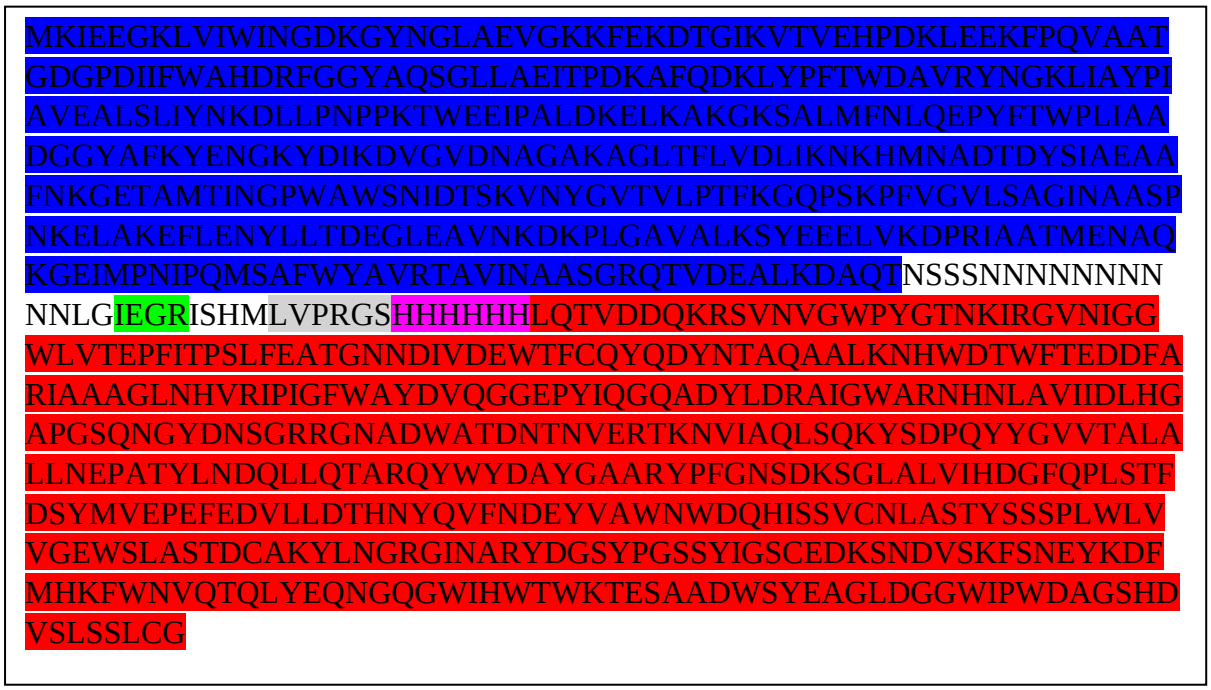

Figure A2: Amino acid sequence of maltose binding protein-cellulase (MBP-cellulase). Blue segment represents the MBP component of the fusion enzyme. The green segment and grey segment represent cleavage sites for Factor Xa and thrombin, respectively. The pink segment represents the hexahistidine tag and the red segment represents the cellulase component of the fusion enzyme.

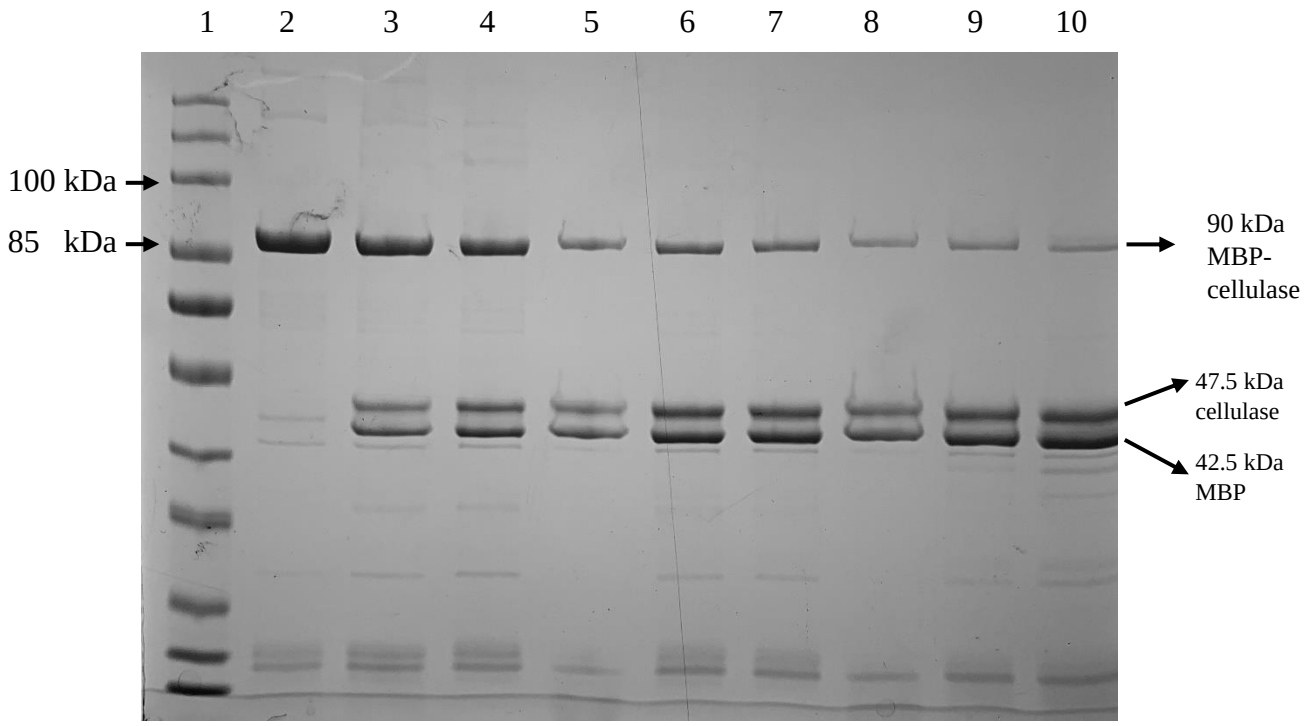

Figure A3: Coomassie stained sodium dodecyl sulphate polyacrylamide (SDS-PAGE) gel illustrating thrombin cleavage trials of maltose binding protein-cellulase at various time points at room temperature (20 ˚C). Lane 1: unstained Protein Standard (New England Biolabs, United States); lane 2: Uncleaved MBP-cellulase; lane 3: cleavage reaction after 5 min; lane 4: cleavage reaction after 10 min; lane 5: cleavage reaction after 30 min; lane 6: cleavage reaction after 1 h; lane 7: cleavage reaction after 2 h; lane 8: cleavage reaction after 4 h; lane 9: cleavage reaction after 8 h; lane 10: cleavage reaction after 24 h.