Investigation of the structure/function relationships of a chitobiase from *Photorhabdus luminescens* through mutagenic strategies

Thesis submitted for the degree of

Doctor of Philosophy

by

Amy Harrington, B.Sc.

Supervised by

Brendan O'Connor, B.Sc., Ph.D

School of Biotechnology

Dublin City University

June 2013

Authors Declaration

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of Doctor of Philosophy is entirely my own work, that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge breach any law of copyright, and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the the text of my work.

Signed:_____ (Amy Harrington)

ID No.: 55499372_____

Date: 16/9/2013_____

Acknowledgements

Firstly, I would like to thank my supervisor Dr. Brendan O' Connor for the opportunity to pursue this PhD. Your enthusiasm, positivity, encouragement and support got me through the toughest of times.

A huge thank you has to go to everyone in the lab and the ISSC, but especially to Paul Clarke, Roisin Thompson and Ruth Larragy. For your guidance and advice, for being a constant source of knowledge, for always being willing to answer any questions I may have and never being too busy to help, without which I would never have made it this far.

I would also like to thank all my great friends, especially Samantha, Johnny and my long-suffering housemates, Lisa, Mark and Andrew. For all the chats, advice and encouragement. Those long lunch-breaks and tea-breaks helped me to keep my sanity.

Finally I would like to thank my family for their constant support and encouragement. Your unwavering faith in me was a great source of comfort. I hope I have made you proud.

Abbreviations

2D	Two dimensional
aa	Amino acids
Ala	Alanine
Amp^r	Ampicillin resistant
APS	Ammonium persulfate
Asp	Aspartate
bp	Base pairs
BSA	Bovine serum albumin
CBM	Carbohydrate-binding module
CBP	Carbohydrate-binding protein
Chb	Chitobiase
CHO	Chinese hamster ovary
dH_2O	Distilled water
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNS	Dinitrosalicylic acid
DSL	*Datura stramonium* Lectin
ECL	*Erythrina cristagalli* Lectin
EDTA	Ethylene diamine tetra acetic acid
ELLA	Enzyme linked lectin assay
EPO	Erythropoietin
ER	Endoplasmic reticulum
FPLC	Fast protein liquid chromatography
GFP	Green fluorescent protein
GlcNAc	N-acetyl-D-glucosamine
Glu	Glutamate
GSL II	*Griffonia simplicifolia* Lectin II
His	Histidine
HPLC	High performance liquid chromatography
IL	Interleukin
IPTG	Isopropyl-β-D-thiogalactopyranoside

kb	Kilobase
Lys	Lysine
MBP	Maltose-binding protein
MS	Mass Spectrometry
MW	Molecular Weight
MSC	Multiple Cloning Site
NMR	Nuclear Magnetic Resonance
NP-HPLC	Normal Phase HPLC
OD	Optical Density
ORF	Open Reading Frame
PAGE	Poly-acrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase Chain Reaction
PNP	para Nitrophenyl
RP-HPLC	Reversed Phase HPLC
rpm	Revolutions per minute
SDS	Sodium dodecyl sulphate
SNA	*Sambucus nigra* Lectin
TEMED	Tetramethylethylenediamine
TLC	Thin Layer Chromatography
TMB	3,3,5,5-Tetramethylbenzidine
Tris	Tris(hydroxymethyl)aminomethane
Trp	Tryptophan
UV	Ultraviolet
WGA	Wheat germ agglutinin
WT	Wild type

Table of Contents

List of Figures

Chapter 1

Chapter 2

Chapter 6

List of Tables

Abstract

Chitin is the second most abundant biopolymer in the natural world and is hydrolysed through the action of chitinases and chitobiases. The hydrolysis of chitin yields commercially important products such as chito-oligosaccharides and N-acetylglucosamine. The hydrolysis of chitin waste is also a major industry. Enzymatic hydrolysis of chitin is the best and most environmentally friendly method of chitin breakdown. There is a constant demand for easy to produce, fast-performing, reliable chitinolytic enzymes.

This study identified and investigated a chitobiase (Chb) from *Photorhabdus luminescens.* The Chb enzyme was cloned, expressed, purified and its activity characterised. The proposed catalytic residues were mutated and the catalytic activity removed to (i) highlight their importance in catalysis and (ii) ascertain whether the specific GlcNAc binding is retained. The binding abilities of this catalytically inactive carbohydrate-binding protein were investigated to determine if it could be used as a "lectin" for the analysis of glycan patterns on glycoproteins. Residues were selected in the active site for site-directed mutagenesis to study their effect on catalytic activity. Three mutants were generated that exhibited enhanced catalytic activity as well as enhanced pH and temperature stability in comparison with the wildtype Chb. HPLC analysis was performed investigating the hydrolytic breakdown of various length chito-oligosaccharides by the Chb enzyme. The results highlighted that the Chb enzyme and subsequent mutant enzymes are industrially applicable for chitin hydrolysis and the production of the biologically relevant sugar N-acetylglucosamine.

Chapter 1

Introduction

1.1 Chitin

Chitin is the second most abundant biopolymer in the natural world. It has an annual turnover of approximately 10 gigatons, a total only exceeded by cellulose (Meibom *et al.*, 2004). Due to its inflexible nature chitin is a structural polysaccharide found in a variety of different organisms, from the exoskeletons of insects and the outer shells of crustaceans to the cell wall of fungi (Khoushab and Yamabhai, 2010; Rinaudo, 2006).

1.1.1 Chitin structure

Chitin is an insoluble, crystalline polysaccharide comprised of linear un-branched molecular chains of β1,4 linked N-D-acetylglucosamine residues. Chitin is structurally similar to cellulose, differing only at the C-2 of each glucose/N-acetylglucosamine residue. Cellulose has a hydroxyl group (OH) at this position while chitin has an N-acetyl amino group ($NHCOCH_3$).

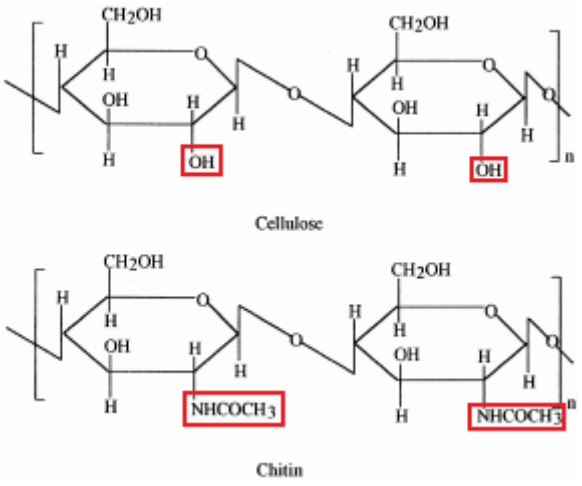


Fig 1.1: The structural differences between cellulose and chitin. The presence of the hydroxyl group at the C-2 position of cellulose and the N-acetyl group at the C-2 position of chitin are highlighted in red (Kumar, 2000).

In nature, chitin is found as crystalline fibrils. It is composed of chains of N-acetylglucosamine, thousands of residues in length. These chains are held together by hydrogen bonds. There are two main forms of chitin: α-chitin and β-chitin. The various forms of chitin may be differentiated between using NMR and X-ray

diffraction. The different forms of chitin are categorised by how the long N-acetylglucosamine chains are organised (Rinaudo, 2006). α-chitin is far more prevalent in nature than β–chitin. While α–chitin is found in an abundance of organisms, from insects and shrimp to fungi and bacteria, β-chitin is found only in squid pens and cocoons (Beckham and Crowley, 2011; Sikorski *et al.*, 2009; Rinaudo, 2006;). The crystal structure of α–chitin was first described by Minke and Blackwell (1978). They described α–chitin as having anti-parallel chains which are held together by hydrogen bonds, formed by interactions between the C=O group of one chain and the amide group of another. The anti-parallel organisation of the chains and the strong inter-chain hydrogen bonds allow the α–chitin chains to be arranged tightly, giving rise to its rigid, insoluble structure (Merzendorfer and Zimoch, 2003). The α–chitin structure was more recently examined by Sikorski *et al.*, (2009) as the presence of anti-parallel chains was a stark contrast to the structure of naturally occurring cellulose, whose chains are aligned in a parallel fashion. Again, the chains were found to be aligned in an anti-parallel fashion and it was noted that they formed a helical configuration (Sikorski *et al.*, 2009).

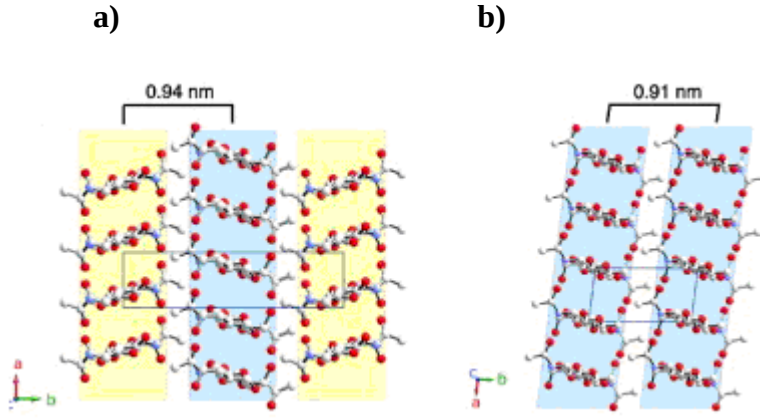


Fig 1.2: The structural differences between a) α–chitin and b) β–chitin. a) The α–chitin chains are organised in an anti-parallel fashion (Sikorski *et al.*, 2009). b) The β–chitin chains are organised in a parallel fashion (Nishiyama *et al.*, 2011).

In comparison with α–chitin, the rarer β-chitin chains were found to be organised in a parallel fashion (Sugiyama *et al.*, 1999). The degree of hydrogen bonding between the chains was determined to be less than that found in α–chitin, allowing hydrogen bonds with water to also form. While β–chitin is also found as crystalline fibrils in

nature, the fibrils themselves are much weaker than in α-chitin (Merzendorfer and Zimoch, 2003; Sugiyama *et al.*, 1999). Jang *et al.*, (2004) investigated the structure of each form of chitin using Fourier transform infrared (FTIR) and Nuclear Magnetic Resonance (NMR) spectroscopy. From this study a third type of chitin, γ-chitin, was proposed, composed of two parallel chains and one anti-parallel chain (Jang *et al.*, 2004). However, the structure of γ-chitin has been debated and it has been proposed that it is just a mixture of α–chitin and β–chitin (Tharanathan and Kittur, 2003).

1.1.2 Chitin synthesis

Chitin is synthesized by the organisms in which it is a structural component and although the chitin synthesis pathway is slightly different in each organism the overall sequence is similar. Glucose is converted to UDP-N-acetylglucosamine, the building block of chitin, in a series of steps. Then, the chitin synthase family of enzymes sequentially add these activated UDP-N-acetylglucosamine residues to the growing chitin chain (Khoushab and Yamabhai, 2010; Merzendorfer and Zimoch, 2003; Tharanathan and Kittur, 2003).

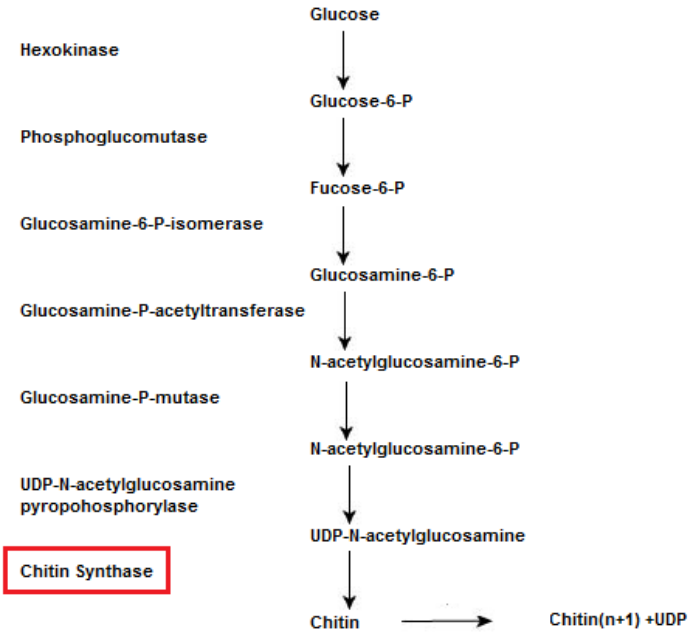


Fig 1.3: The conversion of glucose to UDP-N-acetylglucosamine. Glucose is converted to UDP-N-acetylglucosamine, the building block of chitin, through a series of steps. Chitin synthase then uses this substrate to synthesise chitin (Merzendorfer and Zimoch, 2003).

The production of chitin in fungi is compartmentalised. Chitin synthase is produced as an inactive precursor, and transported to the cell surface in vehicles known as chitosomes where it is then cleaved to yield the active form of the enzyme. It can then catalyze the formation of chitin from the readily available UDP-N-acetylglucosamine (Lenardon *et al.,* 2010; Tharanathan and Kittur, 2003). It is believed that the production of chitin in insects needs a primer, although the exact mechanism is still unknown. It was proposed that an N-acetylglucosamine residue is attached to a lipid carrier such as dolichol phosphate to create dolichyldiphospho-N-acetylglucosamine which may then act as the primer for chitin synthesis (Merzendorfer and Zimoch, 2003; Tharanathan and Kittur, 2003).

1.1.3 Applications

1.1.3.1 Applications of chitin

Because of its wide availability and potential as a renewable bioresource, the use of chitin for various applications has been extensively studied. Processing of crustaceans, such as shrimp and crab, lead to the production of $1x10^9$ metric tons of chitin waste each year (Jeon *et al.*, 2007). Chitin is relatively non-toxic and non-reactive. Its insoluble, rigid structure allows it to be used as a platform on which to immobilise enzymes or whole cells. Other uses include as a wound dressing, as absorbable sutures, as a delivery-system for the controlled release of drugs and as a film/fibre. However its lack of solubility is a major limitation (Rinaudo, 2006; Dutta *et al.,* 2004). A much wider variety of applications are possible for chito-oligosaccharides and N-acetylglucosamine. These shorter polysaccharide chains have bioactive properties that chitin does not possess. This is due to differing physical characteristics, particularly the fact that they are water soluble due to their shorter length.

1.1.3.2 Applications of chito-oligosaccharides

1.1.3.2.1 Potential medical applications of chito-oligosaccharides

Chito-oligosaccharides possess many important biological relevant properties. Their anti-tumour properties have been studied since the 1970s and they have shown great potential as anti-cancer agents because, as opposed to conventional treatments such as chemotherapy, they are non-toxic and non-harmful to the patient. However, these are mainly cell culture and animal-based studies (Aam *et al.,* 2010). The anti-cancer properties of chito-oligosaccharides are proposed to be linked with their ability to stimulate the immune system due to their high molecular weight (Suzuki *et al.,* 2006). Shen *et al.,* (2009) described the inhibitory properties of chito-oligosaccharides on tumour growth and metastasis in hepatocellular carcinoma (liver cancer). They found that treatment of hepatocarcinoma cells with chito-oligosaccharides reduced the rate of cell proliferation and the rate of DNA synthesis *in vitro*. Further *in vivo* studies demonstrated that chito-oligosaccharides reduced the rate of tumour growth and increased the rate of survival in a mouse model (Shen *et al.,* 2009). It is believed that these anti-cancer effects of chito-oligosaccharides were due to the activation of an immune response. Chito-oligosaccharides, such as chitohexaose, were found to impact on the activity of intestinal intraepithelial lymphocytes. Their water-solubility allowed them to be readily absorbed in an *in vivo* mouse model, consequentially enhancing the natural killer activity of intestinal lymphocytes resulting in an inhibition of tumour growth (Maeda and Kimura, 2004).

Chito-oligosaccharides are also known to induce apoptosis in certain cancer cells, as described in a study on liver cancer by Xu *et al.,* (2008). This *in vitro* study on hepatocellular carcinoma cells discovered that the presence of chito-oligosaccharides caused an increase in the levels of Bax, a protein known to induce apoptosis. This resulted in an increase in apoptosis for the hepatocellular carcinoma cells.

Wu *et al.,* (2008) described the anti-angiogenesis properties of chito-oligosaccharides. Angiogenesis, the production of new blood vessels from older blood vessels, is the method by which tumours develop their own blood supply. It allows them to grow and metastasise. Inhibition of angiogenesis has long been a target of anti-cancer therapeutics. Wu *et al.,* (1998) found that chito-oligosaccharides had the ability to

inhibit angiogenesis possibly through the down regulation of pro-angiogenic factors or through the blockage of pro-angiogenic signal transmission.

Additionally, anti-inflammatory effects of chito-oligosaccharides have been observed. *In vitro* studies on macrophages by Lee *et al.,* (2009) described that chito-oligosaccharides inhibited the production of nitric oxide (NO), PGE_2 IL-6 and TNF-α (all pro-inflammatory mediators) via the down-regulation of their transcription. Yang *et al.,* (2010) also noted the inhibition of the expression of various inflammatory mediators, again including NO and PGE_2, in chito-oligosaccharide treated macrophages.

1.1.3.2.2 Potential anti-bacterial/anti-fungal applications of chito-oligosaccharides

Chito-oligosaccharides have also been shown to possess certain anti-bacterial and anti-fungal properties. A critical step in bacterial colonisation is the adherence of the bacterium to the host cell, frequently involving glycan epitopes. Chito-oligosaccharides have been found to bind these bacterial carbohydrate binding proteins, therefore inhibiting adhesion to host cells (Aam *et al.,* 2010). Chito-oligosaccharides were found to inhibit the binding of certain intestinal bacterial strains to colorectal cells. Binding of pathogenic *E. coli* strains was observed to be inhibited, however *Lactobacillus casei* and *L. gasseri* strains were unaffected. This demonstrated the use of chito-oligosaccharides as a selective anti-adhesion therapy (Rhoades *et al.,* 2006).

The use of chito-oligosaccharides as chitinase inhibitors has been described (Aam et al., 2010). *Plasmodium*, the causative bacteria of malaria, requires entry into the mosquito gut as part of its infective life-cycle. To do this the bacteria must first breakdown the chitin-containing periotrophic membrane, and *Plasmodium* expresses and secretes chitinases for this purpose. Chito-oligosaccharides may therefore be used as chitinase inhibitors to prevent infection of the mosquito by *Plasmodium* (Aam *et al.,* 2010).

Park *et al.,* (2008) demonstrated the anti-fungal activity of chito-oligosaccharides. Using confocal microscopy it was observed that an association between the chito-oligosaccharides and lipids in the fungal plasma membrane caused a disruption in this membrane. While lysis was not observed, morphological changes to the fungus did

occur. Chito-oligosaccharides can also be used as anti-fungal treatment for plants. The presence of these oligosaccharides causes the induction of various plant chitinases which break down the fungal cell wall (Tharanathan and Kittur, 2003).
Chito-oligosaccharides have also been used in cosmetics, in food (as an anti-microbial and preservative agent), in the treatment of hypocholesterolemia and to increase the absorption of iron and calcium *in vivo* (Xia *et al.,* 2011).

1.1.3.3 Potential applications of N-acetylglucosamine

N-acetylglucosamine, the smallest unit of the chitin polymer, is also a biologically important monosaccharide. Complexed with N-acetylmuramic acid, N-acetylglucosamine forms murein, a constituent of bacterial cell walls (Chen *et al.*, 2010). Hyaluronic acid, composed of N-acetylglucosamine and glucuronic acid, is a large component of the extracellular matrix, while also being a major part of cartilage, connective tissue and synovial fluid. Cell surface N-acetylglucosamine plays a role in signal transduction, antibody recognition and cell-cell interactions (Ashry and Aly, 2007). Just as there are a large variety of biologically relevant functions of N-acetylglucosamine, there are also a large number of applications for this important monosaccharide.
Glucosamine and N-acetylglucosamine are both used for the treatment of degenerative joint disorders such as osteoarthritis. Osteoarthritis is a type of arthritis where the cartilage around the joints gradually deteriorates over time and the synovial membrane becomes inflamed. The general course of treatment is the management of symptoms with the use of pain-management drugs such as analgesics and anti-inflammatory drugs. However, glucosamine and N-acetylglucosamine are proposed to slow the progression of the disease. Both are components of cartilage and synovial fluid (Shikhman *et al.,* 2005; Moskowitz, 2001). N-acetylglucosamine was patented as a therapeutic for the treatment of joint disorders such as arthritis in 1972 (Rovati *et al.,* 1972). Glucosamine has been found to reduce the progression of the narrowing of the joints in patients and therefore slow the progression of the disorder (Reginster *et al.,* 2001). The ability of N-acetylglucosamine versus glucosamine to inhibit the production of pro-inflammatory mediators was examined by Shikman *et al.,* (2001). N-acetylglucosamine was observed to be a more potent inhibitor of IL1β production by chondrocytes than glucosamine. They furthered this work in 2004 when they found

that N-acetylglucosamine reduced levels of cartilage deterioration in rabbits with experimental osteoarthritis. They also observed anti-inflammatory effects with levels of synovial membrane inflammation reduced (Shikhman *et al.,* 2004). Tamai *et al.,* (2003) studied the affect of N-acetylglucosamine on cartilage injury in rabbits. Administration of N-acetylglucosamine was found to cause an increase in proliferation of chondrocytes around the damaged cartilage, in turn leading to enhanced regeneration of the cartilage in comparison with the controls (Tamai *et al.,* 2003). N-acetylglucosamine now plays a huge role in the functional foods market. It is widely taken as a nutritional supplement for joint stiffness, joint pain and management of disorders such as arthritis.

N-acetylglucosamine has also shown positive results in the treatment of Irritable Bowel Syndrome. In 1993 it was patented as a gastroprotective agent (Burtan and Freeman, 1993). Burtan found that individuals with inflammatory bowel disease lacked N-acetylglucosamine which, as an amino sugar, is part of the protective layer of glycosaminoglycans in the intestinal mucosa. Rhodes (1996) proposed that one of the causes of inflammatory bowel disease was a change in mucosal glycosylation. Salvatore *et al.,* (2000) examined the effect of N-acetylglucosamine in paediatric inflammatory bowel disease. It was found that supplementation of N-acetylglucosamine resulted in an increase in glycosaminoglycans at mucosal surfaces such as the epithelia which resulted in improved patient symptoms.

N-acetylglucosamine has also been used topically in cosmetics (Bissett, 2006). Hyaluronic acid, a glycosaminoglycan comprised of N-acetylglucosamine and glucuronic acid, is a component of connective tissue. Due to its effect on cell proliferation, migration and angiogenesis it is known to promote wound healing. It hydrates the skin and improves its' elasticity, reducing the formation of wrinkles. As a component of hyaluronic acid, N-acetylglucosamine was examined as a cosmetic supplement and was been found to also increase skin hydration and stimulate the healing of wounds (Bissett, 2006). Bissett *et al.,* (2006) also found that the application of N-acetylglucosamine to the skin reduced facial hyperpigmentation. It was proposed that this was through its inhibition of the production of tyrosinase, an oxidase that catalyses a step in the formation of melanin.

1.1.4 Industrial chitin hydrolysis for the production of chito-oligosaccharides and N-acetylglucosamine

Industrial chitin hydrolysis may be achieved through a number of different methods including physical, chemical and enzymatic breakdown (Prashanth and Tharanathan, 2007).

1.1.4.1 Chemical hydrolysis of chitin

Chitin may be hydrolysed into smaller length chito-oligosaccharides using chemical means. Acid hydrolysis is one of the most common chemical industrial methods for chitin breakdown (Aam *et al.,* 2010). There are many steps to this process; removal of contaminants, removal of minerals, acid hydrolysis, neutralisation, fractionation of products and lyophilisation. Chitin is derived from shrimp and crab waste for this purpose. This chitin waste also includes numerous contaminants such as proteins, lipids and minerals which need to be removed prior to chitin hydrolysis. A low concentration of sodium hydroxide (3% NaOH) is added to the chitin waste to eliminate any protein present followed by the addition of a low concentration of acid such as 3% HCl, to remove any minerals, such as calcium, present. Chitin may then either be de-acetylated to produce chitosan through the addition of a higher concentration of NaOH (40%) or immediately treated with a strong acid (up to 40% HCl) to produce various length chito-oligosaccharides through hydrolysis of the glycosidic bond. The temperature and acid concentration must be cautiously chosen to ensure breakdown to chito-oligosaccharides and N-acetylglucosamine without degradation of the final product (Jeon *et al.,* 2000; Kim and Rajapakse, 2005; Prashanth and Tharanathan, 2007; Chen *et al.,* 2010). While chemical hydrolysis is commonly used as it a relatively quick process, there are many disadvantages. The final product is a random mixture of various length chito-oligosaccharides and N-acetylglucosamine monomers. The yields also tend to be quite low, with the production of large amounts of toxic waste leading to environmental concerns (Prasanth and Tharanathan, 2007).

1.1.4.2 Physical hydrolysis of chitin

Radiation methods have been used as a method for the hydrolysis of chitin. This method was proposed as a way of reducing the toxic waste produced by the chemical hydrolysis of chitin. Radiation is applied to chitin in an acetic acid solution to hydrolyse the glycosidic linkages and generate shorter length chito-oligosaccharides. The use of ultrasound and thermal treatments have also been proposed as physical methods for the hydrolysis of chitin (Prashanth and Tharanathan, 2007).

1.1.4.3 Enzymatic hydrolysis of chitin

The enzymatic hydrolysis of chitin has many advantages over the chemical hydrolysis methods. It gives a more defined product and is a more controlled process. There is very little if any toxic waste generated and therefore it is a more environmentally friendly process than chemical/acid hydrolysis. Enzymatic hydrolysis of chitin yields the most specific length chito-oligosaccharide chains and the purest form of N-acetylglucosamine. However, it is a much costlier process. The re-use and immobilisation of the chitinolytic enzymes has been found to reduce this cost (Prashanth and Tharanathan, 2007; Kim and Rajapaske, 2005).

The β-1,4 glycosidic bond between the N-acetylglucosamine residues is hydrolysed through the action of enzymes called chitinases and chitobiases. There are two types of chitinase: endo-acting and exo-acting. The choice of enzyme depends on the final product required. For example, to generate N-acetylglucosamine from chitin, exo-acting chitinases and chitobiases would primarily be used. In comparison, the generation of mid-length chito-oligosaccharides would demand the use of endo-acting chitinases rather than exo-acting chitinases (Dahiya *et al.,* 2006). The reaction conditions, such as incubation time, temperature and enzyme components may be tailored to generate an optimal yield of the specific length chito-oligosaccharides required (Patil *et al.,* 2000). Initially, the enzymatic hydrolysis of chitin was performed in batch reactors. This was a costly process, which did not allow the re-use of the hydrolytic enzymes. The length of chito-oligosaccharides produced was also found to be not tightly controlled. These problems have been solved by enzyme immobilisation, which has allowed the re-use of enzymes. This has significantly reduced the cost of this process and has also facilitated more control over the reaction.

Hydrolysis may be stopped immediately and therefore the length of chito-oligosaccharides generated can be more easily regulated (Kim and Rajapaske, 2005). Ramirez-Coutino *et al.,* (2006) demonstrated the use of chitinases from *Lecanicillium fungicola* in the production of chito-oligosaccharides. They found that these endo-acting chitinases had a higher preference for α-chitin over β–chitin and preferred the de-acetylated form over the acetylated form showing the importance of starting material and enzyme choice in the hydrolysis of chitin. de Assis *et al.,* (2010) employed a fungal chitinase (from *Metarhizum anisopliae*) for the hydrolysis of chitin. This fungal chitinase produced chito-oligosaccharides ranging from 2-6 residues in chain length.

1.1.4.4 Purification and characterisation of chito-oligosaccharides

The final product of hydrolysed chitin is usually a composition of various length chito-oligosaccharides which need to be characterised and separated based on their chain length. The ideal end product is a series of separate fractions, each composed of one specific length chito-oligosaccharide. Various methods have been employed for the separation of chito-oligosaccharides based on size, including filtration methods such as ultrafiltration, mass spectrometry and also chromatographic methods such as Size Exclusion Chromatography (SEC). The chito-oligosaccharides can then be characterised using Nuclear Magnetic Resonance (NMR) or Mass Spectrometry (MS). NMR can elucidate the degree of acetylation of the oligosaccharides while MS can also elucidate the chain length (Aam *et al.,* 2010). Sorbotten *et al.,* (2005) successfully used Size Exclusion Chromatography (SEC) to separate the different length chito-oligosaccharides generated from the enzymatic hydrolysis of chitosan by a chitinase from *S. marcescens*. Ultra-filtration was successfully used by Ngo *et al.,* (2009) to separate chito-oligosaccharides based on their molecular weight.

1.1.5 Chitin active enzymes

1.1.5.1 Chitin Synthase

Chitin synthases are part of the glycosyltransferase family. Glycosyltransferases are involved in the formation of glycosidic bonds through the transfer of a sugar donor to

a sugar acceptor, generally resulting in the formation of a sugar chain (polysaccharide). In other cases, they can catalyse the transfer of a sugar donor to non-sugar acceptor such as a lipid or a protein (Henrissat and Davies, 2000). Chitin synthases are grouped together based on the fact that they synthesise chitin, generally through the addition of the activated donor sugar UDP-N-acetylglucosamine to an N-acetylglucosamine acceptor (Imai *et al.*, 2003). Glycosyltransferases such as chitin synthase can exhibit a modular nature due to the presence of two independent domains; a catalytic domain and a binding domain. While the catalytic domain is involved in the synthesis of glycosidic bonds, the function of the binding domain is to bind insoluble substrates like chitin. The function of this will be explained more under carbohydrate binding proteins (see Section 1.1.5.3). Glycosyltransferases have been industrially useful for the production of oligosaccharides (Palcic, 2011; Merzendorfer, 2006; Henrissat and Davies, 2000).

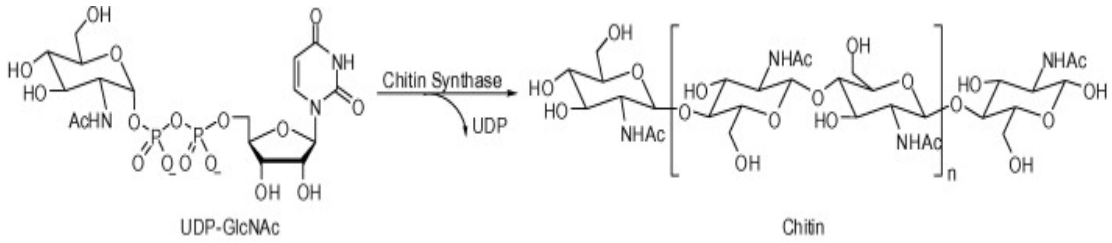


Fig 1.4: Chitin synthase catalysed reaction of chitin synthesis. UDP-N-acetylglucosamine is added to the non-reducing end of the growing chitin chain (Imai *et al.*, 2003).

1.1.5.2 Chitin Deacetylase

Chitin deacetylases are found in fungi, insects and particularly marine bacteria. They catalyse the removal of the N-acetyl group from N-acetylglucosamine residues in chitin, releasing acetic acid. This is critical in the conversion of chitin to chitosan, a deacetylated form of chitin. Chitin deacetylases identify a sequence of four consecutive N-acetylglucosamine residues as their substrate, with one residue then deacetylated. The residue deacetylated depends on the type of deacetylase involved as chitin deacetylases can be endo- or exo-acting. The biological function of chitin deacetylases is to produce chitosan for incorporation into the cell wall and as a defence mechanism (Ghormade *et al.*, 2010). Industrially chitosan has been produced

from chitin using a combination of thermal and chemical methods. However, these conditions tend to be very harsh and damage the final product. Therefore, there is more interest in the use of chitin deacetylases for the commercial generation of chitosan. Chitosan has a variety of applications in the biotechnological industry. For example it has been used as a scaffold for enzyme immobilisation, it has been used in food additives, cosmetics, as a drug delivery system and as a wound dressing (Ghormade *et al.,* 2010; Zhao *et al.,* 2010).

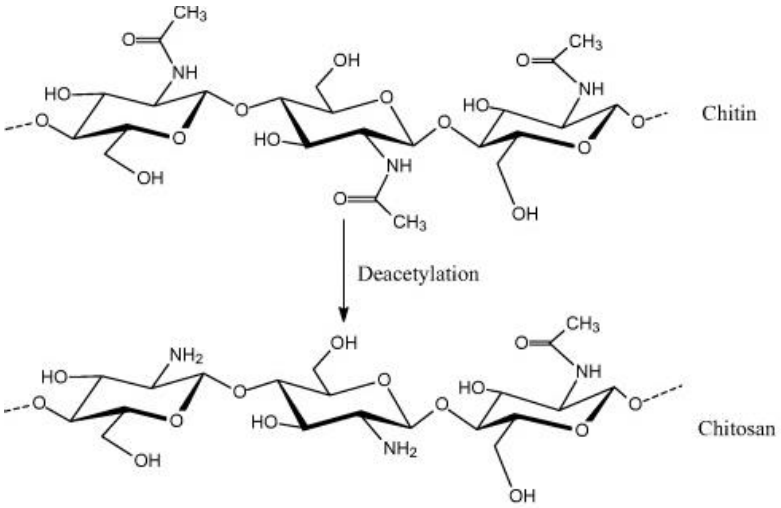


Fig 1.5: Deacetylation of chitin to produce chitosan. Chitin deacetylase removes the N-acetyl group from chitin to produce chitosan (Tran et al., 2011).

1.1.5.3 Chitin-Binding Protein

Carbohydrate-binding proteins (CBPs) belong to seven families in the CAZy database (1, 2, 12, 14, 18, 19, and 37). The CBPs from families 14 and 18 exist as distinct independent proteins with no catalytic activity, while the CBPs belonging to the remaining families exist as domains of larger enzymes, thought to increase the catalytic activity of the enzyme as well as increase the affinity of the enzyme for the substrate (Vaaje-Kolstad *et al.,* 2005). CBPs from families 14 and 18 are chitin-binding proteins that have been found to have certain anti-fungal activity. One of the most characterised CBPs is CBP21 from *S. marcescens*. This CBP works in synergy with the chitinases from *S. marcescens* to enhance the rate of chitin hydrolysis, through the binding of chitin and the subsequent disruption of its' structure, allowing the chitinase enzymes easier access to the chitin substrate. Chitin-binding proteins

have other numerous functions, from defence in plants to cell-wall degradation by bacteria (Vaaje-Kolstad *et al.*, 2005).

1.1.5.4 Chitin Hydrolase

Chitin hydrolases are classified as glycosyl hydrolases and are also known as chitinases or chitobiases. Chitinases and chitobiases are found in all organisms. In fungi and invertebrates they play a role in morphogenesis and promoting the structural integrity of the organism. However, in plants they act in a defensive role against fungal pathogens, in vertebrates they are present in the digestive tract where they also play a defensive role, while in bacteria they primarily undertake a nutritional role (Ubhayasekera, 2011). As already mentioned in Section 1.1.4.3 chitinases and chitobiases form a chitinolytic system comprising endo- and exo-acting enzymes that work synergistically in the hydrolysis of chitin. Firstly, endo-acting enzymes cleave chitin internally generating various length chito-oligosaccharides. Exo-acting chitinases such as chitobiosidases may then break down these chito-oligomers to chitobiose (a dimer of N-acetylglucosamine). Finally exo-acting chitobiases (N-acetylglucosaminidases) hydrolyze chitobiose to single N-acetylglucosamine residues (Bhattacharya *et al.*, 2007; Sahai and Manocha, 1993). Chitinases are divided into two families; Family 18 and Family 19 based on sequence similarity of the catalytic domain. However, each family has structural and mechanistic differences. Chitobiases are part of Family 20 (Synstad *et al.*, 2004).

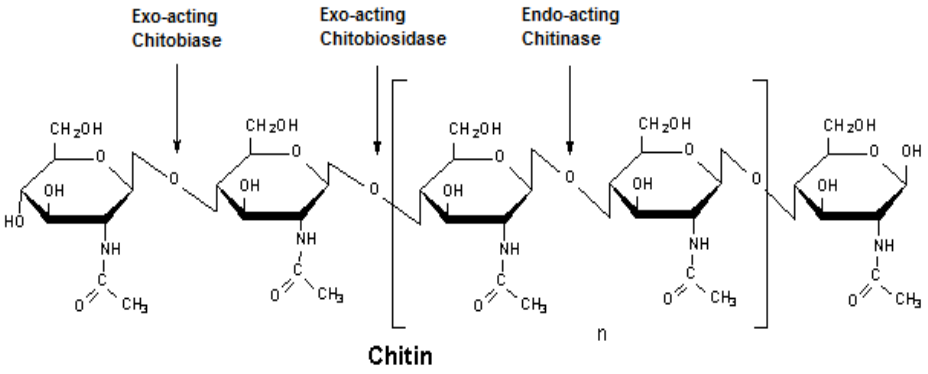


Fig 1.6: The step-wise hydrolysis of chitin. A mixture of endo- and exo-acting enzymes hydrolyse chitin chains to shorter chito-oligosaccharide chains and N-acetylglucosamine monomers. Image obtained from Sigma Aldrich (www.sigmaaldrich.com).

1.2 The Glycosyl Hydrolases (Chitinases and Chitobiases)

1.2.1 Classification - CAZy database – a resource for carbohydrate active enzymes

The huge array of naturally occurring carbohydrates is mirrored by the large variety of enzymes available for their hydrolysis. These enzymes are classified as glycosyl hydrolases in the CAZy database (http://www.cazy.org/). Glycosyl hydrolases are enzymes that catalyse the hydrolysis of glycosidic linkages through the addition of a water molecule. This CAZy database (Carbohydrate-Active Enzymes database) acts as a resource for all carbohydrate active enzymes and classifies these enzymes into families based on amino acid similarity. This takes the structural similarities of the enzymes into account. This system has moved away from the IUB Enzyme Nomenclature system which classifies enzymes based on the type of reaction they catalyse but gave no indication of the structural features of the enzymes or their evolutionary history (Henrissat, 1991; Cantarel *et al.*, 2008). Before entry into the CAZy database the sequences of prospective proteins, obtained from Genbank, are compared and analysed using tools such as BLAST and HMMER. They are categorised into families based on sequence similarities, generally of the catalytic or carbohydrate-binding domains. Once in the database, information is available on the

sequence, structure and function of the protein. Five main types of proteins are covered in this database, all with different activities on carbohydrates; Glycosyl hydrolases, Glycosyltransferases, Polysaccharide lyases, Carbohydrate esterases and Carbohydrate-binding modules. While glycosyl hydrolases are involved in the cleavage of glycosidic linkages, glycosyltransferases are involved in the synthesis of glycosidic bonds. Polysaccharide lyases are also involved in the cleavage of glycosidic bonds, however, with a different mechanism than glycosyl hydrolases. Where glycosyl hydrolases cleave glycosidic linkages through hydrolysis, polysaccharide lyases act on carbohydrates composed of uronic acid and cleave the linkages through a β-elimination mechanism. Carbohydrate esterases are involved in the cleavage of ester-containing groups on polysaccharides. Finally, carbohydrate-binding modules are the only proteins covered in this database with no enzymatic activity. They are involved in the binding of carbohydrates, and through this interaction assist the enzymes in carbohydrate hydrolysis. These carbohydrate-binding modules can be found independently or as a domain in a carbohydrate-active enzyme (Cantarel *et al.,* 2008). There are currently 131 families of glycosyl hydrolases in the CAZy database (Cantarel *et al.,* 2008; Henrissat 1991).

1.2.2 General mechanism of action of Glycosyl Hydrolases

X-ray crystallography and site-specific mutagenesis has allowed the study of the catalytic mechanism of these enzymes (Rye and Withers, 2000). The catalytic mechanism of glycosyl hydrolases has been elucidated through the crystallisation of enzyme-substrate and enzyme-inhibitor complexes. The functions of critical residues have been elucidated through the generation of mutants via site-specific mutagenesis (Rye and Withers, 2000). The glycosidic linkage is known to be one of the most durable linkages in nature. Glycosyl hydrolases are extremely adept enzymes, capable of catalysing the hydrolysis of glycosidic linkages up to 10^{17} times faster than would happen naturally (Davies and Henrissat, 1995). Despite the large number and diverse nature of glycosyl hydrolases available, they all have the same general acid-catalysis mechanism of action, requiring two critical residues in their active site for hydrolysis of the glycosidic bond. These residues act as a proton donor and a nucleophile (a general acid and a general base). Hydrolysis of the glycosidic bond can result in one of two outcomes; inversion or retention of the anomeric configuration (Davies and

Henrissat, 1995). Inverting glycosyl hydrolases result in inversion of the stereochemistry of the anomeric carbon, whereas retaining glycosyl hydrolases maintain the linkage type. The first glycosyl hydrolase to be structurally characterised and its mechanism deciphered was lysozyme (Matthews and Remington, 1975; Blake *et al.*, 1965). In the following years, the catalytic mechanism of many other glycosyl hydrolases has been elucidated along with the residues critical for catalysis (Davies and Henrissat, 1995).

1.2.2.1 Inverting glycosyl hydrolases

Inverting glycosyl hydrolases catalyse the hydrolysis of substrates with the inversion of the anomeric configuration in a single-displacement reaction. Hydrolysis is carried out via a one-step mechanism that involves the creation of an oxocarbenium ion transition state. This mechanism requires a general acid (proton donor) and a general base (nucleophile) that are located approximately 10.5 Å apart. The general acid donates a proton to the anomeric carbon of the sugar (C1). The general base activates the water molecule through the removal of a proton. The nucleophile is a good distance away from the proton donor as it must allow space for the water molecule. The nucleophilicity of the water is then increased, allowing an attack on the anomeric carbon to occur, splitting the glycosidic linkage (Vuong and Wilson, 2010; Zechel and Withers, 2000; Davies and Henrissat, 1995). The mechanism is explained in the diagram below (see Figure 1.7).

Inverting mechanism for a β-glycosidase:

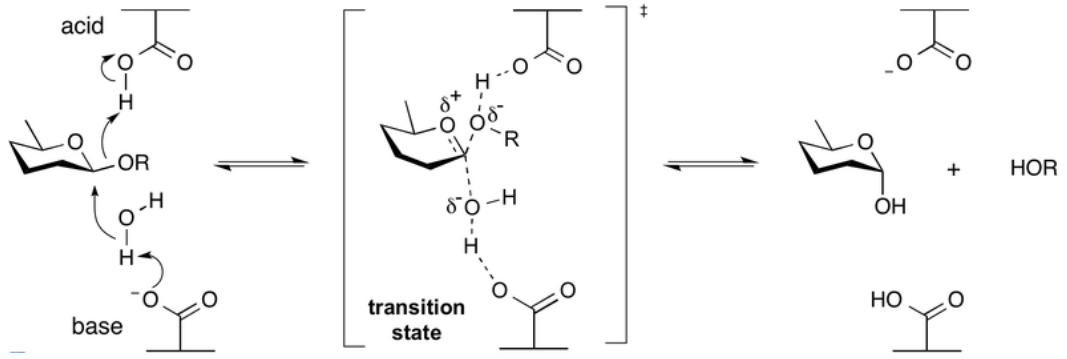


Fig 1.7: Catalytic mechanism of a β-inverting glycosyl hydrolase. The glycosidic bond is cleaved by the glycosyl hydrolase with inversion of the anomeric carbon. Image obtained from CAZypedia. (http://www.cazypedia.org/images/a/a0/Inverting_glucosidase_mechanism.png)

1.2.2.2 Retaining glycosyl hydrolases

The main structural difference between inverting and retaining glycosyl hydrolases is the distance between the catalytic residues in the active site. In both types the proton donor occupies the same position (Zechel and Withers, 2000). In inverting glycosyl hydrolases these groups are approximately 10.5 Å apart. In retaining glycosyl hydrolases these groups are much closer together, approximately only 5.5 Å apart. Hydrolysis by retaining glycosyl hydrolases proceeds with retention of the anomeric configuration. It involves a double-displacement mechanism with the creation of a covalent glycosyl-enzyme intermediate. The first step is the attack of the anomeric carbon by the nucleophile which forms the glycosyl intermediate. The other catalytic residue acts as the proton donor and donates a proton to the glycosidic oxygen bond. The second step involves the hydrolysis of the glycosyl-enzyme intermediate with water, followed by the deprotonation of the water molecule by the catalytic residue (Vuong and Wilson, 2010; Zechel and Withers, 2000; Davies and Henrissat, 1995). The mechanism is explained in the diagram below (see Figure 1.8).

.

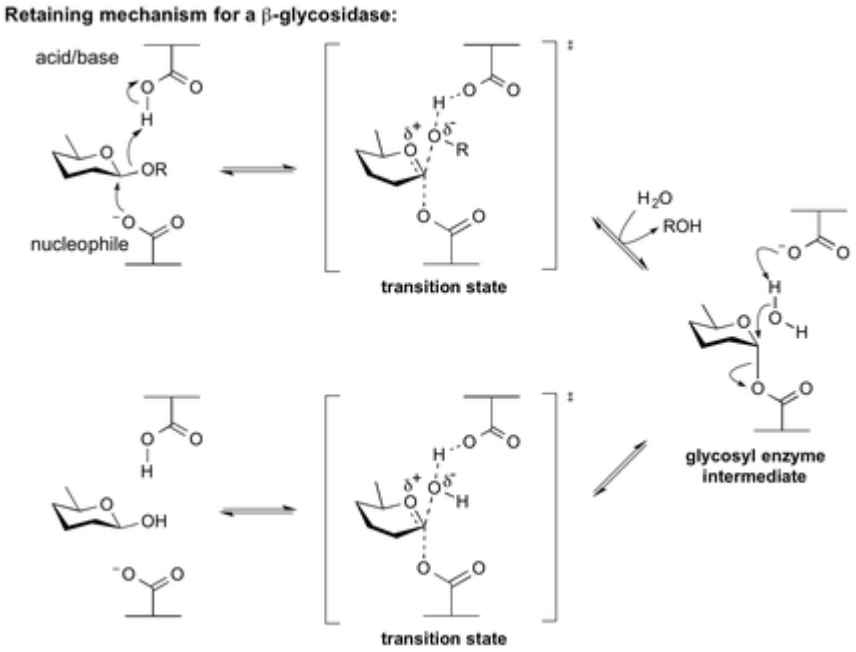


Fig 1.8: Catalytic mechanism of a β-retaining glycosyl hydrolase. The glycosidic bond is cleaved by the glycosyl hydrolase with retention of the anomeric carbon. Image obtained from CAZypedia. (http://www.cazypedia.org/images/7/7f/Retaining_glycosidase_mechanism.png)

1.2.2.3 Substrate-assisted glycosyl hydrolases

Substrate-assisted catalysis is an occurrence in glycosyl hydrolases that hydrolyse substrates containing an N-acetyl group at the C-2 position. These enzymes generally have only a proton donor and no nucleophilic residue. This reaction also proceeds via a double-displacement mechanism, however it is the N-acetyl group that acts as the nucleophile and leads to the formation of an oxazolinium intermediate. The mechanism detailed in the diagram below (see Figure 1.9).

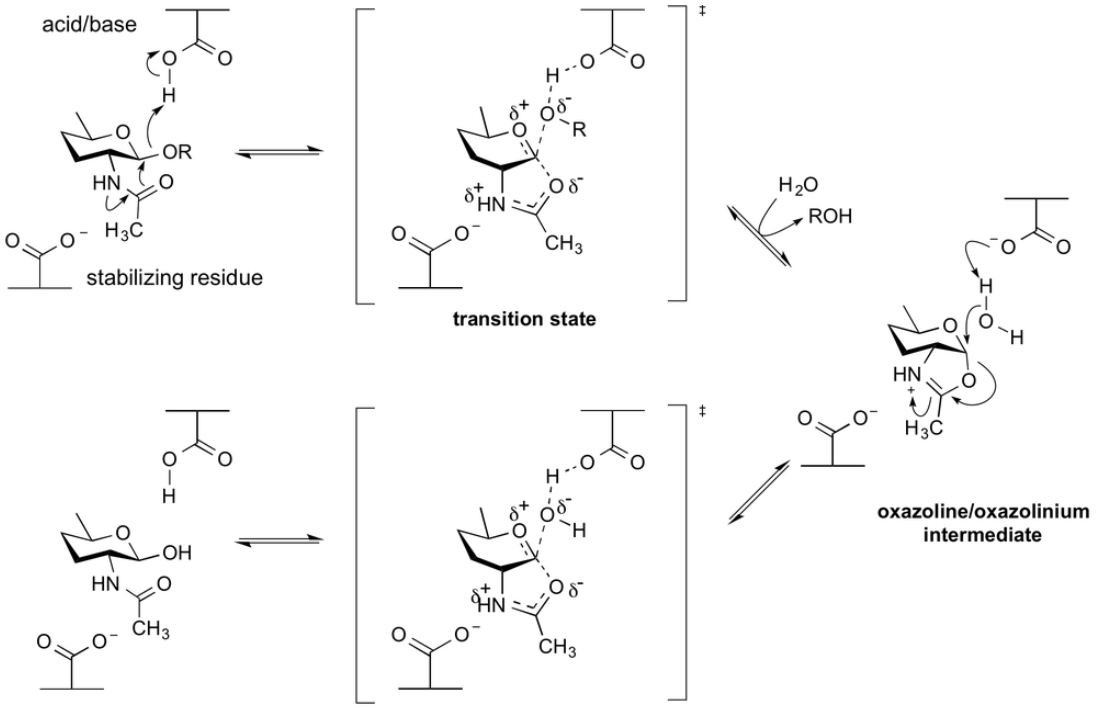


Fig 1.9: Catalytic mechanism of the substrate-assisted glycosyl hydrolases. Image obtained from CAZypedia.
(http://www.cazypedia.org/images/0/0c/Hex_neighboring_mechanism.png)

1.2.3 General topology of the active site of the glycosyl hydrolases

There are three main folds observed in the active site of glycosyl hydrolases; a pocket/crater, a cleft/groove and a tunnel (Davies and Henrissat, 1995). The pocket shape is optimal for the binding of the non-reducing ends of sugars and therefore more favourable towards hydrolysis of shorter chain polysaccharides. It is often observed in exo-acting enzymes. The cleft and groove structures favour the hydrolysis of longer-chain polysaccharides. The accessible cleft site is primarily observed in endo-acting enzymes such as chitinases. The tunnel is less accessible than the cleft fold, and is less common. It has so far only been observed in cellobiohydrolases where the polysaccharide chain is fed through the tunnel and the hydrolysis products are released (Davies and Henrissat, 1995).

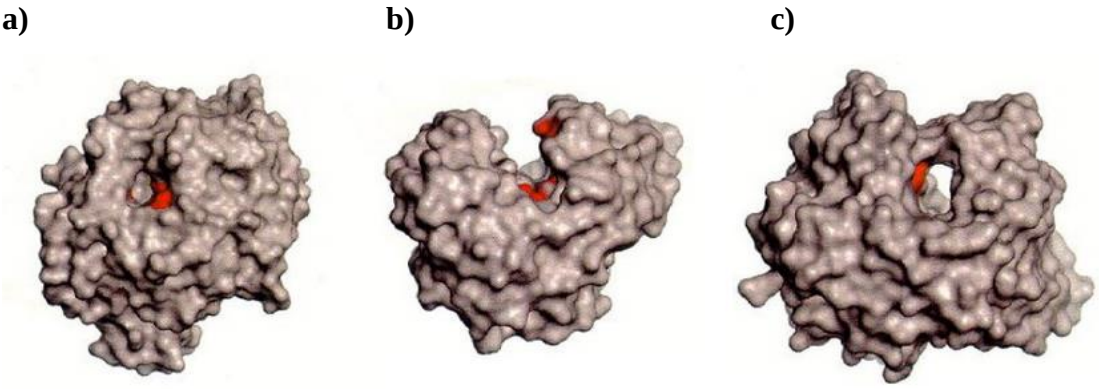


Fig 1.10: The three main folds observed in the glycosyl hydrolase active site a) crater, b) cleft and c) tunnel. The catalytic residues are highlighted in red (Davies and Henrissat, 1995).

1.2.4 Families of the Glycosyl Hydrolases

1.2.4.1 Family 18 Chitinases

Family 18 chitinases are retaining glycosyl hydrolases involved in the breakdown of chitin and chito-oligosaccharide chains. They are found in bacteria, fungi, insects but more interestingly are found in mammals that are not known to possess chitin synthesis and metabolism machinery (Funkhouser and Aronson Jr, 2007). Family 18 glycosyl hydrolases consist of five separate domains. An N-terminal signal peptide is present which transports the enzyme to the periplasm where it is cleaved by peptidases, followed by secretion of the enzyme. The catalytic domain is made up of an $(\alpha/\beta)_8$ barrel containing the active site in a cleft/groove. This groove is largely composed of aromatic residues such as tryptophan, which are important for substrate binding (see Figure 1.10b). A glutamate residue present in the active site is proposed as the catalytically critical proton donor. Unlike the majority of glycosyl hydrolases, there is no nucleophilic residue present. It is believed to catalyse the hydrolysis of chitin in the same substrate-assisted mechanism described in Section 1.2.2.3) (Sorlie *et al.,* 2012; Perrakis *et al.,* 2004; Synstad *et al.,* 2004). There is also a chitin-binding domain that functions in the binding of the insoluble chitin substrate which enhances catalytic activity, a serine/threonine-rich region, and a C-terminal region (Perrakis *et al.,* 2004).

1.2.4.2 Family 19 Chitinases

The first family 19 chitinases to be classified were found in plants. Now, however they have been found in other species such as bacteria, viruses and fungi. Plant family 19 chitinases play a defensive role, while bacterial family 19 chitinases play a role in nutrition with fungal family 19 chitinases playing a developmental and structural role (Ubhayasekera, 2011). In comparison with family 18 chitinases, family 19 chitinases are inverting glycosyl hydrolases. There are two glutamate residues present in the active site required for catalysis, one acting as a general acid and the other acting as a general base (Brameld and Goddard III, 1998). The majority of family 19 chitinases are comprised of an N-terminal signal peptide, a carbohydrate-binding module and a catalytic domain. The carbohydrate-binding module is attached to the rest of the molecule via a linker. These chitinases display an "α-helical bi-lobed" structure, with the active site contained within a wide groove (Ubhayasekera, 2011). Due to their differences in mechanism the family 18 and family 19 chitinases give rise to different hydrolysis products; the family 18 chitinases yield a β-anomer while the family 19 chitinases yield an α-anomer.

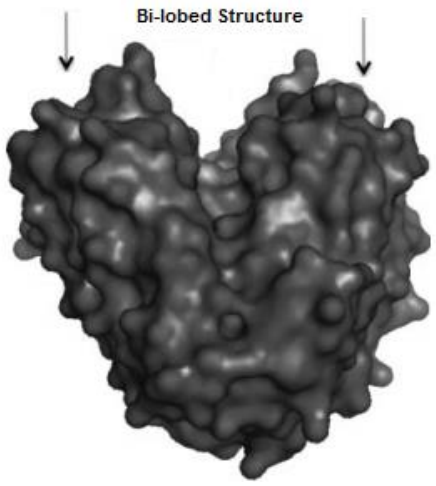


Fig 1.11: The "bi-lobed" structure of the family 19 chitinase. The active site cleft can clearly be seen between the two "lobes" of the family 19 chitinase structure (Ubhayasekera, 2011).

1.2.4.3 Family 20 Chitobiases

Family 20 glycosyl hydrolases encompass enzymes with three different activities; N-acetylhexosaminidases, lacto-N-biosidases and chitobiases (N-acetylglucosamindases). N-acetylhexosaminidases catalyse the hydrolysis of N-aceylglucosamine residues from chitobiose and longer-length chito-oligosaccharides. Lacto-N-biosidases catalyse the hydrolysis of lacto-N-biose (a dimer of galactose and N-acetylglucosamine) from longer oligosaccharides such as lacto-N-tetraose, lacto-N-hexaose etc. While chitinases hydrolyse chitin and chito-oligosaccharide chains, chitobiases hydrolyse chitobiose (a dimer of N-acetylglucosamine) to release two N-acetylglucosamine residues (Intra *et al.,* 2008). Family 20 glycosyl hydrolases are widely found in nature, from bacteria and fungi to insects, plants and mammals. N-acetylhexosaminidases are known to play crucial roles in mammals, as mutations in the HEX genes cause lysosomal storage disorders such as Tay Sachs disease and Sandhoff disease (Tews *et al.,* 1996). In insects they play a role in structural remodelling, in bacteria and fungi they play a role in chitin metabolism while in plants they play a role in defence (Tews *et al.,* 1996). Family 20 glycosyl hydrolases are generally composed of four domains and an N-terminal signal peptide. Functions have been designated to two of the domains; a catalytic domain and a chitin binding domain. The catalytic domain, a $(\beta\alpha)_8$ barrel, contains the active site which is contained in a deep crater. These enzymes have a slightly unusual mechanism of catalysis, a substrate-assisted mechanism, requiring a proton donor such as a glutamate residue in the active site, while the N-acetyl group of the sugar acts as the nucleophile (Prag *et al.,* 2000; Tews *et al.,* 1996).

1.3 Bacterial Chitinases

Chitinolytic systems are found in a variety of bacterial species with the most well characterised system being that from *S. marcescens*. The structures and mechanisms of its chitinases and chitobiases have been well studied.

1.3.1 *Serratia marcescens*

1.3.1.1 *S. marcescens* overview

S. marcescens is an opportunistic pathogen causing nosocomial infections. It is a gram negative, saprophytic bacteria belonging to the genus *Serratia* and is a member of the Enterobacteriaceae family (Hejazi & Falkiner, 1997). Some strains are red-coloured due to the production of a red-pigmented secondary metabolite known as prodigiosin. In nature *S. marcescens* is found primarily in water and soil. In humans the bacteria rarely infects healthy individuals, preferring immuno-compromised individuals as its host (Grimont & Grimont, 2006). *S. marcescens* may cause a number of infections, ranging from sinusitis, urinary tract infections and wound infections to the more lethal pneumonia and septicaemia (Hejazi & Falkiner, 1997). *S. marcescens* is resistant to many commonly used antibiotics including ampicillin, tetracycline, and first and second generation cephalosporins (Grimont & Grimont, 2006; Hejazi & Falkiner, 1997). *S. marcescens* contains many virulence factors that help it to evade the immune system, colonise the host and cause damage. These factors include the production of powerful siderophores for iron aquisition, the production of hemolysin (a molecule capable of lysing host erythrocytes) and the secretion of proteases which can cause tissue damage. Certain strains of *S. marcescens* can even be resistant to the bacteriolytic activity of serum (Grimont & Grimont, 2006).

Fig 1.12: The red colour of *Serratia marcescens* due to the presence of the pigment, prodigiosin. Image obtained from: http://faculty.ccbcmd.edu/courses/bio141/labmanua/lab2/sminsol.html

1.3.1.2 The chitinolytic system from *S. marcescens*

S. marcescens was discovered to be an extremely efficient chitin-degrading bacterium, producing more chitinolytic enzymes than all those tested alongside it (Monreal & Reese, 1969). It was found that *S. marcescens* produced a total of five proteins associated with chitin degradation, all of which have different molecular weights (Fuchs *et al.,* 1986). Further studies deduced that *S. marcescens* contains three chitinases, named ChiA, ChiB and ChiC, a chitobiase and a protein associated with binding chitin named CBP21 (Brurberg *et al.,* 2001). These enzymes comprise the chitinolytic system of *S. marcescens* and their production is induced by the presence of chitin in the environment. All three chitinases are exported, with varying degrees of processing prior to secretion. ChiA contains an N-terminal signal peptide that is cleaved before secretion, ChiB is not processed before secretion and the ChiC N-terminal signal peptide is non-specifically cleaved at various residues to generate various forms of ChiC before secretion (Brurberg *et al.,* 2001). These enzymes each exhibit a modular structure, with one domain designated as a catalytic domain and a second designated as a binding domain. Functions were assigned to the various domains based on sequence homologies with various other enzymes (Brurberg *et al.,* 2001). While ChiA, ChiB and ChiC belong to the family 18 glycosyl hydrolases, the chitobiase belongs to the family 20 glycosyl hydrolases. The chitobiase enzyme also undergoes processing before secretion i.e. the cleavage of its N-terminal signal peptide. The final protein involved in the *S. marcescens* chitinolytic system has no enzymatic function. It is a chitin-binding protein that also has an N-terminal signal peptide cleaved before secretion thought to have a role in disrupting the crystalline structure of chitin and promoting hydrolysis by making enzyme access easier (Brurberg *et al.,* 2001).

1.3.1.3 Chitin hydrolysis by *S. marcescens*

ChiA is the most studied chitinase from *S. marcescens* with its crystal structure being described by Perrakis *et al.,* in 1994. It was found to contain three domains; a β-barrel domain, a β-strand domain and a domain containing a mixture of α and β folds. The structure of the chitinase associated with a chito-oligosaccharide substrate was also studied and this led to the assignment of the β-barrel as the catalytic domain (Perrakis

et al., 1994). Both ChiA and ChiB contain N-terminal signal peptides and are exported proteins. The active site of both ChiA and ChiB is located in a deep cleft in the catalytic domain surrounded by aromatic residues for substrate binding. The main difference in ChiA and ChiB structure is the position of their respective active sites which are located on opposite sides of the catalytic domain, suggesting that they each bind in a different orientation to the chitin substrate (Brurberg *et al.,* 2001). Brurberg *et al.* (1996) studied the catalytic properties of a ChiA and ChiB from *S. marcescens* BLJ200. Incubation of ChiA with $(GlcNAc)_{1-6}$ yielded $(GlcNAc)_2$ and GlcNAc hinting at an exo-chitobiohydrolase activity. Incubation of ChiB with $(GlcNAc)_{1-6}$ also yielded $(GlcNAc)_2$ and GlcNAc. There was no hydrolysis of $(GlcNAc)_2$. Chitobiose ($(GlcNAc)_2$) is then hydrolysed by chitobiases to yield the final N-acetylglucosamine product (Prag *et al.,* 2000; Tews *et al.,* 1996). The hydrolysis of chitin by ChiA and ChiB is enhanced by the action of CBP21, a chitin-binding protein. This protein binds to the insoluble chitin substrate disrupting its structure and thus making it more accessible for the ChiA and ChiB enzymes to bind and hydrolyse the substrate (Vaaje-Kolstad *et al.,* 2005).

1.3.1.4 Mechanism of action of the chitobiase from *S. marcescens*

Drouillard *et al.,* (1997) first described the nature of chitobiose hydrolysis by the chitobiase from *S. marcescens*. They found that the reaction was catalysed with retention of the anomeric configuration. They also proposed a substrate-assisted mechanism of catalysis for this chitobiase. They suggested that the chitobiase employed the N-acetyl group of the sugar as the nucleophile, due to the absence of a conveniently located nucleophile in the active site. Tews *et al.,* (1996) also proposed a substrate-assisted mechanism of hydrolysis where Glu 540 acted as the proton donor while the N-acetyl group of the sugar substrate acted as the nucleophile resulting in retention of the anomeric configuration. Prag *et al.,* furthered this work in 2000 when they performed site-directed mutagenesis on the supposed catalytically important residues. They confirmed that Glu 540 acted as the proton donor while the Asp residue at position 539 held the N-acetyl group of the sugar substrate in place while it acted as the nucleophile. This unusual mechanism of action involving the sugar substrate as the nucleophile is also found in lysozyme and another family 20 enzyme, namely, the N-acetyl-β-hexosaminidase (Rye & Withers, 2000; Drouillard *et al.,*

1997). Davies and Henrissat (1995) described the topology of the active site of chitobiases. The active site is shaped like a pocket or crater. This shape is favourable for the binding of the reducing end of sugars such as that found in chitobiose. This shape is not optimal for the binding of longer-chain polysaccharides due to the lack of available reducing ends.

1.3.2 *Photorhabdus luminescens*

This study has identified a chitobiase enzyme from *P. luminescens* which shows significant similarities with that of the well characterised *S. marcescens* chitobiase (see Section 3.2).

1.3.2.1 *P. luminescens* overview

The genre of bacteria, *Photorhabdus*, is a member of the Enterobacteriaceae family. Fischer *et al.,* (1999) were the first group to classify *Photorhabdus* into three distinct genomic species, *Photorhabdus luminescens*, *Photorhabdus temperata* and *Photorhabdus asymbiotica*, based on numerous studies including 16S rRNA sequencing, ΔTm measurements to determine DNA relatedness and phenotypic characterizations. *P. luminescens* was used in this study and therefore will be the only *Photorhabdus* species discussed further.

Photorhabdus luminescens is a gram negative, motile, facultative anaerobe that forms a symbiotic relationship with soil residing, entomopathogenic nematodes from the Heterorhabditidae family. *P. luminescens* has a life cycle composed of two phases, symbiosis and virulence. During symbiosis *P. luminescens* may be found dwelling in the gut of the nematode during its infective juvenile (IJ) stage (Waterfield *et al.*, 2009; Forst *et al.,* 1997). Infective juvenile nematodes are characterized by the fact that their development is arrested and that they are searching for insect hosts. Infection of the insect by the nematode can be through the digestive tract, respiratory tract or through penetration of the outer skin. The final destination of the nematode is the hemocoel of the insect, a cavity containing blood and hemolymph. Once in the hemocoel, the nematodes release the virulent bacteria present in their system which kills the insect host by inducing septicaemia (Waterfield *et al.,* 2009; Forst & Nealson, 1996).

However, before killing the insect host the bacteria is confronted by the insects' immune system. *P. luminescens* secretes virulence molecules such as toxins in order to evade this immune response. It secretes toxins such as Tcd and Mcf1 which are involved in breaking down the gut epithelium and inducing apoptosis in cells respectively (Waterfield *et al.*, 2009). It also secretes a molecule known as LopT which inhibits uptake and consequently phagocytosis by macrophages. In order to prevent other bacteria and fungi from colonising the insect cadaver, *P. luminescens* secretes various anti-microbial molecules such as the toxic S-type pyocins. Tca is an orally toxic molecule secreted by *P. luminescens* to deter larger animals from eating the insect cadaver (Waterfield *et al.*, 2009).

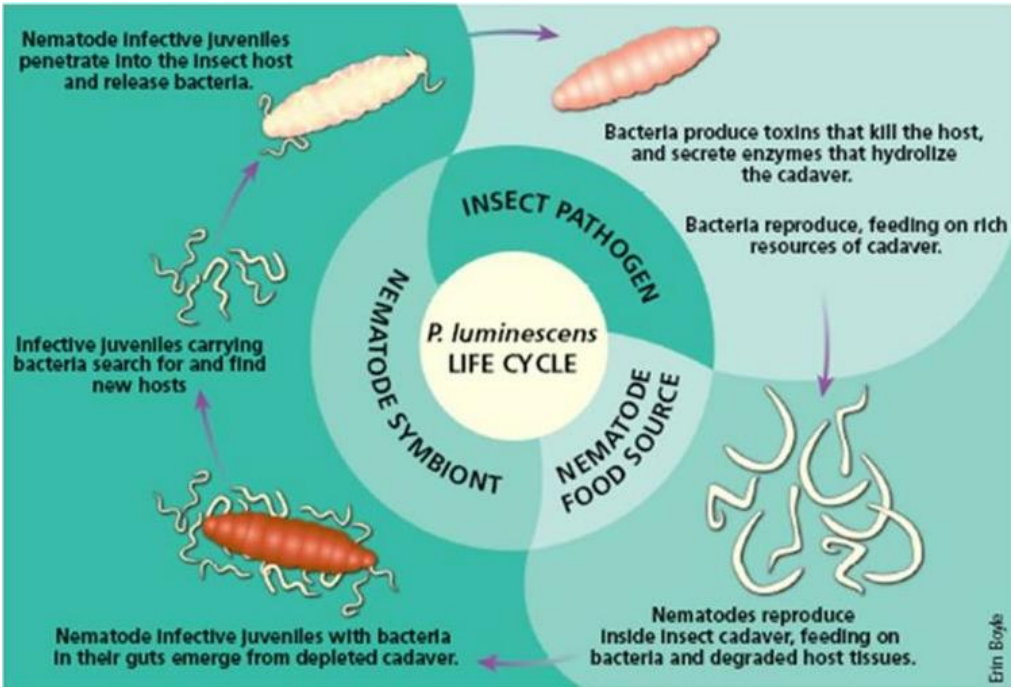


Fig 1.13: The life cycle of *P. luminescens*. The life cycle of the *P. luminescens* bacteria includes two main stages; symbiosis and virulence (Williamson and Kaya, 2003).

The insect carcass plays host to the nematode and bacteria. *P. luminescens* secretes lipases and proteases which breakdown the insect cadaver providing nutrients for the nematodes, allowing them to feed and reproduce. Approximately two weeks post infection, the nematodes, in their infective juvenile stage and in symbiosis again with the bacteria, leave the insect carcass to seek out new prey, repeating the cycle all over again (Forst *et al.*, 1997; Forst & Nealson, 1996;).

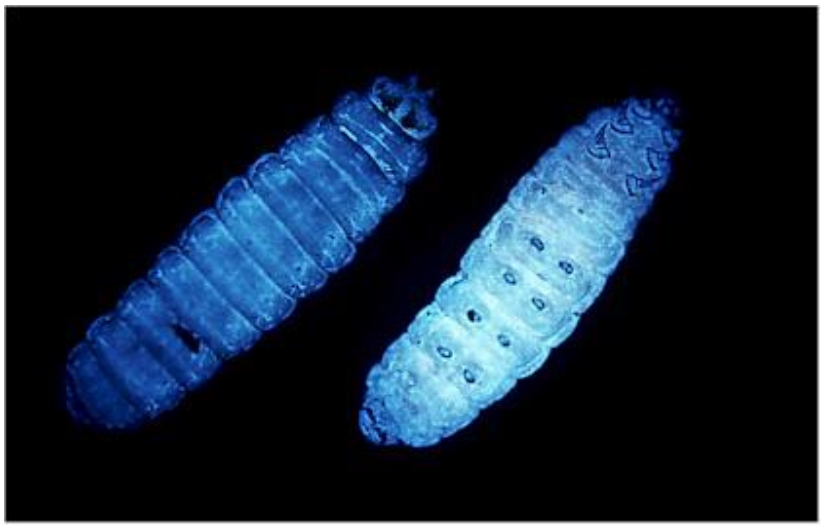


Fig 1.14: Nematode from the Heterorhabditidae family infected with *Photorhabdus luminescens* The glow from the insect is due to the bioluminescent properties of the *P. luminescens* bacteria (University of Bath, 2004).

1.3.2.2 Chitinase activity of *P. luminescens*

Chen *et al.,* (1996) examined the chitinase activity in *Xenorhabdus* and *Photorhabdus* species through the use of N-acetylglucosamine, chitotriose and chitin substrates. They were found to have both endo- and exo-chitinase activity. From a literature search carried out it was observed that there has been very little work carried out on specific chitinases and chitobiases from *P. luminescens*. Part of the *P. luminescens* life cycle involves the killing of the insect host and subsequent colonization of its cadaver. This requires the degradation of the insect cadaver to provide nutrients for the nematode. The outer shell of insects contains chitin as a structural component. Presumably, *P. luminescens* secretes chitinases and chitobiases to aid in the degradation of the insect cadaver, before symbiosis can take place. These chitinases and chitobiases could also play a defensive role during colonisation of the insect, having anti-fungal and anti-microbial effects, preventing colonisation of the insect by other bacteria and fungi (Chen *et al.,* 1996).

The first objective of this work was to carry out a detailed study of the chitobiase enzyme from *P. luminescens*. However, a second objective was also identified. Enzymes have two innate functions; catalysis and binding. While the first objective of

this study was to study the catalytic mechanism of the chitobiase enzyme, a second objective was to study the binding mechanism of the enzyme. Enzymes are extremely specific in the binding of their substrates. The substrate of the chitobiase enzyme is chitobiose, a dimer of N-acetylglucosamine. The chitobiase enzyme has the ability to specifically recognise and bind this sugar before catalysing its hydrolysis. The binding of this sugar by the enzyme was of interest in this study. N-acetylglucosamine is a sugar residue that is a component of glycans found on eukaryotic proteins. It is of commercial importance in the field of glycoprotein biotherapeutics where its presence or absence can greatly affect the pharmacokinetic properties of a biotherapeutic. The glycopatterns on glycoproteins and glycolipids may be analysed using carbohydrate binding proteins such as lectins. There is only one commercially available lectin capable of binding terminal GlcNAc, GSLII. This is difficult to make recombinantly due to its' structural complexity (it is a glycosylated protein with disulphide bridges present). Alternatives are therefore needed that can be produced recombinantly. Recent protein structural studies have used mutagenesis as a tool to remove the catalytic activity of enzymes, while still allowing them to retain their unique specific binding abilities. The ability of the chitobiase enzyme to specifically bind to N-acetylglucosamine is therefore of commercial interest. Therefore, the second objective of this work was the mutagenesis of the chitobiase enzyme in order to engineer it into a novel carbohydrate binding protein that could be used in the analysis of glycoproteins. The mechanism and significance of glycosylation, biologically and industrially, is discussed in detail below (see Section 1.4).

1.4 Glycosylation

The majority of eukaryotic proteins undergo post-translational modifications after synthesis that affect their structure, properties and functions. These covalent modifications basically regulate the proteins behaviour and extend the variety of functions of the proteome (Seo and Lee, 2004). There are many different types of post-translational modifications, some common types include ubiquitination, phosphorylation, sulfation, acylation and glycosylation. While some post-translational modifications are common on intracellular proteins, such as phosphorylation, others are more common on extracellular proteins, such as glycosylation. Glycosylation is

one of the most prevalent and diverse post-translational modifications. It is the process of adding glycans to molecules to form glycoconjugates. Glycans can be added to a variety of molecules, such as proteins and lipids. The addition of glycans to lipids forms biologically important membrane glycolipids such as lipopolysaccharide found in bacterial cell membranes and glycosphingolipids, found in eukaryotic cell membranes. The addition of glycans to proteins is, however, a far more prevalent occurrence, with over two-thirds of human proteins being glycosylated (Apweiler *et al.,* 1999). Glycosylation occurs in the Endoplasmic Reticulum (ER) and Golgi apparatus where specific enzymes catalyze the addition and removal of sugar residues to polypeptide chains. The glycosylation process is not template driven; it contains no proof reading machinery and depends on the physiological status of the cell. Glycans can combine with each other in ways that differ in chain length, sequence and branching points. This leads to enormous variation in the glycan structures produced. Laine (1997) calculated that there were 1.44×10^{15} possible isomers for a hexamer when using 20 sugars. There are two main types of glycosylation; N-linked glycosylation and O-linked glycosylation. Apweiler *et al.,* (1999), through analysis of the SWISS-PROT database, deduced that over 90% of all glycosylated eukaryotic proteins are N-linked. The other 10% are O-linked. Some glycoproteins may contain both O-linked and N-linked glycans.

1.4.1 Eukaryotic glycosylation

Glycosylation is carried out by both eukaryotes and prokaryotes. However, eukaryotic glycosylation was discovered before prokaryotic glycosylation and is better understood and characterized.

1.4.1.1 N-linked glycosylation

Eukaryotic N-linked glycosylation is carried out in the ER and golgi apparatus of the cell. N-linked glycans are linked to the amide nitrogen of asparagine residues (hence N-linked); however the asparagine residue must be present in the following consensus sequence: Asn-X-Ser/Thr, where X denotes any amino acid except proline. The consensus sequence Asn-X-Thr is preferred over Asn-X-Ser. Glycosylation is a stepwise process (Varki *et al.,* 2009). The first step is the production of a core glycan

structure at the membrane of the ER. This core structure is produced on a lipid "anchor" known as dolichol phosphate and then transferred to the nitrogen of an asparagine residue of the consensus sequence by a transferase enzyme, oligosaccharyl transferase, in the lumen of the ER.

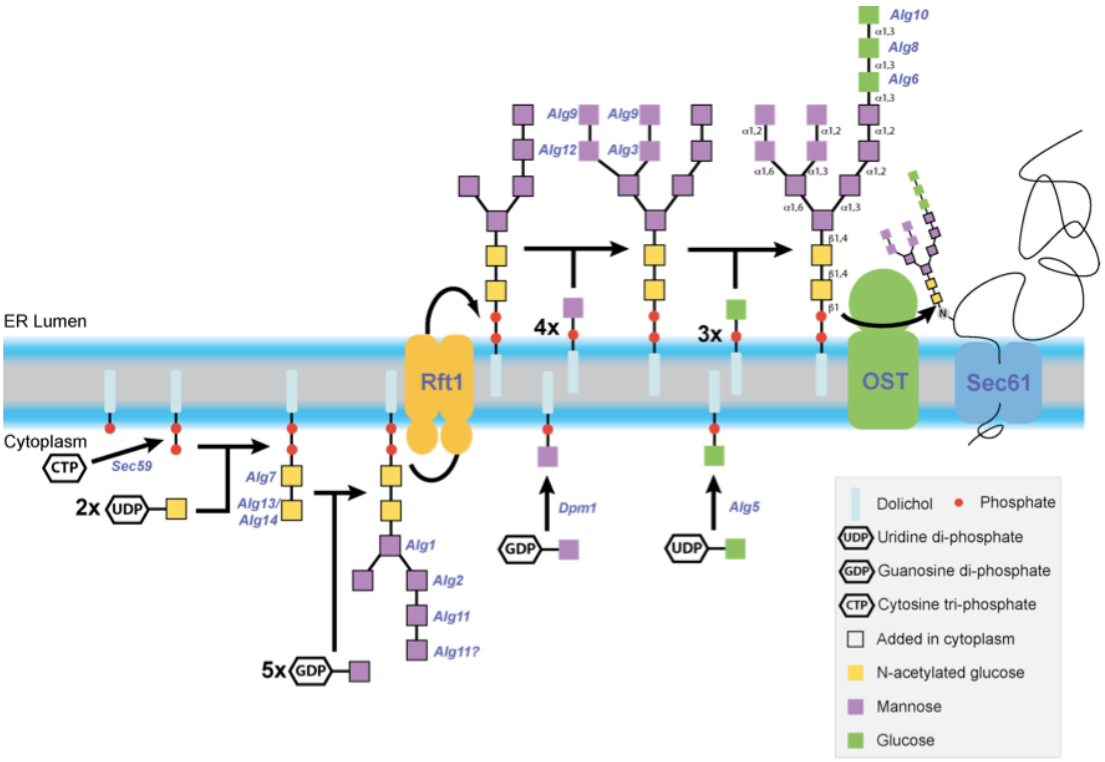


Fig 1.15: The N-linked glycosylation pathway. The stepwise construction of N-linked glycans (Varki *et al.*, 2008).

This core glycan structure is comprised of two N-acetylglucosamine residues, nine mannose residues and three glucose residues. This glycan structure then undergoes modifications as it moves through the ER, including the removal of the three glucose residues through the action of a glucosidase and the removal of up to four of the mannose residues through the action of a mannosidase. Once correctly folded this new glycoprotein, with the glycan core attached, is transported to the golgi apparatus where the glycans are then modified resulting in the production of one of three types of glycan structure; high mannose, complex and hybrid (see Figure 1.16). Transferases are then involved in the completion of the glycan structures through the addition of further glycans. The addition of an N-acetylglucosamine residue catalyses the cleavage of two more mannose residues from the original glycan structure leaving a core structure common to all N-linked glycans of Manα1-6(Manα1-3)Manβ1-

4GlcNAcβ1-4GlcNAcβ1-Asn-X-Ser/Thr. The addition of N-acetylglucosamine, galactose and sialic acid residues are catalysed by N-acetylglycosyltransferases, galactosyltransferases and sialytransferases respectively. These glycans are sequentially attached and can result in branched bi-, tri- and tetra-antennary structures. Sialic acid is generally the terminal glycan residue present on these structures. The complete glycoproteins are then transported from the golgi apparatus to carry out their functions throughout the cell (Schwartz and Aebi, 2011; Varki *et al.*, 2009; Burda and Aebi, 1999; Dwek, 1996).

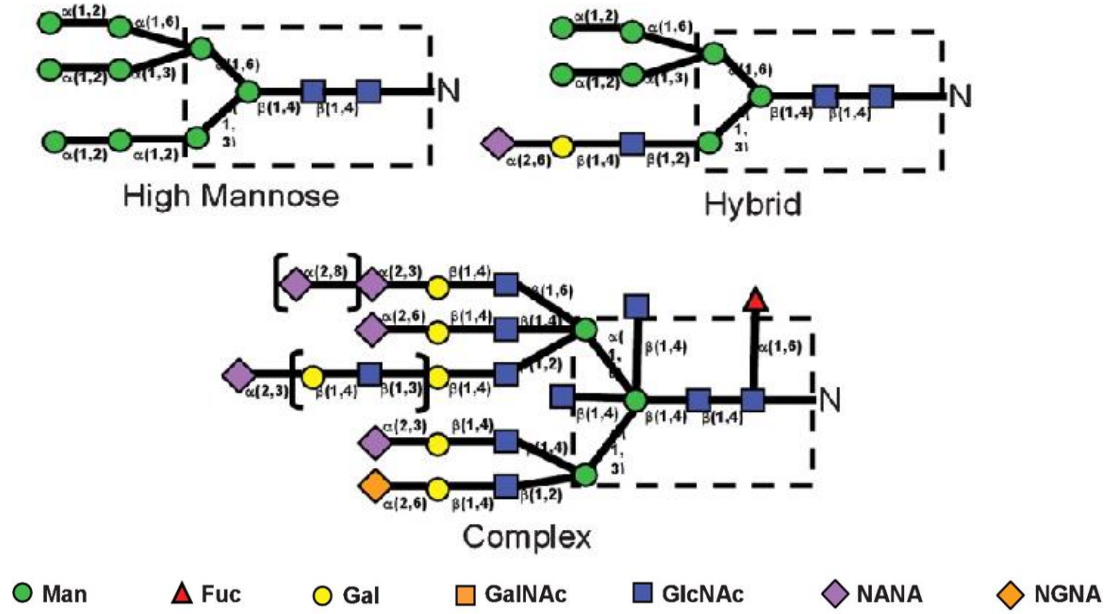


Fig 1.16: Types of N-linked glycosylation. The three main patterns of N-linked glycosylation are high mannose, hybrid and complex (Hossler *et al.*, 2009).

1.4.1.2 O-linked glycosylation

In comparison with N-linked glycans, O-linked glycans do not share the same core glycan structure. A variety of structures are observed for O-linked glycosylation. O-linked glycans are most commonly attached to the hydroxyl group (oxygen – hence O-linked) of a serine or threonine. They can also be attached to tyrosine, hydroxyproline or other hydroxylated amino acids. In contrast with N-linked glycosylation, no attachment consensus sequence is known, therefore the presence of O-linked glycan structures is difficult to predict. While N-linked glycosylation begins in the ER and finishes in the golgi apparatus (Varki *et al.*, 2009), O-linked

glycosylation is solely carried out in the golgi apparatus. O-linked glycans are named by the initiating sugar, i.e. O-linked GalNAc, O-linked Fucose etc. Assembly of the glycan structure most commonly begins with the addition of an N-acetygalactosamine residue to a hydroxyl group of a serine or threonine by an N-acetylgalactosaminyltransferase. However, other residues such as galactose, N-acetylglucosamine, fucose, xylose and glucosamine can also be used. The initiating sugar is attached to a lipid carrier, undecaprenyl pyrophosphate (UDP). These glycan structures may then be elongated by the sequential addition of sugars by various glycosyl transferases and modifications such as acylation and sulfation may also take place. The structure of O-linked glycans is dependent on the enzymes present in the cell, and therefore, the cells metabolic status. Mucins are one of the most commonly studied O-linked glycans. Mucins are glycoproteins that are characterised by an abundance of O-linked glycans due to the presence of a large number of serine and threonine residues. These glycoproteins tend to be saturated with O-linked glycans. Mucins account for a large proportion of mucous, the viscous substance secreted by epithelial cells in areas such as the respiratory, urinary and digestive tracts. Mucin O-linked glycans are highly sialylated which results in them being negatively charged. This allows them to bind water molecules and carry out their functions such as lubrication and acting as a protective barrier for the epithelium against chemical, physical and microbial threats. There are eight main types of core glycan structure, two of which are found commonly in mucins. These structures vary widely in size and composition (Varki *et al.,* 2009; Taylor and Drikamer, 2006; Van den Steen *et al.*, 1998; Dwek, 1996).

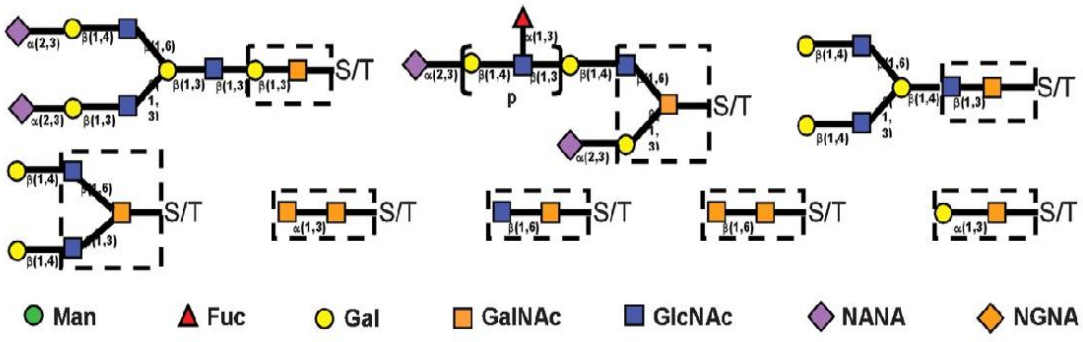


Fig 1.17: The 8 main cores of O-linked glycosylation. There are eight main types of O–linked glycan core which are depicted above (Hossler *et al.,* 2009).

1.4.2 Prokaryotic glycosylation

While eukaryotic protein glycosylation was first described in the 1930s, it wasn't until the 1970s that protein glycosylation in prokaryotes was described (Baker *et al.,* 2013). Until that point, it was thought that prokaryotes did not have the ability to glycosylate their proteins due to the lack of an ER and golgi apparatus in their cellular structure. Glycosylation is now accepted as a widely occurring process in prokaryotes and is even a more complicated process than in eukaryotes due to the larger variety of sugars involved. As with eukaryotes, there are two main types of glycosylation found in prokaryotes; N-linked and O-linked, the glycan linked to the amide nitrogen of an asparagine and the glycan linked to the hydroxyl oxygen of a serine or threonine respectively (Baker *et al.,* 2013; Nothaft and Szymanski, 2010). O-linked glycosylation was the first type of glycosylation discovered in prokaryotes when it was noted that surface layer (S-layer) proteins of *Halobacterium* and *Clostridium* species were modified with O-linked glycans. These glycosylated proteins accounted for up to 50% of the S-layer and covered the surface of the bacterium with the glycan portion of the protein protruding away from the surface (Mescher and Strominger, 1976; Sletyr and Thorne, 1976). Glycans were not only observed on S-layer proteins, as proteins like flagellin and pilli were also found to be extensively glycosylated. Prokaryotic glycosylation was generally found on extracellular proteins and was proposed to have a role in the virulence of the bacterium and the pathogenesis of disease by facilitating interactions with the host. The first comprehensive evidence of an N-linked glycosylation pathway was described for the bacterium *Campylobacter jejuni* by Szymanski *et al.,* in 1999 and Young *et al.,* in 2002. Two proteins, PEB3 and CgpA, were discovered to be glycosylated by this pathway. Now over 65 proteins are known to be glycosylated in *C. jejuni* in this way (Nothaft and Szymanski, 2010).

1.4.3 Biological functions of glycosylation

Due to their presence on the cell surface glycans have major functions in cell-cell interactions, cell recognition, cell signalling and cell trafficking (Dwek, 1996). They also play major roles in protein structure and the immune system which will be discussed below.

1.4.3.1 Structural and protective functions on glycoproteins

Glycans can have profound effects on the structure of the protein and consequently can affect the proteins physio-chemical properties and function. It has been noted that glycosylated proteins have an increased stability in comparison with their non-glycosylated forms. N-linked glycan structures are relatively large, with a sialyated triantennary structure having a molecular weight of around 2851 Da (Imperiali and O'Connor, 1999). Glycans can facilitate the correct folding of the protein in the ER and stabilisation of the tertiary structure. The thermostability and pH stability of a protein are majorly affected by the glycans on its surface. Solubility can also be enhanced due to the hydrophilic nature of the glycans and also their masking of hydrophobic amino acids of the protein. The longevity of the protein is also improved as the glycans cover the protein surface protecting it from proteases. The glycans also mask antigenic domains present on the protein, resulting in decreased immunogenicity (Sinclair and Elliott, 2004; Imperiali and O'Connor, 1999).

1.4.3.2 Immunological functions of glycosylation

Glycans play a major role in mammalian immunity, with their main function being to act as ligands for a variety of immunologically important receptors. Glycans act as ligands for a family of selectins involved in leukocyte trafficking. Selectins are receptors expressed on endothelial cells and leukocytes. There are different types of selectins, the most common being E-selectin (endothelial cells), L-selectin (leukocytes) and P-selectin (platelets). These selectins each recognise a specific glycan ligand, for example P-selectin recognises and binds sialyl-Lewis-X. During inflammation the expression of these receptors is up-regulated. Binding between the selectin and specific glycan allows the leukocytes and platelets to roll along the epithelium, eventually migrating through the epithelial barrier to the site of inflammation (Marth and Grewal, 2008).

Macrophages are phagocytic cells involved in innate immunity. They recognise and engulf pathogens. One way in which they recognise pathogens is through their mannose receptor. As the name suggests this receptor recognises certain mannose patterns, specifically α-linked highly branched mannose structures. Recognition of this pattern, particularly on a pathogen, triggers phagocytosis of the pathogen. This

recognition event can also trigger the release of pro-inflammatory mediators such as cytokines and initiation of the adaptive immune response (Linehan *et al.,* 2000).

The mucous that lines the gastrointestinal tract is composed of mucins which as mentioned before are O-linked highly glycosylated proteins. The glycan patterns on these mucins are recognised by gastrointestinal microbes, both commensal bacteria and pathogenic bacteria. Adhering to these glycans allows the bacteria to colonize the gastrointestinal tract. This is critical for pathogens to initiate an infection but also critical for commensals to carry out their probiotic functions. Bacteria adhere to glycans using adhesins such as fimbriae and pili as well as membrane and cell wall proteins. Mucin glycans also play a protective role in the gastrointestinal tract. Certain glycans have the ability to be released once bound such as those belonging to MUC1. This inhibits colonisation by the pathogen (Juge, 2012; Moran *et al.,* 2011).

Toll-like receptors (TLRs) are pattern recognition receptors. There are 10 in the receptor family (numbered 1-10 eg. TLR1, TLR2 etc) and they are found expressed on immune cells such as dendritic cells. They each recognise different patterns called Pathogen Associated Molecular Patterns (PAMPs). In general these patterns are glycans present on pathogens. After recognition of a PAMP the TLR initiates a signalling cascade in the cell which results in the "switching-on" of the innate immune system. This includes the secretion of cytokines and other pro-inflammatory mediators as well as the activation and differentiation of T cells (van Kooyk and Rabinovich, 2008).

1.4.5 Aberrant glycosylation

Glycosylation is essential for the survival of eukaryotes. The absence of glycosylation in an embryo is fatal and defects in the glycosylation pathway lead to several serious human diseases. In mice a defect in N-acetylglucosaminyltransferase, the enzyme responsible for the transfer of N-acetylglucosamine to the lipid anchor (dolichol phosphate), results in the death of an embryo 5 days after fertilization. This signifies that glycosylation is critical in embryonic development (Marek *et al.,* 1999). Many human diseases of aberrant glycosylation are congenital. There are 12 congenital diseases connected with defects in N-linked glycosylation while there are 7 congenital diseases connected with O-linked glycosylation. The symptoms of these diseases can vary from mild to fatal, with neurological and developmental issues prevalent (Jaeken

and Matthijis, 2007). One of the most studied groups of aberrant glycosylation related diseases are lysosomal storage diseases. These diseases affect the ability of the lysosomes to break down glycoconjugates, specifically a defect in glycosidases leads to an inability to catabolise glycoproteins. Diseases characterised by lysosomal storage disorders include Tay-Sachs and Sandhoff's (Ohtsubo and Marth, 2006). Tay-Sachs is a genetic disease characterised by a mutation in the HEXA gene which causes a hexosaminidase deficiency. This enzyme breaks down G_{m2} type gangliosides. A build-up of these sialyated glycosphingolipids can cause symptoms such as blindness and paralysis and even lead to death. Sandhoff disease is also caused by a hexosaminidase deficiency, which leads to a build-up of gangliosides and the aforementioned symptoms (Tews *et al.,* 1996). Glycosylation is crucial in the functioning of a normal nervous system. Neurodegenerative diseases such as Alzheimers and muscular degenerative diseases such as muscular dystrophy have been shown to have decreased glycosylation or altered glycosylation patterns (Kanninen *et al.,* 2004; Muntoni *et al.,* 2002).

The glycosylation pattern on the surface of a cell may signify the physiological status of the cell. Changes in this pattern are known to be associated with certain diseases. Practically all cancers have been associated with early changes in glycan pattern (Dube and Bertozzi, 2005). Cancer is an umbrella term for a group of diseases all characterised by uncontrolled cell growth. This cell growth is accompanied by phenotypic changes that allow the cells to migrate through tissues and evade the immune system. One mechanism of immune evasion is the alteration of glycan patterns on the surface of the cell. This is achieved by the over-expression of glycans and glycosylation associated enzymes such as glycosyltransferases (Ghazarian *et al.,* 2011). A common alteration observed is an increase in the branching of N-glycans. An alteration of terminal residues is also common. These alterations in glycosylation can act as biomarkers for the diagnosis of cancer as the presence/absence of certain glycans distinguishes then from healthy cells. Certain glycoproteins are also up-regulated in cancer and the detection of these may facilitate a diagnosis. Prostate Specific Antigen (PSA), a glycoprotein produced by prostate cells, is up-regulated during prostate cancer. The diagnosis of prostate cancer now involves the measurement of this antigen. Ovarian cancer is characterised by an upregulation in CA-125, a mucin also known as MUC16. Detection of this glycoprotein can help to

diagnose ovarian cancer. Carcinoembryonic antigen (CEA) is a glycoprotein involved in cell-adhesion. It is thought to be upregulated in colon cancer.

While it is known that altered glycosylation is a hallmark of cancer, it is not known whether the altered glycosylation pattern causes the cancer or is caused by the cancer (Ghazarian *et al.*, 2011; Dube and Bertozzi, 2005).

1.4.6 Glycoprotein therapeutics and the Biopharmaceutical Industry

Originally, proteins for therapeutic use were obtained from natural sources. Historically insulin was obtained and purified from the pancreas of cattle and pigs, however due to a lack of purity, batch to batch variation and the constant risk of contamination this is no longer a viable option (Sethuraman and Stadheim, 2006). More recently, recombinant DNA technologies and protein production have been exploited for the production of new therapeutics. In a move away from the standard chemical drugs made by the pharmaceutical industry, this new industry, the biopharmaceutical industry, produces novel protein-based therapeutics. The industry is growing rapidly, with the biopharmaceutical market estimated to be now worth $100 billion annually. There are over 300 biopharmaceutical therapeutics currently approved and available for use (Langer, 2012). The first biotherapeutic produced and approved for human use was recombinant insulin (humulin) in 1982 (Walsh, 2010).

All therapies based on proteins, nucleic acids, cells and tissues are included in the term biotherapeutics. However, only protein biotherapeutics will be discussed here. Protein biotherapeutics are produced recombinantly using both prokaryotic and eukaryotic expression systems. Prokaryotic systems include the use of bacterial cells such as *E. coli* for the over-expression of the recombinant protein. Eukaryotic systems include the use of cell lines such as Chinese Hamster Ovary (CHO) cells as well as the use of insect cell lines and yeast-based systems. Transgenic animals are used for the production of antibodies. Expression of recombinant proteins in prokaryotic hosts is relatively cheap, simple and fast with large yields of product obtained. However, the use of mammalian cell lines is favored due to their ability to glycosylate proteins in a similar fashion to the native human protein. In 2006 over 70% of therapeutic proteins were observed to be glycosylated (Walsh, 2006; Sinclair and Elliott, 2005). The correct glycosylation of these protein therapeutics is essential as they affect properties including stability, efficacy, immunogenicity and solubility (Werner *et al.*,

2007). As already mentioned glycosylation is largely affected by the metabolic status of the cell. Due to pressures of high level expression and variable growth conditions at large scale, including fluctuations in nutrient composition of the media, heterogeneous glycosylation of the recombinant products often results. The difficulty of tightly regulating all the growth variables during large-scale expression results in a product containing a mixture of glycoforms, differentially glycosylated forms of the same protein (Werner *et al.*, 2007; Walsh, 2006; Sinclair and Elliott, 2005). Different glycoforms of the same therapeutic may have very different pharmacological properties. Erythropoietin (EPO), a hormone involved in the production of red blood cells, is a glycoprotein containing three potential N-linked glycosylation sites (Asn 24, 38, 83) and one potential O-linked glycosylation site (Ser 126) (Delorme *et al.*, 1992). The degree of sialylation on recombinant erythropoietin greatly affects its half-life *in vivo*. Sialylated and non-sialylated versions of erythropoietin were administered to rodents. The sialylated version had a plasma half-life of 5-6 hours in contrast with the non-sialylated version which had a half-life of less than 2 minutes (Erbayraktar *et al.*, 2003). An example of heterogenous glycosylation having an effect on protein half-life was also observed for the drug Lenercept, a "p55 tumour necrosis factor receptor fusion protein" (Jones *et al.*, 2007). This therapeutic was administered intravenously to cynomolgus monkeys and the pharmacokinetic properties were analysed. Blood samples were taken intermittently and the lenercept glycoforms present in these samples were analysed. The glycoforms of lenercept with a high number of terminal GlcNAc residues were rapidly cleared from circulation. It was proposed that the mannose receptor selectively cleared these glycoforms (Jones *et al.*, 2007). N-acetylneuraminic acid, when present at the terminal end of glycan structures, has been shown to increase the half-life of therapeutics in circulation. Hyper-glycosylation of therapeutics has also been shown to increase the half-life of therapeutics (Hossler *et al.*, 2009). The consensus sequence for N-linked glycosylation may also be engineered into the protein sequence, increasing the overall glycosylation of the protein. This is known to increase the activity and half-life of the therapeutic protein. Such was the case with erythropoietin (Hossler *et al.*, 2009).

CHO cells may add N-glycolylneuraminic acid (NGNA) to glycoproteins, a glycan not present on human glycoproteins. EPO expressed from CHO cells can contain various amounts of NGNA. When present in low amounts, <1% of total terminal residues there are no adverse effects observed. However, when present in higher

quantities, such as >7%, immune responses have been observed (Noguchi *et al.*, 2005).

Efforts have been made to alter the glycosylation system of the expression systems to either try ensure more homogenous glycosylation of the resulting therapeutics or to alter the glycosylation patterns for more favorable characteristics of the protein therapeutics. In 2005 Roche bought a glycoengineering company called GlycArt for $190 million. This company specializes in the modification of mammalian cell lines for optimal glycosylation of protein therapeutics. They engineered a cell line that was deficient in fucosyltransferases but over produced n-acetylglycosaminyltransferases. When used for the production of monoclonal antibodies, an extra N-acetylglucosamine residue was added to the Fc portion of the antibody, increasing the efficiency of its interaction with the Fc receptor (Sheridan, 2007).

As the glycosylation pattern of a protein therapeutic has a significant affect on its immunogenicity and efficacy, regulatory bodies increasingly require the glycoproteins to be thoroughly characterized initially and throughout production to ensure batch to batch consistency. Any alteration in the manufacturing process of biotherapeutics requires the resulting products to be extensively characterized to ensure changes in glycosylation have not occurred that could result in varied efficacy and safety. This has led to the industry wide need for fast, high-throughput glycoprotein analysis techniques (Werner *et al.*, 2007; Walsh, 2006; Sinclair and Elliott, 2005).

1.4.7 Analysis of glycans

Analysis of the glycan structures attached to glycoproteins is a difficult task due to the diversity of glycans and branched nature of the structure. There is a need for rapid and accurate methods for glycoprotein analysis for both the diagnostic and biopharmaceutical industries. There are many techniques currently used including 2-D Gel Electrophoresis, Mass Spectrometry (MS), High Performance Liquid Chromatography (HPLC), Nuclear Magnetic Resonance (NMR) and Lectin Microarrays. Certain techniques such as MS require enzymatic release of the glycan prior to analysis. Techniques that can analyse the glycoprotein as a whole, with no need for prior enzymatic treatment, are preferred.

1.4.7.1 2-D Gel Electrophoresis

Two-dimensional gel electrophoresis can be used as an initial step in the analysis of glycoproteins. An advantage of this technique is that intact glycoproteins can be analysed. However, it is difficult to separate various glycoforms in a mixture. 2-D gel electrophoresis separates proteins based on their molecular weight and isoelectric point. Slight differences in glycosylation do not cause much of a change in the molecular weight of the protein, so these different glycoforms are not fully separated from each other. Sialylation has a big effect on the pI of the protein, but it is very difficult to separate neutral sugars. Many glycoforms could be represented by one spot on the gel. This method does not allow for the characterisation of glycans, it only acts as a representation of the various glycoforms present (Geyer and Geyer, 2006).

1.4.7.2 Release of glycans from glycoproteins

The majority of glycoprotein analysis techniques require the release of glycans from the protein or at least the hydrolysis of the protein into smaller peptide fragments. Glycans can be released from glycoproteins using chemical or enzymatic means. Enzymatic methods are more specific than chemical methods, cleaving off specific glycans. Enzymes allow the cleavage of peptide fragments with glycans attached, allowing the site of glycosylation to be deciphered. However, the use of enzymes also has disadvantages due to the fact that not all glycans are cleaved due to the specificity of the enzyme. Chemical methods are not specific and can therefore cleave all glycans from the glycoprotein; however the use of chemicals can also lead to the degradation of the protein. This results in information being lost on the number and location of glycosylation sites. Trypsin and PGNaseF are commonly used in the enzymatic treatment of N-linked glycoproteins. Trypsin cleaves the C-terminal end of a Lys-Arg sequence while PGNaseF cleaves the majority of N-linked glycans through the hydrolysis of the GlcNAc attached to the asparagine residue of the protein. The release of O-linked glycans is more complicated than N-linked glycans due to the high variety in their structure. The enzymes used for this purpose are limited in their specificity, such as an O-glycanase. A commonly used chemical for glycan release is sodium borohydride, which results in the release of N-linked and O-linked glycans.

One disadvantage of this method is O-linked glycans can be damaged during this process (Geyer and Geyer, 2006).

1.4.7.3 Mass Spectrometry (MS)

MS has been used for glycoprotein analysis due to its relatively high levels of sensitivity and throughput. It can be used to detect different glycans in a sample as well as providing information on their structures. MS measures the mass of a sample through the measurement of its charge to mass ratio (Hitchen and Dell, 2006). The samples are converted to gas-phase ions in an ioniser and are accelerated to a mass analyzer where they are detected. This produces a mass spectrum. Glycoprotein analysis by MS may be carried out on intact or cleaved glycans. The mass of intact glycans can be deciphered. However structural analysis requires the glycans to be cleaved from the surface of the glycoprotein. These glycoprotein samples are subject to enzymatic or chemical cleavage prior to analysis. An issue with this method is that not all glycans will be removed from the glycan and although the structures of the glycans can be deciphered there is no information elucidated on the individual glycosylation sites (Dalpathado and Desaire, 2008; Geyer and Geyer, 2006;). Another disadvantage is the isomeric nature of carbohydrates. MS cannot distinguish between isomers due to their identical molecular mass. There is also no information on linkage or branching obtained (Pilobello and Mahal, 2007).

1.4.7.4 High Performance Liquid Chromatography (HPLC)

HPLC is a method than can separate samples in a mixture using a wide range of various adsorbtion and separation materials. The wide range of materials available allows for the separation of isomeric compounds, which is not possible using MS. Glycans are released from the glycoprotein for analysis. Labelling of the individual glycans can also be carried out which allows individual components to be quantified. There are a variety of different types of HPLC, each separating components based on different characteristics. Anion exchange chromatography separates molecules based on their charge. Reversed-phase HPLC (RP-HPLC) separates molecules based on their hydrophobicity while normal-phase HPLC (NP-HPLC) separates molecules based on their hydrophilicity. Both RP-HPLC and NP-HPLC also generally separate

the glycans by size. The structures of the glycans can be proposed when compared with the elution profiles of known glycans (Geyer and Geyer, 2006).

1.4.7.5 Nuclear Magnetic Resonance (NMR)

NMR analyses glycans based on the extent that a glycan distorts a magnetic field. This technique has the ability of completely characterising the entire glycan, from its mass and structure to its anomeric configuration and linkage. However, there are also some major disadvantages associated. It is an expensive technique that requires skilled personnel, specialised equipment and large concentrations of purified glycans. It is also better suited to the analysis of single glycan samples rather than mixtures (Geyer and Geyer, 2006).

1.4.7.6 Analysis of glycoproteins using lectins

1.4.7.6.1 Functions of lectins

Lectins are carbohydrate-binding proteins (CBPs) and have been described in the literature since the 19th century. The function of lectins depends on their source. In general, the main function of lectins is recognition. While plant lectins were the first to be discovered their exact functions are still not completely clear. Two assumptions made in the 1970s are still being reported and discussed today. It is widely believed that plant lectins play a defensive role, due to their ability to bind carbohydrates on the surface of invading pathogens such as bacteria and fungi. Lectins can also be insecticidal as they are found in large numbers in the seeds of plants that are poisonous to insects such as wheat germ agglutinin (WGA). The second function of lectins in plants is the most debated. It was proposed that lectins facilitate the binding of rhizobia to plants forming a symbiotic relationship. Rhizobia are nitrogen-fixing bacteria that can provide the leguminous plant with all its nitrogen needs. However, there has been conflicting evidence for this theory and it has neither been confirmed nor disproved. The function of microbial and viral lectins is more defined. Lectins facilitate the binding of microbes and viruses to host cells/tissues through recognition of carbohydrate patterns on their surface. The addition of an appropriate inhibiting sugar can block binding and therefore prevent infection. There are numerous functions described for animal lectins, it depends on the type of lectin. Collectins such

as Mannose binding proteins (MBP) turn on innate immunity through the initiation of the complement pathway caused by the recognition of surface mannose patterns on bacteria. MBPs have a modular nature; their C-terminus contains a carbohydrate recognition domain (CRD) which binds to the invading microbe while the N-terminus triggers complement activation through its interaction with the cell (Ambrosi *et al.,* 2005). Selectins facilitate the binding of leukocytes to blood vessels, allowing them to migrate into tissues to sites of inflammation (Nilsson, 2007; Sharon and Lis, 2004; Lis & Sharon, 1998). The family of galectins have various functions, from controlling development of tissues through their differential expression to regulating the immune system through their action on T cells (Camby *et al.,* 2006). Expression of galectin-1 and galectin-3 on certain tumour cells such as colon and breast has been noted, leading to the proposition of either as cancer biomarkers (Ambrosi *et al.,* 2005; Takenaka *et al.,* 2004).

1.4.7.6.2 Lectin analysis

Lectins have been exploited for the analysis of glycans due to their innate binding specificities. They have the ability to recognise and bind very specific glycan structures found on glycoproteins. Lectins can be used in a variety of ways for glycoprotein analysis such as Lectin Affinity Chromatography (LAC), Enzyme-Linked Lectin Assays (ELLA) and Lectin Microarrays (Geyer and Geyer, 2006).

LAC exploits the specificity of lectin binding. The lectins can be attached to a stationary phase in a column to separate a mixture of glycoproteins. The glycoproteins of interest may either be retained on the column, then requiring elution, or are free to flow through the column. This method may be used to retain the sugar of interest or to remove certain glycoproteins from a mixture. A series of columns with different specificities can be used to separate an array of glycoproteins from a mixture (Geyer and Geyer, 2006). ELLAs follow the same format as the commonly used ELISA. In this method the glycoprotein is immobilised on the surface of a 96-well plate where it can be probed with a selection of labelled lectins of different specificities. Binding is detected using colorimetric techniques (Thompson *et al.,* 2011). A panel of lectins of known specificities can be used which allow the pattern of glycans on the protein to be elucidated. Another format in which lectins are useful in glycoanalysis is the lectin microarray which consists of a panel of lectins immobilised on a chip. Samples

containing glycoproteins, generally fluorescently tagged, are added to the chip and allowed to hybridise. The pattern of spots generated allows the surface glycans of the glycoprotein to be elucidated (Pilbello and Mahal, 2007).

There are many advantages associated with the use of lectins for glycoprotein analysis. Primarily, the glycoproteins can be analysed in an intact state and there is no requirement for prior denaturation of the protein. Terminal glycan residues can be elucidated from the intact glycoprotein, while underlying residues can be exposed using specific enzymatic treatments, such as the use of glycosidases that remove the terminal glycan residues. Large concentrations of the glycoprotein are not required due to the specificity of the reaction. Lectins can be employed for *in vitro* but also *in vivo* work if conjugated to fluorescent proteins such as GFP (Varki *et al.,* 2008; Pilbello and Mahal, 2007; Geyer and Geyer, 2006).

Lectins are available from a wide-range of sources as discussed above (see Section 1.4.7.6.3). However, lectins commercially used in glycan analysis are generally obtained from plants. There can be large batch-to-batch variations and limited availability. In 2006, there were only 60 commercially available fully characterised lectins (Pilobello and Mahal, 2007). A way to circumvent this issue is through the use of lectins from different sources. The use of recombinant lectins from microbial sources is fast gaining prominence. These recombinant lectins can be produced cheaply, quickly, in large quantities and there are no issues with batch-to-batch variation.

1.5 Engineering of enzymes to potentially generate novel carbohydrate-binding proteins for glycoprotein analysis

Enzymes have two major innate functions; binding and catalysis. In order to catalyse a reaction they must first bind the substrate. This chitobiose binding of the enzyme and its actions are very specific. Binding of the substrate occurs through Van der Waals forces along with ionic and hydrogen bonds. If the catalytically active residues of the enzyme are known, then, in theory, the catalytic activity of the enzyme may be eliminated/reduced using site-directed mutagenesis. This would potentially make a unique carbohydrate binding protein. Vital information for this work is the identification of catalytically important residues (Plapp, 1995).

Before the 1970s, active site residues were generally modified using chemical means to elucidate their function. In this way the catalytically important residues were discovered. Each amino acid has a different side group which can react with chemicals in a different way. A chemical was chosen that would affect an amino acid in a specific way. Modification of the amino acid would then result in catalytic changes depending on the function of the amino acid (Ferhst, 1999). Now techniques such as site-specific mutagenesis and structural modelling are used. This allows for the mutation of proposed catalytic residues and generation of mutants with varying degrees of catalytic activity (Plapp, 1995).

The generation of lectins from enzymes exploits the specificity of the enzyme for its substrate. Various studies have tried to generate carbohydrate binding proteins from enzymes by removing their catalytic activity. Laine *et al.,* (2007) have generated an inactive endoglycosidase that binds chitin while Jokilami *et al.*, (2007) have generated an inactive polysialidase that binds to polysialic acid. Woods *et al.,* have a patent pending on Lectenz. They are generating lectin-like proteins from carbohydrate-processing enzymes. *In silico* analysis is performed to visualise and improve the enzyme-substrate binding using molecular dynamic simulations.

This study attempts to generate a carbohydrate-binding protein through the mutagenesis of a chitobiase enzyme from *P. luminescens*. The main substrate of this enzyme is chitobiose, a dimer of N-acetylglucosamine. It was proposed that the removal of the catalytic activity of this enzyme would result in the generation of a protein capable of detecting and binding to N-acetylglucosamine. This carbohydrate-binding protein would have huge potential in the fields of diagnostics and glycoanalysis due to the prevalence of N-acetylglucosamine residues on cell surfaces and biotherapeutics and the lack of choice in commercial GlcNAc binders available. This novel carbohydrate-binding protein could be incorporated or immmoblised into ELLAs, lectin microarrays and lectin affinity chromatography.

1.6 Objectives

This project aimed to study a homologue of the *S. marcescens* chitobiase enzyme that was identified in *Photorhabdus luminescens* (Chb). There were two main objectives of this project. The first aim was to study the catalytic activity of the enzyme. This included the cloning, expression and purification of the enzyme for the first time. Characterisation of the physio-chemical properties of the purified recombinant enzyme was then carried out, including the use of oligosaccharide substrates to determine its kinetics.

The second aim of this work was to take a closer look at the binding capabilities of the enzyme through the elimination of the catalytic activity of the enzyme using mutagenesis. This was in the interest of generating a protein that bound to N-acetylglucosamine for the purpose of glycoprotein analysis. The binding abilities of the catalytically inactive enzyme were determined using techniques such as ELLAs. This mutagenesis was carried out with the goal to potentially generate a novel carbohydrate binding protein with important binding specificities that may have uses in glycoprotein analysis.

A third and related aim was to carry out site-specific mutagenesis on the rChbL enzyme and rChb binding protein. In the case of the rChbL enzyme this was to provide information of the importance of certain chosen active site residues in catalysis, while also being an attempt to improve the catalytic activity of the enzyme. In the case of the rChb binding protein the mutagenesis was carried out on the same active site residues in order to determine the importance of these resdidues in substrate binding and to improve the binding abilities of the protein.

Finally the catalytic activity of the rChbL enzyme was tested on natural substrates such as chitin to determine if it would be a viable option for commercial chitin breakdown.

Chapter 2:

Materials and Methods

2.1 Vectors, Primers, Constructs and Bacterial Strains:

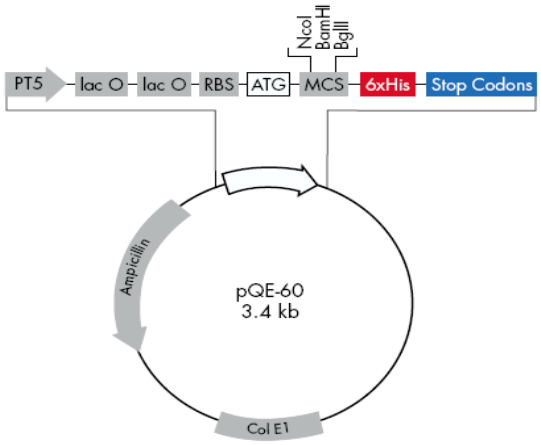


Fig 2.1: pQE60 vector from Qiagen. This 3.4kb vector from Qiagen contains the following features; a multiple cloning site (MCS) followed by a His_6 tag which enables the protein to be expressed with an N-terminal His_6 tag, a T5 promoter/lac operon for inducible expression of the protein and ampicillin resistance in the form of a beta lactamase gene. The vector image was obtained from the Qiagen website (www.qiagen.com).

Table 2.1 Primer Sequences

Primer Name	Sequence (5′-3′)
Cloning:	
chb+_F	GTCAGCCCATGGACAAGAAATTTAAATTGCAT
chb+_R	GTCAGCGGATCCCTTTACATTATCAATTCG
chb-_F	GTCAGCCCATGGCGCAGCAAGTGGTT
chb-_R	GTCAGCGGATCCCTTTACATTATCAATTCG
Mutagenesis:	
chb+cbp_F	[P]GGGGAGCTGACGCGAAGAA
chb+cbp_R	[P]AAAATGCCAAGTCTCCAATGGTC
chb+cbplnk_F	[P]GGTTCTCATCACCATCACCATCACTAAGCTTAATTA
chb+cbplnk_R	[P]GCCACCAGATCCTGCCTTTACATTATCAATTCG
63A_F	[P]GGGGCAGATGCGGCTAAATGTAATCG
63A_R	[P]CAGTACCGCACAGTTAACGCCGTG
648_F	[P]GGGACACTCTCTACGCGGGCGGC

648_R	[P]AAAAATTCACCCCAACGCGTTTGGTGGC
545A_chb_F	[P]GGAGATGAAGCGGCGAATATCCGTTTG
545A_cbp_F	[P]GGAGCTGACGCGGCGAATATCCGTTTG
545A_R	[P]CCCAAAATGCCAAGTCTCCAATGGTG
551A_F	[P]ATCCGTTTGGGAGCGGGATTCCAGG
551A_chb_R	[P]ATTCTTCGCTTCATCTCCCCCAAAATG
551A_cbp_R	[P]ATTCTTCGCGTCAGCTCCCCCAAAAT
545A551A_chb_F	[P]GGAGCGGGATTCCAGGATAAAAATGGTCCTATC
545A551A_chb_R	[P]CCTTCACTTCGCGCGTTATAGGCAAAC
545A551A_cbp_F	[P]CGTTTGGGAGCGGGATTCCAGGAT
545A551A_cbp_R	[P]GATATTCGCCGCGTCAGCTCCCCC

*[P] denotes 5′ phosphorylation of primer

Table 2.2 Constructs

Construct Name	Description	Source
pQE60	Expression vector for C-terminally tagged His_6 proteins, T5 promoter/*lac* operon, amp^R, ColE1 origin	Qiagen
pQE60_Chb+	pQE60 plasmid containing *chb* ORF with N-terminal signal peptide	This work
pQE60_Chb-	pQE60 plasmid containing *chb* ORF without N-terminal signal peptide	This work
pQE60_Chb_CBP	pQE60 plasmid containing *chb* ORF with N-terminal signal peptide with D539A and E540D mutations	This work
pQE60_Chb_CBPlnk	pQE60 plasmid containing *chb* ORF with N-terminal signal peptide, D539A and E540D mutations and C-terminal linker of 3xGlycine and 1xSerine	This work
pQE60_Chb _63A	pQE60 plasmid containing *chb* ORF with N-terminal signal peptide and W63A mutation	This work
pQE60_Chb _648A	pQE60 plasmid containing *chb* ORF with N-terminal signal peptide and D648A	This work

	mutation	
pQE60_Chb _545A	pQE60 plasmid containing *chb* ORF with N-terminal signal peptide and K545A mutation	This work
pQE60_Chb _551A	pQE60 plasmid containing *chb* ORF with N-terminal signal peptide and D551A mutation	This work
pQE60_Chb _545A/551A	pQE60 plasmid containing *chb* ORF with N-terminal signal peptide and K545A and D551A mutations	This work
pQE60_Chb _CBP_63A	pQE60 plasmid containing *chb* ORF with N-terminal signal peptide with D539A, E540D and W63A mutations	This work
pQE60_Chb _CBP _648A	pQE60 plasmid containing *chb* ORF with N-terminal signal peptide with D539A, E540D and D648A mutations	This work
pQE60_Chb _CBP _545A	pQE60 plasmid containing *chb* ORF with N-terminal signal peptide with D539A, E540D and K545A mutations	This work
pQE60_Chb _CBP _551A	pQE60 plasmid containing *chb* ORF with N-terminal signal peptide with D539A, E540D and D551A mutations	This work
pQE60_Chb_CBP _ 545A/551A	pQE60 plasmid containing *chb* ORF with N-terminal signal peptide with D539A, E540D, K545A and D551A mutations	This work

Table 2.3 Strains of bacteria used

Strain	Description	Use	Source
E. coli JM109	FΔtraD36, *pro*AB+ *lac*Iq, Δ*lac*Z M15 *end*A1 *rec*A1 *hsd*R17 (rk-, mk+) *mcr*A *sup*E44 – *gyr*A96 re*l*A1 Δ (*lacproAB*) *thi-1*	All purpose cloning strain. Produces stable plasmid DNA.	Sigma
E. coli XL10 Gold	Δ(*mcr*A)183 Δ(*mcr*CB-*hsd*SMR -mrr)173 *end*A1 *rec*A1 *rel*A1 *gyr*A96 *sup*E44 *thi-1* *lac*[FΔ*pro*AB *lac*IqZΔM15 ::Tn10 (tetR) Amy (CmR)]	All purpose cloning strain. High transformation efficiency. Produces stable plasmid DNA.	Stratagene
E coli KRX	*tra*D36, Δ*amp*P, *pro*A+B+, *lac*Iq, Δ(*lacZ*)M15] Δ*omp*T, *end*A1, *rec*A1, *gyr*A96 (Nalr), *thi-1*, *hsd*R17 (rk-mk+), *rel*A1, *sup*E44, Δ(*lac-pro*AB), Δ(*rha*BAD) ::T7 RNA polymerase	Protease deficient. High transformation efficiency. Expression strain	Promega
Photorhabdus luminescens	TT01 Wildtype		Dr. David Clarke (UCC)

2.2 Microbiological Media:

- **Luria Broth (LB) Broth**

Tryptone	10 g/L
Yeast Extract	5 g/L
Sodium Chloride	10 g/L

Media was autoclaved before use.

- **Luria Broth (LB) Agar Plates**

Tryptone	10 g/L
Yeast Extract	5 g/L
Sodium Chloride	10 g/L
Agar	15 g/L

Agar was autoclaved and 15 mL was poured aseptically into sterile Petri dishes.

- **Terrific Broth (TB) Medium**

Base Medium:	
Tryptone	12 g
Yeast Extract	24 g
Glycerol	4 mL

Made up to 900 mL with dH_2O and sterilised by autoclaving. 100 mL of sterile 1 M potassium phosphate buffer, pH 7.2, was added aseptically before use.

- **Potassium Phosphate Buffer (1 M)**

KH_2PO_4 (0.17 M)	23.1 g/L
K_2HPO_4 (0.72 M)	12.4 g/L

The buffer was sterilised by autoclaving.

2.3 Buffers and Solutions:

- **Buffers for the preparation of competent cells**

TBF1 :

RBCl	100 mM
$MnCl_2$*	50 mM
$CaCl_2$	10 mM
CH_3CO_2K	30 mM
Glycerol	15%
pH	5.8

Made with ultrapure water. $MnCl_2$ was added after adjusting pH. Solution was sterile filtered through a 0.45 μM filter. Stored at 4^oC.

TBF2:

MOPs	10 mM
RbCl	10 mM
$CaCl_2$	75 mM
Glycerol	15%
pH (KOH)	6.8

Made with ultrapure water. Solution was sterile filtered through a 0.45 μM filter and stored at 4^oC.

- **Phosphate Buffered Saline 10x (PBS)**

Na_2HPO_4	10.9 g/L
NaH_2PO_4	3.2 g/L
NaCl	90 g/L

For PBST, the detergent Tween-20 was added to a final concentration of 0.1% (v/v).

- **Tris Buffered Saline (TBS)**

Tris	20 mM
NaCl	150 mM
$CaCl_2$	1 mM
$MgCl_2$	1 mM
pH	7.6

The pH was adjusted to 7.6 with HCl before addition of the $MgCl_2$ for a final concentration of 1 mM. For TBST the detergent Tween-20 was added to give a final concentration of 0.1% (v/v).

- **SDS PAGE Buffer (5x)**

Tris	15 g/L
Glycine	72 g/L
SDS	5 g/L
pH	8.3

- **SDS-PAGE Sample Buffer (5x)**

Glycerol	5 mL
Tris	1.25 mL, 0.5 M pH 6.8 stock
SDS	2 mL 10% stock
β-mercaptoethanol	0.5 mL
Bromophenol Blue	0.5% (w/v)
dH_20	1 mL

- **Coomassie Blue Stain Solution**

Methanol	40% (v/v)
dH_2O	50% (v/v)
Acetic Acid	10% (v/v)
Brilliant Blue	0.25% (w/v)

- **Coomassie Blue Destain Solution**

Methanol	40% (v/v)
dH_2O	50% (v/v)
Acetic Acid	10% (v/v)

- **Tris Acetate EDTA Buffer 50x (TAE)**

Tris	242 g/L
EDTA	100 mL/L, 0.5M Stock
Glacial Acetic Acid	57.1 mL/L
pH	8.0

- **0.7% Agarose Gel**

TAE Buffer	1 L
Agarose	7 g/L

Gel was boiled until all agarose was dissolved. Stored at 60^{o}C.

- **Agarose Gel Loading Dye**

Bromophenol Blue	0.25% (w/v)
Xylene Cyanol	0.25% (w/v)
Ficoll Type 100	15% (w/v)

- **Ethidium Bromide Stain**

A 10 mg/mL stock solution in dH_2O was stored at 4ºC in the dark. For the staining of agarose gels, 100 µL of stock solution was mixed into 1 L of dH_2O. The staining solution was kept in a plastic tray and covered to protect against light. Used ethidium bromide stain was collected and the ethidium bromide was extracted by mixing with a de-staining bag (GeneChoice, Gentaur) overnight. The clear liquid was disposed of routinely, and the ethidium bromide was incinerated.

- **Lysis Buffer**

NaH_2PO_4	20 mM
NaCl	0.5 M
Imidazole	20 M
pH	7.4

- **TMB Solution :**

Citric Acid	1.37 mL, 0.1 M Stock
Sodium Citrate	3.63 mL, 0.1 M Stock
dH_2O	5 mL

1 TMB Tablet (Sigma T3405) was dissolved in the TMB solution and 2 µl of H_2O_2 was added directly before use.

- **1,2,3 Method**

Solution 1	
Glucose	0.5 M
EDTA	0.1 M
Tris HCl	1 M
Solution 2	
NaOH	1 M
SDS	10%
Solution 3	
KOAc	3 M
pH	4.8

11.5 mL of glacial acetic acid and 28.5 mL dH_2O were added to 60 mL of 5 M potassium acetate to give 3 M potassium and 5 M acetate.

- **TE Buffer**

Tris	10 mM
Na_2-EDTA	1 mM
pH	8.0

- **DNS Assay**

Solution A

Sodium Potassium Tartrate	300 g

Dissolved in 500 mL dH_20 and heated to 40^oC.

Solution B

NaOH	16 g
DNS	10 g

The NaOH was dissolved in 200 mL dH_20 and heated to 90^oC before adding the DNS.

Solution A and B were mixed together and allowed to cool. Phenol was added at a final concentration of 0.2%. The solution was made up to 1 L with dH_20 and then stored in the dark at room temperature.

- **Western blot transfer buffer**

Tris	60 mM
Glycine	149 mM
SDS	0.055%
Methanol	10% (v/v)

- **Methanol acetate borate solution (MABS)**

Sodium acetate	240 mg
Boric acid	200 mg

The chemicals were dissolved in 10 mL methanol. The solution was stored for one week at 4^{o}C.

2.4 Antibiotics:

Ampicillin was prepared in dH_2O at a concentration of 100 mg/mL and was stored at -20^{0}C. It was used at a working concentration of 100 μg/mL for solid and liquid media for growth of *E. coli*.

2.5 Storing of Bacteria:

Bacteria were stored as 33% glycerol stocks. 1 mL of an overnight culture (late log phase) was aseptically added to 500 μL of sterile 80% glycerol. If the bacteria contained a plasmid with antibiotic resistance the appropriate antibiotic was added to the overnight culture. These stocks were stored at -20^{o}C and -80^{o}C, as working stocks and permanent stocks, respectively.

2.6 Culturing of Bacteria:

A streak was taken from the working stock onto a fresh LB agar plate containing antibiotics, where appropriate. The plate was incubated overnight at 37^{o}C. These could be stored at 4^{o}C after growth for up to one week.

2.7 Isolation of DNA:

2.7.1 Isolation of Genomic DNA

Genomic DNA was obtained using a Genomic DNA Purification Kit from Promega. The kit was used according to the manufacturer's instructions. 1 mL of an overnight bacterial culture was centrifuged at 13,000 rpm for 2 minutes. The culture media was removed and the pellet was gently resuspended in 600 μL nuclei lysis solution. This was incubated at 80^{o}C for 5 minutes in a heating block to lyse cells and was then

cooled to room temperature. 30 μL of RNase solution was added to the cell lysate and inverted 5 times to mix. This was incubated at 37^{o}C for 30 minutes and then cooled to room temperature. 200 μL of protein precipitation solution was added and the solution was vortexed vigorously at high speed for 20 seconds. It was kept on ice for 5 minutes and then centrifuged at 13,000 rpm for 3 minutes. The culture media was removed and added extremely slowly to an eppendorf containing 600 μL room temperature isopropanol. This was gently mixed by inversion until strands of thread-like DNA could be seen. This was centrifuged at 13,000 rpm for 2 minutes. The culture media was removed and 600 μL of room temperature 70% ethanol was added to the pellet. This was gently inverted to wash the pellet. It was centrifuged at 13,000 rpm for 2 minutes. The ethanol was aspirated off and the pellet was air dried in a vacuum at maximum speed for 10 minutes. 100 μL of DNA rehydration solution was added and it was incubated at 4^{o}C overnight. The DNA was then stored at 4^{o}C.
The presence of the plasmid DNA was then verified by visualisation on a 0.7% agarose gel (see Section 2.8 for an example).

2.7.2 Isolation of Plasmid DNA

Two methods for the isolation of plasmid DNA were carried out during this work. The method used depended on the purpose of the experiment. The 1, 2, 3 Method first described by Birnboim and Doly (1979) yielded good quality DNA and was cheaper than the kit method; however it was more time-consuming. It was used for the screening of multiple clones from transformations.
The HiYield Plamid miniprep kit (RBC Biosciences) yielded high quality, pure DNA and was less time-consuming. It was used for the preparation of plasmid DNA for transformations and sequencing.

2.7.2.1 Isolation of Plasmid DNA - 1, 2, 3 Method

This method allows covalently closed circular DNA to remain double stranded while selectively denaturing the high molecular weight chromosomal DNA with an alkaline treatment. Bacteria was removed from an agar plate using a sterile loop and added to an eppendorf containing 200 μL of Solution 1 (see Section 2.3) and resuspended. Solution 1 contained EDTA which had two purposes; firstly to chelate any divalent

cations such as Mg^{2+} and Ca^{2+} which are crucial for DNase activity and secondly to destabilise the bacterial membrane. Alternatively 1.5 mL of a bacterial culture was centrifuged at 13,000 rpm for 2 minutes and the culture media discarded. The pellet was resuspended in 200 μL of Solution 1. The re-suspension was left at room temperature for 5 minutes.

200 μL of Solution 2 (see Section 2.3) was added to lyse the cells through the action of SDS and to break down any dsDNA due to the presence of NaOH. This was mixed by inversion for several seconds. It was then left on ice for 5 minutes.

200 μL of Solution 3 (see Section 2.3) was added to neutralise the solution. The smaller plasmid DNA was able to renature and solubilises in these conditions however the larger chromosomal DNA and proteins precipitated out of solution. This solution was mixed by inversion for several seconds. It was left on ice for 10 minutes until a clot of chromosomal DNA formed. This was pelleted by centrifugation at 13,000 rpm for 10 minutes.

The following steps were performed in order to clean-up the plasmid DNA sample:

The culture media was transferred to a fresh tube to which 450 μL of phenol chloroform isoamylalcohol (25:24:1) was added. This was mixed by vortexing and then centrifuged at 13,000 rpm for 5 minutes.

The aqueous layer was carefully transferred to a fresh tube to which an equal volume of isopropanol was added. This was then mixed by inversion for several seconds and left on ice for 10 minutes.

It was centrifuged at 13,000 rpm for 10 minutes to pellet the plasmid DNA. The culture media (isopropanol) was removed and the pellet was washed with 70% ethanol. The pellet was not resuspended. It was centrifuged again for 4 minutes. The ethanol was removed and the plasmid DNA was dried using a vacuum dryer (SpeedVac) for 5 minutes. The plasmid DNA was resuspended in 50 μL TE Buffer and was stored at $4^{o}C$. The presence of the plasmid DNA was then verified by visualisation on a 0.7% agarose gel.

2.7.2.2 Isolation of Plasmid DNA - Kit Method

High quality plasmid DNA was obtained using a HiYield plasmid mini kit (RBC Biosciences). This kit follows the same principles as the 1, 2, 3 method above and was used according to the manufacturers instructions. Bacteria was removed from an agar

plate using a sterile loop and resuspended in 200 µL PD1 buffer. Alternatively 1.5 mL of an overnight culture was centrifuged at 13,000 rpm for 2 minutes and the pellet was resuspended in 200 µL PD1 buffer. 200 µL of PD2 buffer was added to the tube re-suspension and gently inverted to mix. This was allowed to stand at room temperature for 2 minutes until the lysate cleared. 300 µL of PD3 buffer was added and mixed immediately by inversion. This was centrifuged at 13,000 rpm for 2 minutes. The clear lysate was removed and added to a PD column that was placed in a 2 mL collection tube. This was centrifuged at 13,000 rpm for 30 seconds. The flow through was discarded. 600 µL of W1 buffer was added to the column and centrifuged at 13,000 rpm for 30 seconds. The flow through was removed. 500 µL of wash buffer (+ ethanol) was added and centrifuged at 13,000 rpm for 30 seconds. The flow through was discarded. This step was repeated. The column was then centrifuged at 13,000 rpm for 10 minutes to ensure removal of all residual ethanol. The column was transferred to a clean 1.5 mL eppendorf. 50 µL of elution buffer was added directly to the centre of the membrane and was allowed to stand for 2 minutes to ensure it was absorbed. The column was then spun at 13,000 rpm for 2 minutes to elute the purified plasmid DNA. The flow through was kept and stored at 4^{o}C.

The presence of the plasmid DNA was then verified by visualisation on a 0.7% agarose gel.

2.8 Agarose Gel Electrophoresis:

DNA was analysed by electrophoresis in BioRad gel apparatus. Agarose was made by adding the required concentration of agarose (between 0.7% and 2%) to 1X TAE Buffer (see Section 2.3) and dissolved by boiling. This solution was stored at 60^{o}C. When needed, 30 mL-50 mL (depending on the size of the gel needed) was poured into the gel casing and a plastic comb was slotted on to provide wells for the addition of samples. This was left to set at room temperature. The set gel was then placed in the gel rig, 1X TAE running buffer was added and the comb was removed. Samples were mixed with agarose gel loading dye (see Section 2.3) in a 9:1 ratio. The loading dye was to facilitate the loading of samples and to give an indication of sample migration during electrophoresis. 6 µL of 1 Kb Plus DNA Ladder (Invitrogen) was added to each gel as a molecular weight marker. Gels were run at 120 volts for 20-40

minutes depending on the size of the gel. Gels were then removed and immersed in an ethidium bromide bath (see Section 2.3) for 20 minutes. The gel was then visualised and the image captured on a UV transilluminator coupled with an image analyser.

a) **b)**

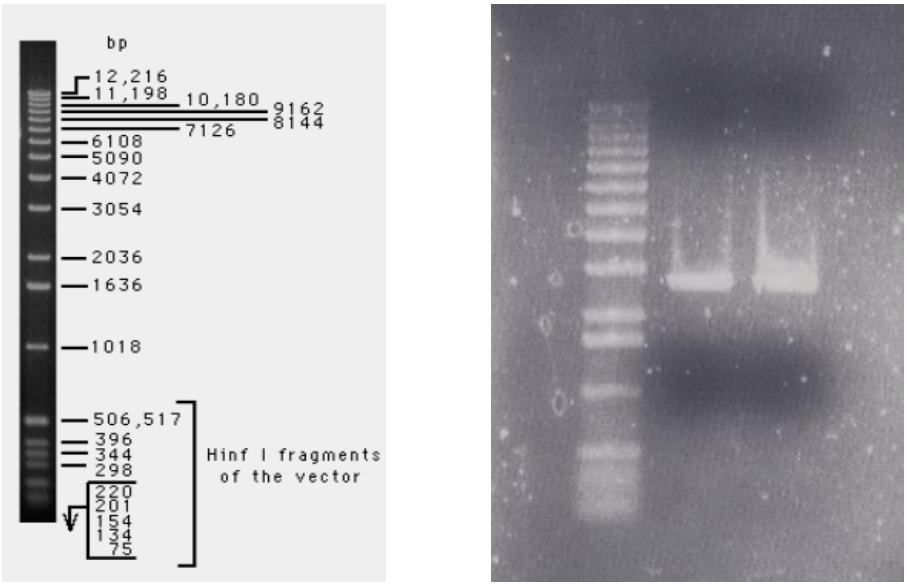


Fig 2.2: a) 1 kb Molecular Marker, b) Example of an ethidium bromide stained agarose gel. a) The DNA ladder (Invitrogen) used as a molecular weight marker. Image obtained from the Invitrogen website (www.invitrogen.com). b) An example of a DNA agarose gel. Lane 1; Ladder, Lane 2; Plasmid sample, Lane 3; Plasmid sample.

2.9 Gel extraction procedure for the isolation of DNA from agarose gels

DNA was isolated from agarose gels using a Hi Yield^{TM} Gel/PCR DNA extraction kit from RBC Bioscience. The kit was used as per the manufacturer's instructions. The gel was stained with 30 mL of Sybr Safe DNA gel stain (Invitrogen) for 30 minutes in the dark. The gel was then put under blue light to visualise the bands of DNA. The appropriate bands were excised from the gel using a scalpel and transferred to a centrifuge tube. 700 µL of Detergent Free (DF) buffer was added to the gel slice which was then mixed by vortexing. This was incubated at 55^{o}C for 10-15 minutes until the gel slice was dissolved. A DF column was placed in a collection tube. 800 µL of the sample mixture was added to the column and centrifuged at 13,000 rpm for

30 seconds. The flow through was discarded. The remainder of the sample mixture was added to the column and centrifuged at 13,000 rpm for 30 seconds. The flow through was again discarded. 600 µL of wash buffer was then added to the column and spun at 13,000 rpm for 30 seconds. The flow through was discarded. This wash step was repeated. The column was then centrifuged at 13,000 rpm for 10 minutes to ensure complete removal of the wash buffer as it contained ethanol. The flow through was discarded. The DF column was transferred to a sterile centrifuge tube. 20-50 µL of elution buffer was added to the column and allowed to stand for 2 minutes until all the buffer had absorbed. This was centrifuged at 13,000 rpm for 2 minutes. The purified DNA in the flow through was collected and stored at 4^{o}C.

2.10 DNA sequencing

Recombinant clones and potential mutants were verified by DNA sequencing. Commercial sequencing services were provided by MWG Biotech/Eurofins and Source BioScience gene service. Suitable sequencing primers for standard vectors were sent with samples to be analysed. Samples were sent as plasmid preparations prepared in dH_2O.

2.11 PCR

Standard PCR reaction mixture:

Template DNA	1 µL
dNTPs	1 µL
Primers (10 µM)	1 µL of each
Buffer (10X)	5 µL
dH_2O	40 µL
Taq polymerase	1 µL

Standard PCR programme cycle for Phusion *Taq* polymerase.

Reactions:

Stage 1: Step 1: 95ºC for 10 minutes

Stage 2: Step 1: 95ºC for 30 seconds

Step 2: Annealing temperature for 30 seconds

Step 3: 72ºC for 30 seconds per kb to be synthesised

Stage 3: Step 1: 72ºC for 10 minutes

2.12 Enzymatic reactions

All enzymes and relevant buffers were obtained from Invitrogen Life Technologies®, New England Biolabs® or Sigma Corporation and were used according to the manufacturer's instructions.

*Dpn*I Restriction Digest

PCR Product	10 µL
*Dpn*I	2 µL
Relevant Buffer (10X)	2 µL
BSA (10X)	2 µL
dH_2O	4 µL

The reaction mix was incubated in a 37^{o}C water-bath for 16 hrs.

Restriction digest

PCR Product	15 µL
Restriction Enzyme	3 µL
Relevant Buffer (10X)	3 µL
BSA (10X)	3 µL
dH_2O	6 µL

The reaction mix was incubated in a 37^{o}C water-bath for 16 hrs.

Ligation	
PCR Product	12 μL
Vector	4 μL
T4 DNA Ligase	1 μL
Ligase Buffer (10X)	2 μL
dH_2O	1 μL

The reaction was left to incubate overnight at room temperature.

2.13 Site-specific mutagenesis

Mutations were introduced into open reading frames (ORF) on plasmid constructs by PCR amplification using complementary phosphorylated primers carrying the desired mutation. A standard PCR reaction mixture, as described in Section 2.11, was set up. The PCR was carried out using high fidelity Phusion *Taq* polymerase and the standard program cycle for Phusion *Taq* polymerase (see Section 2.11). An extension time of 2 minutes was used to amplify plasmids of 3.5-4 kb. The template DNA was then eliminated by digestion with *Dpn*I restriction endonuclease (see Section 2.12). *Dpn*I selectively digests the template DNA from *dam+ E.coli* strains over the newly synthesized DNA as it is biased towards a methylated recognition sequence. The amplified plasmid, with the incorporated mutation, was circularised by ligation (see Section 2.12) which was facilitated by the use of phosphorylated primers.

2.14 Transformations

2.14.1 Preparation of highly efficient competent cells

The Rubidium Chloride (RC) Method:
This follows the method first described by Hanahan (1985). 5 mL of LB with the appropriate antibiotics was inoculated with a single colony of the relevant bacteria and allowed to grow overnight at 37^{o}C in a shaking incubator at 220 rpm. 2.5 mL of this overnight culture was added to 250 mL LB in a 1 L flask and allowed to grow to an OD_{600} of 0.5 (early-mid exponential phase) at 37^{o}C at 220 rpm. The flask was

placed on ice for 10 minutes to cool. All subsequent steps took place at 4°C. The culture was centrifuged to collect the cells at 3,000 rpm for 5 minutes. The culture media was decanted and the cell pellet was resuspended very gently in 75 mL of ice-cold TBF1 buffer (see Section 2.3). The suspension was then left on ice for 90 minutes. The cells were again collected by centrifugation as before. The culture media was again discarded and the pellet was resuspended very gently in 10 mL ice-cold TBF2 (see Section 2.3). The cell suspension was aliquoted (200 µL/tube) into sterile microfuge tubes (1.5 mL) which were then flash frozen in -80°C ethanol. The competent cells were then stored at -80°C.

2.14.2 Transformation of Competent Cells

The competent cells were taken from the -80°C freezer and thawed on ice. 200 µL of competent cells were mixed with plasmid DNA or ligation reaction. This mixture was kept on ice for 30 minutes. It was then heat shocked at 42°C in a water bath for 30 seconds and quickly returned to ice for 2 minutes. 800 µL of LB Broth was added aseptically and incubated at 37°C in a water bath for 1 hour. 200 µL was spread on agar plates with the appropriate antibiotics and incubated overnight at 37°C. The rest was concentrated by centrifugation at 13,000 rpm for 1 minute. The culture media was discarded and the pellet was resuspended in 100 µL sterile saline solution This was spread on an agar plate and incubated overnight at 37°C as previously described.

2.14.3 Determination of cell efficiency

Competent cell efficiency is defined in terms of the number of colony forming units obtained per µg of transformed plasmid DNA. A 10 ng/µl stock of pBR322 plasmid DNA was diluted to 1 ng/µL, 100 pg/µL and 10 pg/µL. 1 µL of each dilution was transformed as described above (see Section 2.14.2). The cell efficiency was calculated from the number of colonies obtained, taking into account the dilution factor and the fraction of culture transferred to the spread plate.

2.15 Expression

Single colonies of the bacteria containing the expression plasmid of interest were prepared from a working glycerol stock. A single colony was selected and used to inoculate a sterilin containing 5 mL of LB (and antibiotic if required). The culture was incubated overnight at 37^{o}C at 220 rpm. A 1 L baffled flask containing 500 mL TB media and appropriate antibiotic was inoculated with 5 mL of this overnight culture. The culture was grown at 37^{o}C, with shaking at 220 rpm, until an OD_{600} of 0.5-0.6 was reached (mid-late exponential phase). At this point IPTG was added to a final concentration of 100 μM to induce expression of the protein. The culture was then incubated at 30^{o}C for a defined period of time before the cells were collected by centrifugation at 4,000 rpm for 20 minutes.
At this stage the pellet can be stored at -20^{o}C and the culture media can be stored at 4^{o}C depending on requirements.

2.15.1 Generation of soluble and insoluble cellular fractions from expression cultures

5 mL samples were taken from expression cultures and the absorbance was read. The samples were normalised to ensure a uniform number of cells present on each cellular sample. Each sample was centrifuged at 13,000 rpm for 10 minutes. The culture media was discarded and the cell pellet was resuspended in 1X lysis buffer (see Section 2.3). The samples were sonicated as described in Section 2.16.1 and centrifuged at 13,000 rpm for 10 minutes. The culture media (soluble fraction) was removed and 5X protein loading dye was added to it. The pellet (insoluble fraction) was resuspended in 1X protein loading dye and both were boiled at 100^{0}C for 10 minutes before addition to a gel.

2.16 Preparation of cleared lysate for protein purification

Cells can be lysed by numerous methods. The methods employed here were sonication, cell disruption or by the addition of lysozyme.

2.16.1 Cell Lysis by Sonication

A cell pellet from a 500 mL expression culture was re-suspended in 50 mL lysis buffer containing 20 mM imidazole, pH 7.6 (see Section 2.3). The cells were disrupted with a 3 mm micro-tip sonicator (Sonics & Materials Inc.) using 2.5 s, 40 kHz pulses for 60 s. The lysate was then centrifuged at 4,000 rpm for 20 minutes at 4ºC (using a Rotanta 460 R rotor) to pellet any insoluble material. The cleared lysate was transferred to a fresh container and passed through a 0.2 μm filter to remove any remaining cell debris and stored at 4°C for protein purification.

2.16.2 Cell Lysis using Lysozyme

A cell pellet from a 500 mL culture was resuspended in 50 mL lysis buffer containing 20 mM imidazole, 10 mg/mL lysozyme and 10 μg/mL DNase at pH 7.6. The cells were incubated in a water bath at 37^{o}C for 30 minutes. The cell lysate was centrifuged at 4,000 rpm for 20 minutes at 4^{o}C to remove any insoluble debris. The lysate was then transferred to a fresh container and passed through a 0.2 μm filter to remove any remaining debris. It was then stored at 4^{o}C for protein purification.

2.16.3 Cell Lysis by Cell disruption

A cell pellet from a 500 mL culture was resuspended in 200 mL lysis buffer containing 20 mM lysis buffer (see Section 2.3) and 0.01% antifoam. The cells were fully homogenised in this buffer until they were the consistency of yoghurt. The cells should be approximately 250 g/L wet weight or 25 g/L dry weight. The pressure was selected according to the organism being disrupted. For *E. coli* a pressure of 15,000 psi was chosen. The sample reservoir was filled with lysis buffer and run through the system at the selected pressure. The sample was then loaded into the reservoir and disrupted. The disruption step could be repeated if the sample was particularly viscous. 20 mL of buffer was chased through the instrument at the selected pressure to harvest any residual sample. The unit was thoroughly cleaned after each use by running 150 mL of dH_2O, 150 mL of ethanol followed by a final 150 mL dH_2O through the machine. The disrupted sample was then spun at 13,000 rpm for 30

minutes at 4°C to pellet any insoluble debris. The culture media which was the cell lysate was stored at 4°C for protein purification.

2.17 Preparation of culture media containing protein

Conversely if the protein is secreted from the cell as was the case with the protein expressed in this work it can be collected straight from the culture media.
The culture media was centrifuged again after the initial cell harvest at 13,000 rpm for 30 minutes. The sample was then passed through a 0.2 μm filter to remove any debris and stored at 4°C for protein purification.

2.18 Isolation of different cell fractions by the water lysis method

This method was first described by Henderson and MacPherson (1986). The method was performed in this study as described by Ward *et al.* (2000). The cells were lysed using osmotic pressure. The periplasmic, cytoplasmic and membrane fractions were isolated in this way to determine which cellular fraction contained the most protein.
50 mL of an expression culture was centrifuged at 4,000 rpm for 10 minutes at 10°C. The culture media was discarded and the pellet was resuspended in 10 mL 0.2 M Tris at pH 8.0. This re-suspension was stirred at room temperature for 20 minutes. Once the stirring was complete this was called Time Zero. At Time Zero 4.85 mL of 1 M Sucrose/0.2 M Tris/1 mM EDTA was added. At Time 1.5 minutes 65 μL of 10 mg/mL lysozyme prepared in 1 M Sucrose/0.2 M Tris/1 mM EDTA was added. At Time 2 minutes 9.6 mL of dH_20 was added and the mixture was allowed to stir constantly for 20 minutes. This mixture was spun at 18,000 rpm for 20 minutes at 4°C. The culture media was retained and stored at -20°C. This was the periplasmic fraction. The pellet was resuspended in 15 mL dH_2O until fully homogenised. The homogenised mixture was then stirred for 30 minutes at room temperature. The mixture was then spun at 18,000 rpm for 20 minutes at 4°C. The culture media was retained and stored at -20°C. This was the cytoplasmic fraction. The pellet was washed twice with 30 mL 0.1 M Sodium Phosphate buffer containing 1 mM mercaptoethanol. Each time the pellet was resuspended using a homogeniser and sedimented at 18,000 rpm for 20 minutes at 4°C. The final pellet was resuspended in

an ice-cold re-suspension buffer of 0.1 M Sodium Phosphate containing 1 mM mercaptoethanol at pH 7.2 using a 1 mL homogeniser and stored at -20^{o}C. This was the mixed membrane fraction.
Each fraction was analysed by SDS-PAGE as described previously to determine protein location.

2.19 Purification of the recombinant protein

Immobilised Metal Affinity Chromatography (IMAC) was used to purify the recombinant proteins as they contained C terminal His_6 tags. IMAC is based on the covalent binding event between proteins and metal ions. Here a resin containing Ni^{2+} was used, as this metal ion binds to histidine residues. Therefore the resin bound the His_6 tagged recombinant protein.

2.19.1 IMAC purification

Ni-Sepharose resin (GE Healthcare) was used. 5 mL of resin was added to a 20 mL column. The storage buffer was allowed to flow through and was washed with 5-10 column volumes (CV) of dH_2O until all the resin had settled. The column was equilibrated with 10 CV of lysis buffer containing 20 mM imidazole. The cell lysate or filtered culture media was then added to the column and allowed to slowly flow over the resin to ensure optimum capture of the protein by the resin. The resin was subsequently washed with 10 CV of lysis buffer containing a defined concentration of imidazole. The recombinant protein was then eluted with a high concentration of imidazole (100-500 mM) and this fraction was collected. The column was then washed with 10 CV of dH_2O followed by 5 CV of 20% ethanol in which the resin was then stored. Fractions were taken at each step of the process and run on SDS-PAGE as described before.

2.19.2 Stripping and Recharging the IMAC Resin

The resin was stripped and recharged prior to use. It was initially washed with 2 CV of dH_2O followed by 2 CV of 50% ethanol. The resin was then stripped of metal ions

by washing with 2 CV of 100 mM EDTA, pH 8.0. Any remaining impurities were removed with 2 CV of 200 mM NaCl, a rinse with 2 CV of dH_2O followed by a 10 CV wash with 30% Isopropanol over 30 minutes. The resin was then washed with 10 CVof dH_2O and recharged with ½ a CV of 100 mM $NiSO_4$. Finally the column was washed with 10 CV of dH_2O and stored in 20% ethanol.

2.20 Desalting/Buffer exchange of purified protein using a HiPrep 26/10 desalting column

The proteins discussed in this work were buffer exchanged into suitable buffers using a Fast Performance Liquid Chromatography (FPLC) system with a HiPrep 26/10 desalting column. The HiPrep column is a gel filtration column that contained Sephadex G-25 Fine. This gel allowed the high molecular weight proteins to elute first as they were excluded from the gel. The smaller substances entered the gel pores and therefore eluted later.

The FPLC back pressure limit was set to 0.35 MPa and the flow rate to 1 mL/min before attachment of the HiPrep 26/10 desalting column. The top of the column was attached to pump outlet No.1 and the bottom of the column attached to the FPLC using green peek tubing (0.75 mm I.D.). Before the first sample application the column storage buffer (water:ethanol 80:20) was removed, and the column was equilibrated with sample buffer. This was done by washing with 2 CV of water (106 mL) followed by 5 CV (265 mL) of PBS. Equilibration was not necessary between runs using the same buffer. Washing was carried out at the flow rate intended for chromatography (5 mL/min). Samples were collected in 1 mL fractions in 96-well plates.

2.21 Sodium dodecyl Sulfate PolyAcrylamide Gel Electrophoresis (SDS-PAGE)

Gel Preparation:

Protein samples were analysed by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE), based on the method outlined by Laemmli (1970).

The gels were made as per the recipes below (see Tables 2.4 and 2.5) without the addition of TEMED as this immediately catalyzed the polymerisation of the gel. When the gel apparatus was set up (Vertical mini electrophoresis system, ATTO), the TEMED was added to the resolving gel which was immediately poured and overlaid with isopropanol. After polymerisation, the isopropanol was thoroughly washed off. TEMED was added to the stacking gel and this was poured on top of the resolving gel. A comb was inserted to allow the formation of wells for the loading of samples.

Table 2.4 Resolving Gel Recipe

	7.5%	10%	12.5%	15%
30% Acrylamide (mL)	1.875	2.5	3.12	5.0
1.5 M Tris-HCl pH 8.8 (mL)	1.875	1.875	1.875	1.875
10% SDS (µL)	75	75	75	75
H_2O (mL)	3.63	3.01	2.38	1.75
10% APS (µL)	37.5	37.5	37.5	37.5
TEMED (µL)	3.75	3.75	3.75	3.75

Table 2.5 Stacking Gel Recipe

	4%
30% Acrylamide (µL)	325
0.5 M Tris-HCl pH 6.8 (µL)	625
10% SDS (µL)	25
H_2O (mL)	12.5
10% APS (µL)	1.54
TEMED (µL)	2.5

2.21.1 Sample Preparation

40 µL of sample and 10 µL of 5X SDS-PAGE sample buffer (see Section 2.3) were added to a microfuge tube. This mixture was boiled at 100^{o}C on a heating block for 10 minutes. They were stored on ice until loading onto the gel.

2.21.2 Sample Application

20 μL of each sample was added to each well leaving one well spare for the relative molecular weight protein marker (Broad range protein marker, NEB). 20 μL of the ladder was added to the spare well for Coomassie Staining. If the gel was to be used for Western Blot analysis, the NEB pre-stained or colour plus pre-stained protein markers were used.

The Broad range protein marker consisted of Rabbit Muscle Myosin (212 kDa), MBP-β-galactosidase from *E.coli* (158 kDa), β-galactosidase from *E.coli* (116 kDa), Rabbit muscle Phosphorylase B (97 kDa), Bovine Serum Albumin (66 kDa), Bovine liver Glutamic dehydrogenase (55 kDa), MBP2 from *E.* coli (42 kDa), Thioredoxin reductase from *E. coli* (34 kDa), Soybean Trypsin inhibitor (20 kDa), Chicken egg white lysozyme (14 kDa), Bovine lung Aproptinin (6.5 kDa), Bovine pancreas insulin (3.4 kDa) and Bovine pancreas B chain (2.34 kDa). The prestained protein marker consisted of MBP-β-galactosidase from *E. coli* (175 kDa), MBP-paramyosin from *E. coli* (80kDa), MBP-CBD from *E. coli* (58 kDa), CBD-Mxe Intein-2CBD from *E. coli* (46 kDa), CBDBmFKBP13 from *E. coli* (25kDa), Chicken egg white lysozyme (17 kDa), Bovine lung aprotinin (7 kDa).

The colour plus protein marker consisted of MBP-β-galactosidase from *E. coli* (175 kDa), MBP-truncated-β-galactosidase *E.coli* (80 kDa), MBP-CBDE *E. coli* (58 kDa), CBD-*Mxe* Intein-2CBDE *E. coli* (46 kDa), CBD-*Mxe* Intein *E. coli* (30 kDa), CBD *E. coli* ParE (23 kDa), Chicken egg white lysozyme (17kDa), Bovine lung aprotinin (7 kDa).

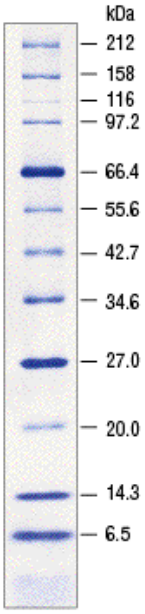


Fig 2.3: Broad range protein marker (NEB). Image from obtained from http://www.neb.com.

2.21.3 Gel Visualisation

Polyacrylamide gels were removed from the electrophoresis chamber and washed with dH_2O. The gels were stained for one hour with Coomassie blue stain solution (see Section 2.3). The stain was removed and the gels were washed with dH_2O. Coomassie blue destain solution was then added and left overnight to destain the gel. After the overnight destain it was washed again with dH_2O to enhance the protein bands further.

Silver staining was used for gels that required a higher degree of sensitivity for visualisation.

Table 2.6 Silver Staining of SDS-PAGE gels

Step	Reagent	Time
Fix	30% ethanol	>1 hour
	10% Acetic Acid	
Wash 1	20% Ethanol	15 minutes
Wash 2	Water	15 minutes
Sensitisation	0.1% Sodium Thiosulphate	1 minute
Rinse	Water	2 x 20 seconds
Silver Nitrate	0.1% AgNO3	30 minutes
	0.026% Formaldehyde	
Rinse	Water	2 x 20 seconds
Development	3% Sodium Carbonate	Until bands appear
	0.019% Formaldehyde	
	0.002% Sodium Thiosulphate	
Stop	5% Trizma Base	1 minute
	2.5% Acetic Acid	

*50mL per gel per step

2.22 Western Blot analysis

SDS-PAGE gels were run as described in Section 2.21 with 20 μL of NEB pre-stained molecular weight ladder. Two western blot methods were employed; a dry method using an iBlot (Invitrogen) and a semi-dry method using an electroblotter.

For the dry method an iBlot Dry Blotting System from Invitrogen was used. The gel was set-up as seen in Figure 2.3 and allowed to run on the machine for 7 minutes to allow transfer to take place.

For the semi-dry method, four pieces of blotting paper and a piece of nitrocellulose membrane were cut to the same dimensions as the SDS-PAGE gel. The blotting paper, nitrocellulose and gel were then soaked in transfer buffer (see Section 2.3) and arranged on the semi-dry electroblotter as per Figure 2.4. Transfer then occurred at a constant 20 V for 15 minutes.

To detect the His_6 tagged recombinant proteins the membrane was blocked overnight in 3% BSA in TBS with 0.05% tween (TBST). Afterwards the blot was washed 4 times for 5 minutes with TBST. The blot was then incubated at room temperature for 1 hour with a 1:10,000 anti-His antibody diluted in 1% BSA TBST. This was then washed 4 x 5 minutes with TBST. For visualisation one sigma fast dab tab and one Sigma fast urea H_2O_2 tab were dissolved in 15 mL dH_2O. This was poured over the blot and soaked until bands appeared. The blot was rinsed with dH_2O to stop the reaction.

When detecting biotinylated proteins the same procedure was followed except that an anti-Biotin antibody was used in place of the anti-His antibody.

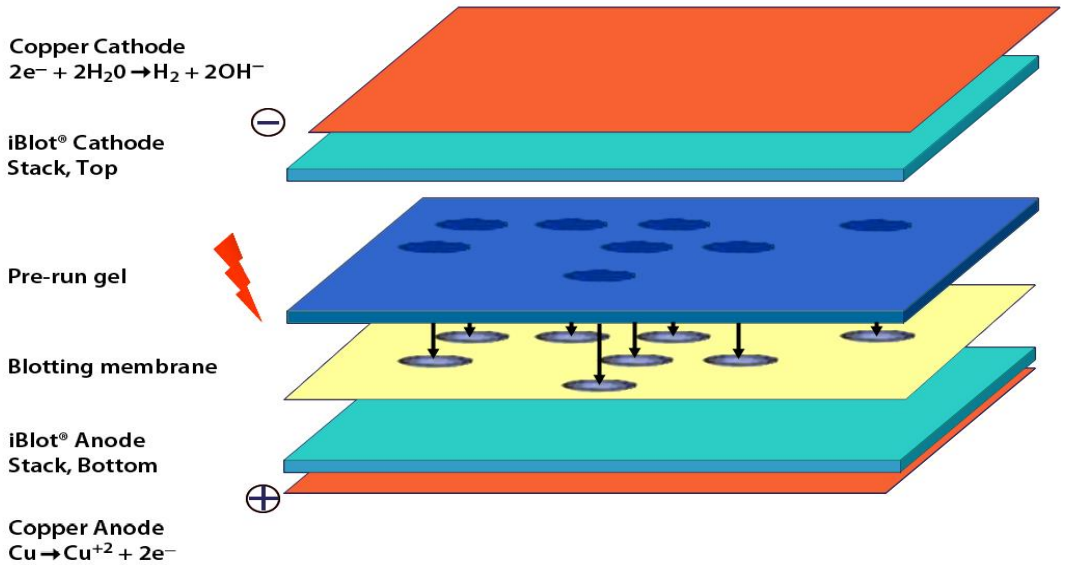


Fig 2.4: Schematic of western blot layout using the iBlot (dry method). The apparatus was set up as in diagram. Image obtained from Invitrogen. (www.invitrogen.com).

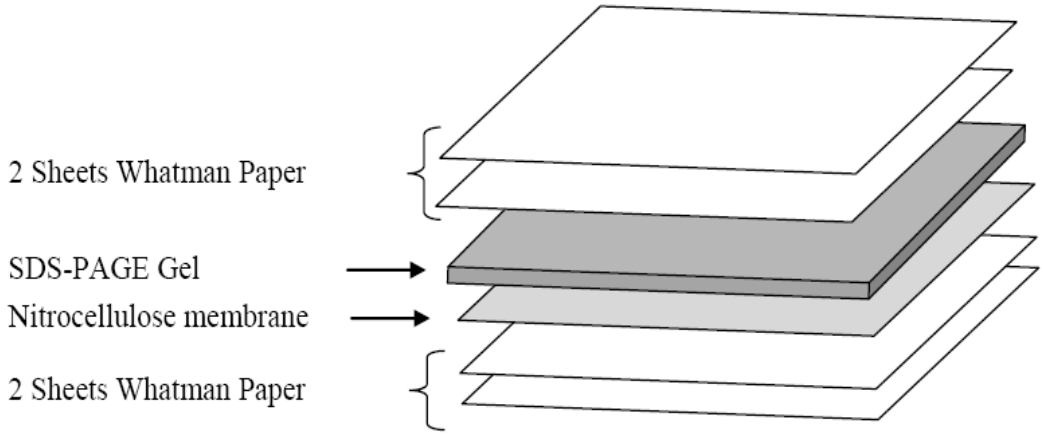


Fig 2.5 Schematic of the western blot layout using semi-dry method. The apparatus was set up as in diagram.

2.23 Protein quantification

Protein was quantified using two methods, the Bradford Assay and by reading its absorbance at 280 nm. The Bradford Assay was described by Bradford (1976). It is a colorimetric method in which the Coomassie dye binds to the protein in an acidic medium and immediately a shift in absorption maximum occurs from 465 nm to 595 nm in which a colour change takes place from brown to blue. Proteins can also be quantified reading their absorbance at 280 nm as proteins have the ability to absorb ultraviolet light in solution with the absorbance maxima at 280 nm.

2.23.1 Quantification of protein using the Bradford Assay

The Bradford assay using BradfordUltra^{TM} reagent (Expedeon) was used to quantify total protein in the range of 1 μg/mL – 1.5 mg/mL. For low range protein samples (1-25 μg/mL) 150 μL of sample was mixed with 150 μL Bradford Ultra. For high protein range samples (0.1 -1.5 mg/mL) 20 μL of sample was mixed with 300 μL of BradfordUltraTM reagent. All samples, standards and blanks were prepared in triplicate. The absorbance was read at 600 nm. Bovine serum albumin (BSA) was used as the reference protein. BSA standards were prepared in PBS and assayed in triplicate to yield a standard curve. Protein concentration of samples was determined from this standard curve.

2.23.2 Quantification of protein using 280nm readings

Protein samples were spun at 13,000 rpm for 10 minutes to remove any debris in the solution. This was to ensure nothing interfered with the absorbance readings. The UV lamp was set to 280 nm and was zeroed with the same buffer the protein samples were in. The protein samples were read in triplicate. The concentration of the protein was then calculated using the Beer-Lambert law:

$$A = \varepsilon cl$$

Beer-Lambert law used to determine protein concentration using 280 nm absorbance readings. A = absorbance, ε = extinction coefficient of the protein in question, c = concentration (in units corresponding to ε) and l = path length (cm).

2.24 Immobilisation of protein onto cyanogen bromide-activated sepharose

The protein to be immobilised was buffer exchanged into Coupling buffer (0.1 M $NaHCO_3$ with 0.5 M NaCl at pH 8.5 or PBS at pH 7.2). The capacity for the gel was 5-10 mg per mL of gel. The cyanogen bromide-activated resin was washed and swelled in cold 1 mM HCl for 45 minutes at 200 mL/g of gel. The resin was washed with 10 CV of dH_2O and then one CV of coupling buffer. The protein solution was immediately added at 5-10 mg of protein/mL of gel. The solution was left to mix by inversion overnight at 4^oC. The gel was extensively washed with coupling buffer to remove any unreacted ligand. 2 CVs of 0.2 M Glycine were added to block the unreacted groups. This was left to mix by inversion at 4^oC for 16 hours. This was washed extensively with coupling buffer and then 0.1 M acetate buffer with 0.5 M NaCl at pH 4.0. This wash cycle was repeated 4-5 times until the rinsings were devoid of protein. Resin to be used immediately was equilibrated in coupling buffer. Resin not immediately used was stored in 1 M NaCl at 4^oC with a suitable bacteriostat such as sodium azide (2 mM).

2.25 Quantitative determination of reducing sugars by DNS assay

Reducing sugars, under alkaline conditions, reduce 3,5-dinitrosalicylic acid to 3-amino-5-nitrosalicylic acid. A red colour is produced on boiling, which absorbs maximally at 540-550 nm (Miller 1959). The DNS assay was used to quantify reducing sugars in the range 0.2-1 mg/mL. Standards for a standard curve were prepared using GlcNAc at concentrations between 0-1 mg/mL. The carbohydrate solutions (standards and samples) were mixed with DNS reagent (see Section 2.3) in a ration of 1:2. The samples were then boiled at 100^{o}C for 10 minutes. The samples were allowed to cool and then 250 µL of each sample was aliquoted into a microtitre plate. All samples, standards and blanks were prepared in triplicate. The plate was read at 570 nm.

2.26 Assay to determine activity of β-N-acetylglucosaminidase

An assay had to be devised and optimised to quantitatively measure the activity of the enzyme on N-acetylglucosamine. 4-Nitrophenyl N-acetyl-β-D-glucosaminide (PNP-GlcNAc) is a chromogenic substrate that can be broken down by the β-N-acetyl-glucosaminidase enzyme (chitobiase enzyme). Upon breakdown 4-Nitrophenol is released, whose absorbance can be read at 405 nm. Standards in the range of 0-1 mM were made from 4-Nitrophenol. 210 µL of each standard was added to a NUNC Maxisorp 96-well plate in triplicate. 100 µL of 1 mM substrate made up in 50 mM Sodium Phosphate pH 8.0 was added to the wells of the 96-well plate in triplicate. 10 µL of enzyme was added to each well containing substrate and incubated at 37^{o}C for 45 minutes. 100 µL of 1 M Sodium Carbonate at pH 11.5 was added to stop the reaction and release the product. The absorbance was read on a plate reader at 405 nm.

2.26.1 Determination of V_{max} and K_m

The rate of catalytic activity (V) of the enzyme was measured over a range of increasing PNP-GlcNAc substrate concentrations. A plot of the rate of activity (V) versus substrate concentration (S) generated a Michaelis Menten curve. The V_{max} was

the maximum rate of activity achieved by the enzyme. The K_m was determined as the substrate concentration which gave rise to the half the maximum V_{max}.
Hanes-Woolf plot – A plot of S/V versus S was generated. The slope of the line gave $1/V_{max}$ and the x-axis intercept gave $-K_m$.

2.26.2 Assay to determine specificity of β-N-acetyl-glucosaminidase

Assay as above (Section 2.26) except different 4-nitrophenyl conjugated substrates were used. The other substrates used were PNP-α-Glucose, PNP-α-Galactose, PNP-α-Fucose, PNP-α-Mannose, PNP-β-Glucose, PNP-β-Galactose and PNP-Sialic Acid.

2.26.3 Sugar Inhibition of β-N-acetyl-glucosaminidase

Assay as above (Section 2.26) except different concentrations (0-100 mM) of free sugars were incubated at 37^{o}C with the enzyme for 15 minutes before addition to the 96-well plate. The sugars used were GlcNAc, glucose, galactose, lactose, fucose, mannose and xylose.

2.27 Enzyme Linked Lectin Assay (ELLA)

50 µL of a Glycoprotein solution 5 µg/ml (such as asialo agalacto fetuin) made in PBS was added to the wells of a NUNC Maxisorp 96 well plate and incubated overnight at 4^{o}C. The unbound glycoprotein was then removed by inversion and the plate was then blocked with 200 µL of a 0.5% PVA solution for 2 hours at 25^{o}C. The plate was then washed four times with TBST. 50 µL of a 5-10 µg/mL lectin solution, diluted in TBST was then added to the plate and incubated for an hour at 25^{o}C. The plate was washed again four times with TBST. 50 µL of a 1:10000 HRP-labelled murine anti-biotin or anti-histidine antibody in TBST was added to each well. The plate was again incubated for 1 hour at 25^{o}C. Following this incubation the plate was washed 4 times with TBST followed by a single PBS wash. The assay was developed by adding 100 µL TMB solution to each well and left to react for 10 minutes. The reaction was stopped by adding 50 µL of 10% H_2SO_4. The absorbance was read on a plate reader at 450 nm.

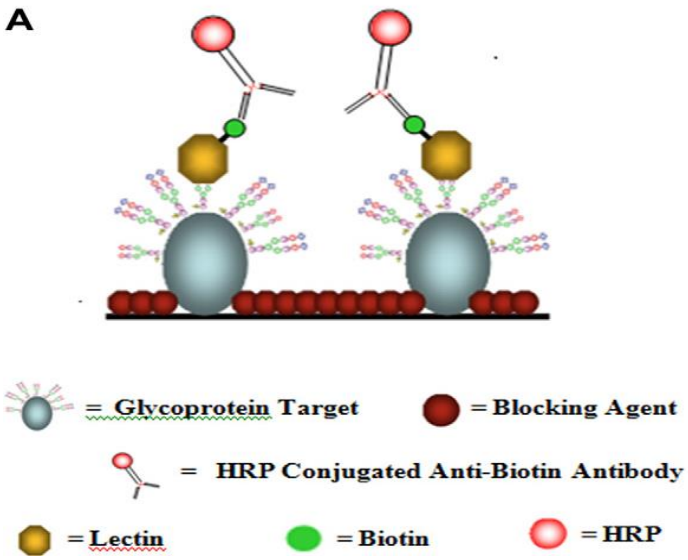


Fig 2.6: Schematic of an ELLA. The ELLA follows the same principle as an ELISA. The glycoprotein target is placed on the surface of the plate and probed with a labelled lectin which is detected colormetrically. Image taken from Thompson *et al*. (2011).

2.28 Biotinylation of Carbohydrate Binding Protein (CBP)

A biotinylation kit from Thermo Scientific (Sulfo-NHS-LC-Biotin 21327) was used as per the manufacturer's guidelines. The biotin was removed from the freezer and equilibrated to room temperature. 180 µL of ultrapure water was added to the 1 mg tube of biotin to give a 10 mM biotin stock solution. The protein to be biotinylated was resuspended in an amine free buffer such as PBS. An aliquot of the 10 mM biotin stock was added to the protein solution to give a 20 fold molar excess of biotin to protein. The reaction was left to continue on ice for 2 hours. The sample was then buffer exchanged to remove any excess biotin with PD MidiTrap G-25 columns (GE Healthcare).

2.29 Insoluble Substrate Assays

The method used was modified from that described by Suginta *et al* (2007). Chitin (α and β), chitosan and cellulose are all insoluble and therefore they cannot be used in an ELLA format. Instead this insoluble substrate assay was used which allows the binding of proteins to these substrates be visualised on SDS-PAGE gels. Avicel (crystalline cellulose), crustacean shell chitosan and chitin (α and β) were used in these experiments.

2.29.1 Insoluble Substrate Assay –rChbL Enzyme

0.05 g of the insoluble substrate was added to a 1.5 mL microfuge tube along with 1 mL of 50 μg/mL rChbL enzyme or rChbL enzyme mutant in 50 mM sodium phosphate buffer pH 8.0. This was allowed to incubate for one hour at 4^{o}C while being continuously rotated. 100 μL samples were taken at time-points (5, 15, 30 and 60 minutes). Each sample was centrifuged at 13,000 rpm for 5 minutes. 50 μL of the culture media was removed and 10 μL of 5X protein sample buffer (see Section 2.3) was added to this. The sample was boiled at 100^{o}C for 10 minutes. The insoluble substrate was washed twice with PBS and 50 μL of 1X protein sample buffer was added to it. It was also boiled at 100^{o}C for 10 minutes. All samples were run on SDS-PAGE gels as described in Section 2.21 and stained using Silver Staining as described in Table 2.6.

2.29.2 Insoluble Substrate Assay – rChb CBP

0.05 g of the insoluble substrate was added to a 1.5 mL microfuge tube along with 1 mL of 50 μg/mL rChb CBP or rChb CBP mutant in 50 mM sodium phosphate buffer pH 8.0. This was allowed to incubate for one hour at 45^{o}C while being continuously rotated. 100 μL samples were taken at time-points (5, 15, 30 and 60 minutes). Each sample was centrifuged at 13,000 rpm for 5 minutes. 50 μL of the culture media was removed and 10μL of 5X protein sample buffer (see Section 2.3) was added to this. The sample was boiled at 100^{o}C for 10 minutes. The insoluble substrate was washed twice with PBS and 50 μL of 1X protein sample buffer was added to it. It was also boiled at 100^{o}C for 10 minutes. All samples were run on SDS-PAGE gels as described in Section 2.21 and stained using Silver Staining as described in Table 2.6.

2.30 High Performance Liquid Chromatography (HPLC) to analyze chitin hydrolysis

2.30.1 Instrumentation

The chromatographic separation was performed using an Ultimate 3000 HPLC system from Dionex. The derivatised samples were separated on an XTerra RP18 6x150mm column of 3.5 μm from Waters at a flow rate of 700 μL/min. The column temperature was kept constant at 20^{o}C using a column oven. 50 mM sodium acetate (pH 4.0) was used as mobile phase A while acetronitrile was used as the organic solvent (mobile phase B). The gradient program used was: 0-12 minutes (2% B), 12-27 minutes (2-10% B), 27-35 minutes (10-2% B) and 35-40 minutes (2% B). The separations were read using a UV detector at 254 nm. Chromeleon software was used to plot the resulting chromatograms.

2.30.2 Sample preparation

Standards used were N-acetylglucosamine, Chitobiose and Chitohexose all at 1 mM. Hydrolysis of chitin and chito-oligosaccharides (1 mM) were performed at 37^{o}C in 1 mL reactions with 1 μg/mL enzyme in sodium phosphate buffer at pH 8.0. 200 μL samples were taken at time-points (0, 0.5, 1, 3, 6 and 24) and kept on ice or at 4^{o}C. Each sample was centrifuged at 13,000 rpm for 5 minutes to remove any insoluble material. The samples were then run through vivaspins (Sartorius) to remove the enzyme present. The flow-through was retained. 100 μL of each sample was derivatised.

2.30.3 Derivatisation of samples and standards

The derivatisation reagent was composed of a methanol acetate borate solution, MABS (see Section 2.3). 30 mg of anthranilic acid and 20 mg of sodium cyanoborohydride were dissolved in 1 mL MABS prior to use. 100 μL of the samples and standards were mixed with 100 μL of the derivatisation agent in a 2 mL screw-top polypropylene plastic tube. This mixture was heated at 80^{o}C for 80 minutes in a water

bath. After allowing the sample to cool it was diluted with 800µL Mobile phase A and centrifuged at 13,000 rpm for 3 minutes. 120 µL of the sample was transferred to a HPLC vial for analysis.

2.32 Statistical Analysis and Data Analysis

The student T-test was used to determine if differences measured between samples were significantly different from one another. The level of statistical significance was indicated by *($p<0.5$), **($p<0.01$) and ***($p<0.001$).
All experiments were carried out in triplicate. The results shown are the means ±SD of three samples.

2.33 Data Mining and Protein Modelling

To identify homologous protein and DNA sequences deposited in Genbank, the BLAST program available at NCBI (www.ncbi.nlm.nih.gov) was used. DNA sequences for the strains used were obtained from the following sequencing studies; *Serratia marcescens* (Q54468) and *Photorhabdus luminescens* (NP_928634.1). DNA and protein sequences were aligned using ClustalW available at http://www.ebi.ac.uk/Tools/clustalw and Genedoc program, available to download at http://www.nrbsc.org/. DNA sequences were analysed for restriction sites using the Webcutter 2.0 programme (http://rna.lundberg.gu.se/cutter2/).
Protein modelling was carried out using Swiss-Model available from http://swissmodel.expasy.org/. Homology modelling was carried out using the automated mode to predict the structure of the *P. luminescens* Chb based on already determined crystal structures of the *S. marcescens* Chb. The *S. marcescens* structures can be found in the Protein Data Bank (http://www.rcsb.org/pdb/home/home.do) labelled 1QBA and 1QBB. Pymol from Delano Scientific was used as the protein imaging software, available for download at http://pymol.sourceforge.net.

Chapter 3:

The cloning, expression, purification and characterisation of a chitobiase (rChbL) enzyme from *Photorhabdus luminescens*.

3.1 Overview

This chapter describes the discovery, cloning, expression, purification and characterisation of the recombinant Chb enzyme for the first time from *P. luminescens.* As stated previously, *S. marcescens* is reported to be one of the most efficient bacteria involved in the degradation of chitin, capable of expressing at least 3 chitinases, a chitobiase and a chitin binding protein (Brurberg *et al*., 2001). The protein sequence for the extensively characterised chitobiase (Chb) from *S. marcescens* was used as a reference sequence to investigate the presence or absence of a gene encoding a potential chitobiase in the genomic sequence of *P. luminescens.*

3.2 Identification of a chitobiase from *P. luminescens*

A BLAST search (available from www.ncbi.nlm.nih.gov) was used to compare the *S. marcescens* Chb protein sequence with the *P. luminescens* proteome. The BLAST resulted in a list of 6 proteins showing various levels of homology to Chb. The most promising candidate for further study was a novel *N*-acetyl-β-glucosaminidase (NP_928634.1) that displayed 66% identity to the Chb sequence from *S. marcescens.*

Table 3.1 Sequences producing significant alignments to the *S. marcescens* chitobiase

Protein	Query Cover	E value	Max ident.
N-acetyl-beta-glucosaminidase	99%	0.0	66%
Hypothetical protein plu3801	7%	3.9	29%
Hypothetical protein plu1131	11%	4.7	27%
Hypothetical protein plu2394	3%	7.8	41%
IS630 family transposase	7%	8.7	28%
Hypothetical protein plu3021	8%	9.4	27%

Further sequence alignments (see Figure 3.1) between the Chb protein from *S. marcescens* and ChbL from *P. luminescens* were performed using the online tool, ClustalW (available from: http://www.ebi.ac.uk/Tools/clustalw/) and the software program Genedoc (available to download from: http://www.nrbesc.org/). The

sequences were obtained from Genbank, *S. marcescens* (Q54468) and *P. luminescens* (NP_928634.1).

This analysis revealed that the Chb from *P. luminescens* (ChbL) shared approximately 80% sequence similarity and 66% identity with the Chb from *S. marcescens*. Most notably high levels of amino acid conservation were observed around the site of the catalytically important residues D539 and E540. The least sequence similarity was observed at the proteins N-terminus and between residues 200 and 300. The *P. luminescens* sequence was marginally longer than the *S. marcescens* sequence, comprising 891 amino acids compared to 884 amino acids. The relatively high degree of sequence similarity between the *P. luminescens* Chb and *S. marcescens* Chb, combined with the relatively high levels of conservation around the proposed active site, made the *P. luminescens* Chb (ChbL) a favourable target enzyme for further study.

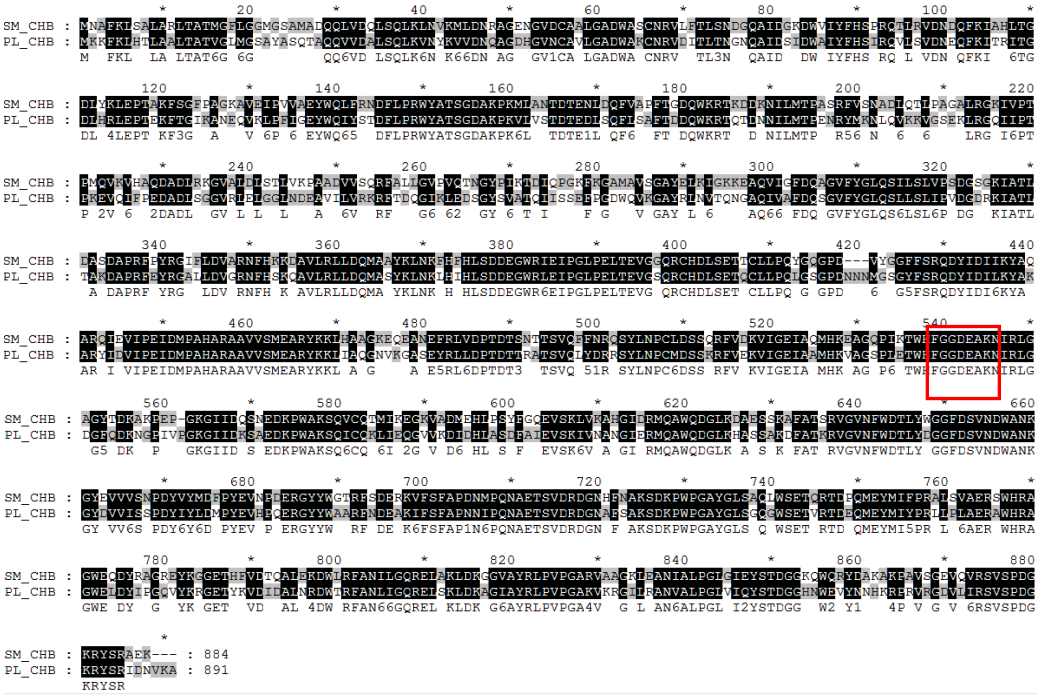


Fig 3.1: Sequence alignment of the well-characterised chitobiase from *S. marcescens* (SM_CHB) and the chitobiase from *P. luminescens* (PL_CHB). The sequence alignment was generated with the SM_CHB (Accession number Q54468) and PL_CHB (Accession number NP_928634.1) gene sequences using ClustalW (available from http://www.ebi.ac.uk/Tools/msa/clustalw2/) and visualised using Genedoc software. The catalytically important residues are highlighted in red and the conserved residues are highlighted in black.

3.3 Cloning of the *chb* gene from *P. luminescens*

Overexpression of a protein of interest is usually difficult to achieve using the natural prokaryotic host. Typically proteins may be produced recombinantly using a number of different cell systems. The method chosen will depend on the protein to be expressed. Factors to be considered include whether or not the native protein is glycosylated, the size/complexity of the protein and the presence of any other post-translational modifications.
In this study the recombinant protein of interest was expressed in the prokaryotic host *Escherichia coli* as these cells are relatively easy and cheap to culture and grow rapidly. In addition, there are no complex purification steps (Baneyx, 1999).

The protein of interest was produced recombinantly in *E. coli* as over expression in its native host *P. luminescens* was unattainable. *E. coli* is one of the most commonly used hosts for recombinant protein expression as its genetics are well defined, it grows rapidly, is inexpensive to culture and is amenable to large scale protein production (Baneyx, 1999). One of the disadvantages of the *E. coli* expression system, compared to eukaryotic cells, is its lack of a glycosylation system (Makrides, 1996). This can become a problem when expressing proteins of plant or human origin as they are generally modified post-translationally by the addition of glycans. The ChbL protein however is also of prokaryotic origin and was therefore suitable for over-production in *E. coli*.

The gene encoding the ChbL protein was identified from the NCBI database, accession no NC_005126.1.

atgaagaaatttaaattgcatacattagcagcactgactgcaactgttggacttatgggttctgcatat
gcttcccaaacagcgcagcaagtggttgatgcactgagtcaattaaaagttaattataaagttgttgat
aatcaggcaggagatcacggcgttaactgtgcggtactgggggcagattgggctaaatgtaatcgagtg
gatattactttaaccaatggtaatcaggcgattgactctattgactgggcgatttatttccacagtatt
cgtcaggtcctgagcgtggataatgaacagtttaaaatcacgcgtattacgggcgatttacatagactg
gagccaaccgaaaaattcaccggtattaaagctaatgagcaggtaaaactgccgtttattggagaatac
tggcagatttacagcactgattttctacctcgctggtatgcaacgtcgggtgatgcgaagcctaaagtg
cttgttagtacggatacagaagatctaagtcaattcttatctgctttcaccgatgatcagtggaaacgc
acacagacagataacaatattttgatgacgcctgaaaatcgttatatgaagaatttgcaggtcaaaaag
gttggatcagagaagttgcgtgggcaaattatcccgacaccaaaagaggttcagatttttcccgaggat
gctgatttatctggcggggtgagactggaattgggtggtttaaatgacgaggcagttattcttgtcaga
aaacgttttactgatcaaggcattaaattagaggatagtggttattcagttgcgacccagattatctct
tcagagtttcctggcgattggcaggttaaaggggcttaccgtctgaatgtgacccagaatggggcacaa
attgttgcttttgaccaatctggggtgttttacggtcttcaatctctgttgtcattaattcctgttgat
ggcgataggaagattgctactttgacggcaaaagatgcgccccgctttgaatatcgtggtgcactcctt
gatgtagggcgtaatttccacagtaaacaggctgttctgcgtttgcttgaccaaatggcaagttacaaa
ctgaataaactccatattcatctgtctgatgatgaaggttggcgtttggaaattcctggattacctgag
ttgacggaagttggcagccagcggtgtcatgatttaagtgaaactcaatgtttattaccacaattgggt
tcaggccccgataacaacaacatggggagcggatatttcagccgtcaagattacattgatattctgaaa
tatgctaaagctcgttatatcgacgttattcctgaaattgatatgccagcccatgctcgtgccgcggtt
gtttctatggaagcccgctataaaaaattaattgctcaggggaatgtgaaaggggcgagtgaatatcgt
ttgctcgatccaacggataccaccagagcaacatcggttcagctctatgatcgccgcagttatttgaac
ccgtgtatggattcttcaaaacggtttgttgagaaagtcataggtgaaattgctgctatgcataaagta
gcaggttcaccattggagacttggcattttgggggagatgaagcgaagaatatccgtttgggagacgga
ttccaggataaaaatggtcctatcgtaccgggtaagggcatcattgataaaagtgcggaagataaaccg
tgggctaagtctcagatttgtcagaagttgatagaacagggtgttgtgaaagacattgatcatcttgcc
agcgacttcgcaattgaagtgagcaaaattgtgaatgctaatggtatcgagaggatgcaggcatggcag
gatggtttaaaacatgctagcagtgctaaagattttgccaccaaacgcgttggggtgaatttttgggac
actctctacgatggcggctttgattcagtcaatgattgggcgaataagggctatgacgtggtgatatct
agcccagattatatttatctggatatgccatatgaagttcatccacaagaacgtggctattattgggcg
gctcgttttaacgatgaagctaagatctttagttttgcgcccaataacataccgcaaaatgctgaaact
tccgttgaccgtgatggtaatgcattcagtgcgaaaagtgataaaccgtggccgggtgcctatgggctt
tcaggtcaaggatggagtgaaacggttcggactgatgagcagatggaatatatgatttatccgcgcttg
ttgccgctagcggaacgggcatggcaccgggccggttgggagttggattatattccgggccaggtatat
aagcgtggagaaacgtataaagtcgatatagatgcgcttaatcgtgactggacccggttcgctaacctg
atagggcaacgtgagcttagtaaactggataaagcgggcattgcttaccgtctaccggttccgggggct
aaggtaaaacggggtattttgcgggctaatgttgctctgcctggattggtaatccagtattcgacggat
ggcggtcataactgggaagtttataacaaccataaacgcccaagagtgagaggtgatgtgttgatccgt
tctgtcagcccggacggtaaacgttacagtcgaattgataatgtaaaggcatga

Fig 3.2: *chbl* gene sequence from *P. luminescens*. The *P.luminescens chb* gene sequence was obtained from Genbank (Accession number NC_005126.1). The proposed signal peptide is highlighted in red.

Primers were designed to amplify the *chbl* gene from *P. luminescens* genomic DNA for ligation with the pQE60 vector from Qiagen (as described in Sections 2.11 and 2.12). This company produce a variety of pQE vectors which enable high-level production of recombinant proteins. pQE vectors have numerous features that make them attractive for cloning. The vectors allow for the addition of N- or C-terminal His_6 tags to the protein of interest, which facilitates downstream IMAC affinity-based purification. The vectors contain extensive multiple cloning sites and an optimised promoter sequence which includes a phage T5 transcriptional promoter and a lac

operon sequence for inducible expression of the protein. The pQE vectors also contain a β-lactamase gene that confers ampicillin resistance to the cell.

The Chb protein from *Serratia marcescens* contains an N-terminal 27 amino acid signal peptide that is believed to be cleaved during post translational processing of the protein (Tews *et al*., 1996). The ChbL protein from *P. luminescens* is predicted to contain a signal peptide as the area was highly conserved between the two homologous proteins (see Figure 3.3).

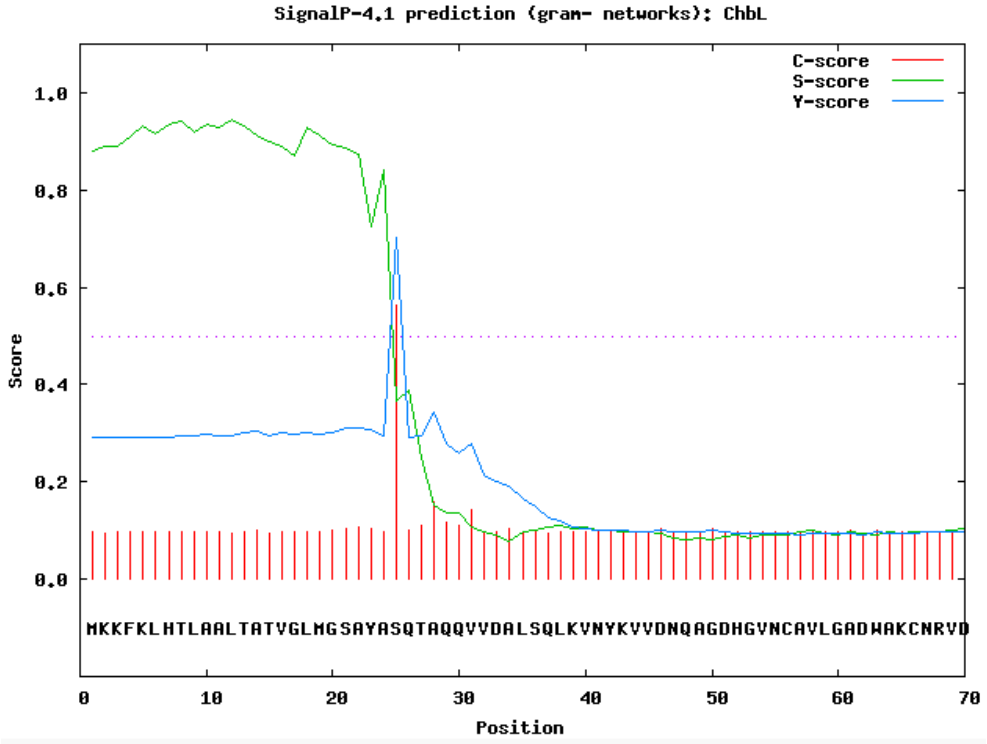


Fig 3.3: Prediction of a signal peptide sequence in the rChbL sequence. SignalP 4.1 (available from http://www.cbs.dtu.dk/) was used to predict the presence/absence of a signal peptide in the ChbL protein sequence (NP_928634.1).

Table 3.2 Signal Peptide Prediction

Measure	Position	Value	Cutoff	Signal Peptide
max. C	25	0.562		
max. Y	25	0.703		
max. S	12	0.946		
mean S	1-24	0.899		
D	1-24	0.796	0.570	Yes

A signal peptide was predicted to be present on the ChbL protein from residue 1 to 24. The mature protein sequence was predicted to begin at residue 25. In order to examine the effects the signal peptide had on expression of the protein, the ChbL protein was amplified and cloned into the pQE60 vector both as an immature protein with its signal peptide (Chb+) and as a mature protein without its signal peptide (Chb-). The pQE60 vector was used as it adds a C-terminal His_6 tag to the recombinant proteins. This was pertinent as Chb+ contained an N-terminal signal peptide which, if cleaved, would also remove the His tag.

3.3.1 Cloning the immature rChbL protein (Chb+)

The gene sequence was amplified as an NcoI-BamHI fragment for cloning into the pQE60 vector. An NcoI site was incorporated using the forward primer chb+_F. In the NcoI sequence the internal start codon is followed by a 'G' and lysine cannot be encoded by any codons beginning with 'G'. Therefore, an additional codon was incorporated to facilitate the NcoI site. The best candidate was alanine as it is the smallest and simplest amino acid. It is also found at the beginning of the *S. marcescens* signal peptide, which shows a high level of similarity to that of the *P. luminescens* sequence. It was believed this would not dramatically impact the function of the signal peptide.

Met Ala Lys Lys Phe Lys Leu His
CC ATG GAC AAG AAA TTT AAA TTG CAT
NcoI

Rather than the wildtype;

Met --- Lys Lys Phe Lys Leu His
ATG --- AAG AAA TTT AAA TTG CAT

A BamHI restriction site was incorporated through the reverse primer chb+_R. No additional amino acids were required. The resultant Nco-BamHI PCR product was cloned into an NcoI-BglII digested pQE60 vector to make the pQE60_Chb+ construct.

3.3.2 Cloning the mature rChbL protein (Chb-)

Comparison with the *S. marcescens* protein sequence showed high conservation between the sequences even in the signal peptide.

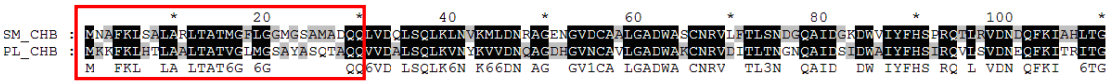


Fig 3.4: Partial sequence alignment between the SM_CHB and the PL_CHB highlighting sequence similarity in the signal peptide region. The sequence alignment was generated using the ClustalW online tool and Genedoc software. The predicted signal peptide is highlighted in red. The conserved residues are highlighted in black.

The *chb-* gene sequence was also amplified as an NcoI-BamHI fragment for cloning into the pQE60 vector. An NcoI site was incorporated using the forward primer chb-_F. The first residue of the mature ChbL protein is a serine. In the NcoI restriction site the internal start codon is followed by a 'G'. Serine cannot be encoded by any codons beginning with 'G'. Therefore an additional codon was incorporated to facilitate the NcoI site. Again the best candidate was alanine as it is the smallest and simplest amino acid. The reverse primer chb-_R was the same as that used for the cloning of the immature Chb+ protein. The resultant NcoI-BamHI PCR product was cloned into an NcoI-BglII digested pQE60 vector to make the pQE60_Chb- construct.

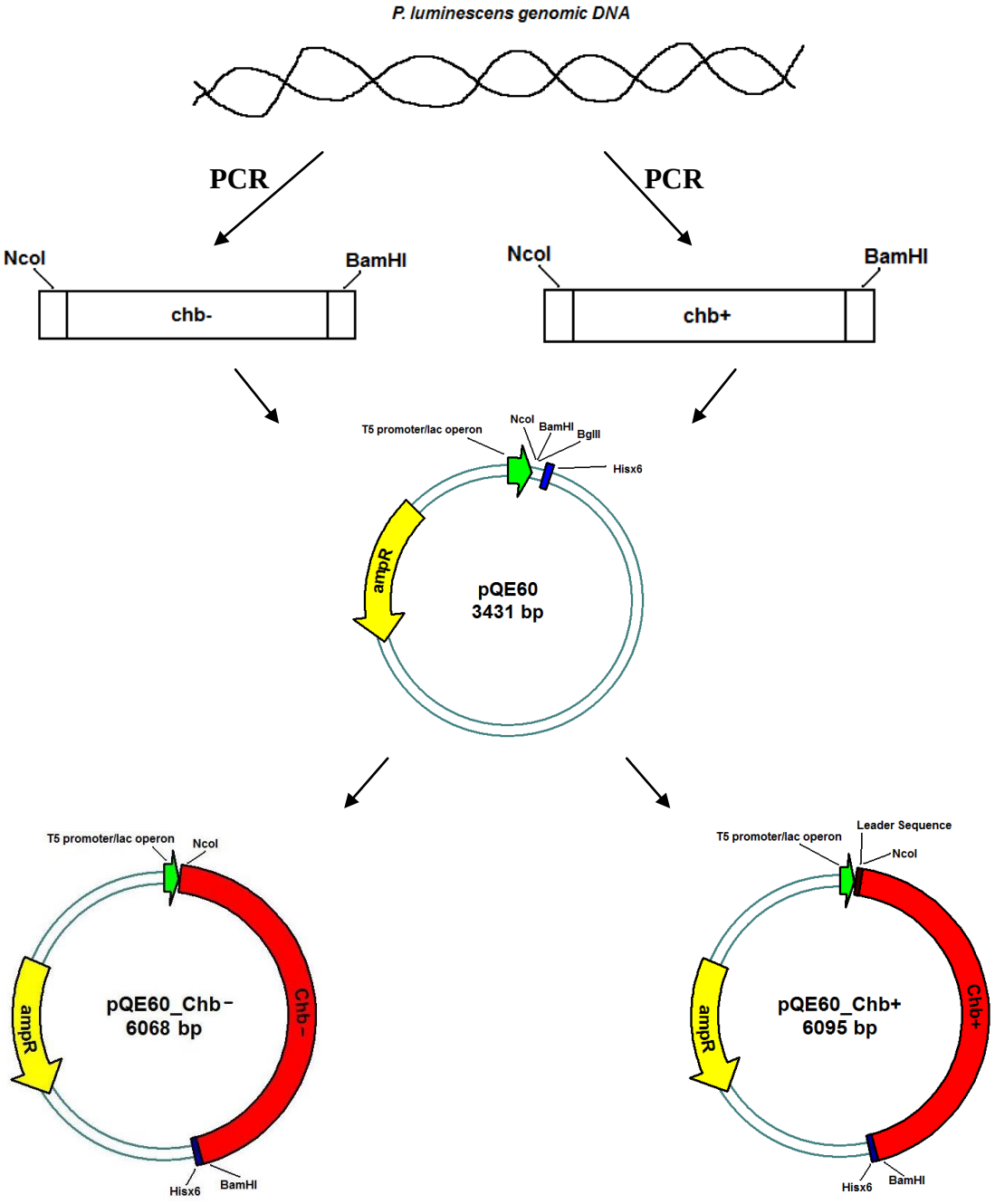


Fig 3.5: Schematic describing the production of the pQE60_Chb- and pQE60_Chb+ constructs. These constructs were created from the primers chb+_F, chb+_R, chb-_F and chb-_R. The chb+ and chb- primers were used to amplify the *chb* gene with and without its signal peptide respectively, into a pQE60 vector. The gene was cloned as an NcoI-BamHI fragment into an NcoI-BglII digested vector. The red insert represents the *chb* gene, the yellow insert represents the β-lactamase gene, the green insert represents the promoter, the purple insert represents the signal peptide (leader sequence) and the blue insert represents the C-terminal His tag.

The *chb* gene was successfully cloned into the pQE60 vector as a mature (Chb+) and immature peptide (Chb-). The sequences were verified using commercial sequencing from Eurofins (see Section 2.10). The vector was then transformed into *E. coli* JM109 cells for stable storage of the plasmid and into *E. coli* KRX competent cells for expression of the recombinant Chb protein.

3.4 The effect of the signal peptide on rChbL protein expression

Expression of the immature and mature rChbL proteins was undertaken to examine if the protein was expressed in *E. coli* and the effect, if any, of the signal peptide on cell growth and soluble protein production. The cultures were monitored using OD_{600nm} readings over a 16 hour period and the soluble and insoluble protein fractions were analysed by SDS-PAGE to determine the extent of protein expression and solubility (see Figure 3.6).

a) **b)**

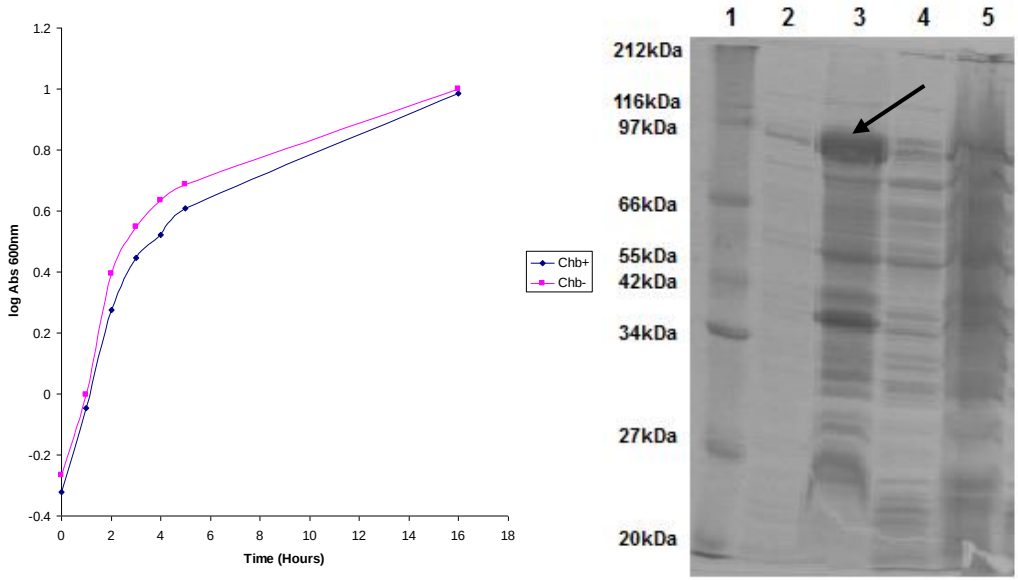


Fig 3.6: a) Analysis of the growth rate of Chb+ (immature) and Chb- (mature) cultures b) Analysis of protein expression from the soluble and insoluble fractions from Chb+ and Chb- cultures. Cultures were grown and absorbance readings at 600nm were taken at different time-points. Soluble and insoluble fractions were prepared from Chb+ and Chb- cultures and analysed by 12.5% SDS-PAGE (see Section 2.15.1). Lane 1; Protein ladder Lane 2; Soluble Chb+ Lane 3; Insoluble Chb+ Lane 4; Soluble Chb- Lane 5; Insoluble Chb-.

As seen in Figure 3.6 the growth rates of Chb+ and Chb- cultures were similar. The observed cell density of both Chb+ and Chb- cultures was similar at all time-points, suggesting that the presence/absence of the signal peptide did not affect the growth rate of the cells. It was evident from SDS-PAGE analysis that higher levels of rChbL protein expression resulted from the Chb+ construct. The majority of the Chb+ protein however was found in the insoluble cellular fraction, with a lesser amount visible in the soluble fraction. The analysis revealed very little evidence of Chb- protein expression in either the soluble or insoluble protein fractions. Therefore, in all further experiments the immature Chb+ protein was expressed and from here on will be referred to as rChbL.

3.5 Optimisation of the expression of rChbL in *E. coli*

As reported in Section 3.4 the majority of rChbL protein expressed to the insoluble protein fraction. While not impossible, the purification of insoluble proteins is more difficult than that of soluble proteins and involves the denaturation and renaturation of the target protein. This can lead to large losses of protein as well as reduced protein activity. This section explores the optimisation of the expression of soluble rChbL. Time course analysis of both soluble and insoluble protein fractions was undertaken to determine the solubility of ChbL over the course of the expression experiment.

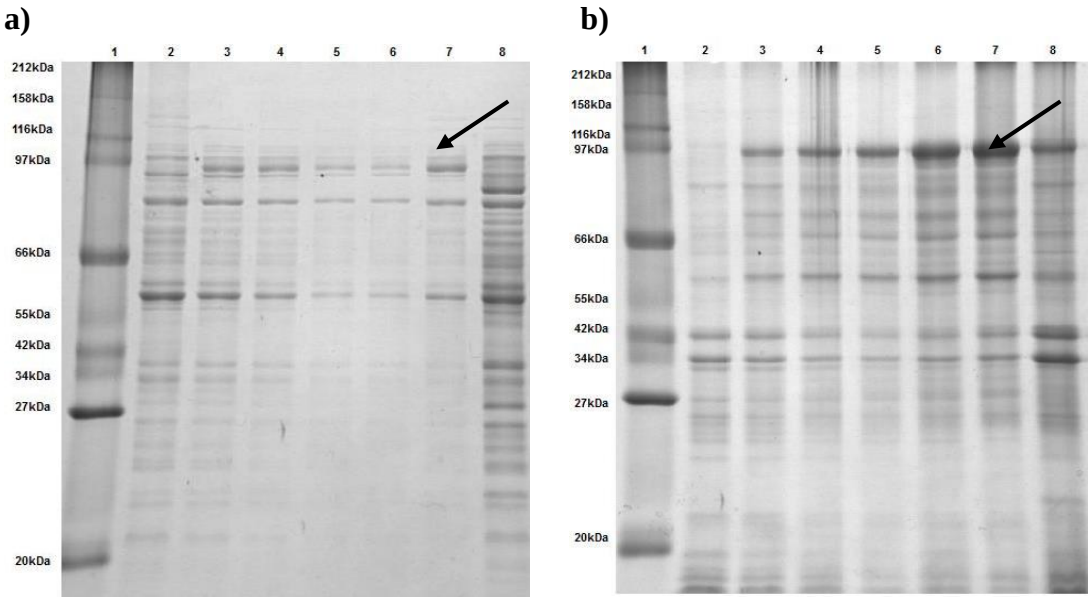


Fig 3.7: Time-course expression analysis of a) soluble and b) insoluble fractions from rChbL cultures. Once the culture had reached an OD_{600} of 0.5 and had been induced, 1 mL samples were taken every hour for 5 hours followed by an overnight sample. Samples from the soluble and insoluble cell fractions were obtained (see Section 2.15.1) and then analysed by 12.5% SDS-PAGE. Lane 1; Protein ladder Lane 2, T=0 hours Lane 3, T=1 hour Lane 4, T= 2 hours Lane 5, T=3 hours Lane 6, T=4 hours Lane 7, T=5 hours Lane 8, Overnight (T=16 hours). The largest protein concentration observed is highlighted by an arrow.

As observed in Figure 3.7, time-course analysis revealed that the majority of the rChbL protein was found in the insoluble fraction. However, as observed previously there was also a small amount of protein present in the soluble fraction, with a similar

level of rChbL protein observed in the soluble fraction at all time-points, while the highest level of protein observed in the insoluble fraction was 5 hours post induction.

Knowing that rChbL was expressed to the soluble protein fraction but did not accumulate there over time, meant that the water lysis method (as described in Section 2.18) was utilised to pinpoint the exact cellular location of the expressed rChbL. Cytoplasmic, periplasmic, outer membrane and extracellular fractions were analysed by SDS-PAGE following an overnight expression experiment (see Figure 3.8).

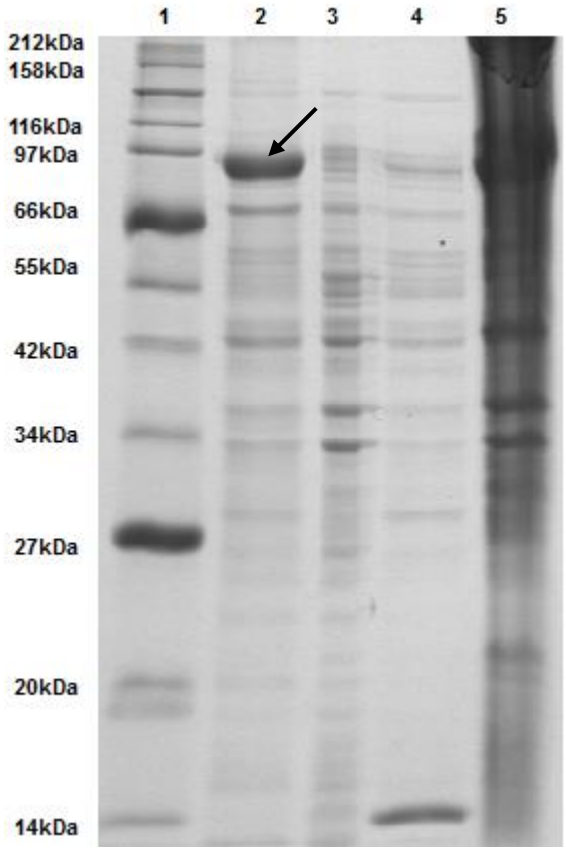


Fig 3.8: Cellular compartment analysis of the soluble fraction of rChbL cultures. The soluble fraction from rChbL was analysed using the water lysis method (see Section 2.18) to examine where the recombinant protein was located in the cell. The culture media was also examined (see Section 2.16). The fractions were then run on a 12.5% SDS gel. Lane 1, Protein ladder Lane 2, Culture media Lane 3, Cytoplasmic fraction Lane 4, Membrane fraction Lane 5, Periplasmic fraction. The largest protein concentration observed is highlighted by an arrow.

From analysis by the water lysis method (see Figure 3.8) it became apparent that that soluble rChbL was being transported across the cell membrane to the culture media. This result indicated that the protein could be purified from the culture media for experimental use.

Expression and export of rChbL to the culture media was monitored by time-course analysis (see Figure 3.9). Culture media samples were taken (see Section 2.17) and analysed by 12.5% SDS-PAGE. The rChbL protein began to appear in the culture media two hours post-induction. As with the protein from the insoluble cellular fraction, the optimal time for protein harvest was five hours following induction.

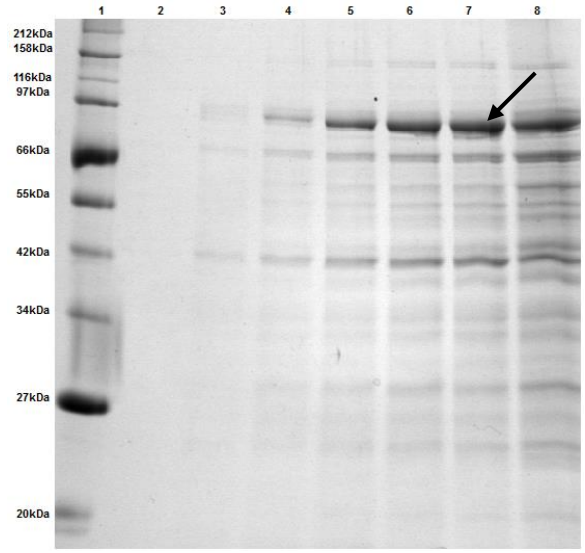


Fig 3.9: Time-course expression analysis of the rChbL protein culture media samples. Recombinant protein expression was induced at OD_{600nm} = 0.5. 5 mL samples were taken from the culture media every hour for five hours, following induction and an overnight sample was taken 16 hrs post induction (see Section 2.15). The samples were then analysed by 12.5% SDS-PAGE. Lane 1; Protein ladder, Lane 2; T=0 hours, Lane 3; T=1 hour, Lane 4; T=2 hours, Lane 5; T=3 hours, Lane 6; T=4 hours, Lane 7; T=5 hours, Lane 8; Overnight (T=16 hours). The largest protein concentration observed is highlighted by an arrow.

3.6 Optimisation of conditions for the expression of the rChbL protein

Different culture conditions can have a huge impact on the levels of recombinant protein expression. Various conditions for growth, including the concentration of the inducer IPTG, temperature and media were examined in order to optimise the expression of the rChbL protein.

The *E. coli* strains KRX, JM109 and XL10 Gold were examined in order to determine the most favourable expression host for the production of rChbL (see Figure 3.10). Gene transcription from the pQE60 plasmid is under the control of a T5 promoter/lac operon. This is a highly controllable expression system, in which the T5 promoter is repressed by the presence of a LacI protein. An inducer molecule known as Isopropyl β-D-1-thiogalactopyranoside (IPTG) inactivates the LacI repressor allowing transcription and subsequent protein translation to proceed (Reece, 2004).
Induction of the *E. coli* strains KRX, JM109 and XL10 Gold, harbouring the pQE60_Chb+ plasmid was achieved by the addition of 100 μM IPTG. Culture media samples were analysed at 5 hours post-induction for the presence of rChbL (see Figure 3.10). While the rChbL protein was expressed in all strains tested, unsurprisingly, the recombinant protein was most strongly expressed in the endonuclease deficient *E. coli* KRX strain.

Fig 3.10: Analysis of ChbL expression in the *E.coli* strains KRX, JM109 and XL10 gold. The ChbL protein was expressed in three different strains of *E. coli*; KRX, XL10Gold and JM109. Culture media samples were taken 5 hours after induction (Section 2.15) and analysed by 12.5% SDS-PAGE. Lane 1; Protein ladder Lane 2; KRX Lane 3; XL10 Gold Lane 4; JM109. The largest protein concentration observed is highlighted by an arrow. The largest protein concentration observed is highlighted by an arrow.

Further optimisation was undertaken by analysing the optimal concentration of IPTG for maximal recombinant protein expression in *E. coli* KRX (see Figure 3.11). The use of alternative media (see Figure 3.12) and growth temperatures (see Figure 3.13) were also examined.

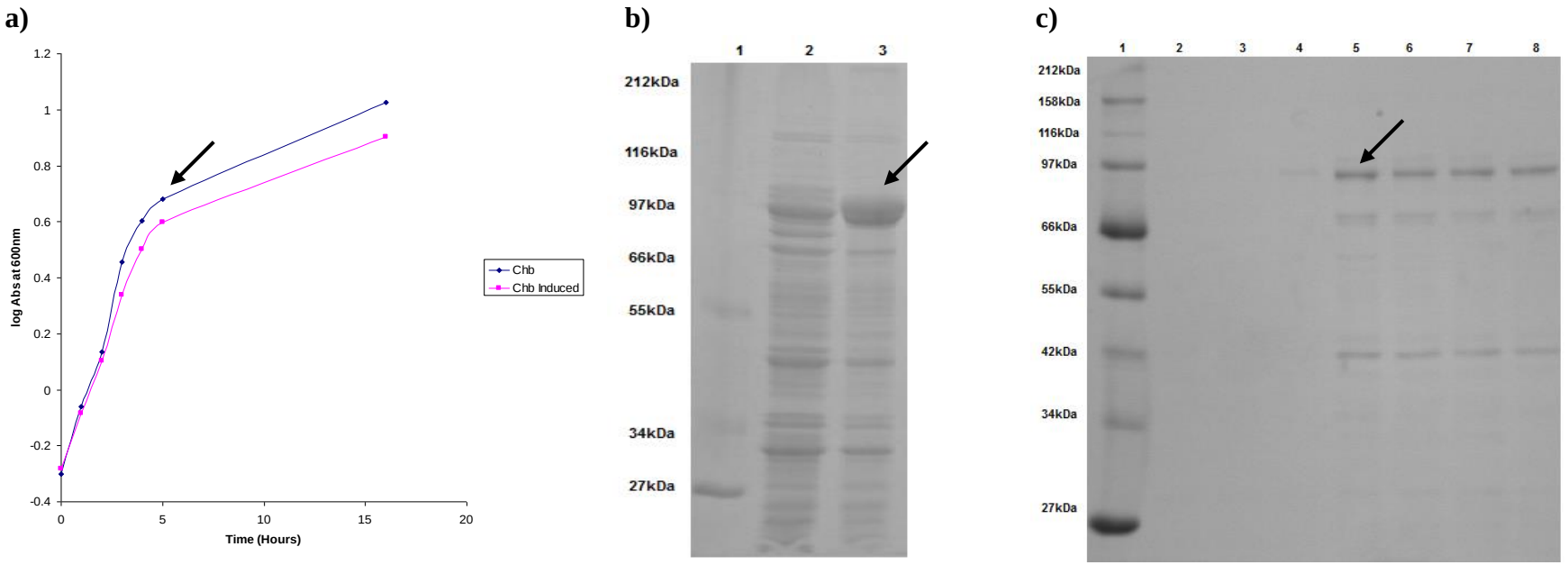


Fig 3.11: Examining the effect of IPTG a) on growth rate b) on recombinant protein expression and c) concentration on rChbL protein expression. a) and b) IPTG was added to a 100 mL rChbL culture at a concentration of 100 µM, a second ChbL culture was left uninduced. The absorbance of each culture was monitored at 600 nm every hour for five hours post-induction and again following overnight incubation. The culture media was collected 5 hours post-induction (as described in Section 2.15) and the samples were analysed by 12.5% SDS-PAGE. Lane 1; Protein Ladder Lane 2; Chb not induced Lane 3; Chb induced.

c) IPTG was added at 0, 5, 50 100, 250, 500 µM and 1M concentrations to determine the optimum IPTG concentration for maximum expression of ChbL. Culture media samples were collected and analysed at 16 hours post induction. Lane 1; Protein Ladder Lane 2; 0 µM Lane 3; 10 µM Lane 4; 50 µM Lane 5; 100 µM Lane 6; 250 µM Lane 7; 500 µM Lane 8; 1 M.

Addition of IPTG to the rChbL culture resulted in a marginal drop in growth rate. The biggest difference in cell density between the two cultures was observed after 16 hours (overnight incubation).

Analysis of the effect of IPTG addition on protein expression demonstrated that expression of the recombinant protein could be tightly controlled. When no IPTG was added there was very little protein produced. However, when IPTG was added the protein was expressed at quite high levels. Therefore, the protein expression system was inducible and tightly controlled. The optimum concentration of IPTG for maximum protein yield was observed to be 100 μM. Higher concentrations of IPTG did not result in higher protein yields, therefore 100 μM IPTG was used for induction in all future experiments.

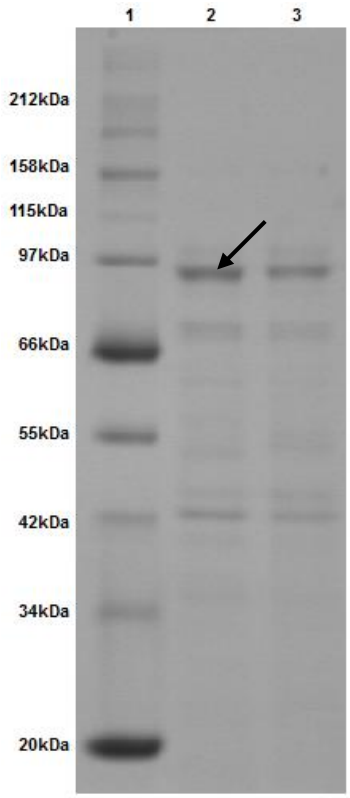


Fig 3.12: Analysis of rChbL expression in different media; TB and LB. The rChbL protein was expressed in both TB or LB (see Section 2.3) and culture media samples were taken for analysis 5 hours post-induction. These samples were analysed by 12.5% SDS-PAGE. Lane 1; Protein ladder Lane 2; Culture media TB Lane 3; Culture media LB. The largest protein concentration observed is highlighted by an arrow.

There was a discernible difference observed in the expression levels of rChbL depending on the media used (see Figure 3.12). A higher level of protein expression resulted from the use of TB compared to LB. TB was used in all subsequent experiments.

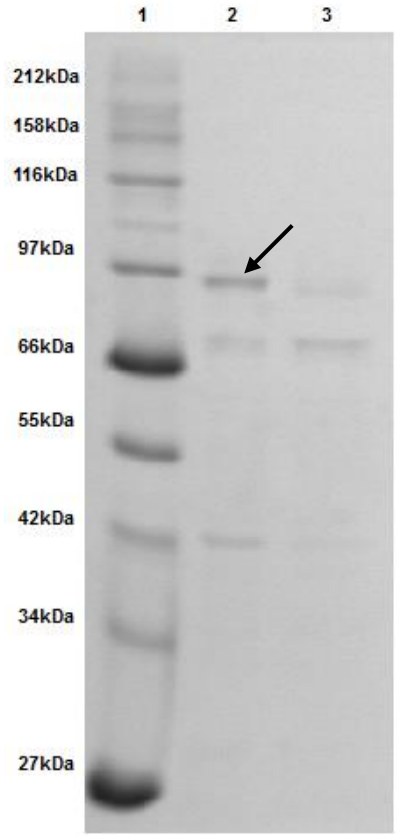


Fig 3.13: Analysis of rChbL protein expression at different temperatures. The cultures were incubated at 37^{o}C until an OD_{600nm} of 0.5 was reached. They were subsequently induced with 100 µM IPTG, then incubated at either 30^{o}C or 37^{o}C for 5 hours. A culture media sample was then evaluated by 12.5% SDS-PAGE. Lane 1; Protein ladder Lane 2; 30^{0}C Lane 3; 37^{0}C. The largest protein concentration observed is highlighted by an arrow.

The study of the effect of incubation temperature on rChbL expression post induction (see Figure 3.13) revealed that the recombinant protein was more highly expressed following incubation at 30°C than at 37°C and was therefore grown at 30^{o}C in all subsequent experiments.

3.7 Comparison of the rChbL protein purifications from the soluble cellular fraction and the culture media fraction

Immobilised Metal Affinity Chromatography (IMAC) is a method used to purify a protein of interest from a complex mixture. It exploits the affinity between certain amino acids and metal ions. In this instance the strong interaction between histidine residues and nickel ions (Ni^{2+}) was exploited (Reece, 2004). As discussed previously the rChbL protein was genetically tagged with 6 consecutive histidine (His_6) residues using the pQE60 vector from Qiagen. The His_6 tag which has a high affinity for Ni^{2+} will facilitate purification by IMAC. IMAC sepharose, to which chelating groups had been covalently coupled was purchased as slurry and charged using $NiSO_4$. This resin was packed into a column (to a height of 2.5 mL) and the protein samples (100 mL) were passed over it slowly to ensure maximum time for binding between the His_6 tag and the immobilised Ni^{2+}. As histidine residues are often found on the surface of many other cellular proteins, wash steps must be employed to remove any weakly binding contaminants. These wash steps were performed with various concentrations of free imidazole, a 5-membered ring belonging to the histidine side chain. It facilitates the removal of any weakly bound proteins from the IMAC resin (Ni^{2+}) due to its higher affinity for free imidazole than the weakly bound cellular proteins. Stringency washes composed of increasing concentrations of imidazole were employed to test the strength of the interaction between the protein of interest and the IMAC column. Imidazole washes were also used for the elution of the recombinant protein of interest using a much higher concentration.

This section outlines the purification of protein from the culture media and from the soluble protein fraction to compare both yield and purity of the recombinant protein. IMAC optimisation was undertaken to determine which protein had superior binding to the Ni^{2+} resin and the wash stringencies that could be employed (see Figure 3.14). The optimised purification was then carried out to examine the yield and purity of the rChbL protein (see Figure 3.15).

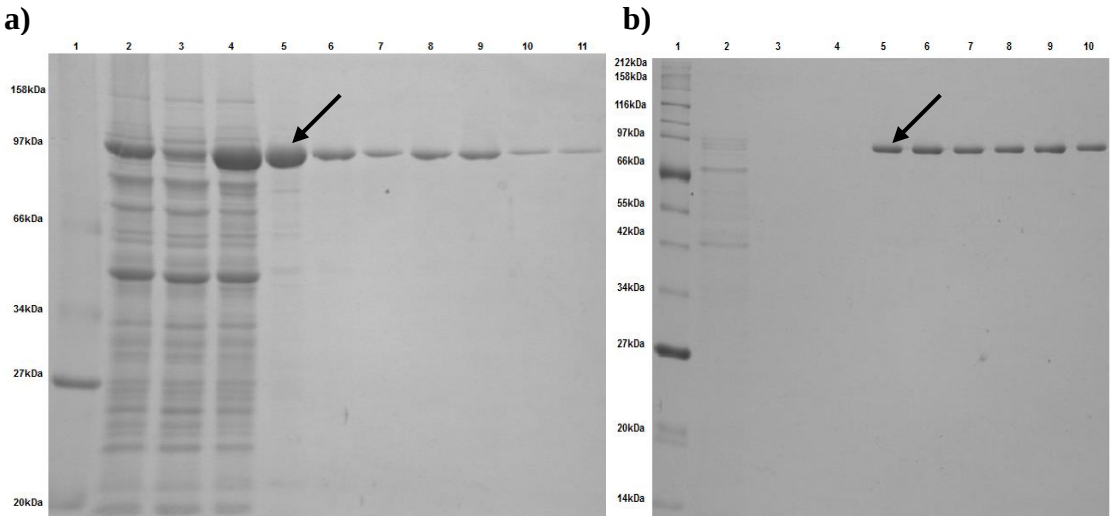


Fig 3.14: IMAC optimisation for the purification of rChbL from a) the soluble protein fraction and b) the culture media. Stringency washes were performed with an increasing imidazole concentration gradient from 20 mM to 250 mM, to determine when ChbL was eluted from the IMAC column. The fractions were then examined by 12.5% SDS-PAGE.

a) Lane 1; Protein ladder Lane 2; Lysate Lane 3; Flow through one Lane 4; Flow through 2 Lane 5; 10 mM Imidazole Lane 6; 20 mM Imidazole Lane 7; 40 mM Imidazole Lane 8; 60 mM Imidazole Lane 9; 80 mM Imidazole Lane 10; 100 mM Imidazole Lane 11; 200mM Imidazole.

b) Lane 1; Protein ladder Lane 2; Flow through Lane 3; 10 mM Imidazole Lane 4; 20 mM Imidazole Lane 5; 40 mM Imidazole Lane 6; 60 mM Imidazole Lane 7; 80 mM Imidazole Lane 8; 100 mM Imidazole Lane 9; 200 mM Imidazole.

The rChbL protein as isolated from the culture media displayed superior binding to the IMAC column over that obtained from the soluble protein fraction. It can be seen from Figure 3.14 that the soluble cellular protein eluted from the column upon the application of 10 mM imidazole. In comparison the protein isolated from the culture media was able to withstand imidazole washes of up to 40 mM before eluting from the resin.

The ability of the culture media form of rChbL to withstand higher levels of imidazole in the wash buffers is very advantageous. Purification of rChbL from the soluble protein fraction would result in a highly contaminated product as 10 mM imidazole is not a strong enough concentration to wash away all contaminating proteins and to ensure the elution of pure protein. This capacity of the culture media form of rChbL to

withstand high stringency imidazole washes ensures the elution of a pure homogenous fraction of protein as seen in Figure 3.15.

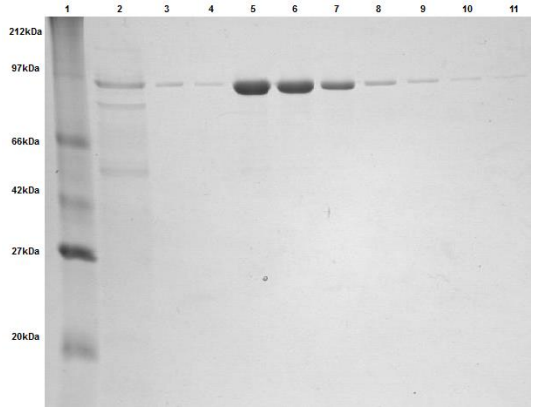


Fig 3.15: Purification of the recombinant ChbL protein from the culture media. The culture media was collected after cell harvest from an rChbL expression culture. The protein samples were run over an IMAC column, washed with 40 mM imidazole and eluted with 250 mM imidazole. Fractions were analysed by 12.5% SDS-PAGE.
b) Lane 1; Protein ladder, Lane 2; Flow through, Lane 3; Wash 1 (40 mM), Lane 4; Wash 2 (40 mM), Lane 5; Elution one, Lane 6; Elution three, Lane 7; Elution five, Lane 8; Elution 8, Lane 9; Elution 11, Lane 10; Elution 14, Lane 11; Elution 18.

Following the isolation of pure protein it is necessary to carry out a number of 'polishing' steps to remove the imidazole and exchange the protein into a buffer suitable for downstream assays and storage. Two methods of buffer exchange were compared (see Figure 3.16). The first method involved the use of size exclusion membranes (vivaspins columns from Sartorius). The second method involved the use of the 26/10 desalting column (as described in Section 2.20) on the FPLC (see Figure 3.17).

Vivaspins are centrifugal filter units with a cellulose membrane. The membranes come in various sizes that allow for the removal of salts and proteins with lower molecular weights that the membrane cut off. If the correct membrane size is selected the target protein will be retained while allowing the buffer salts to flow through the membrane freely. This technique was utilised to dilute out the imidazole and to exchange the protein into a buffer of choice, which in this case was PBS.

From the graph below (see Figure 3.16) it can be seen that the majority of the protein was lost using the vivaspin columns. This was thought to be due to the protein binding to the cellulose membrane, possibly due to the carbo-binding properties of the protein. The FPLC method employed used a gel filtration column which allowed the high molecular weight recombinant proteins to elute first as they were excluded from the gel. The smaller substances such as imidazole entered the gel pores and therefore eluted later. The protein yield resulting from FPLC was almost three times higher than that obtained using the vivaspin columns so this was the method chosen for all future buffer exchange of the protein.

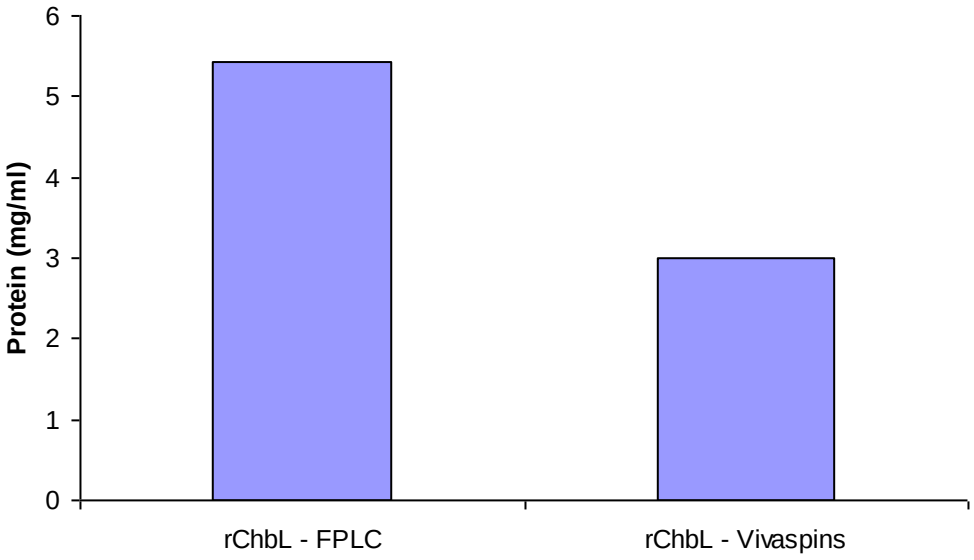


Fig 3.16: Comparison of the buffer exchange (desalting) of the purified rChbL protein using Vivaspins and FPLC. The elution fractions from the culture media purification containing protein as determined by absorbance readings (see Section 2.23.2) were pooled together following IMAC purification. The pooled fractions were then either buffer exchanged using Vivaspins (Sartorius) or using FPLC (see Section 2.20). The final protein concentration was determined using 280 nm readings as described in Section 2.23.2.

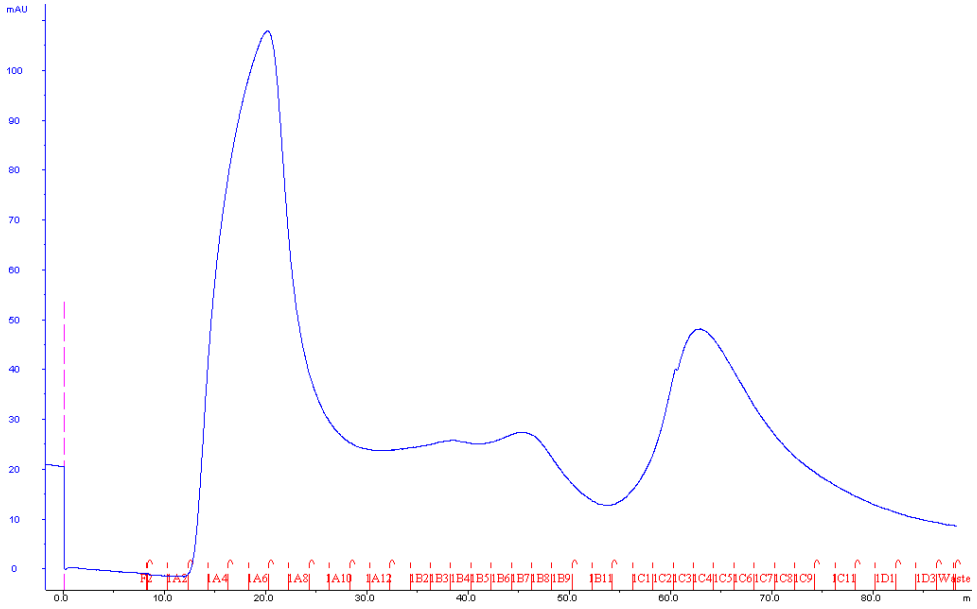


Fig 3.17: Desalting/Buffer exchange of purified protein using a HiPrep 26/10 desalting column on an FPLC system. Buffer exchange was carried out as per Section 2.20. The first large peak in blue signifies the protein being eluted from the column. The other smaller peaks are impurities and smaller molecules such as the imidazole being removed. Fractions A2-A10 were collected and the protein quantified via absorbance readings (see Section 2.23.2).

Approximately 5 mg per 250 mL TB culture of pure rChbL protein was obtained after purification and buffer exchange. This indicated that the majority of the rChbL protein was being secreted from the cell after production.

3.8 Characterisation of rChbL catalytic activity

The catalytic activity of an enzyme may be determined by measuring either the rate of formation of a product or the rate of depletion of a substrate. The catalytic activity of the enzyme was determined by measuring the rate of product formation using a commercial synthetic chromogenic substrate, 4-Nitrophenyl N-acetyl β-D-glucosaminide (PNP-GlcNAc) from Sigma. This compound is comprised of an N-acetylglucosamine molecule linked to a fluorescent compound, 4-Nitrophenol. Enzymatic cleavage of PNP-GlcNAc releases 4-PNP, which forms the yellow coloured 4-Nitrophenolate under basic conditions. Na_2CO_3 was used to stop the reaction and develop the phenolate ion which was measured at 405nm. Activity on natural compounds such as chitobiose and other chito-oligosaccharides were explored using electrophoresis and HPLC assays as discussed in Chapter 6.

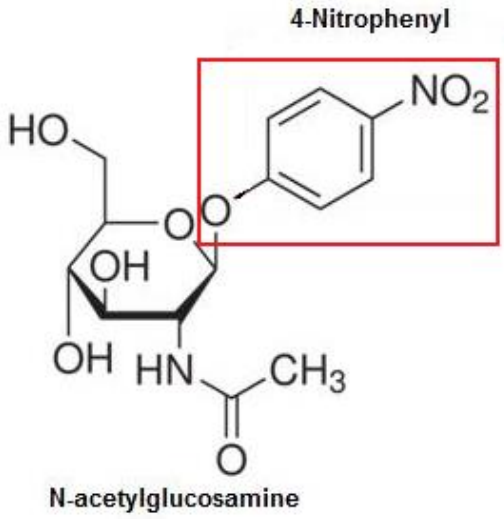


Fig 3.18: 4-Nitrophenyl N-acetyl β-D-glucosaminide (PNP-GlcNAc). The synthetic substrate used for the chitobiase activity assay. Image obtained from Sigma Aldrich and modified (http://www.sigmaaldrich.com/ireland.html).

3.8.1 Development of the rChbL activity assay

A modification of the assay was developed by examining the optimum concentration at which to use the rChbL enzyme (see Figure 3.20) and the time it took for the rChbL – PNP-GlcNAc reaction to reach completion (see Figure 3.21). A standard curve was developed using commercial 4-Nitrophenol to determine the amount of substrate cleaved in the Chb assay.

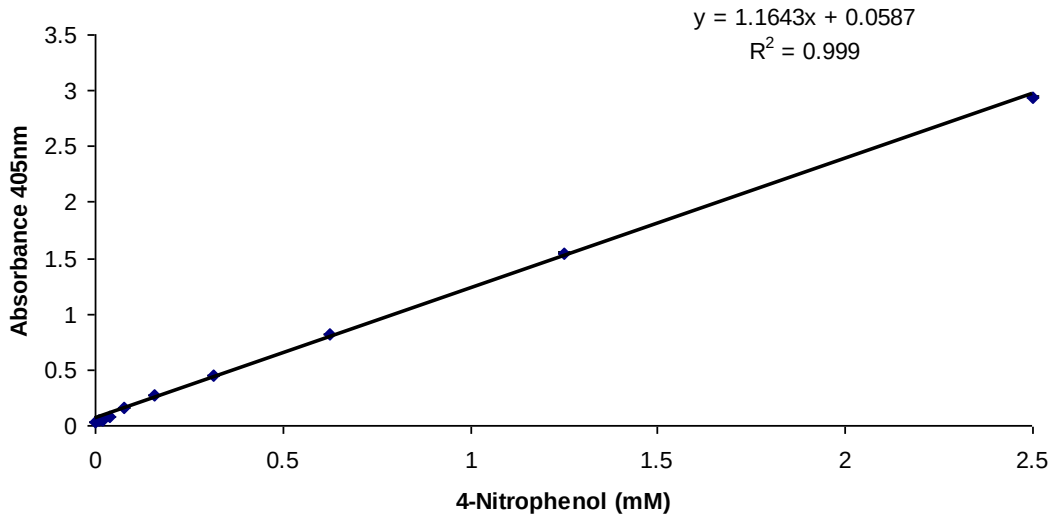


Fig 3.19: Standard curve of 4-Nitrophenol. A standard curve was constructed using 4-Nitrophenol (Sigma), the chromogenic molecule that is released upon cleavage of 4-Nitrophenyl N-acetyl β-D-glucosaminide. A range of concentrations of 4-Nitrophenol was made up in sodium phosphate buffer (50 mM, pH 8.0). 110 µL of each concentration was added to a 96-well plate in triplicate, 100 µL Na_2CO_3 was added to ionise all the 4-Nitrophenol to 4-Nitrophenolate and the absorbance was read at 405 nm. Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

This standard curve was used to determine the millimolar (mM) concentration of 4-Nitrophenol released through enzymatic cleavage of the PNP-GlcNAc substrate in further activity assays.

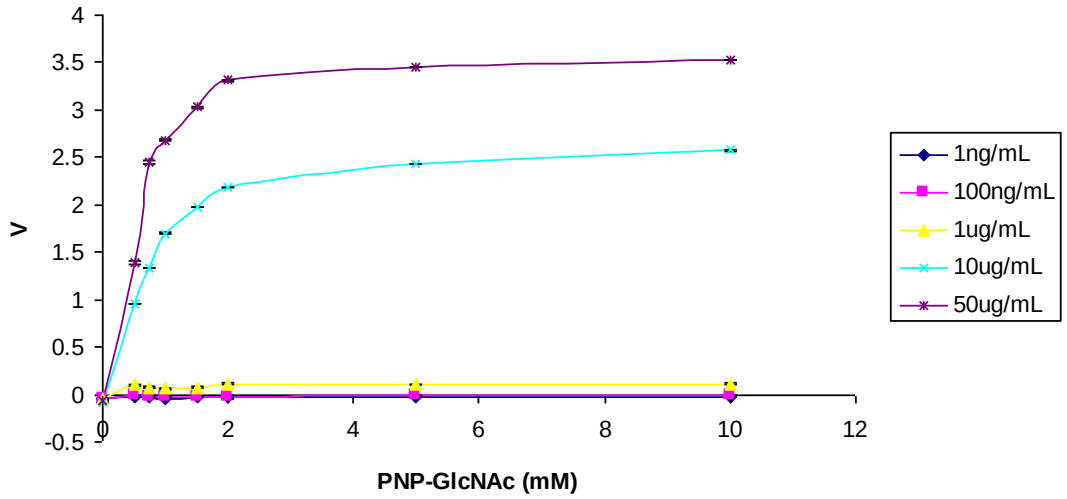


Fig 3.20: Determination of the optimum rChbL concentration for the Chb activity assay. The assay was carried out as described in Section 2.26 with a variety of enzymatic concentrations (1 ng/mL – 50 µg/mL) over a wide range of PNP-GlcNAc concentrations (0-10 mM) to determine the optimum concentration at which to use the enzyme. The rate, V (mM PNP released/min/mL enzyme), of the reaction was plotted. Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

There was very little catalytic activity observed when the lower concentrations of enzyme were used, i.e. 1 ng/mL, 100 ng/mL and 1 µg/mL. The most catalytic activity was observed when concentrations of 10 µg/mL and 50 µg/mL were used. When 10 µg/mL of rChbL was used the absorbance values were in the range of 0.1-1.0 for the lower substrate concentrations (0-2 mM PNP-GlcNAc). This concentration achieved the most reliable results and was therefore used for future experiments.

Enzymatic activity also peaked at a substrate concentration of approximately 2 mM over all enzyme concentrations used. A concentration of 2 mM PNP-GlcNAc was used in all further experiments in order to ensure substrate saturation. This was to guarantee that substrate was not a limiting factor in the reaction and, thus, would not affect the kinetic data.

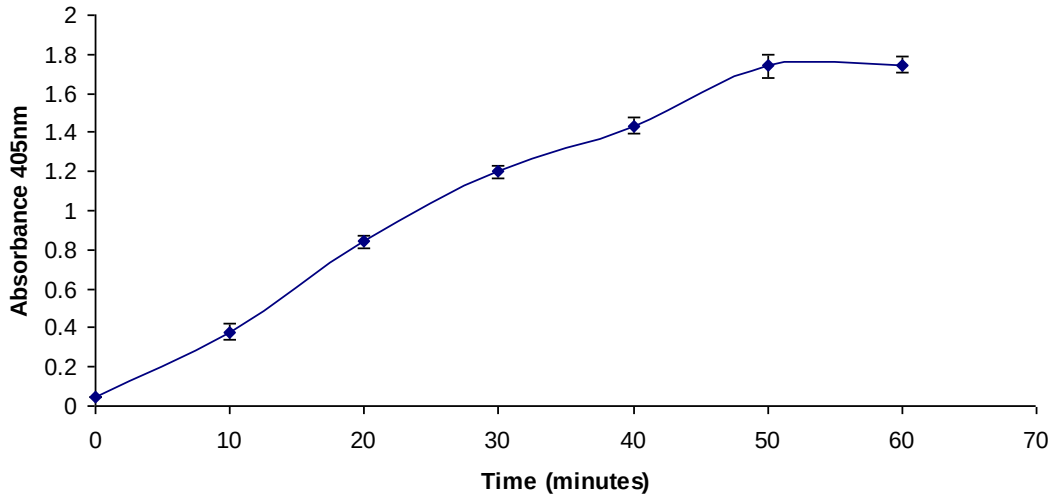


Fig 3.21: Determination of the endpoint of the rChbL catalytic reaction. The Chb activity assay was carried out as described in Section 2.26 in 50 mM sodium phosphate at pH 8.0. It was terminated at various time-points by the addition of 1 M sodium carbonate to determine when the catalytic activity had been exhausted. The rate, V (mM PNP released/min/mL enzyme), of the reaction was plotted. Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

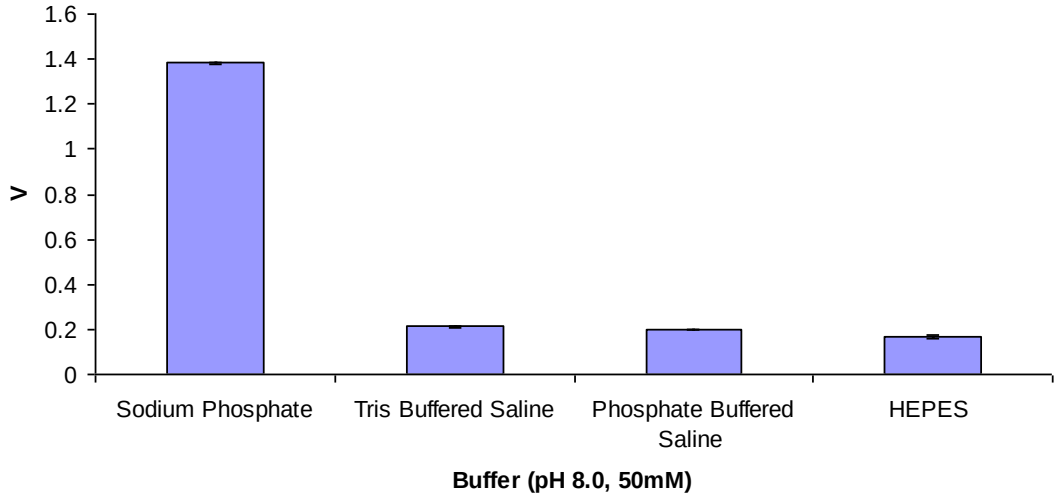
Figure 3.21 demonstrates that breakdown of the substrate peaked after approximately 50 minutes. At 60 minutes the reaction had become saturated. No additional PNP-GlcNAc had been broken down as signified by the lack of change in absorbance/rate of reaction. Therefore, 45 minutes was chosen to ensure that catalytic activity was measured under initial-rate conditions and the rate of product formation was linear throughout the duration of the assay. Catalytic activity was not limited by a lack of substrate.

3.8.2 Characterisation of the rChbL activity assay

The conditions of the assay were altered to determine the environment that optimises rChbL activity. The parameters of the assay examined were the pH at which the reaction took place (see Figure 3.22) and the temperature at which the reaction was carried out (see Figure 3.23). The stability of the enzyme was also determined; its stability at high temperatures was tested (see Figure 3.24), the half-life of the enzyme

was calculated (see Figure 3.25) and the long term stability of the enzyme was tested over a range of time-points and under various storage conditions (see Figure 3.26).

a)



b)

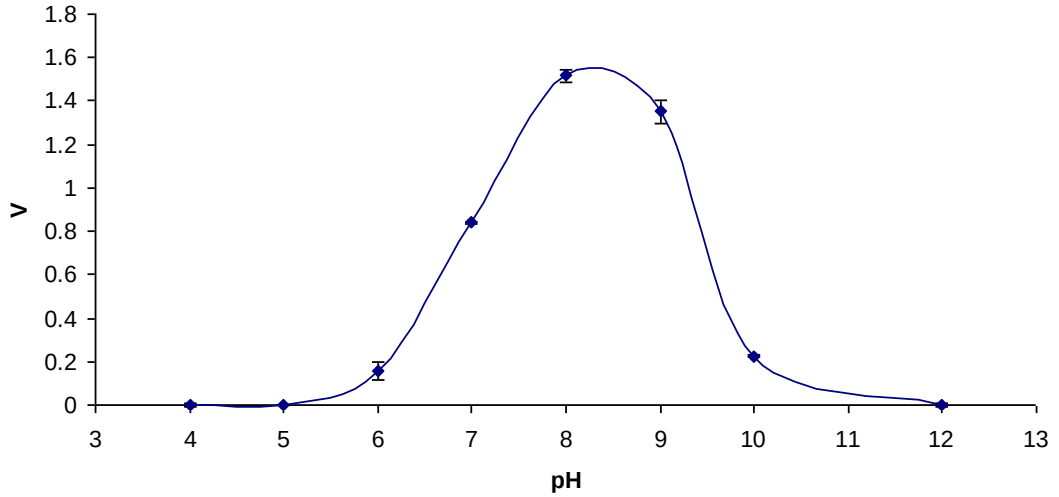


Fig 3.22: Determination of optimum buffer and pH conditions for the Chb activity assay. The Chb activity assay was carried out as described in Section 2.26 with (a) different buffers all at a concentration of 50 mM and all at pH 8.0 and (b) in sodium citrate buffer (50 mM) over a wide range of pH values from 4-12. The rate, V (mM PNP released/min/mL enzyme), of the reaction was plotted. Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

The highest rate of catalytic activity was observed in sodium phosphate buffer. A much lower rate of catalytic activity was observed in Phosphate buffered saline (PBS) and Tris buffered saline (TBS). The lowest rate of activity was observed in HEPES. A bell shaped curve was observed when catalytic activity was examined over a wide range of pH concentrations. There was relatively little catalytic activity observed at low or high pH values. The highest rate of catalytic activity was observed around neutral pH values. Catalytic activity peaked at approximately pH 8.0 after which the activity steadily declined. The optimum buffer conditions for enzyme activity were determined to be 50 mM sodium phosphate buffer at pH 8.0.

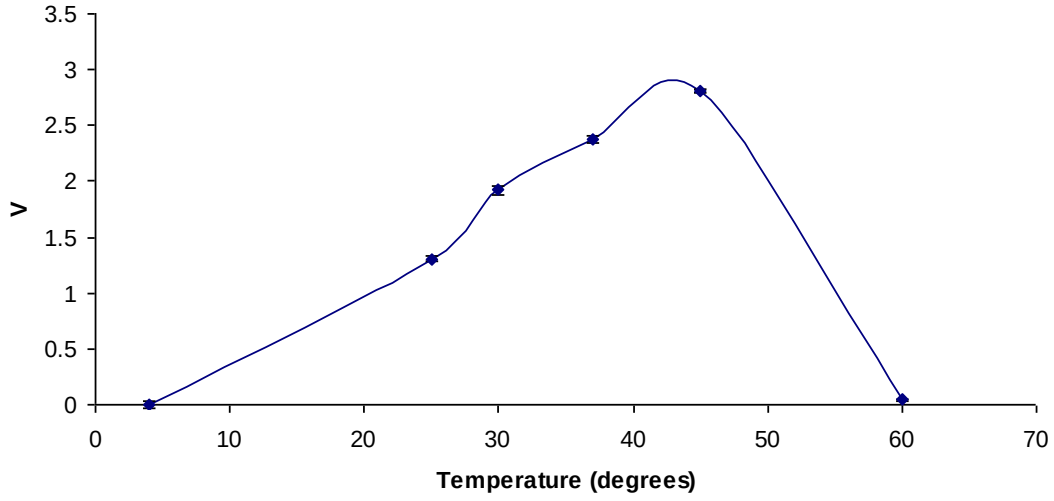


Fig 3.23: Determination of the optimum temperature for the Chb activity assay. The Chb activity assay was carried out as described in Section 2.26 in 50 mM sodium phosphate buffer at pH 8.0 over of range of temperatures, from 25^{o}C to 60^{o}C. The enzyme, substrate and buffer were all maintained at the specific reaction temperature. The rate, V (mM PNP released/min/mL enzyme), of the reaction was plotted. Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

A skewed bell-shaped curve was observed when the rate of catalytic activity was plotted against temperature. The rate of catalytic activity increased until it reached a maximum at 45^{o}C after which there was a dramatic decrease in activity. The optimum temperature for catalytic activity was approximately 45^{o}C. Under 45^{o}C the enzymatic activity had not yet reached its peak. Over 45^{o}C the catalytic activity dramatically decreased. Above 60^{o}C there was negligible activity.

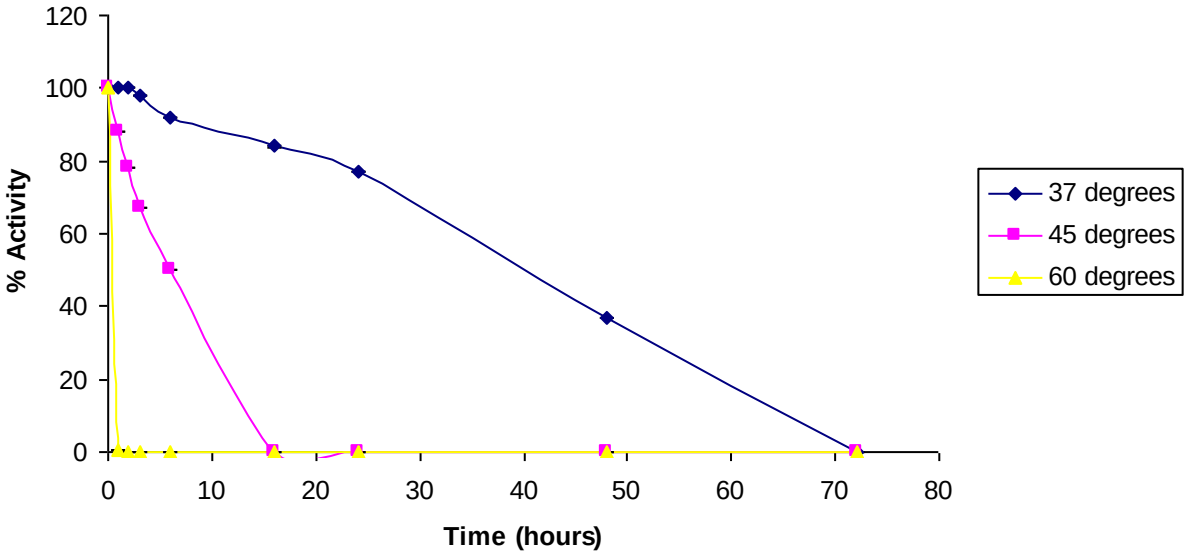
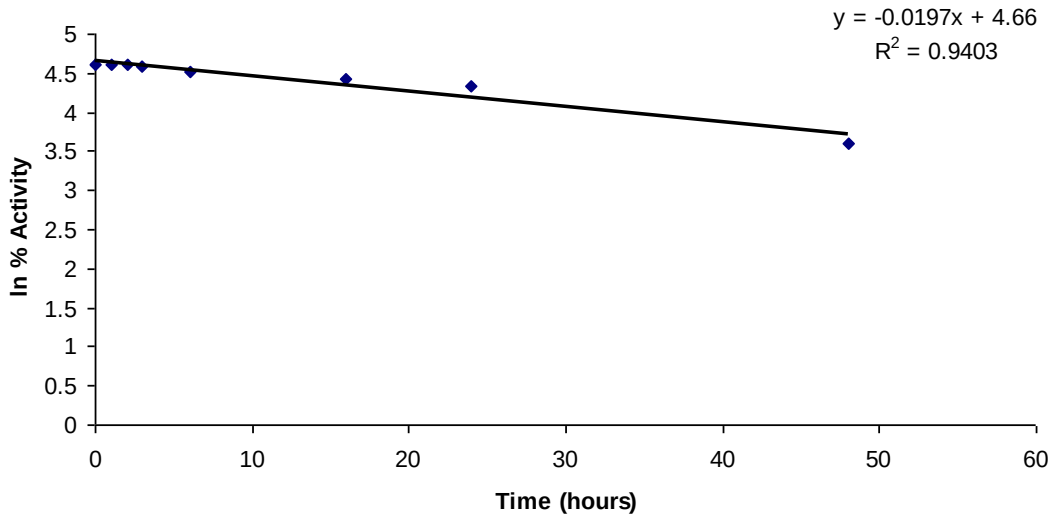


Fig 3.24: Determination of the rChbL enzyme temperature stability at 37°C, 45°C and 60°C. The rChbL enzyme was incubated at 37°C, 45°C and 60°C. Samples were taken at various time-points. The samples were allowed to cool and then the Chb activity assay was carried out as described in Section 2.26 under optimal conditions. Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

The rChbL enzyme was most stable for the longest duration when incubated at 37°C. Catalytic activity remained at 100% for almost 3 hours. It took 3 days to completely inactivate the enzyme. The ChbL enzyme was inactivated after 16 hours when incubated at 45°C. The most severe effect on catalytic activity was observed when the rChbL enzyme was incubated at 60°C. The enzyme was inactivated after only 15 minutes.

a)



b)

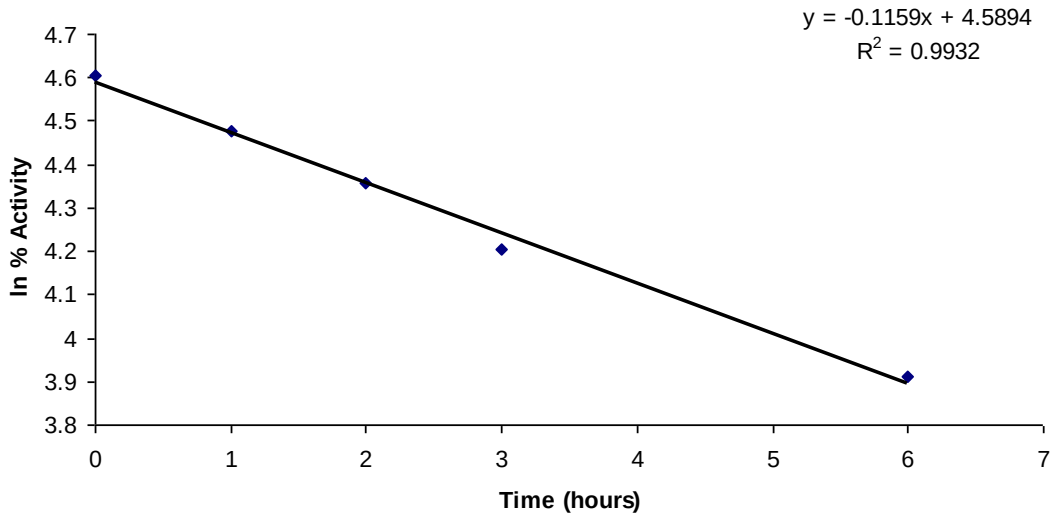


Fig 3.25: Determination of the half-life of the rChbL enzyme at a) 37^{o}C and b) 45^{o}C. Using the results from Fig 3.24, the ln(% Activity) was plotted against time to generate a linear graph. The slope of the line was used to calculate a half for the enzyme using the formula: Half-life = 0.693/k.

The linear plots generated in Figure 3.24 allowed the half-life to be calculated from the slope of the line (k). The half-life at 37^{o}C was calculated to be 35 hours. The enzyme retained the majority of its' catalytic activity for up to 24 hours. Total activity was lost after 72 hours. The half-life at 45^{o}C was calculated to be 6 hours. It took 16 hours for the enzyme to lose all of its' catalytic activity.

3.8.3 Optimisation of the stable storage of rChbL

After purification proteins are no longer in their native environment and may become unstable, are prone to aggregation and proteolysis and may lose their structural integrity. Enzymes may lose their catalytic activity as structure is related to function. Therefore, it was critical to examine the optimal storage conditions for the enzyme to ensure it was kept stable and viable for as long as possible. Time-course analysis was carried out to determine the long-term viability of the rChbL enzyme under different storage conditions. Cold temperatures are generally optimal for the storage of proteins, as they tend to promote protein stability. When freezing proteins, their structure may be disrupted by the formation of ice crystals. Therefore, glycerol was routinely used as a cryoprotectant to prevent ice crystal formation and hence, stabilise the structure of the protein.

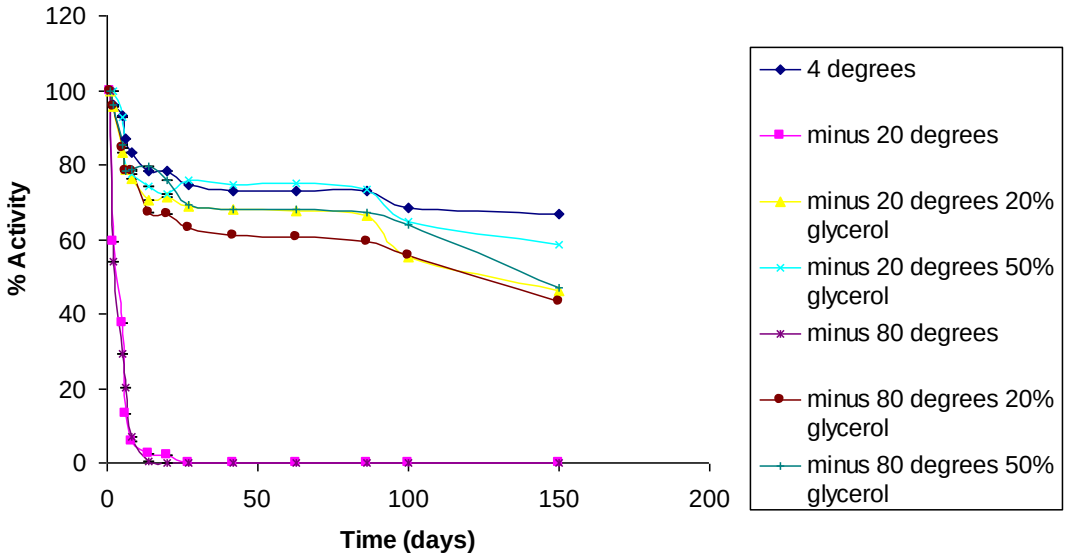


Fig 3.26: Examination of the long-term viability of the rChbL enzyme under various conditions. The purified rChbL enzyme was freshly stored at 100 µg/mL with and without glycerol as a cryoprotectant over a range of temperatures. The Chb activity assay was carried out as described in Section 2.26 at the various time-points. The rate, V (mM PNP released/min/mL enzyme), of the reaction was plotted against time (days). Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

Figure 3.26 clearly highlights that storage at -20^{o}C in the absence of a cryoprotectant was very unsuitable for the rChbL enzyme. After two days the enzyme had lost over half of its catalytic activity with total catalytic activity lost after a week. Storage at -20^{o}C in the presence of varying concentrations of a cryoprotectant significantly increased the storage life of the enzyme. In the presence of 20% glycerol, half of the catalytic activity was lost after 150 days while increasing the concentration of glycerol to 50% reduced the activity loss further with 66% of the activity still present after 150 days. Again, at -80^{o}C it was observed that storage without a cryoprotectant caused a dramatic decrease in the rate of catalytic activity while addition of a cryoprotectant greatly protected against the loss of this activity. The difference in activity between these storage conditions highlighted the fact that if the rChbL enzyme was to be frozen it needed a cyroprotectant to protect its structural integrity and therefore its catalytic activity. Overall, however, the ideal condition for storage of the rChbL enzyme was at the milder temperature of 4^{o}C. At this temperature the ChbL enzyme retained the most catalytic activity over the longest period of time. It lost roughly 20% of its activity over the first week in storage but in the following months the catalytic activity remained at a constant stable level. In general proteins are stored at high concentrations above 1 mg/mL. This is thought to prevent the proteins from sticking to the surface of the storage vessel and promotes the stabilisation of the protein structure. However, the ChbL enzyme had to be stored at a concentration of 100 μg/mL as higher concentrations were found to promote aggregation of the protein and caused precipitation of the protein from solution.

3.8.4 Preliminary kinetic study of rChbL

It is important to calculate the reaction rates of an enzyme so that enzymes under different conditions or from different sources may be compared. Parameters that can be measured from kinetics plots include the catalytic efficiency of the enzyme, the specific activity of the enzyme, the affinity of the enzyme for its substrate and the effect of inhibitors on the catalytic activity of the enzyme. There are many different plots that can be used to describe the kinetic values of an enzyme; Michaelis-Menten equation, Lineweaver-Burk plot and Hanes-Woolf plot. The Michaelis-Menten equation is the simplest way to describe the rate of a reaction. It plots the rate of the reaction against the substrate concentration. In this way it describes the relationship

between the rate of substrate conversion and the concentration of that substrate. Lineweaver-Burk plots are a more sophisticated version of the Michaelis-Menten equation. However, they tend to be unreliable. While the Michaelis-Menten equation uses a curve the Lineweaver-Burk plot generates a straight line. This is possible by plotting the inverse of the rate of the reaction against the inverse of the substrate concentration. The kinetic values are then calculated from the equation of the line. However, by plotting reciprocal values any small discrepancy in the data can result in large errors and therefore they will not be used in these kinetic studies. Lineweaver-Burk plots are, however, useful when studying enzyme inhibition, allowing the various types of inhibitions to be distinguished. This will be discussed more in Section 3.87. Hanes-Woolf plots are generally regarded as the best plot to use. The Hanes-Woolf plot also generates a linear graph and allows the kinetic data to be calculated from the equation of the line. The relationship between the substrate and rate of the reaction (S/V) is plotted against the substrate concentration (S). Here the Michaelis-Menten equation and Hanes-Woolf plot were used to describe the results.
Below are the values that can be calculated from Michaelis-Menten and Hanes Woolf plots:

V_{max}: The maximum enzyme velocity achieved in terms of substrate concentration.
K_m: The Michaelis-Menten constant. It is a measure of the affinity of the enzyme for the substrate.
k_{cat}: Turnover number (measure of the rate of product formation).
k_{cat}/K_m: A measure of how efficiently the enzyme converts the substrate to product, (catalytic efficiency).

The enzymatic activity was measured over a range of substrate concentrations and the values were determined using the above methods.
However, these plots and calculations are only applicable to first order reactions. This had to be verified by determining the dependency of the catalytic activity on enzyme concentration (see Figure 3.27).

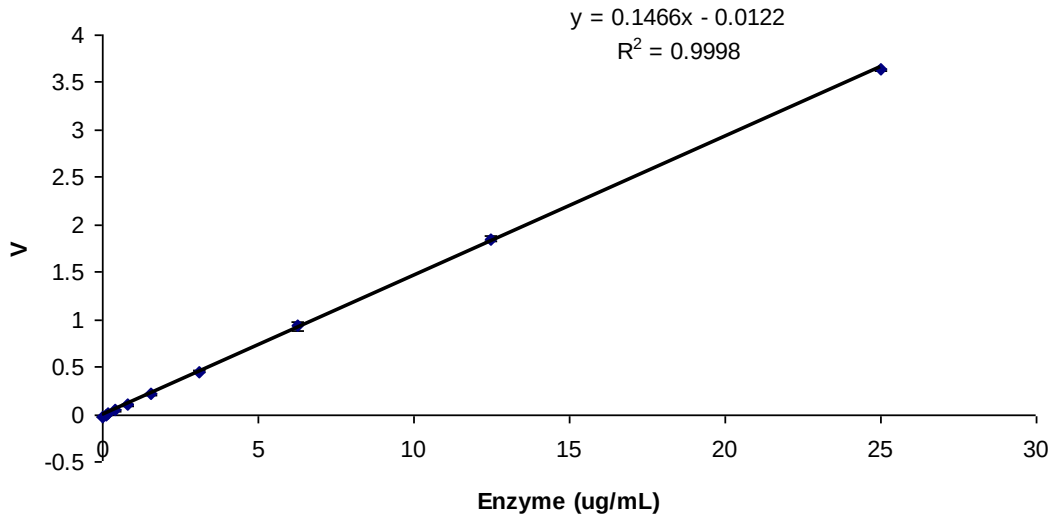
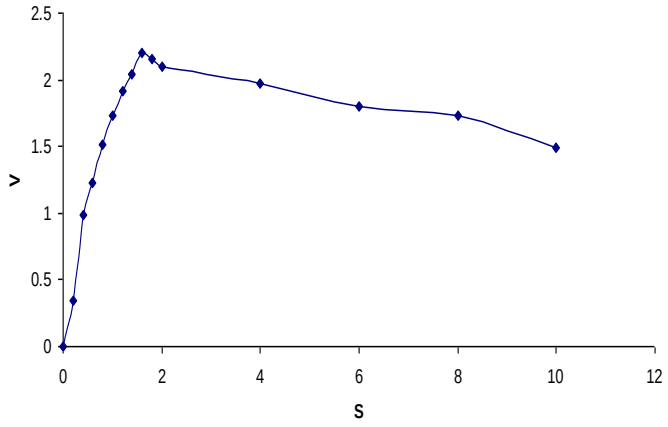


Fig 3.27: Determination of the dependency of catalytic activity on enzyme concentration. The Chb activity assay was carried out as described in Section 2.26. A constant concentration of PNP-GlcNAc was incubated with increasing concentrations of ChbL, from 0-25 μg/mL. The rate of the reaction, V (mM PNP released/min/mL enzyme), was plotted against enzyme concentration (μg/mL). Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

The rate of reaction increased with enzyme concentration in a linear fashion, showing this to be a first order reaction.

a)



b)

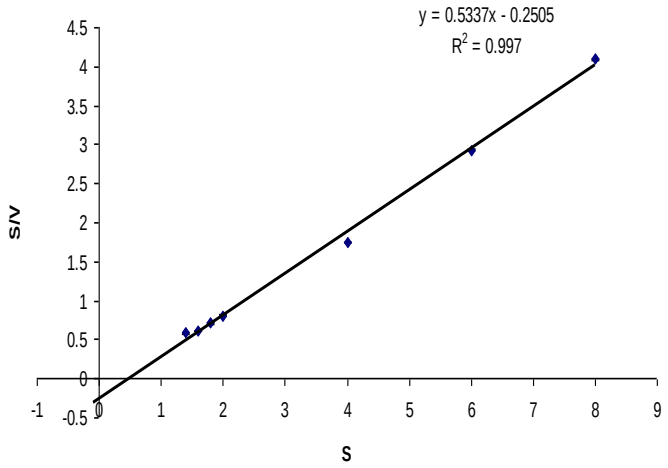


Fig 3.28: rChbL kinetics were examined and compared using two different methods a) Michaelis-Menten curve and b) Hanes-Woolf plot. The Chb activity assay was performed as per Section 2.26 over a range of PNP-GlcNAc concentrations from 0 to 10 mM. The rChbL enzyme was used at a concentration of 10 μg/mL. The rate, V (mM PNP released/min/mL enzyme), of the reaction was plotted against substrate concentration (mM). Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

Table 3.3 Kinetic values of the rChbL enzyme as determined by a Michaelis Menten curve and a Hanes-Woolf plot

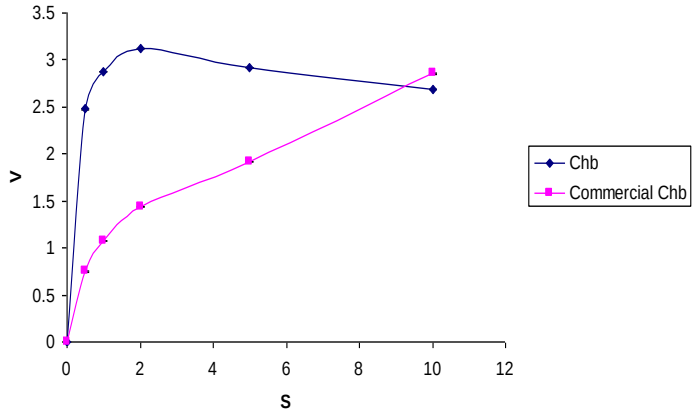
Rate Constants	Michaelis Menten	Hanes-Woolf
V_{max} (μM)	2610	1870
K_m (mM)	0.65	0.47
k_{cat} (s^{-1})	$4.3x10^5$	$3.1x10^5$
k_{cat}/K_m ($mM^{-1}.s^{-1}$)	$6.6x10^5$	$6.56x10^5$
SD (±)	0.00562	0.00562

The kinetic values obtained from the Michaelis-Menten curve and Hanes-Woolf plot correlated quite closely (see Table 3.3). The high V_{max} values obtained for the rChbL enzyme represent the high velocity reached by the rChbL enzyme. The low K_m values observed for the ChbL enzyme demonstrate the high affinity of the rChbL enzyme for PNP-GlcNAc. The ChbL enzyme was discovered to be very catalytically efficient as described by the high k_{cat} and k_{cat}/K_m values.

3.8.5 Comparison of rChbL with a commercial chitobiase

A chitobiase (β-N-acetylglucosaminidase) was obtained from New England Biolabs (NEB). This enzyme had been cloned from *Xanthomonas manihotis* and expressed in *E. coli*. The enzymatic activity was identical to that of the recombinant Chb in that it cleaves N-acetylglucosamine residues from the terminal non-reducing end of oligosaccharides. The activity of the commercial Chb was calculated (see Figure 3.29).

a)



b)

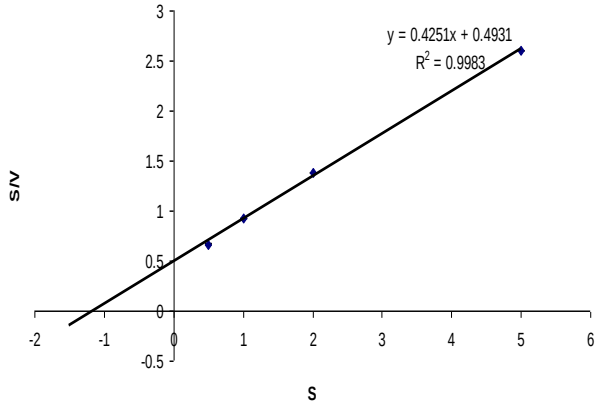


Fig 3.29: Commercial Chb kinetics measured using a a) Michaelis-Menten curve and b) Hanes Woolf plot. The Chb activity assay was performed as described in Section 2.26. Kinetic values were calculated and compared for both enzymes. The rate, V (mM PNP released/min/mL enzyme), of the reaction was plotted. Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

Table 3.4 Kinetic values for the commercial Chb as determined by a Michaelis-Menten curve and Hanes-Woolf plot.

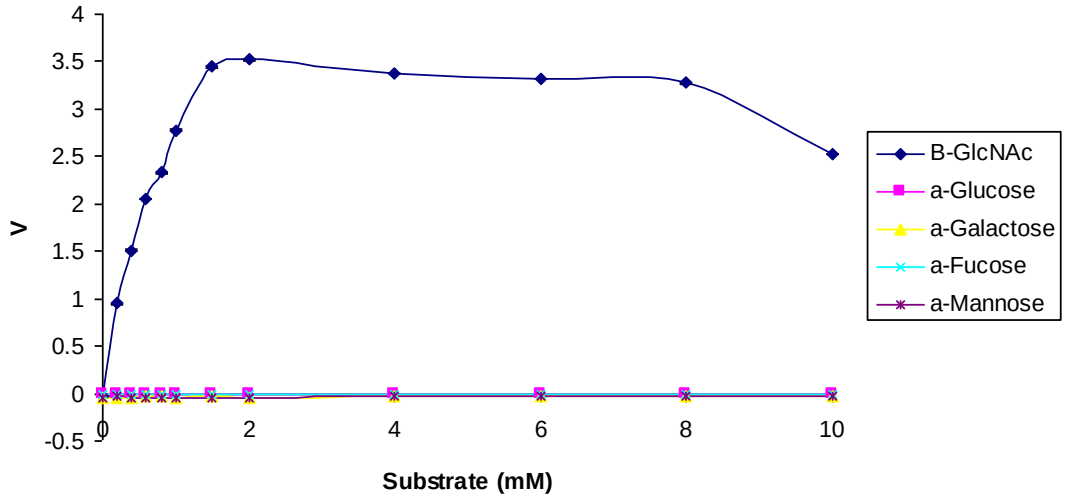
Rate Constants	Michaelis Menten	Hanes Woolf
V_{max} (μM)	2856	2350
K_m (mM)	1.9	1.16
k_{cat} (s^{-1})	$3.43x10^5$	$2.82x10^5$
k_{cat}/K_m ($mM^{-1}.s^{-1}$)	$1.8x10^5$	$2.44x10^5$
SD (±)	0.00212	0.00212

The commercial Chb was found to have a slightly higher velocity than the rChbL from this study. However, the K_m of the commercial Chb was found to be higher, signifying a lower affinity for its substrate as compared with the lower K_m of the rChbL from this study. The k_{cat} values of both enzymes were similar; each value was 10^5 in magnitude. The k_{cat}/K_m values were also similar; both again were 10^5 in magnitude. The rChbL from this study had slightly higher k_{cat} and k_{cat}/K_m values than the commercial Chb, demonstrating a higher catalytic efficiency than the commercial Chb.

3.8.6 rChbL Specificity

A number of different PNP-linked sugar substrates were tested against the rChbL enzyme to investigate its specificity for PNP-GlcNAc as described in Section 2.26.2. The sugars used were PNP linked α-glucose, α-galactose, α-fucose, α-mannose, sialic acid, β-glucose and β-galactose.

a)



b)

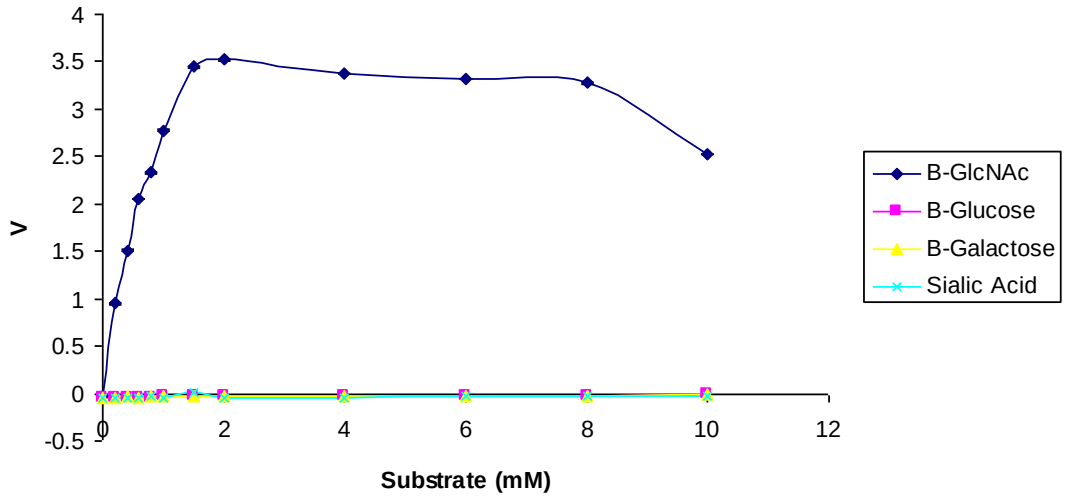


Fig 3.30: Examination of the substrate specificity of the rChbL enzyme. The Chb activity assay was performed as per Section 2.26.2 using different a) PNP α -linked substrates and b) PNP β -linked substrates and PNP-Sialic Acid. The substrates were used over a concentration range from 0-10 mM. The rate V (mM PNP released/min/mL enzyme) of the reaction was plotted against substrate concentration (mM). Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

Figure 3.30 highlights the specificity of the rChbL enzyme for PNP-GlcNAc. The rChbL enzyme had relatively no affinity for any of the other substrates. No activity was expected for the α-sugars, however a small amount of activity wouldn't have

been unexpected for the β-sugars as the native substrate of the rChbL enzyme is a β-linked sugar.

3.8.7 Sugar Inhibition of rChbL

Various free sugars were added in increasing concentrations to the rChbL - PNP-GlcNAc reaction to determine their inhibitory effect as described in Section 2.26.3. N-acetylglucosamine, glucose, galactose, lactose, mannose, fucose and xylose were chosen as they covered a wide variety of structures, which allowed the affinity of the rChbL enzyme towards PNP-GlcNAc to be tested.

There are three types of reversible enzyme inhibition, competitive, non-competitive and uncompetitive, which may be distinguished using Lineweaver-Burk plots. Competitive inhibition occurs when the inhibitor binds to the active site of the enzyme thereby blocking this site for the substrate. The inhibitor and substrate are structurally similar and must compete for the same site. The 1/V intercept stays the same however the slope of the line increases with increasing inhibitor. Non-competitive inhibition is when the inhibitor binds to the enzyme at a site other than the active site. It does not inhibit the substrate binding to the active site but has changed the conformation of the enzyme so it is unable to catalyse the reaction. The slope and 1/V intercept increase with increasing inhibitor. Uncompetitive inhibitors only bind to the enzyme-substrate complex. This inhibits the enzyme from catalyzing the reaction. The slope of this line stays the same however the 1/V intercept increases with increasing inhibitor.

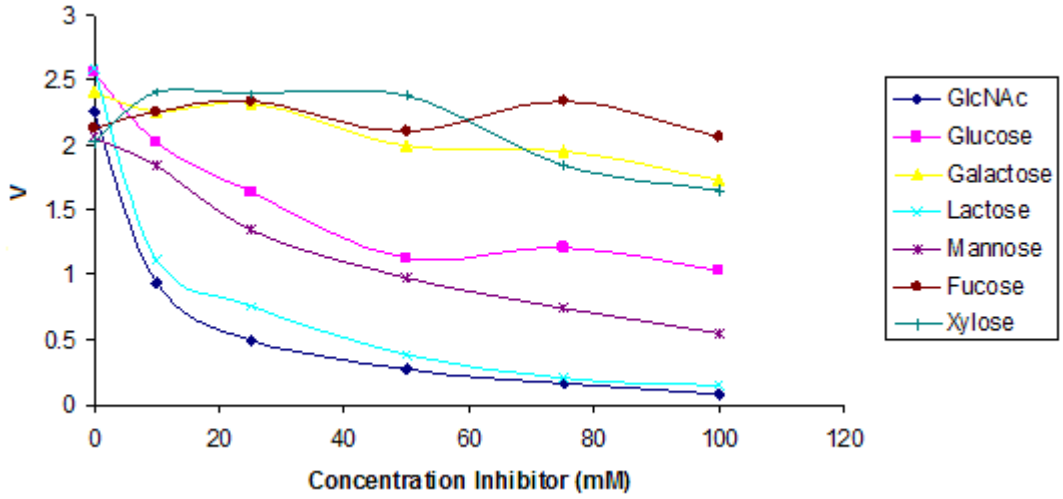
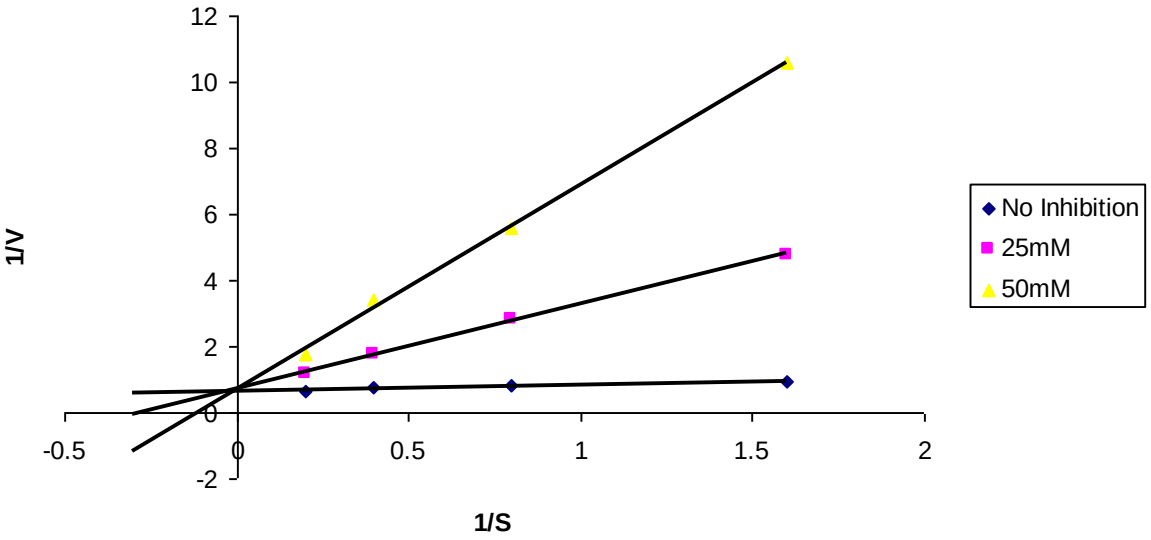


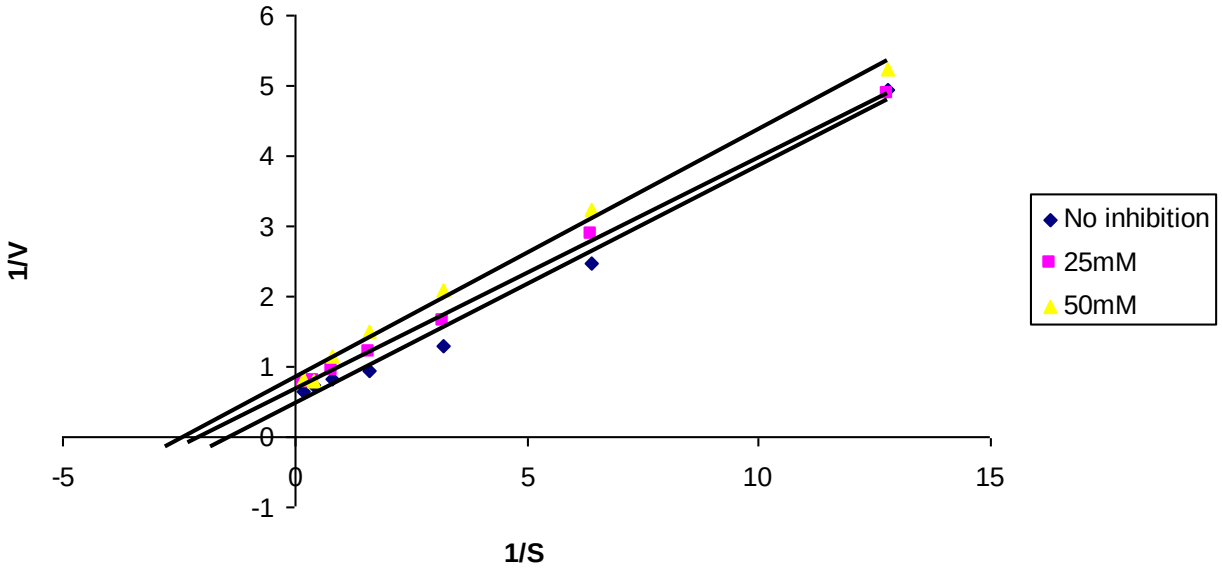
Fig 3.31: The inhibitory effect of free sugars on the rate of rChbL catalytic activity on PNP-GlcNAc. The Chb activity assay was performed as described in Section 2.26 with 1 mM PNP-GlcNAc and increasing concentrations of GlcNAc, glucose, galactose, lactose, mannose, fucose and xylose. The catalytic rate, V (mM PNP released/min/mL enzyme), of the reaction was plotted against substrate concentration (mM). Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

High levels (100 mM) of N-acetylglucosamine (GlcNAc) demonstrated almost complete inhibition of rChbL activity on PNP-GlcNAc (see Figure 3.31). There was also inhibition observed when increasing concentrations of glucose were added. As N-acetylglucosamine contains the basic glucose structure this was also to be expected, as it was similar to its natural substrate. There was relatively little inhibition observed when galactose was added to inhibit the reaction, however, when high concentrations of lactose were added there was a surprisingly high level of inhibition noted (up to 80%). Not only is its structure very different from that of GlcNAc, but it is a dimer of galactose, which did not cause any inhibition. When high concentrations of mannose were added a high level of inhibition was also observed (up to 75%). Mannose has some structural similarity to GlcNAc. There was very little inhibition observed on rChbL enzymatic activity on PNP-GlcNAc when high concentrations of fucose and xylose were used. The structures of these sugars do not resemble that of GlcNAc so this was to be expected.

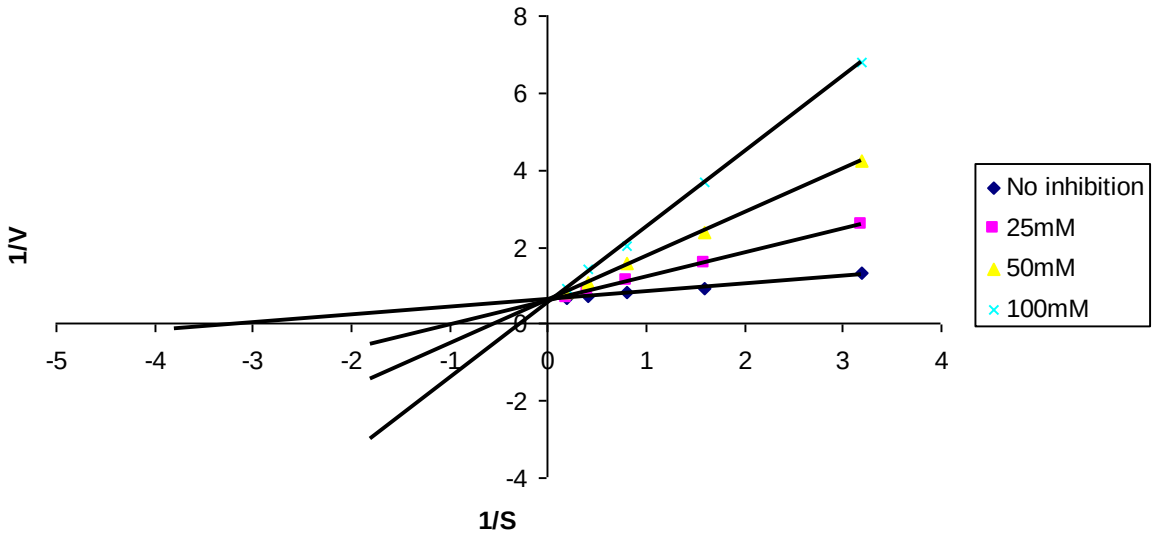
a)


12
10
8
6
4
2
0
-2
-0.5
0
0.5
1
1.5
2
1/V
1/S
No Inhibition
25mM
50mM

b)


6
5
4
3
2
1
0
-1
-5
0
5
10
15
1/V
1/S
No inhibition
25mM
50mM

c)



d)

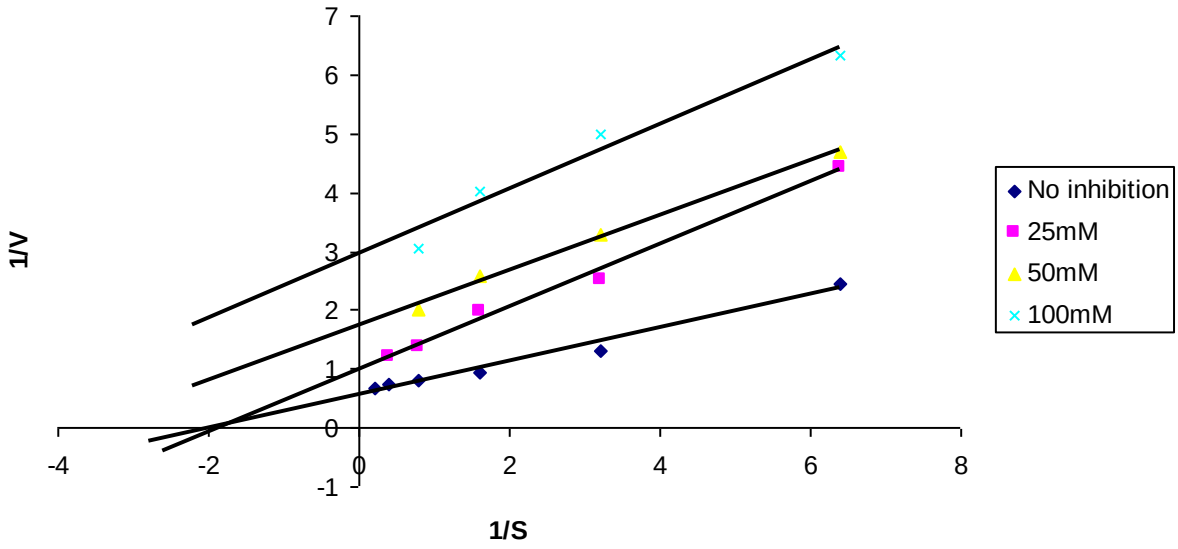


Fig 3.32: Lineweaver-Burk plots describing the inhibition occurring with a) GlcNAc, b) glucose, c) mannose and d) lactose. The Chb activity assay was performed as described in Section 2.26. Increasing concentrations of PNP-GlcNAc were incubated with a constant concentration of enzyme and either 25 mM, 50 mM or 100 mM inhibiting sugar. The reciprocal rate of the reaction was plotted against the reciprocal of the substrate concentration. Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

Table 3.5 Kinetic values describing the inhibition occurring between the rChbL - PNP-GlcNAc reaction and GlcNAc, glucose, mannose and lactose.

	1/V intercept	Slope
GlcNAc (No sugar)	0.646	0.187
GlcNAc (25 mM)	0.713	2.550
GlcNAc (50 mM)	0.705	6.190
Glucose (No sugar)	0.464	0.338
Glucose (25 mM)	0.673	0.329
Glucose (50 mM)	0.884	0.352
Mannose (No sugar)	0.634	0.206
Mannose (25 mM)	0.573	0.628
Mannose (50 mM)	0.614	1.134
Mannose (100 mM)	0.542	1.956
Lactose (No sugar)	0.477	0.300
Lactose (25 mM)	0.993	0.529
Lactose (50 mM)	1.747	0.466
Lactose (100 mM)	2.956	0.550

The values observed in Table 3.5 were calculated using the plots described in Figure 3.32. Addition of an increasing concentration of GlcNAc increased the slope of the line but the 1/V intercept value remained similar over both concentrations. This trend indicated that GlcNAc was inhibiting the reaction in a competitive manner. Increasing the concentration of glucose resulted in the slope of the line remaining unchanged; however, the 1/V intercept increased signifying uncompetitive inhibition. Addition of an increasing concentration of mannose resulted in an increased slope of the line and a mainly unchanged 1/V intercept indicating that it was a competitive inhibitor of the reaction. Finally, when an increased concentration of lactose was added to the reaction the slope of the line and the 1/V intercept increased. This indicated that lactose was inhibiting the reaction in a non-competitive manner.

3.9 Immobilisation of rChbL to cyanogen bromide-activated sepharose

Immobilised enzymes can be defined as "enzymes physically confined or localised in a certain defined region of space with retention of their catalytic activities, and which can be used repeatedly and continuously" (Katchalski-Ktazir, 1993). Advantages of immobilised enzymes over those in solution include increased stability of the enzyme, the ability to re-use the enzyme until its catalytic activity has been exhausted, the capacity to use the enzyme in large reactors and the absence of enzyme in the final product solution. One major disadvantage may be a loss of catalytic activity due to immobilisation (Katchalski-Ktazir, 1993).

Insoluble, inert matrices are generally used for enzyme immobilisation. In this study the rChbL enzyme was immobilised on the commercially available cyanogen bromide-activated sepharose (Sigma). This matrix is a porous material comprised of cross-linked agarose beads. Cyanogen bromide reacts with the agarose hydroxyl group to form cyanate esters or imidocarbonates. Cyanate esters are highly reactive with the primary amines found on the surface of proteins forming an isourea bond thus coupling the enzyme to the matrix. These covalent bonds are irreversible and therefore the matrix was not reusable once the enzymes' catalytic activity had been exhausted.

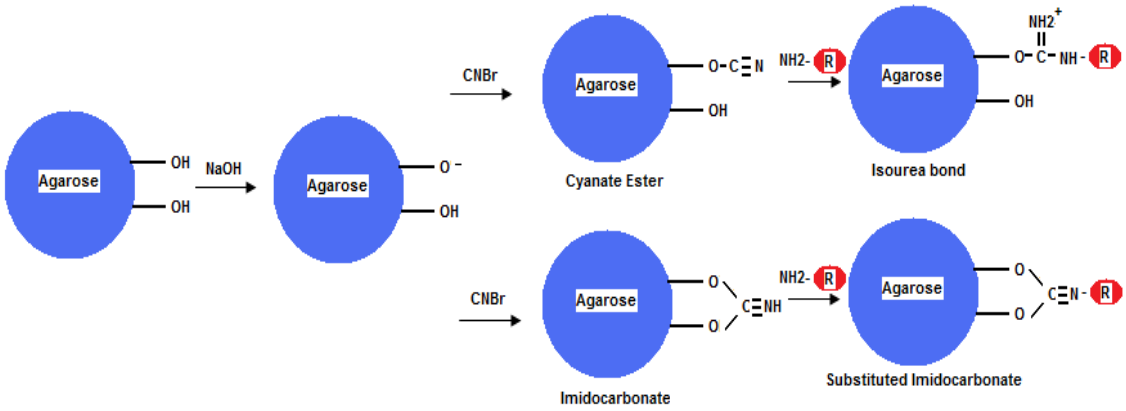


Fig 3.33: Reaction schematic describing the formation of the covalent bond during enzyme immobilisation to cyanogen bromide activated sepharose. The formation of the bond between the cyanogen bromide-activated sepharose and the primary amine of the protein under basic conditions is described in the schematic. The cyanate ester is highly reactive towards the primary amine of the protein compared with the mild reactivity of the imidocarbonate.

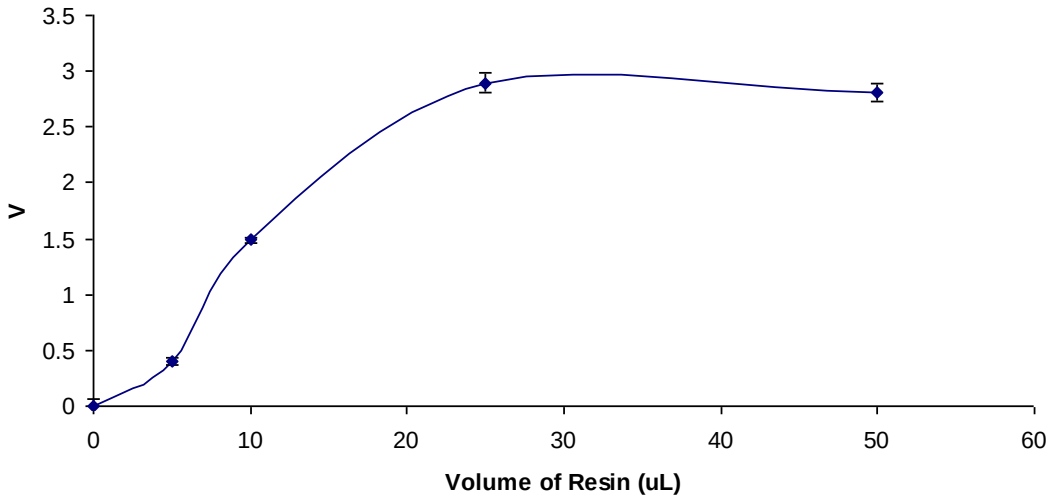
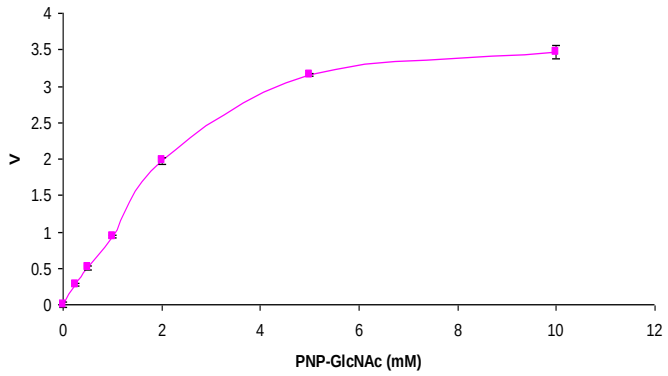


Fig 3.34: Determination of the volume of rChbL resin that exhibited the highest rate (V) of activity. The Chb activity assay was performed as described in Section 2.26. Increasing volumes of resin at a concentration of 10 µg/mL (enzyme) were added to a constant concentration of PNP-GlcNAc. The rate, V (mM PNP released/min/mL enzyme), of the reaction was plotted against the resin volume (µL). Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

The rChbL resin (1.2 mg/mL) was prepared as described in Section 2.24 and diluted to a concentration of 10 μg/mL for the purpose of Figure 3.34. The highest rate of activity was observed when 30 μL of rChbL resin was added to the PNP-GlcNAc substrate. This was in contrast with the native Chb whose activity was at approximately the same rate at 10 μL of a 10 μg/mL solution (see Figure 3.20). Three times as much rChbL resin was needed compared with the native rChbL to elicit the same reaction rate.

a)



b)

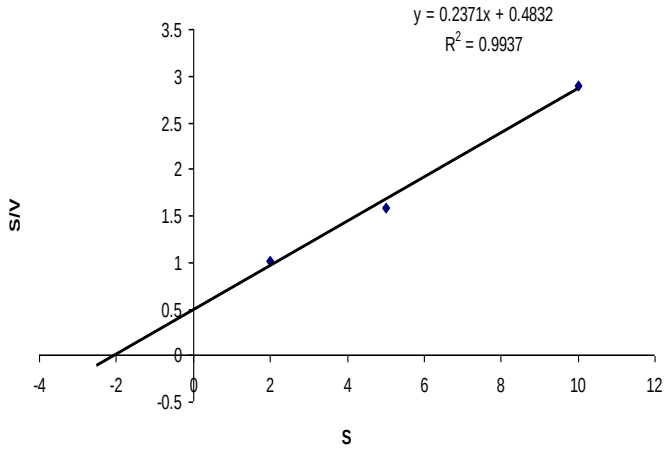


Fig 3.35: Examination of the of the immobilised rChbL kinetics using two different methods a) Michaelis-Menten curve and b) Hanes-Woolf plot. The enzyme assay was performed as described in Section 2.26 with an increasing concentration of PNP-GlcNAc (mM) to compare the rates of reaction between the immobilized rChbL and native rChbL. The rate, V (mM PNP released/min/mL enzyme), of the reaction was plotted against substrate concentration (mM). Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

Table 3.6 Kinetic values for the immobilised rChbL as determined by a Michaelis-Menten curve and Hanes-Woolf plot

Rate Constants	Michaelis-Menten	Hanes-Woolf
V_{max} (μM)	3463	2070
K_m (mM)	1.7	0.49
k_{cat} (s^{-1})	$1.92x10^5$	$1.15x10^5$
k_{cat}/K_m ($mM^{-1}.s^{-1}$)	$1.13x10^5$	$2.35x10^5$
SD (±)	0.00444	0.00444

Table 3.6 highlights the fact that the V_{max} of the immobilised enzyme was slightly higher than that of the native rChbL enzyme. The immobilised rChbL enzyme reached a higher rate than the native rChbL enzyme. There is a difference between both K_m values so as Hanes-Woolf plots are deemed to be the most accurate only these values will be discussed. The K_m was found to be similar to that of the native enzyme. Immobilisation of the enzyme to resin did not appear to change the affinity of the enzyme for its PNP-GlcNAc substrate. The catalytic efficiency of both the native rChbL and immobilised rChbL are similar, each are in the range of 10^5.

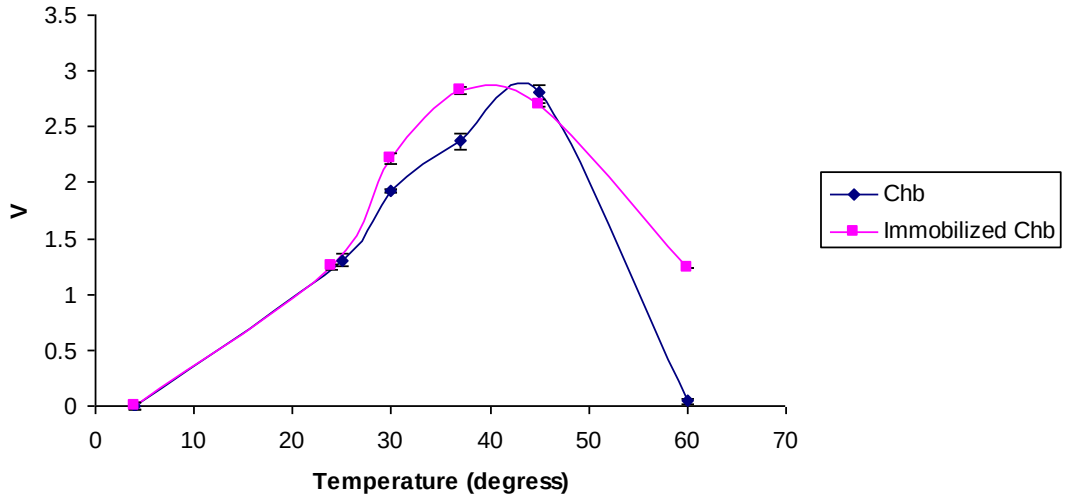


Fig 3.36: Comparison of the temperature stability between the immobilised rChbL enzyme and the native rChbL enzyme. Samples of the immobilised and native enzyme were incubated at various temperatures before cooling. The enzyme assay was then performed as described in Section 2.26. The rate, V (mM PNP released/min/mL enzyme), of the reaction was plotted against increasing temperature. Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

As identified in Figure 3.36 the immobilised rChbL enzyme was slightly more stable at the higher temperature of 60^{o}C than the native Chb enzyme. The native rChbL enzyme was inactivated after incubation at 60^{o}C; however the immobilised rChbL retained almost half of its activity. Immobilisation of the rChbL enzyme to resin slightly increased its stability at higher temperatures.

3.10 Discussion

A chitobiase enzyme (rChbL) from *P. luminescens* TT01 was successfully cloned, expressed, purified and characterised for the first time. The pQE60_Chb+ vector was transformed into two strains of *E. coli*, JM109 and KRX, and of the two strains, the JM109 cells stored the DNA safely for the longest period of time. The JM109 strain is routinely used for stable plasmid DNA storage as it contains recA1 and endA1 mutations which minimise recombination events and promote high quality plasmid DNA preparations. At the end of the study these cells were still producing active and stable enzyme.

KRX cells are a commonly used expression strain. They contain a somewhat defective restriction system in the form of hsd and e14 mutations and there is an absence of a common endonuclease (endA-). The KRX cells were discovered to be unsuitable for stable storage of the vector. After three months it was observed that the cells produced a much lower yield of enzyme that was also inactive. The pQE60_Chb+ vector had to be taken from plasmid stocks and re-transformed into fresh KRX cells every three months. This reduction in protein yield after storage using this strain was also reported by Vethanayagam and Flower (2005). They found that protein production decreased over time when utilizing the T7 promoter/polymerase system. DNA sequencing verified that this problem did not arise due to mutations in the plasmid containing the gene of interest but was discovered to be due to the low levels of active T7 polymerase present in the cell. The T7 polymerase lost its function due to mutations acquired on the bacterial chromosomal DNA. Much lower levels of protein expression were observed due to the lack of functional T7 promoter present in the cell (Vethanayagam and Flower, 2005). The decrease in rChbL protein expression in this study may also have been due to a lack of fully functioning polymerase. The *E. coli* polymerase may have picked up mutations during long-term storage. Lack of protein expression also signifies that there may have been a problem with the promoter sequence. Problems with promoter deletion from pQE vectors by KRX cells have been seen before within the group (unpublished results). To overcome this problem in future studies an alternative *E. coli* strain such as BL21 could be used. The expression levels of both strains are comparable, however

BL21 cells have leaky expression of the protein of interest and therefore induction cannot be as stringently controlled as with KRX cells. It would be advantageous in future studies to use BL21 cells to ensure a constant supply of protein with no risk of a reduction in protein expression over time or the production of inactive protein.

The major issue encountered while exploring the expression of the rChbL protein was determining the location to which it was expressed. During initial studies it was observed that the enzyme was found as an insoluble protein in the cell lysate. Proteins can be insoluble due to misfolding of the protein or to the incorrect formation of disulphide bonds causing the protein to aggregate in an insoluble mass. Isolation and purification of insoluble proteins is difficult, time consuming and the quality of the protein can be compromised. Insoluble proteins can be purified in the presence of denaturents such as guanadine hydrochloride or urea followed by renaturation of the protein. Denaturation and renaturation of proteins is obviously not ideal as misfolding can occur which can impact on protein function. Ensuing studies investigated the compartmental location of any soluble protein. Using the water lysis method rChbL was identified in the periplasmic fraction. Following further investigation it was discovered that the majority of the protein was secreted from the cell into the media. This result was not unexpected as the enzyme contained a signal peptide and in nature this enzyme would be employed in the breakdown of chitin outside the cell. Signal peptides are generally found on the N-terminus of proteins, where they guide the newly synthesized immature proteins to the periplasm where the signal peptide is cleaved to generate the mature form of the protein (Reece, 2004). The periplasm can also facilitate correct protein folding with the help of molecular chaperones and proteases (Miot and Betton, 2004). N-terminal signal sequences and are very common among chitin-active proteins, which are frequently extracellular enzymes (Tews *et al*., 1996).

While investigating the hydrolysis of chitobiose by a chitobiase from *S. marcescens* 2170, Toratani *et al*., (2008) discovered that the chitobiase was present not only in the culture medium but also in the periplasm when the bacteria was cultured in media containing chitin, similar to the findings of this study. Tews *et al*., (1996) purified a recombinant *S. marcescens* Chb homologue from the periplasm of *E. coli*. However, only 50% of the total protein was recovered. Whether the culture media was tested for

the presence of Chb protein was not mentioned in this paper. If protein was present in the culture media the isolation procedure would have been more straightforward and relatively little protein would have been lost in comparison. Drouillard *et al*., (1997) also purified a Chb protein from the periplasmic fraction of *E. coli* again with no mention of any secreted protein. The Chb protein purified by Tews *et al*., (1996) was sequenced and compared with the original sequence. The purified protein was 27 amino acids shorter than the original sequence. They proposed that this was due to the removal of the signal peptide which was also 27 amino acids long. In the case of the rChbL protein, it is likely that the signal peptide guided the immature protein to the periplasm, where the signal peptide was removed, the protein potentially underwent some post-translation modification and was secreted from the cell.

A significant finding was that the rChbL enzyme was secreted from *E. coli* after production. Previous papers investigating the Chb protein from *S. marcescens* produced in *E. coli* contained no mention of checking the culture media for protein. Instead the protein was isolated from the periplasmic fraction. Isolation from the culture media would be less complicated and time-consuming.

Purifications of the rChbL protein from the different fractions highlighted the differences between the proteins' structure. While the protein from the culture media had a relatively strong interaction with the nickel resin (40 mM imidazole was needed for elution), the cellular protein from the soluble fraction had a very weak interaction with the nickel (only 10 mM imidazole was needed to elute the protein). It was believed that the His_6 tag was more accessible on the rChbL protein from the culture media than that isolated from the soluble cellular fraction. The His_6 tag on the rChbL protein from the soluble cellular fraction may have been inaccessible due to the tertiary structure of the protein. Again, this provided some evidence that the cellular fraction contained the immature form of the protein which had not been fully processed in the periplasm. The correctly folded, fully mature form of rChbL is thought to be generated in the periplasm through the correct formation of disulphide bonds (the *S. marcescens* Chb contained 3 disulphide bonds) facilitated by the oxidising environment and the presence of enzymes such as disulphide isomerase and disulphide oxidoreductase. This mature form was then secreted.

The catalytic cleavage of natural substrates is often hard to analyse. For the purpose of developing a quick, reliable, efficient assay for the analysis of the catalytic activity of the rChbL from *P. luminescens*, 4-Nitrophenyl N-acetyl β–D-glucosaminide was used. Most kinetic characterisation studies on glycosyl hydrolases are performed with synthetic substrates. Suginta *et al*., (2010) utilized PNP-GlcNAc as a substrate in their study on novel N-acetylglucosamindases from *Vibrio harveyi 650*, Liu *et al.*, (2011) studied the catalytic activity of an insect β-N-acetyl-D-hexosaminidase on a 4-methylumbelliferyl N-acetyl β–D-glucosaminide (4MU GlcNAc) substrate, Drouillard *et al*., (1997) studied the action of a chitobiase from *S. marcescens* also utilizing a *p*-Nitrophenyl N-acetyl-β-D-glucosaminide substrate, while Toratani *et al.*, (2008) also studied the action of chitobiases from *S. marcescens* but used a 4-methylumbelliferyl N-acetyl β–D-glucosaminide substrate. These substrates provide relatively quick results that are easy to read due to the presence of the linked molecule. However, these substrates do not entirely resemble the enzymes' natural substrate due to the presence of the artificial leaving groups (see Figure 3.37). These leaving groups are much smaller than the natural sugar leaving group and molecules such as PNP and 4MU can be negatively charged at higher pHs (Krokeide *et al*., 2007).

a) **b)**

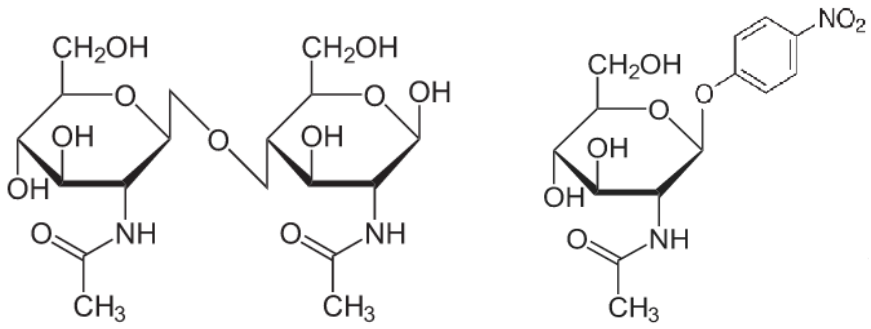


Fig 3.37: Difference between (a) the natural substrate chitobiose and the (b) synthetic substrate PNP-GlcNAc. As demonstrated in the diagram the PNP leaving group is much smaller than the second GlcNAc residue present on the chitobiose sugar.

However, investigating rChbL activity on its natural substrate (chitobiose) requires a lot more time and optimisation. Hydrolysis of the chitobiase sugar can be measured

through the use of methods such as High Performance Liquid Chromatography (HPLC) or Thin Layer Chromatography (TLC). The hydrolysis of chitobiose and other chito-oligosaccharides was investigated and reported in Chapter 6 using HPLC. The conditions for the PNP-GlcNAc assay had to be optimised to ensure the rChbL was working to its optimal potential, as enzymatic activity is sensitive to its environment. Enzymes are especially affected by the pH and temperature of their environment.

Comparison of the optimum reaction conditions of the *P. luminescens* Chb with the *S. marcescens* Chb homologue revealed similarities. The optimum temperature of the *P. luminescens* Chb was identified as 45^{o}C while a Chb homologue from *S. marcescens* was found to exhibit optimum catalytic activity at the higher temperature of 52^{o}C (Tews *et al*., 1996). Another study conducted by Drouillard *et al.,* (1997) determined the optimum temperature of a Chb from *S. marcescens* to be 40^{o}C. The optimum activity of the Chb at similar temperatures is explained by the fact that *P. luminescens* and *S. marcescens* are both mesophilic bacteria, preferring moderate temperatures. Using circular dichroic spectroscopy the T_m of the *S. marcescens* Chb was calculated to be 60^{o}C (Tews *et al*., 1996). While that technology was not available for this study, the *P. luminescens* Chb was found to be inactivated after incubation at 60^{o}C. It can be hypothesized that T_m of the *P. luminescens* Chb was 60^{o}C as the inactivation was probably due to denaturation of the enzymes' tertiary structure. Both the *P. luminescens* Chb and *S. marcescens* Chb (Tews *et al*., 1996; Drouillard *et al*., 1997) were most catalytically efficient in 50 mM sodium phosphate buffer. While the *P. luminescens* Chb was most catalytically active at pH 8.0 the *S. marcescens* Chb was found to be optimally catalytically active over a wider pH range, from pH 6.0 to pH 8.0 (Tews *et al*., 1996). Sodium phosphate buffer is commonly used in many life science applications and tends to favourably suit proteins due to its neutral pH and the presence of salt to stabilize the protein structure. Phosphate buffers are also known to have a good buffering capacity, preventing fluctuations in pH that may occur during the enzymatic reaction.

Comparison of the rates of activity between the rChbL from this study and from a commercially available Chb demonstrated the similarity of the catalytic activity of both enzymes. While the commercial Chb had a higher velocity than the rChbL, the rChbL had greater affinity for its substrate. The efficiency of both enzymes was

comparable, with the rChbL being slightly more efficient than the commercial version. This may have been due to the higher affinity for the substrate.

Table 3.7 Comparison of the rate constants of chitobiase enzymes from different sources.

Source	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m (s^{-1}.mM)	Reference
Photorhabdus luminescens	0.47	$3.1x10^5$	$6.56x10^5$	This study
Serratia marcescens	0.165	-	-	Tews (1996)
Serratia marcescens	$6.3x10^{-5}$	$8.27x10^2$	$1.3x10^4$	Prag (2000)
Serratia marcescens	56.7	$1.11x10^2$	-	Drouillard (1997)
Streptococcus pneumoniae	0.132	-	-	Clarke (1995)
Vibrio harveyi 650	0.172	0.08	$4.65x10^5$	Suginta (2010)
Ostrina furnacalis	0.107	$4.35x10^2$	$4.063x10^3$	Liu (2011)
Trichoderma harzianum	0.0806	-	-	Marco (2004)
Candida albicans	0.77	$7.4x10^2$	-	Cannon (1994)

Table 3.7 highlights the wide variety in the rates of activity obtained for Family 20 chitobiases from a variety of different sources including bacterial chitobiases, fungal chitobiases and an insect chitobiase. All of kinetic studies were performed with PNP-GlcNAc. The *P. luminescens* Chb had a slightly higher K_m than the other chitobiases from bacterial sources, signifying a lower affinity for its substrate. However, its K_{cat} and K_{cat}/K_m values were much higher indicating that the Chb from *P. luminescens* is a more efficient enzyme. The *P. luminescens* Chb also had a lower substrate affinity than that of an insect chitobiase, but again it was found to be more catalytically efficient. Finally, the Chb from *P. luminescens* had a higher substrate affinity in comparison with one fungal chitobiase (*Trichoderma harzianum*) but a lower affinity compared with another fungal chitobiase (*Candida albicans*). The efficiency of the chitobiase from *Candida albicans* was lower than that of the Chb from *P. luminescens*. Overall, while not having the highest substrate affinity of the chitobiases investigated in other studies, the Chb from *P. luminescens* was much more efficient.

The rChbL enzyme was found to be specific, failing to hydrolyse any of the substrates it encountered except its natural substrate GlcNAc. However, when inhibition studies

were performed, the enzyme was inhibited by a number of free sugars. Inhibition of catalytic activity by N-acetylglucosamine was expected, as this is a monomer of the rChbL enzymes' natural substrate. This sugar was found to competitively inhibit the enzymes catalytic activity towards the PNP-GlcNAc substrate. Presumably, the N-acetylglucosamine was bound to the active site of the rChbL, thereby blocking binding of the PNP-GlcNAc substrate. Unexpectedly, mannose was also found to competitively inhibit catalytic activity. The structure of mannose is slightly similar to that of N-acetylglucosamine, however, it has an axial –OH at C-2 rather than an equatorial N-acetyl group. As a competitive inhibitor it can also be presumed that it bound to the active site of the ChbL enzyme inhibiting its action on PNP-GlcNAc. Glucose was found to be an uncompetitive inhibitor of ChbL activity. N-acetylglucosamine is a derivative of glucose, the only structural difference being an N-acetyl group at carbon 2. Despite the similarity glucose does not bind to the active site. Instead it was found to be un-competitively inhibiting the enzymes catalytic activity through binding to the enzyme-substrate complex. The glucose sugar bound to the enzyme-substrate complex in some way to cause inhibition. Most surprisingly, lactose was found to inhibit binding. Lactose is a dimer of galactose (which does not cause any inhibition). The inhibition caused was not due to a similarity in the structures between the natural substrate and lactose. Instead it binds to a different site causing a conformational change and therefore preventing the rChbL enzyme from cleaving its substrate.

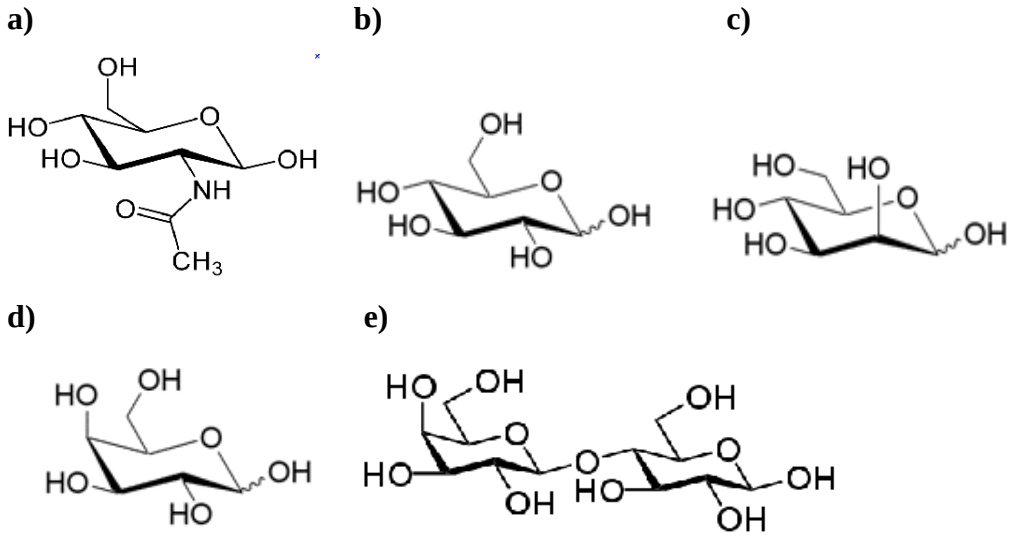


Fig 3.38: Sugar structures a) N-acetylglucosamine, b) glucose, c) mannose, d) galactose and e) lactose. The structures of the inhibiting sugars used highlighting their similarities and differences.

Enzymes have been immobilised to different surfaces for use in bioreactors with various consequences on their activities and temperature stability. The immobilisation of a β-glucosidase on a polysullfone capillary for the breakdown of oleuropein into aglycon and glucose increased the enzymes affinity for its substrate (K_m), albeit slightly decreasing the catalytic activity (Mazzei *et al.*, 2009). Nagatomo *et al*., (2009) also immobilised a β-glucosidase, however, they immobilised it onto a gelatine gel. This surface was found to promote thermostability, increasing its stability at higher temperatures as compared with the native enzyme. The immobilised enzyme was also reusable and storage was improved with no decrease in activity after one month. Chitinases have also been immobilised to various surfaces for the large-scale hydrolysis of chitin. A chitinase from *S. marcescens* was immobilised onto a hydroxypropyl methylcellulose acetate succinate polymer surface by Chen *et al*., (1994). By using an immobilised enzyme that could be removed from the product solution by precipitation they were able to move from a batch reaction to a semi-batch reaction which increased product yield 1.4 fold.

The purified rChbL from *P. luminescens* was immobilised to sepharose in this study. Immobilisation of the rChbL enzyme to resin was found to increase its velocity compared with the native rChbL. However, it was found to lower the affinity and

efficiency of the enzyme which is why a higher concentration of enzyme was needed to achieve the same catalytic rate as the native enzyme. Immobilisation on an alternative surface or using different immobilisation chemistry may improve the activity of the rChbL as there is a lysine at the binding site in *S. marcescens* that is conserved in *P. luminescens*. The rChbL enzyme may have been immobilised in such a way that it was more difficult for the substrate to access the active site. An advantage of the immobilisation was the increased stability of the rChbL at higher temperatures. While the native enzyme was catalytically inactive at 60^{o}C, the immobilised enzyme retained some of its activity. Immobilisation increased the rChbL enzymes' temperature stability which is important when using enzymes in bioreactor formats.

A chitobiase (rChbL) enzyme from *Photorhabdus luminescens* has been cloned, expressed and purified for the first time. An activity assay has been devised and optimised to determine the enzymes' catalytic activity. The rates of catalytic activity of the rChbL enzyme were found to be comparable with that of a commercial chitobiase and chitobiases from other studies. The rChbL enzyme was specific for the GlcNAc substrate. The rChbL has also been immobilised to a commercial sepharose resin with only a small loss in catalytic activity and increased temperature stability. The rChbL enzyme can be stored for over 6 months at 4^{o}C and is stable for up to three days at 37^{o}C. There properties make this rChbL enzyme a viable commercial prospect.

Chapter 4:

Site-specific mutagenesis of the *Photorhabdus luminescens* rChbL enzyme to generate a novel Carbohydrate Binding Protein (rChb CBP)

4.1 Overview

As discussed previously, Tews *et al.,* (1996) and Prag *et al.,* (2000) have reported the cloning, expression, purification and characterisation of a Chb enzyme from *S. marcescens.* It was observed that mutagenesis of an aspartate residue to an alanine at position 539 and a glutamate residue to an aspartate at position 540 abolished all catalytic activity. However, this work was taken no further. The main aim of this chapter was to generate a novel carbohydrate binding protein (rChb CBP) from rChbL through mutagenesis of D539A and E540D as described by Prag *et al.,* (2000). It was hypothesized that the catalytic activity would be destroyed from the rChbL enzyme in the same way it was abolished in the *S. marcescens* Chb resulting in a non-catalytically active carbohydrate binding protein. This chapter also examines the protein sequence and the tertiary structure of the *P. luminescens* Chb more closely. Chitobiase sequences from a variety of different bacterial species were aligned to study the conservation in the sequences of these proteins. The structure of the *P. luminescens* Chb was modeled on that of the *S. marcescens* Chb to examine its tertiary structure and the effects the mutations would have on this structure.

4.2 Chitobiase sequence analysis

BLAST (available from www.ncbi.nlm.nih.gov) was used to search for Family 20 chitobiase enzymes similar to the *S. marcescens* and *P. luminescens* Chbs. A sequence alignment was performed with chitobiase sequences displaying high levels of sequence homology obtained from the BLAST search. Chb protein sequences from twelve bacterial sources were aligned; *S. marcescens* (SM), *Serratia plymuthica* (SP), *P. luminescens* (PL), *Photorhabdus asymbiotica* (PA), *Xenorhabdus nematophila* (XN), *Enterobacter sp.* (E), *Yersinia enterocolitica* (YE), *Yersinia mollaretii* (YM), *Providencia stuartii* (PS), *Vibrio cholerae* (VC), *Vibrio sinaloensis* (VS) and *Vibrio furnissii* (VF). The sequence alignment (see Figures 4.1 and 4.2) highlighted the high level of conservation occurring in these Chb enzymes. Approximately 60% similarity was observed. The catalytically important residues (Aspartate 539 and Glutamate 540) are conserved across all bacterial species examined.

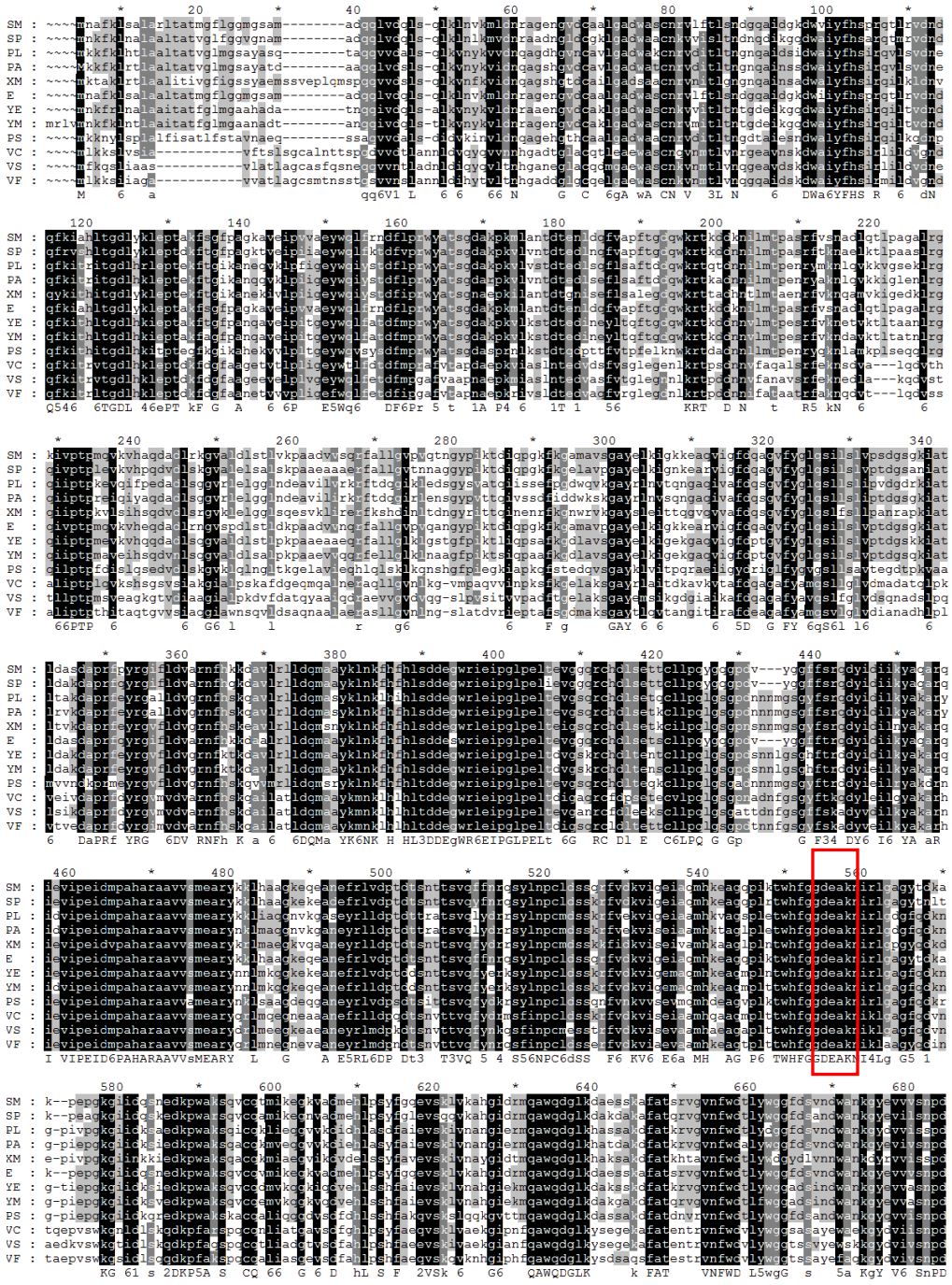


Fig 4.1: Sequence alignment of chitobiase enzymes from 12 different bacterial species. The conserved residues are highlighted in black and the conserved catalytic residues are highlighted by a red box. The sequences used are *S. marcescens* (SM), *Serratia plymuthica* (SP), *P. luminescens* (PL), *Photorhabdus asymbiotica* (PA), *Xenorhabdus nematophila* (XN), *Enterobacter sp*. (E), *Yersinia enterocolitica* (YE), *Yersinia mollaretii* (YM), *Providencia stuartii* (PS), *Vibrio cholerae* (VC), *Vibrio sinaloensis* (VS) and *Vibrio furnissii* (VF). The sequences were obtained from

Genbank. The alignment was performed using ClustalW (online tool) and Genedoc software.

4.3 Protein Modelling of the rChbL and rChb CBP

The crystal structure of the *S. marcescens* Chb was solved by Tews *et al.*, (1996) to a resolution of 1.9 Å using X-ray crystallography. The Chb was also co-crystallised with its substrate chitobiose. The structures can be found in the Protein Data Bank (available from: http://www.rcsb.org/pdb/home/home.do) labelled 1QBA and 1QBB respectively. The tertiary structure of the Chb from *P. luminescens*, however, is unknown. Homology modelling allows the unknown tertiary structure of a protein to be suggested. The ChbL tertiary structure was modelled using SWISS-MODEL (available from: http://swissmodel.expasy.org/workspace), with the *S. marcescens* Chb (1QBB) as the protein template.

a) **b)**

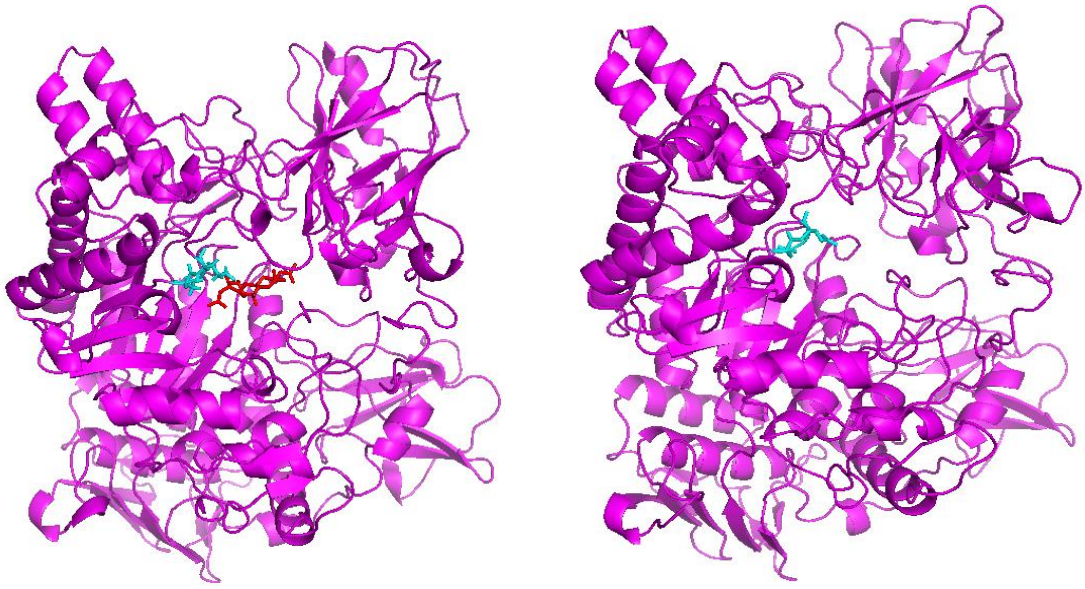


Fig 4.2: a) The *S. marcescens* Chb and b) the *P. luminescens* Chb. The *P. luminescens* Chb tertiary structure was modelled on that of the Chb from *S. marcescens* using Swiss-Model. The structures were visualised using Pymol software (available from http://www.pymol.org/). The active site residues, D539 and E540, are highlighted in blue. The chitobiose sugar is highlighted in red.

The active site of the Chb is located in a deep crater of the enzyme structure. The *P. luminescens* Chb may have a more accessible active site, since the homology model shows that there are fewer large amino acids at the site entrance.

The mutations of the catalytic residues (D539A E540D) were modelled (see Figures 4.5 and 4.6), allowing the interaction with the chitobiose sugar to be visualised. The distance between the sugar and the amino acid at position 540 (now aspartate) was increased and there was no contact observed with the chitobiose sugar. The aspartate at position 539 was mutated to the smaller amino acid alanine, leaving space in the active site that the side chain of the aspartate used to occupy.

a) **b)**

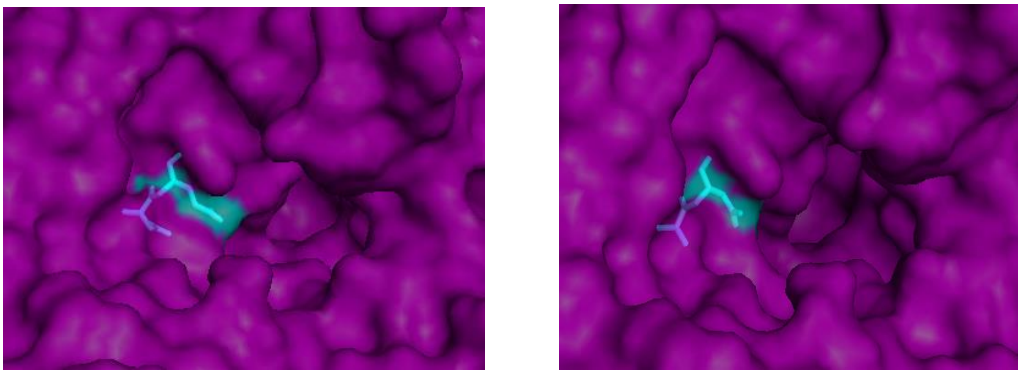


Fig 4.3: Active site of the *P. luminescens* a) ChbL and b) Chb CBP. Image (a) contains residues D539 and E540 while image (b) depicts the mutations D539A and E540D. The images were generated using Pymol software. The mutations were also simulated using Pymol. The catalytic residues (D539 and E540) are highlighted in blue.

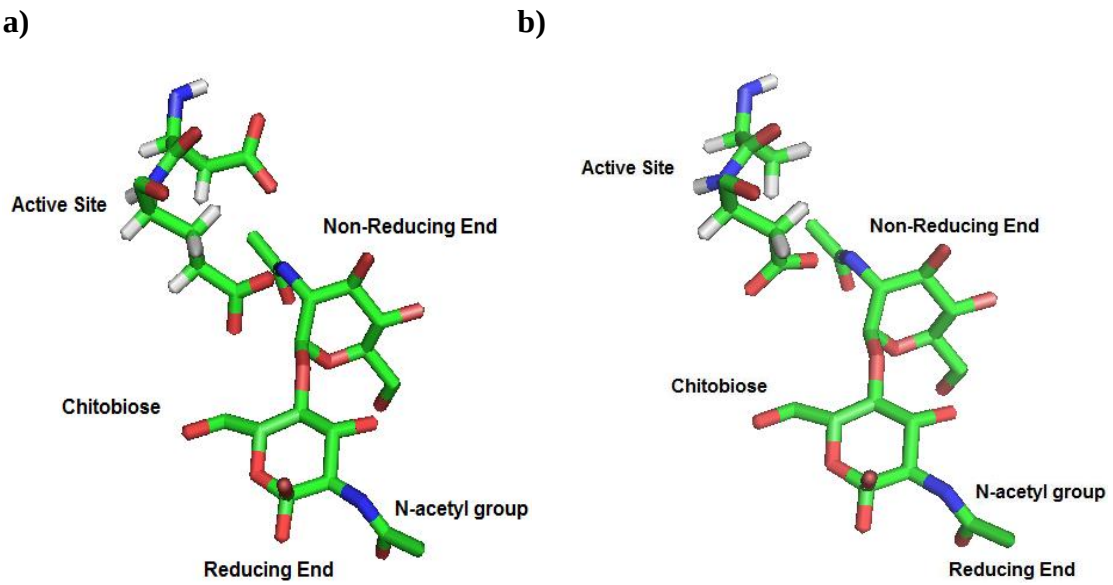


Fig 4.4: Visualisation of the interactions of a) ChbL and b) Chb CBP with chitobiose. The active site residues, a) D539 and E540 and b) D539A and E540D, are highlighted in blue. On the chitobiose sugar, the green denotes carbon, the red denotes oxygen, the white denotes hydrogen and the blue denotes nitrogen.

The Chb from *S. marcescens* has been described as a monomeric protein containing four domains; Domain I, Domain II, Domain III and Domain IV (Tews *et al.*, 1996). Domain I was assigned as the binding domain while the active site was contained in Domain III. There has been no function assigned to Domains II and IV, it can be presumed that they are purely structural. Figure 4.5 highlights the four potential domains of rChbL as modelled on those from the *S. marcescens* Chb. The conserved catalytic residues are present in Domain III. The rChbL protein sequence was analysed using the Simple Modular Architecture Research Tool (SMART) (available from http://smart.embl-heidelberg.de/) which assigned domains and annotations to ChbL based on sequence similarity to known protein domains. Domain I from ChbL was found to be similar to the cellulose-binding domain of cellulases, as was the *S. marcescens* Chb. Domain I of ChbL was therefore also assigned as a carbohydrate-binding domain.

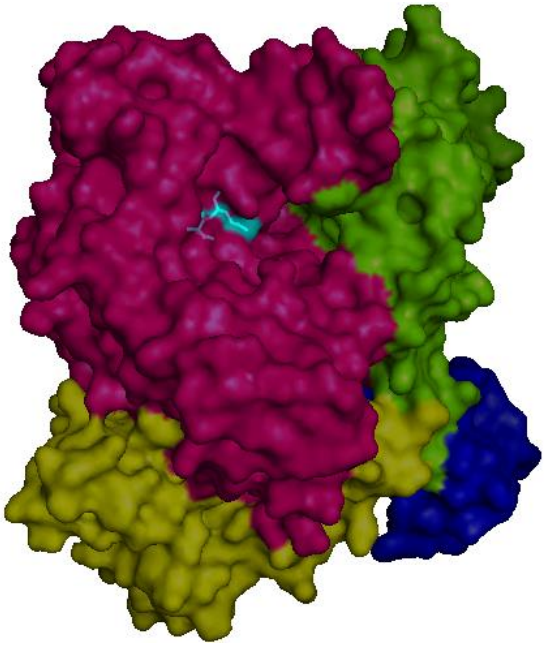


Fig 4.5: The four protein domains of ChbL. Domain I is highlighted in green, Domain II is highlighted in yellow, Domain III is highlighted in pink and Domain IV is highlighted in blue. The structures were visualised using Pymol software (available from http://www.pymol.org/).

4.4 Cloning of the rChb CBP

The immature Chb+ protein (rChbL) was used as the template for the construction of the new recombinant carbohydrate binding protein rChb CBP. The pQE60_Chb+ vector was used as the template to construct the new pQE60_Chb_CBP vector, inclusive of the signal peptide. The site-specific mutations (D539A and E540D) were introduced into open reading frames of the *chb* gene by whole vector PCR amplification using complementary phosphorylated primers carrying the desired mutations. Both mutations were introduced on the same forward primer (chb+cbp_F). The primers used were chb+cbp_F and chb+cbp_R (see Table 2.1). The pQE60_Chb_CBP vector was transformed into *E. coli* JM109 for stable plasmid (DNA) storage and *E. coli* KRX for recombinant protein expression.

gg gga **gct gac** gcg aag aa

cca ttg gag act tgg cat ttt ggg gga **gat gaa** gcg aag aat atc cgt ttg

gga **gac** gga ttc cag gat aaa aat ggt cct atc gta ccg ggt aag ggc

Fig 4.6: The forward primer (chb+cbp_F) used to introduce site-specific mutations to the *chbl* gene in the construction of the pQE60_Chb_CBP vector. The gene sequence is shown in black, the forward primer sequence is shown in red and the site-specific mutations are red, underlined and bold.

4.5 Expression and purification of the rChb CBP

The culture conditions as optimised in Section 3.6 for the expression of the rChbL enzyme were also employed here for the expression of the rChb CBP. The rChb CBP was expressed in *E. coli* KRX, in TB, at 30^{o}C and induced with 100 µM IPTG. The protein was harvested five hours post-induction. The rChb CBP was collected from the culture media (see Section 2.16) and purified via IMAC (see Section 2.19.1).

a) **b)**

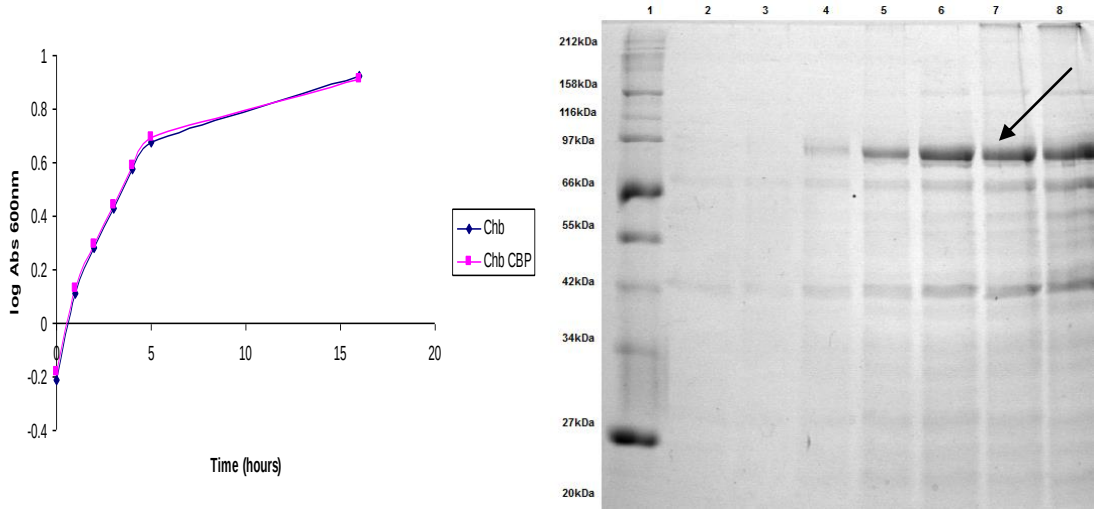


Fig 4.7: a) Comparison between growth rates of rChbL and rChb CBP cultures b) Time-course analysis of culture media samples from rChb CBP expression cultures. Recombinant protein expression was induced at OD_{600nm} = 0.5. 5 mL samples were taken from the culture media every hour for five hours, following induction and an overnight sample was taken 16 hours post-induction (see Section 2.15). The samples were then analysed by 12.5% SDS-PAGE. Lane 1, Protein ladder; Lane 2, T=0; Lane 3, T=1 hour; Lane 4, T=2 hours; Lane 5, T=3 hours; Lane 6, T=4 hours; Lane 7, T=5 hours; Lane 8, Overnight (T=16 hours).

The growth rate of the rChb CBP culture did not vary from that of the rChbL culture. The observed cell density of both rChbL and rChb CBP cultures was similar at all time-points, suggesting that the site-specific mutations did not affect the growth rate of the cells. Expression of the rChb CBP was similar with that of the native rChbL protein. The mutations did not affect the expression of the rChb CBP to the culture media. As with rChbL, the rChb CBP began to appear in the culture media two hours post induction. Protein production was optimal five hours after induction. The rChb CBP was therefore collected five hours post-induction.

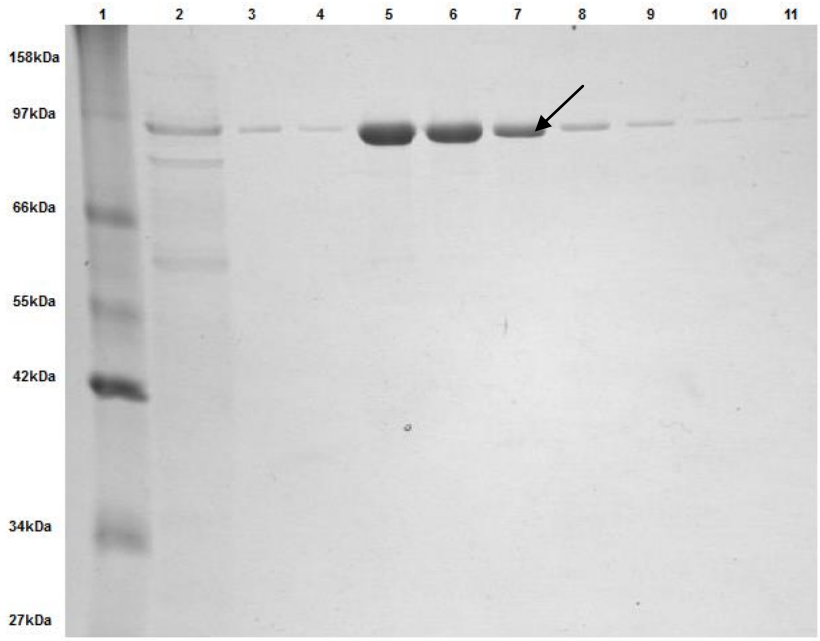


Fig 4.8: Purification of the recombinant rChb CBP from the culture media. The culture media was collected after cell harvest from a rChb CBP expression culture (see Section 2.17). Purification was carried out as described in Section 2.19. The protein sample was run over an IMAC column, washed with 40 mM imidazole and eluted with 250 mM imidazole. Fractions were run on 12.5% SDS gels. Lane 1, Protein ladder; Lane 2, Flow through; Lane 3, Wash 1 (40 mM); Lane 4, Wash 2 (40 mM); Lane 5, Elution 1; Lane 6, Elution 3; Lane 7, Elution 5; Lane 8, Elution 9; Lane 9, Elution 12; Lane 10, Elution 15; Lane 11, Elution 18.

The purification profile of the rChb CBP was similar to that of the native rChbL enzyme. As in the purification of rChbL rChb CBP was observed to elute from the resin at a 40mM concentration of imidazole concentrations but respectable yields of protein were recovered, as much as 5 mg per 250 mL culture.

4.6 Characterisation of the rChb CBP

Following the successful cloning, expression and purification of rChb CBP the proteins catalytic activity was examined. Using the activity assay as described in Section 2.26 it was demonstrated that the catalytic activity was removed from the rChbL enzyme through the introduction of the two mutations, D539A and E540D (see Figure 4.9). A catalytically inactive protein had been successfully generated, rChb CBP.

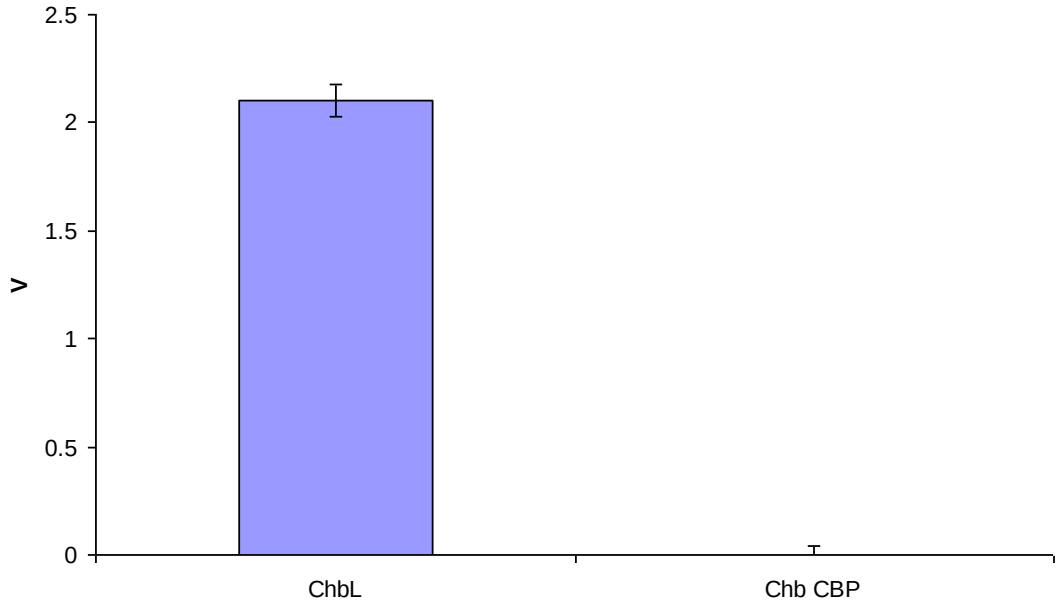


Fig 4.9: Analysis of the catalytic activity of rChbL compared with rChb CBP. The activity assay was carried out as per Section 2.26 under optimum conditions (50 mM sodium carbonate, pH 8.0) using 10 µg/mL rChbL and 10 µg/mL rChb CBP. Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

Although it was demonstrated that the rChb CBP was catalytically inactive, it remained to be determined if the rChb CBP retained its binding activity. As discovered in Chapter 3, rChbL was capable of cleaving PNP-GlcNAc, suggesting that rChb CBP may be capable of binding to GlcNAc residues on glycoproteins. The binding abilities of the rChb CBP were tested against a variety of glycoproteins using Enzyme Linked Lectin Assays (ELLAs) as described in Section 2.27. The glycoprotein target of interest was probed with the carbohydrate binding protein.

Detection of binding was facilitated by the engineered C-terminal His_6 tag on the rChb CBP with the use of a HRP labelled anti-His antibody.

Three glycoproteins were used in this study; fetuin, transferrin and thyroglobulin. Fetuin is a protein found in the blood. The fetuin utilised in this work was obtained from fetal calf serum. Bendiak *et al.*, (1989) described the heterogeneous nature of the glycan patterns on fetuin as a mixture of bi-, tri-, tetra- and penta-sialyl patterns. The most common glycan pattern is shown in Figure 4.10(a). Transferrin is a glycoprotein involved in the transport of iron. The transferrin utilised in this work was of human origin. The structure of the most common glycan pattern found on transferrin was elucidated by Spik *et al.*, (1975 and 1985) and is shown in Figure 4.10(b). Thyroglobulin is a high molecular weight thyroid glycoprotein (Spiro, 1965). The thyroglobulin used in this study was of porcine origin. The pattern of the most abundant glycan pattern on porcine thyroglobulin as elucidated by Kamerling *et al.*, (1988) is shown in Figure 4.10(c). Each of these glycoproteins had varying degrees of glycosylation allowing the binding capabilities of the rChb CBP to be thoroughly examined.

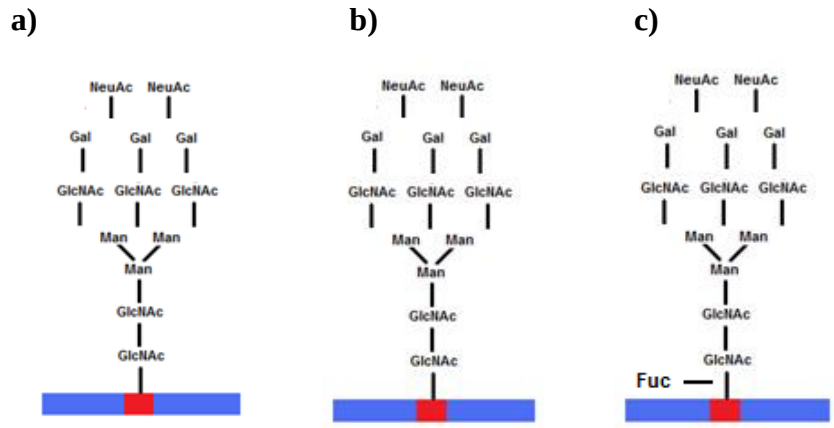


Fig 4.10: The most common glycan patterns found on a) Fetuin b) Transferrin and c) Thyroglobulin. The most common glycan patterns on the glycoproteins used in this study.

The glycoproteins described above all terminate in sialic acid. As it was hypothesised that rChb CBP would bind to structures with terminal GlcNAc and there was no access to glycoproteins naturally ending in this sugar it was therefore necessary that they were created. Glycosyl hydrolases are widely employed for the sequential

hydrolysis of glycan structures (Wongmadden and Landry, 1995). Commercially available glycosyl hydrolases were utilized for the sequential hydrolysis of the glycoproteins to expose the terminal glycan residue of interest. The glycoprotein remained untreated when examining binding to terminal sialic acid residues. When examining binding to terminal galactose residues a neuraminidase from *Clostridium perfringens* was used to cleave terminal sialic acid residues from the glycoprotein exposing underlying galactose residues of N-linked glycans. When analysing binding to terminal GlcNAc residues, a galactosidase from *Bacillus fragilis* was used to cleave terminal β-galactose residues from the glycoproteins.

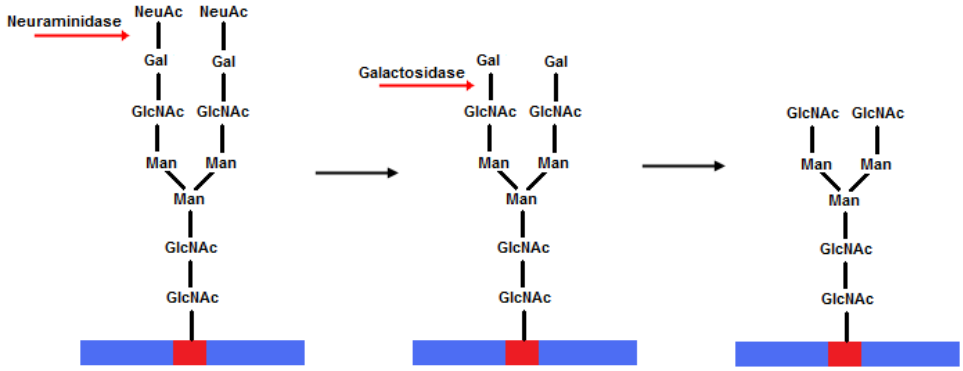
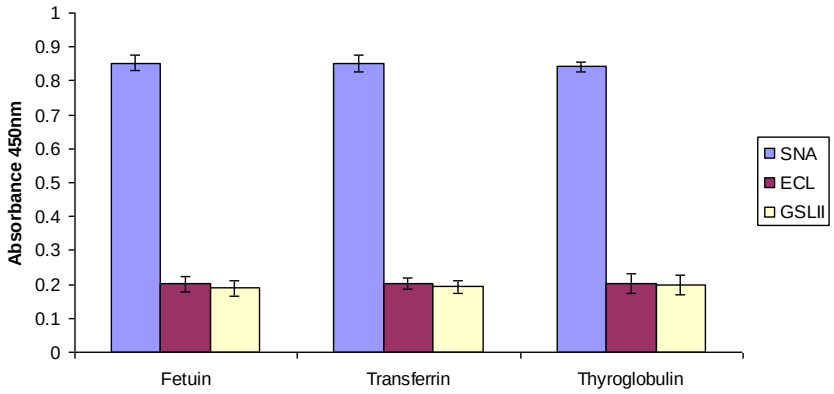


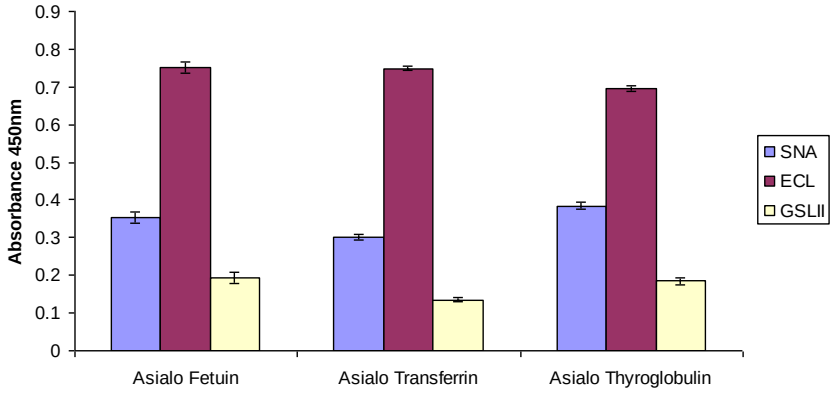
Fig 4.11: Glycan structures following sequential enzymatic hydrolysis with glycosyl hydrolases. This is the general pattern of N-linked glycosylation found on glycoproteins. The first structure is the untreated glycoprotein with terminal sialic acid. The second structure is the neuraminidase treated glycoprotein with terminal galactose (asialo glycoprotein) and the third structure is the galactosidase treated glycoprotein with terminal GlcNAc (asialo agalacto glycoprotein).

Initially, the effectiveness of the glycoprotein treatments had to be examined to ensure that the correct terminal residues were being exposed. Each treated glycoprotein was probed with a commercially available lectin, known to be specific for the terminal residue of interest. SNA was used to detect terminal sialic acid, ECL was used to detect terminal galactose and GSLII was used to detect terminal GlcNAc.

a)



b)



c)

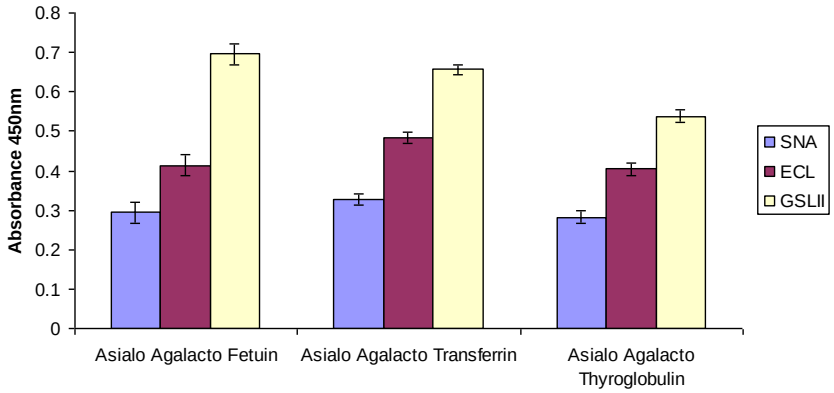


Fig 4.12: Analysis of the effectiveness of glycoprotein treatments to generate a) terminal sialic acid b) terminal galactose and c) terminal GlcNAc. The generation of various terminal sugar residues on glycoproteins were examined by probing with specific lectins, SNA to detect terminal sialic acid, ECL to detect terminal galactose and GSLII to detect terminal GlcNAc. Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

The binding of rChb CBP to terminal GlcNAc, galactose and sialic acid residues on these newly created glycoproteins was tested (see Figures 4.13, 4.14 and 4.15).

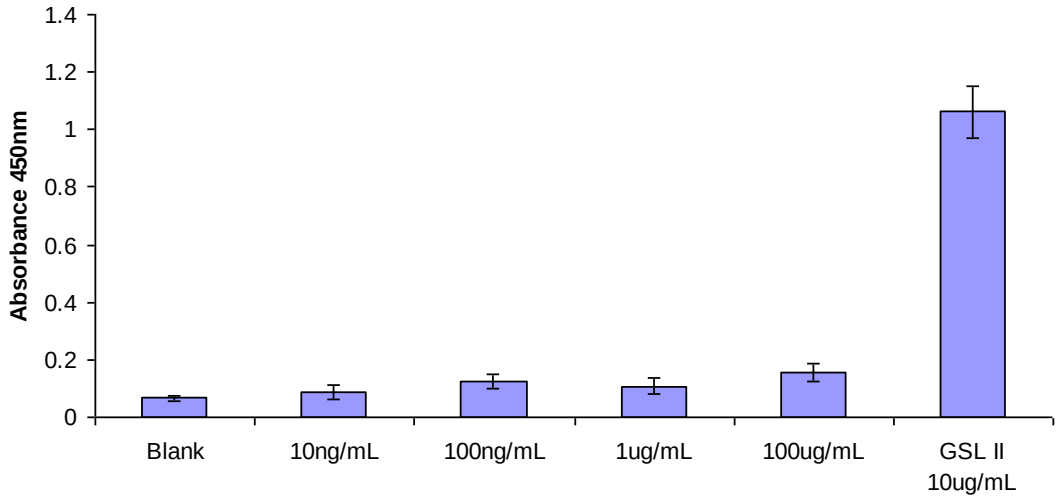


Fig 4.13: ELLA examining the binding of the rChb CBP to GlcNAc residues on asialo agalacto fetuin. An ELLA was carried out as per Section 2.27. Different concentrations of rChb CBP were used, from 10 ng/mL to 100 µg/mL. Commercial GSLII was used as a positive control and was detected using an anti-biotin antibody. The rChb CBP was detected using an anti-His antibody. Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

There was very little binding detected between the rChb CBP and GlcNAc residues in this ELLA format at all concentrations examined. The positive signal from GSLII verified the assay was working correctly. The highest absorbance obtained was 0.16 which was negligible compared with the positive control and similar to the blank. The binding of rChb CBP to asialo agalatco transferrin and thyroglobulin was subsequently examined as they contained varying degrees of glycosylation on their surface (see Figures 4.14 and 4.15). Binding of the rChb CBP to terminal GlcNAc residues on these glycoproteins was also not observed. Binding of the rChb CBP to sialic acid or galactose was also not observed (see Figure 4.15). These results signified that either the rChb CBP was not binding to the GlcNAc residues or that its binding could not be detected with the anti-His antibody.

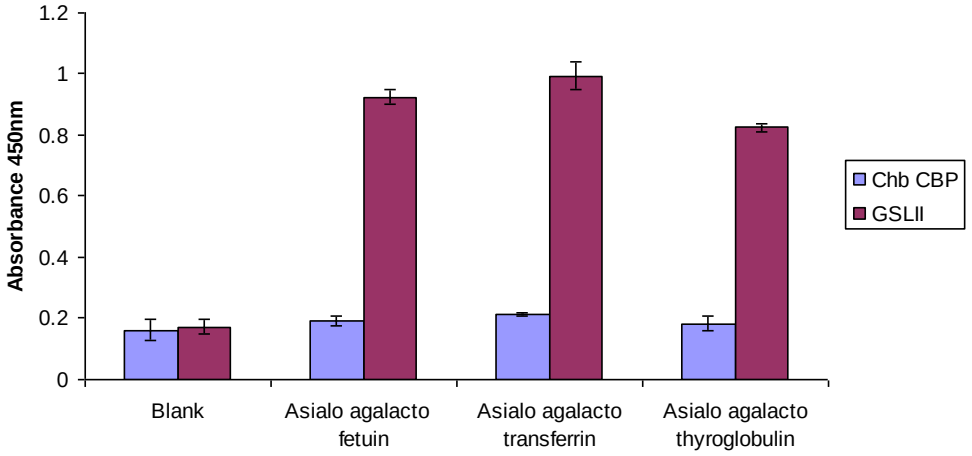


Fig 4.14: ELLA to investigate the binding of the rChb CBP to GlcNAc residues on various glycoproteins. An ELLA was carried out as per Section 2.27. rChb CBP was used at a concentration of 10 µg/mL. Commercial GSLII was used as a positive control at a concentration of 10 µg/mL and was detected using an anti-biotin antibody. The rChb CBP was detected using an anti-His antibody. Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

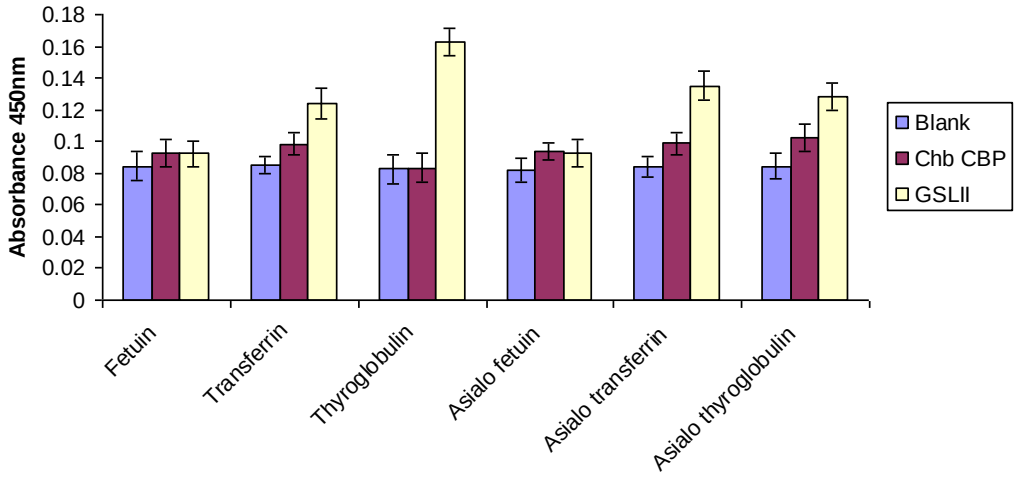


Fig 4.15: ELLA to investigate the binding of rChb CBP to sialic acid and galactose residues on various glycoproteins. An ELLA was carried out as per Section 2.27. rChb CBP and GSLII were used at a concentration of 10 µg/mL and detected with anti-His and anti-Biotin antibodies respectively. Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

4.7 Examination of the accessibility of the His_6 tag on rChb CBP

In order to rule out the issue of detection of the rChb CBP, the accessibility of the His tag was examined. Western blot analysis was carried out on the purified rChb CBP (see Section 2.22) using an anti-His antibody to detect the His_6 tagged recombinant rChb CBP. Although a protein of correct molecular weight was present after IMAC purification no His_6 tagged protein was detected on the blot. It was concluded that the His tag was not accessible to the anti-His antibody. The rChb CBP protein was also directly immobilised on an ELLA plate and probed with the anti-His antibody. The difficulty encountered when detecting the His_6 tag on the rChb CBP was clearly demonstrated in Figure 4.16. Over all the concentrations examined only a very low signal was obtained. The His_6 tag was not available to the anti-His antibody in this format. Measures were taken to improve the detectability of the rChb CBP via the His_6 tag through the addition of a linker sequence (see Section 4.8).

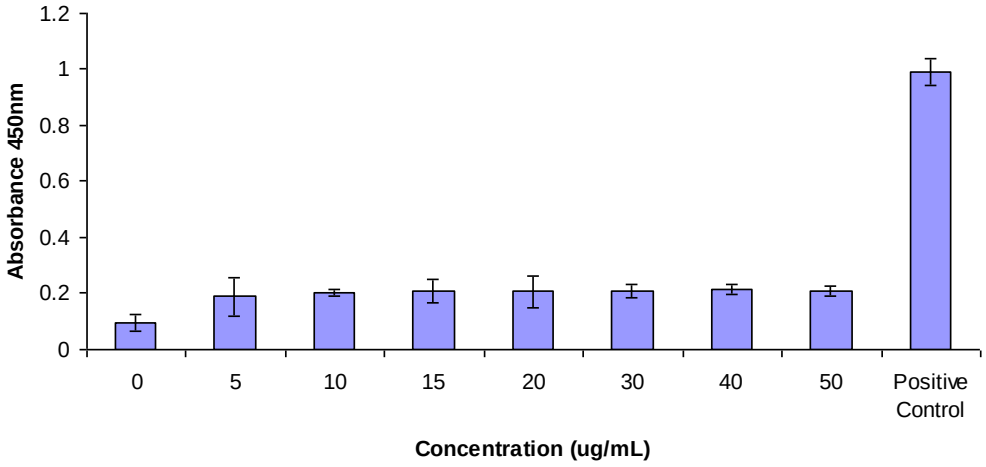


Fig 4.16: Investigation of the limit of detection of the His_6 tag on the rChb CBP. 50 µL of the rChb CBP was laid down on a 96-well plate at various concentrations in triplicate. 50µL of a His_6 tagged positive control (10µg/mL) was also laid down on the plate. The plate was blocked as per Section 2.27 and the protein was probed with an anti-His antibody. Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

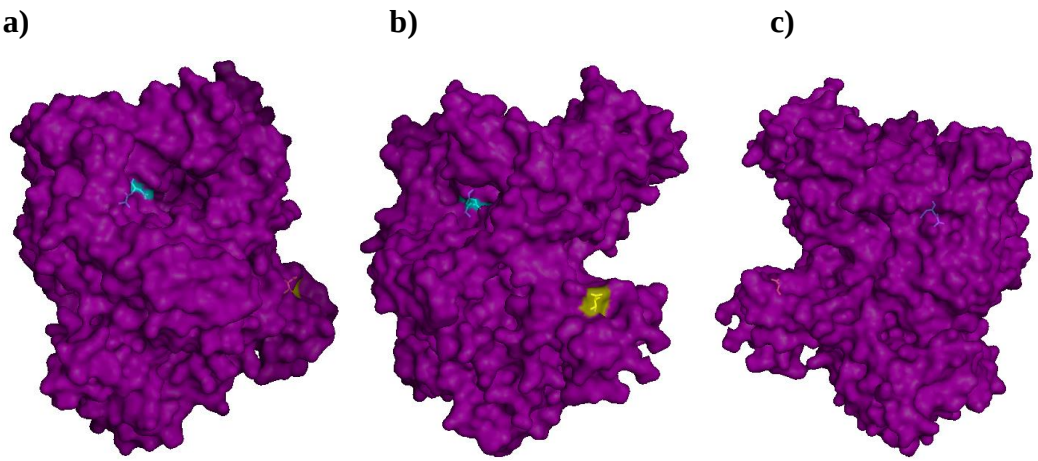


Fig 4.17: Chb CBP His_6 tag highlighted a) front b) side and c) back view. The images were generated using Pymol. The catalytic residues are highlighted in blue. The His_6 tag is highlighted in yellow.

Figure 4.17 illustrates why there are issues being encountered when trying to detect the His_6 tag in the ELLA format. The His_6 tag is present on the side of the protein. If the rChb CBP is bound to the GlcNAc residues on the surface of the plate as in Figure 4.17 (c), the His_6 tag is hidden by the large bulky structure of the protein. The anti-His antibody would have difficulty accessing the tag.

4.8 Addition of a linker sequence to the rChb CBP

Many proteins contain multiple domains, which are often joined by short peptide fragments also known as linkers (George & Heringa, 2003). Linkers are made up of different residues depending on their function. Some contain small flexible residues such as serine and glycine to allow the domains to interact with one another, while others contain rigid residues such as prolines to keep the domains in a specific conformation. In recent times multi-functional proteins have been engineered comprising domains from different proteins. These domains are attached to one another via linkers. These allow each domain to act independently, hence the multi-functional nature of the protein.

4.8.1 Cloning of the rChb CBP to include a linker peptide

In this study a linker sequence comprising four small residues (three glycine residues and one serine residue) was used. This linker was genetically engineered in between the main protein and the His_6 tag (see Figure 4.18). The objective of the linker was to allow the His_6 tag to protrude from the protein and, therefore, facilitate detection with an anti-His antibody.

Fig 4.18: Schematic of the rChb CBP DNA fragment with the linker inserted. The rChb CBP was amplified to include a linker which was comprised of three glycine residues and one serine residue.

GCT TAA CTA TTA CAT TTC CGT

AGC CCG GAC GGT AAA CGT TAC AGT CGA ATT GAT AAT GTA AAG GCA

CCT AGA **CCA CCG GGT TCT** CAT CAC CAT CAC CAT CAC TAA GCT

GGA TCT **GGT GGC GGT TCT** CAT CAC CAT CAC CAT CAC TAA GCT

TAA TTA

TAA TTA GCT GAG CTT GGA CTC CTG

Fig 4.19: The linker sequence inserted into the pQE60_Chb_CBP construct. The linker sequence was split between the forward and reverse primers. The primers used were chb+cbplnk_F and chb+cbplnk_R (see Table 2.1). The gene sequence is in black, the reverse primer is in blue, the forward primer is in red and the linker sequence is bold and underlined. Again the primers were 5' phosphorylated and whole vector amplification was used. The resulting construct was named pQE60_Chb_CBPlnk.

4.8.2 Expression and purification of rChb CBPlnk

There was very little difference observed in the growth rates of the rChb CBP and rChb CBPlnk cultures. The observed cell density of both rChb CBP and rChb CBPlnk cultures was similar at all time-points. The addition of the linker peptide did not affect the cell growth rate. Once again it was noted that the protein was expressed and exported to the culture media. As with the other recombinant proteins (rChbL and rChb CBP) the optimal time to harvest the protein from the culture media was five hours post-induction (see Figure 4.20 (b)).

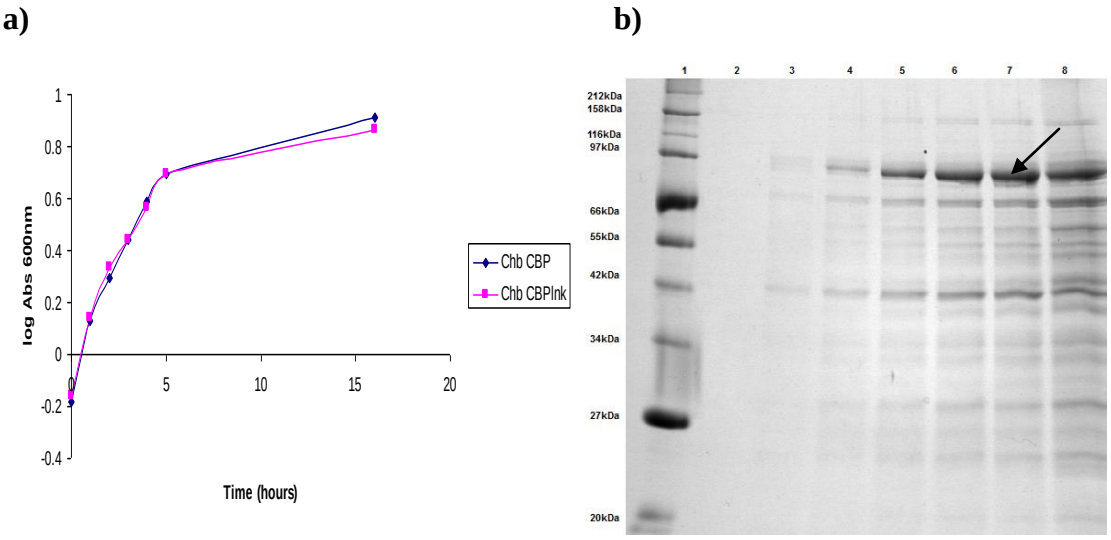


Fig 4.20: a) Comparison of the growth rates of rChb CBP and rChb CBPlnk cultures b) Time-course analysis of rChb CBPlnk protein expression from culture media samples. Recombinant protein expression was induced at OD_{600nm} = 0.5. 5 mL samples were taken from the culture media every hour for five hours, following induction and an overnight sample was taken 16 hours post induction (Section 2.16). The samples were then analysed by 12.5% SDS-PAGE. Lane 1, Protein ladder; Lane 2, T=0 hours; Lane 3, T=1 hour; Lane 4, T=2 hours; Lane 5, T=3 hours; Lane 6, T=4 hours; Lane 7, T=5 hours; Lane 8, Overnight (T=16 hours).

It was observed that the addition of the linker improved the ability of the protein to sustain imidazole wash concentrations of up to 100 mM, during IMAC purification (see Figure 4.21). This result highlighted that the addition of the linker sequence succeeded in projecting the His_6 tag further from the protein surface making it more accessible to the IMAC column and it therefore bound more tightly. The loss of less

protein in the wash fractions meant that the yields of rChb CBPlnk were slightly higher than those of rChb CBP, approximately 5.5-6.5 mg per 250 mL culture.

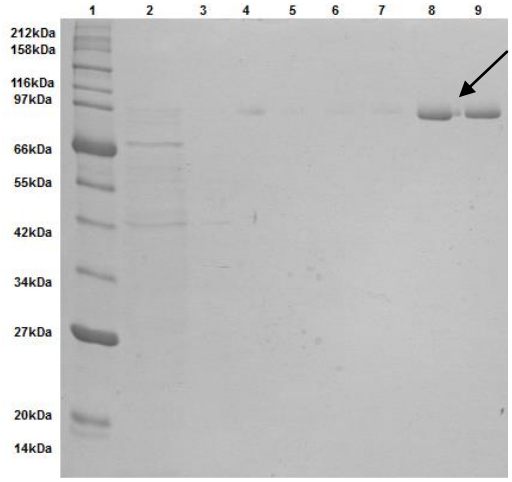


Fig 4.21: IMAC optimisation for the purification of rChb CBPlnk from the culture media. Stringency washes were performed with an increasing imidazole concentration gradient from 20 mM to 250 mM, to determine when rChb CBPlnk was eluted from the IMAC column. The fractions were then examined by 12.5% SDS-PAGE. Lane 1, Protein ladder; Lane 2, Flow through; Lane 3, 20 mM Imidazole; Lane 4, 40 mM Imidazole; Lane 5, 60 mM Imidazole; Lane 6, 80 mM Imidazole; Lane 7, 100 mM Imidazole; Lane 8, 150 mM Imidazole; Lane 9, 250 mM Imidazole.

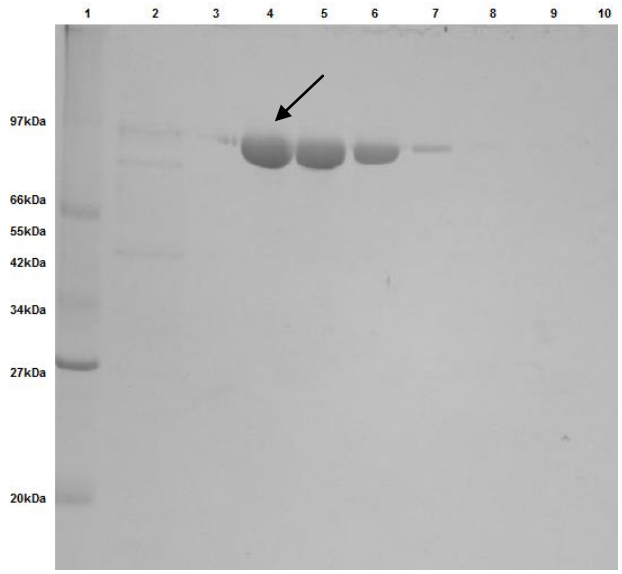


Fig 4.22: Purification of the recombinant rChb CBPlnk from the culture media. Purification was carried out as per Section 2.19. Washes were carried out with 80 mM imidazole and elution was carried out at 250mM imidazole. Lane 1, Protein ladder; Lane 2, Flow through; Lane 3, Wash; Lane 4, Elution 1; Lane 5, Elution 2; Lane 6, Elution 3; Lane 7, Elution 4; Lane 8, Elution 5; Lane 9, Elution 6; Lane 10, Elution 7.

In order to confirm that the His tag was more accessible a western blot on both rChb CBP and rChb CBPlnk was carried out. The rChb CBP and rChb CBPlnk samples were analysed under native conditions by SDS-PAGE to allow the proteins to retain their natural conformation. β-mercaptoethanol was not added to the sample buffer and the samples were not boiled prior to addition to the gel. This was to ensure that the proteins were detected as they would be in the ELLA format, if the proteins were denatured then the results would not reflect the ease of detection of the His_6 tag of each protein. It was clear from the blot result (see Figure 4.23) that the linker sequence dramatically increased the ability of the rChb CBPlnk to be detected by an anti-His antibody while the rChb CBP was not so easily detected. Similarly it was hypothesised that the linker allowed the anti-His antibody more access to the His_6 tag on rChb CBPlnk.

Fig 4.23: Western blot analysis comparing the detection of rChb CBP and rChb CBPlnk. 12.5% SDS gels were run as per Section 2.22 with two exceptions, β-mercaptoethanol was not added to the sample buffer and the samples were not boiled prior to addition to the gel. Western blotting was carried out as per Section 2.22. Lane 1; rChb CBP Lane 2; rChb CBP Lane 3; Blank Lane 4; rChb CBPlnk Lane 5; rChb CBPlnk.

4.8.3 Characterisation of the binding ability of the rChb CBPlnk

Once it had been confirmed that the addition of the linker sequence significantly improved the detectability of the rChb CBPlnk proteins His_6 tag, ELLA analysis was again undertaken. First, it was confirmed if direct detection of the rChb CBPlnk was possible (see Figure 4.24). It was observed that the limit of detection was slightly improved with the addition of the linker peptide. 10 μg/mL of rChb CBPlnk allowed the highest level of detection. However the signal was still relatively low. It was unknown whether any potential glycoprotein binding would be detectable in an ELLA.

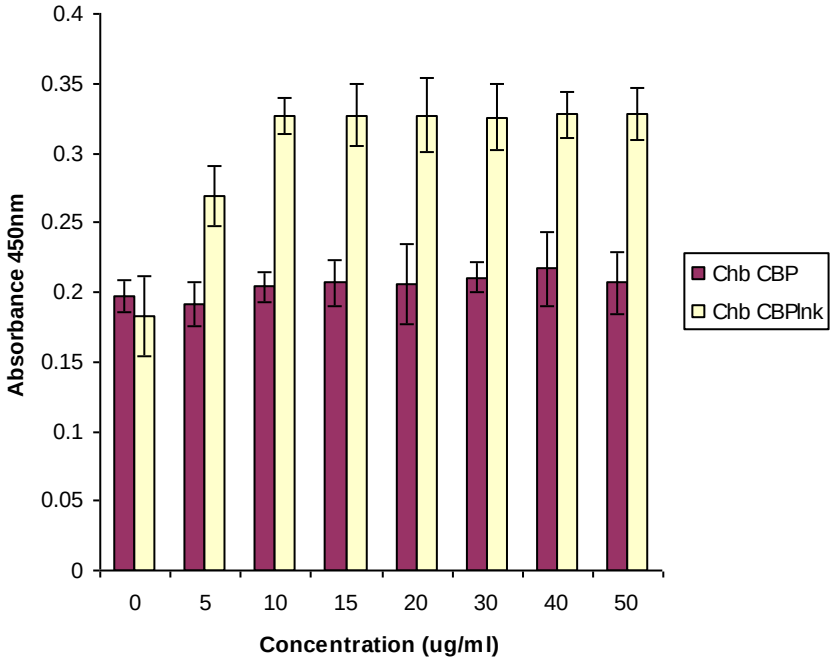


Fig 4.24: Comparison of the limit of detection of the His_6 tag between the rChb CBP and rChb CBPlnk. 50 µL of the rChb CBP and rChb CBPlnk were laid down on a 96-well plate at various concentrations (0-50 µg/mL) in triplicate. The plate was blocked as per Section 2.27 and the proteins were detected with an anti-His antibody. Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates. An unpaired T-test was used to determine if the difference in detection was significant (***$p<0.0003$).

The binding of the rChb CBPlnk to GlcNAc, galactose and sialic acid terminal residues was examined over an array of glycoproteins, as previous, to investigate whether the linker improved detection of the rChb CBP in an ELLA format (see Figure 4.25). Binding was again determined using a HRP labelled anti-His antibody.

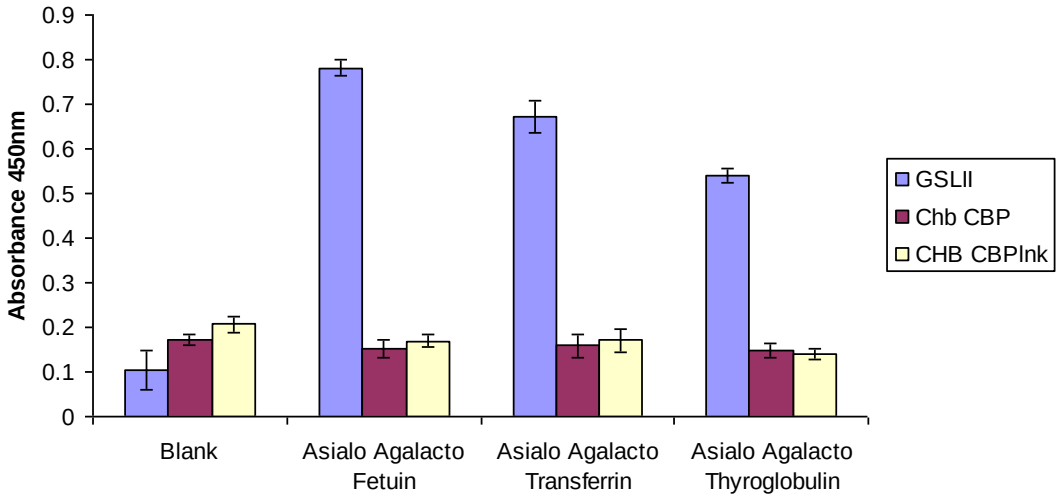


Fig 4.25: ELLA investigating the binding of rChb CBP and rChb CBPlnk to GlcNAc residues on an array of glycoproteins. An ELLA was performed as per Section 2.27. rChb CBP and rChb CBPlnk were used at a concentration of 10 µg/mL. Fetuin, transferrin and thyroglobulin were treated to expose GlcNAc residues. GSLII was detected with an anti-Biotin antibody while rChb CBP and rChb CBPlnk were detected with an anti-His antibody. Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

No binding of rChb CBP or rChb CBPlnk to GlcNAc residues was observed for any of the glycoproteins tested. The absorbance values received were as low as those of the blanks. Very little difference was observed between the signals from rChb CBP and rChb CBPlnk. This may be as a result of no rChb CBP binding to the exposed GlcNAc residues or it could potentially still be a problem of detection. Further testing was subsequently carried out against a synthetic GlcNAc substrate and bovine serum mucin. Commercial BSA-GlcNAc contains a much higher concentration of GlcNAc residues on its surface than enzymatically treated fetuin, transferrin or thyroglobulin. The GlcNAc residues are separated from the BSA by a 14-atom spacer. Bovine mucin contains only O-linked glycans, with up to 80% of the proteins weight account for by sugars. It was anticipated that higher levels of glycosylation, specifically with more exposed terminal GlcNAc, would result in more of an opportunity for binding between the rChb CBPlnk and the GlcNAc residues. However, while the commercial plant lectin GSLII bound very well to these substrates no binding of rChb CBP was detected (see Figure 4.26).

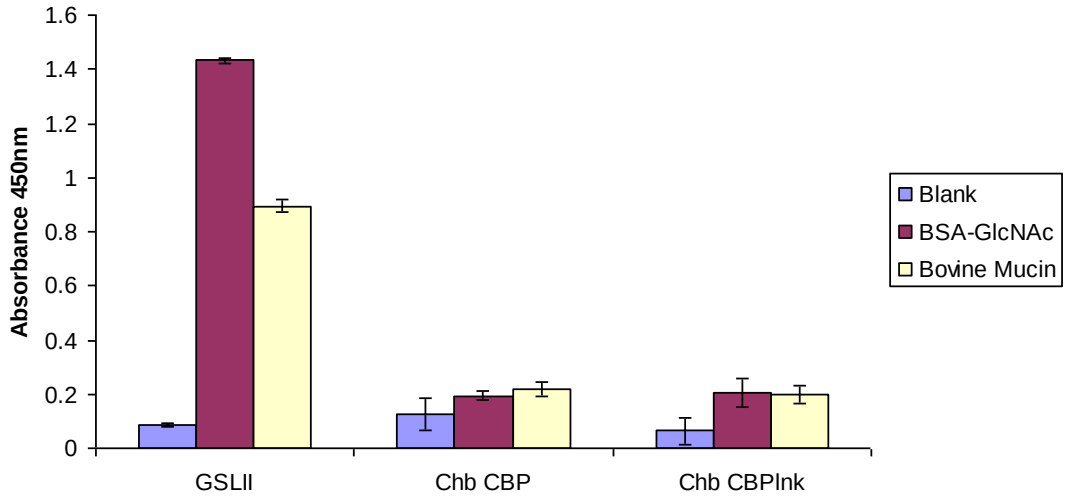


Fig 4.26: ELLA investigating the binding of rChb CBP and rChb CBPlnk to BSA-GlcNAc and Bovine Serum Mucin. An ELLA was carried out as described in Section 2.27. A commercial BSA-GlcNAc conjugate and bovine serum mucin (Sigma) were used. GSLII, rChb CBP and rChb CBPlnk were used at a concentration of 10 μg/mL. GSLII was detected using an anti-biotin antibody and rChb CBP and rChb CBPlnk were detected using an anti-His antibody. Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

4.9 Indirect ELLA methods

Due to the difficulties encountered with the detection of the rChb CBP protein, alternative indirect detection methods using the ELLA were employed to definitively conclude if the rChb CBP molecule could act as a GlcNAc binding protein. These methods included the use of a competitive ELLA and a sandwich ELLA.

The competitive ELLA was used to study the effective competition between the binding of rChb CBP and GSLII to exposed terminal GlcNAc on asialo agalacto fetuin. Asialo agalacto fetuin was immobilised on a Nunc maxisorp plate and the plate was blocked. Subsequently an increasing gradient of rChb CBP was added to the wells containing glycoprotein and incubated at 24^{o}C for 30 minutes, followed by the addition of a set concentration of GSLII. Bound GSLII was detected using an anti-biotin antibody. If rChb CBP was binding to the glycoprotein it was expected that the

binding signal for GSLII would decrease as the concentration of bound rChb CBP increased. Figure 4.27 highlights that binding of rChb CBP to GlcNAc was observed from the competitive ELLA. The higher the concentration of rChb CBP there was present, the less signal observed from GlcNAc bound GSLII. This was thought to be due to rChb CBP binding to the terminal GlcNAc residues, however it may also have been due to some interaction between the rChb CBP and GSLII.

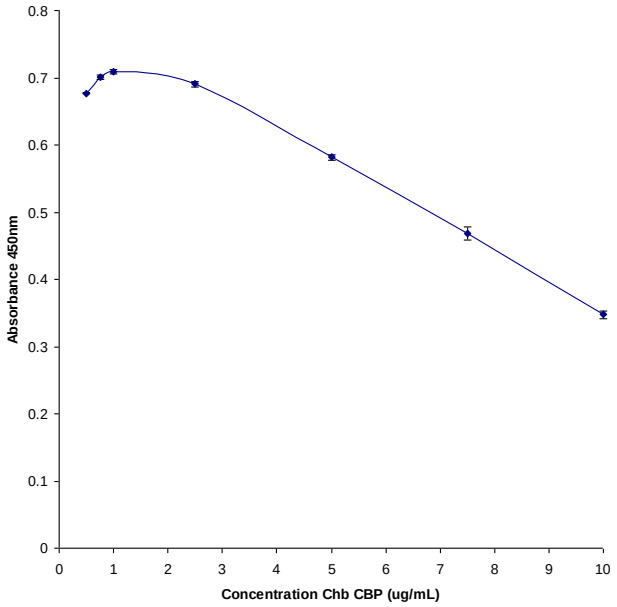


Fig 4.27: Competitive ELLA between rChb CBP and GSLII investigating binding to GlcNAc residues. Asialo agalacto fetuin was used. Increasing concentrations of rChb CBP (from 0.5 µg/mL to 10 µg/mL) were added to the plate and incubated at 24^{o}C for 30 minutes before addition of a constant concentration of GSLII (10 µg/mL). GSLII was detected using an anti-biotin antibody. Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

The sandwich assay involved the capture of asialo agalacto fetuin by plate immobilised rChbCBP, with the detection of bound glycoprotein using GSLII. Simply, rChb CBP was laid down on a Nunc 96 well plate which was subsequently blocked. Asialo agalacto fetuin was then added and allowed to bind with the rChb CBP. Finally biotinylated GSLII was added and allowed to interact with any bound fetuin creating a sandwich structure. Anti-biotin antibody was used to detect the GSLII and a signal was expected to only be obtained if both the rChb CBP/WGA and

GSLII bound to the asialo agalacto fetuin. Figure 4.28 demonstrates that a signal was obtained for the rChb CBP–GSLII sandwich but not for the WGA-GSLII sandwich. WGA is a plant lectin characterised as capable of binding to GlcNAc and chitobiose. It was expected that this would act as a positive control for the assay however no signal was obtained for this sandwich.

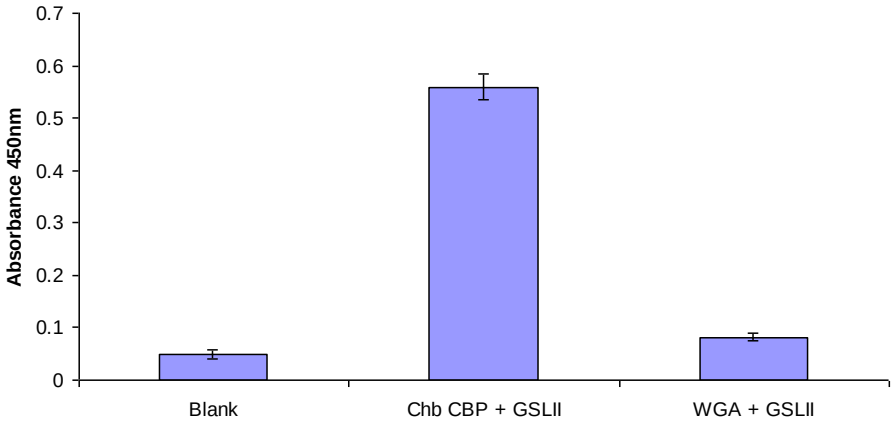


Fig 4.28: Sandwich ELLA between rChb CBP and GSLII investigating binding to GlcNAc residues. rChb CBP was laid down on the plate which was then blocked as per Section 2.27. Asialo agalacto fetuin was then added and incubated for 60 minutes at 24^{o}C. GSLII was added and also incubated at 24^{o}C for 60 minutes. Finally an anti-biotin antibody was added and the ELLA was completed as per Section 2.27. Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

As a result of the observations made from the experimental results shown in Figures 4.27 and 4.28 further investigation was undertaken to determine whether the results obtained were due to an interaction between the rChb CBP and the glycoprotein or the rChb CBP and GSLII. In order to do this the sandwich assay was repeated with the inclusion of free GlcNAc during the incubation of the asialo agalacto fetuin. It was hypothesised that the addition of free sugar would inhibit the binding of rChb CBP to the glycoprotein resulting in a decrease in the amount of bound GSLII. If there was no observed decrease in bound GSLII signals between sugar replete and deplete samples then it could be concluded that rChb CBP was interacting directly with the GSLII.

The addition of increasing concentrations of free GlcNAc (sugar) did not result in less bound GSLII. Only a slight loss of signal was observed at the high end of the GlcNAc concentrations (see Figure 4.29). As stated previously this observation implied that there was an interaction occurring between both carbohydrate-binding proteins rather than between the carbohydrate-binding proteins and the glycoprotein.

.

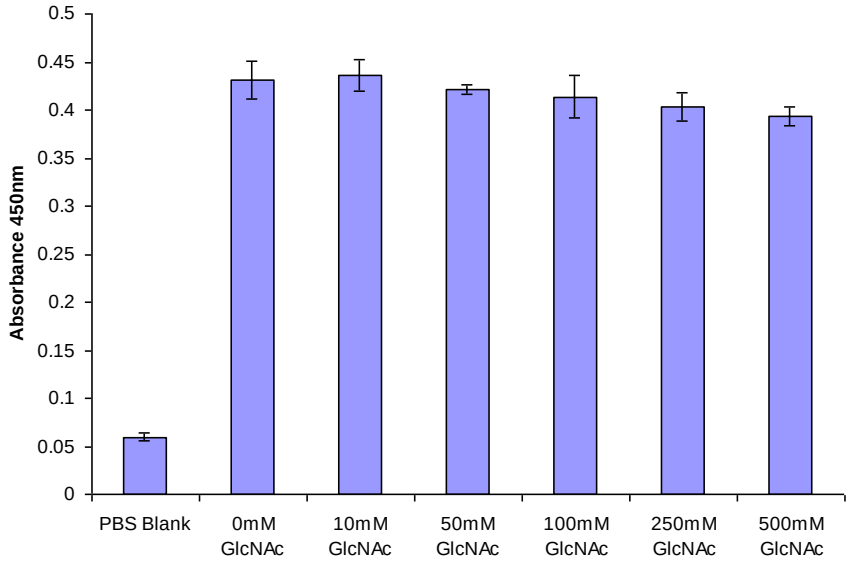


Fig 4.29: Examining the specificity of the sandwich ELLA. rChb CBP was laid down on the plate which was then blocked as per Section 2.27. Asialo agalacto fetuin was then added, along with increasing concentrations of free GlcNAc and incubated for 60 minutes at 24^{o}C. GSLII was added and also incubated at 24^{o}C for 60 minutes. Finally an anti-biotin antibody was added and the ELLA was completed as per Section 2.27. Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

To definitively prove that the interaction was taking place between the two carbohydrate-binding proteins, rChb CBP and GSLII, an ELLA was carried out as described in Figure 4.30. rChb CBP was laid down on the plate and probed with GSLII. In the absence of any glycoprotein a signal was obtained for bound biotinylated GSLII, as the concentration of rChb CBP increased so did the amount of bound GSLII. Therefore it was concluded that an interaction was occurring between the rChb CBP and GSLII. GSLII itself is glycosylated, hinting that the rChb CBP may have been binding to glycans on its surface.

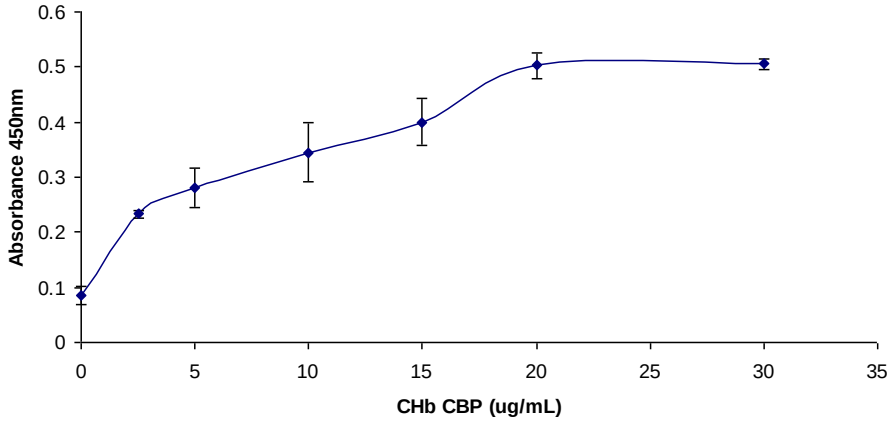


Fig 4.30: Interaction between rChb CBP and GSLII. Increasing concentrations of rChb CBP was immobilized on the plate which was then blocked as per Section 2.27. The rChb CBP was probed with GSLII and incubated at 24^{o}C for 60 minutes. Finally an anti-biotin antibody was added and the ELLA was completed as per Section 2.27. Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

As chitobiose was thought to be the natural substrate of rChbL, ELLA analysis was undertaken using a biotinylated PAA-chitobiose substrate. Binding between the rChb CBP and chitobiose was examined via the immobilisation of rChb CBP and detection of any bound biotinylated chitobiose. A definite interaction between rChb CBP and the chitobiose substrate was observed (see Figure 4.31).

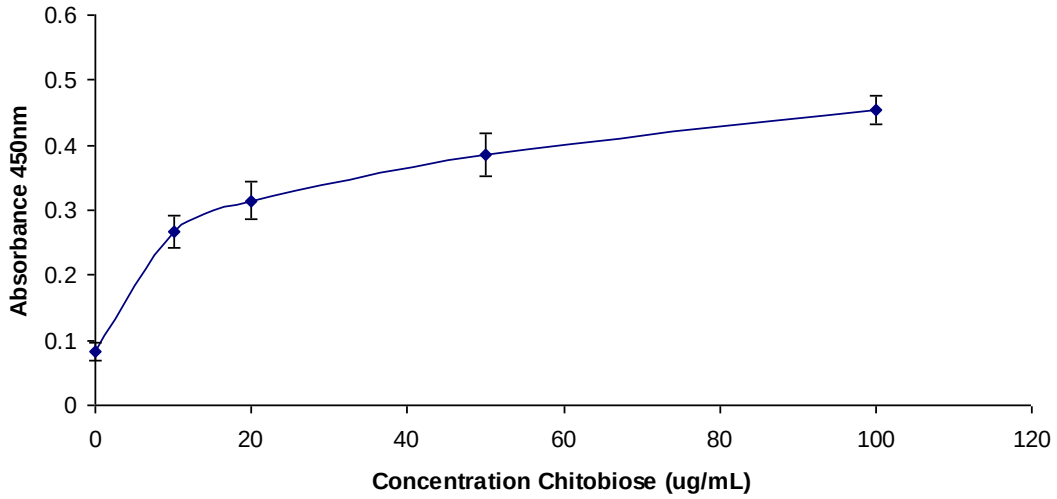


Fig 4.31: ELLA examining the binding of rChb CBP to chitobiose. A constant concentration of rChb CBP was laid down on the plate at 10 µg/mL. An increasing concentration of biotinylated chitobiose was added and probed for with an anti-biotin antibody as described in Section 2.27. Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

4.10 Biotinylation of the rChb CBPlnk

A final effort was made to enable direct detection of the rChb CBPlnk in an ELLA format. This was to ensure that detection of the rChb CBPlnk was not the issue and that the problem was due to a lack of interaction between the rChb CBP and GlcNAc residues. As manipulation of the His_6 tag did not drastically improve its accessibility an alternative system of protein labelling and detection was investigated. A commercial biotinylation kit was used to label the rChb CBPlnk which could then be detected using an anti-Biotin antibody. The biotin molecules are functionalised to contain NHS esters, which are highly reactive towards the primary amine groups found in lysine side-chains, allowing the biotin molecules to covalently link to the rChb CBP surface.

It was observed (in Figures 4.32 and 4.33) that the biotinylated rChb CBP was readily detectable with an anti-Biotin antibody. Figure 4.32 highlights the ability to detect the rChb CBP protein on a PVDF membrane but more importantly it can be seen from

Figure 4.33 that the biotinylated protein could be directly detected in the ELLA format.

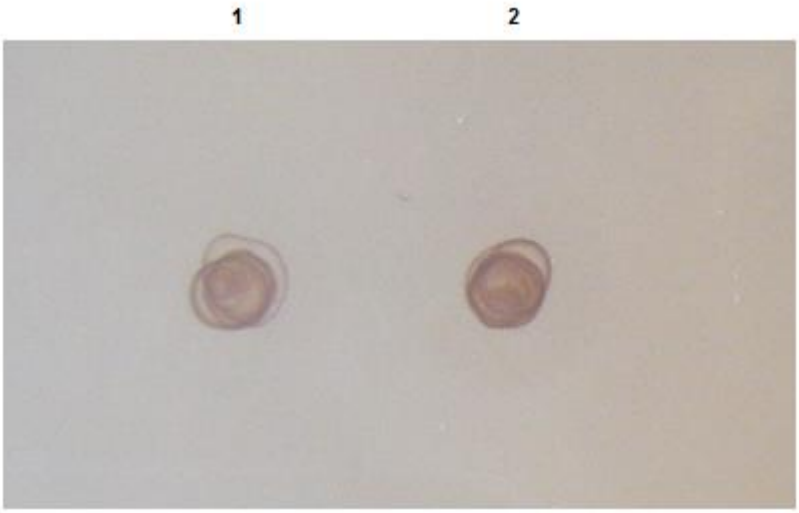


Fig 4.32: Detection of biotinylated rChb CBP on a PVDF membrane. 1 µL of the biotinylated rChb CBPlnk was spotted onto a PVDF membrane at a concentration of 10 µg/mL and allowed to dry. This was repeated three times. 10 µg/mL GSLII was used as a positive control. The PVDF membrane was then washed with TBST, probed with an anti-Biotin antibody and visualised as in Section 2.22. Lane 1; biotinylated rChb CBPlnk Lane 2; biotinylated GSLII (positive control).

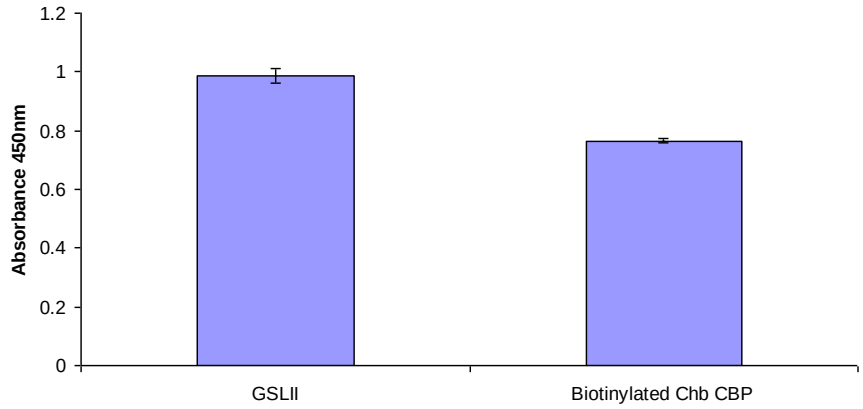


Fig 4.33: Detection of the biotinylated rChb CBPlnk. 50 µL of the rChb CBPlnk was laid down on a 96-well plate a concentration of 10 µg/mL in triplicate. The plate was blocked as per Section 2.27 and the proteins were detected with an anti-Biotin antibody. GSLII at a concentration of 10 µg/mL was used as a positive control and was also detected with an anti-Biotin antibody. Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

As it had now been shown that it was possible to directly detect the biotinylated rChb CBP in the ELLA format, the glycoprotein binding assays were repeated. The biotinylated rChb CBP was tested for binding against exposed terminal GlcNAc residues on the various glycoproteins (fetuin, transferrin and thyroglobulin), the BSA-GlcNAc conjugate and bovine serum mucin. The background signal from the biotinylated protein was relatively high in comparison with the GSLII blank. The signal obtained from the rChb CBP, when tested against asialo agalacto fetuin, transferrin and thyroglobulin, was similar to the blank. There was no binding detected between the biotinylated rChb CBP and GlcNAc residues (see Figures 4.34 and 4.35). The high background was probably due to the presence of residual biotin in the sample. Biotin is generally removed using centrifugal units (vivaspins) to filter out the biotin, however this was not possible with the rChb CBP, as large volumes were lost when previously using vivaspins, presumably due to binding to the cellulose membrane (see Figure 3.16). Instead, gel filtration units were used which allowed the larger rChb CBP to flow through the gel while retaining the smaller biotin molecules. These were not very efficient for this purpose as residual biotin was still present in the sample, resulting in a high background signal.

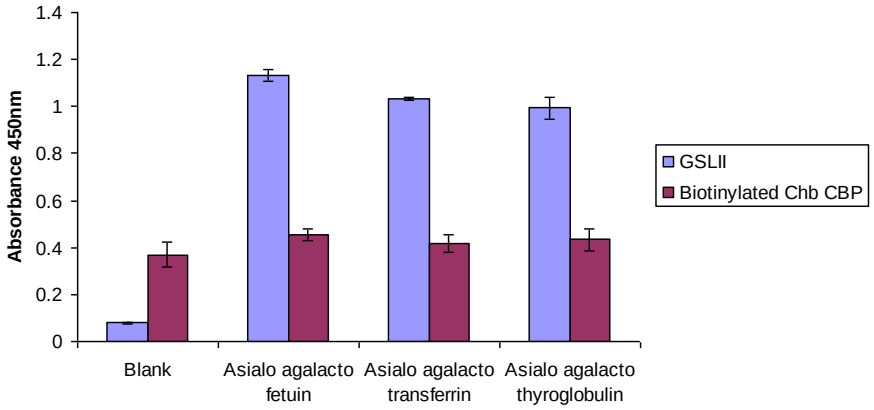


Fig 4.34: ELLA investigating the binding of the biotinylated rChb CBPlnk to GlcNAc residues on different glycoproteins. An ELLA was performed as per Section 2.27. Biotinylated rChb CBPlnk was used at a concentration of 10 µg/mL. Asialo agalacto Fetuin, transferrin and thyroglobulin were used. GSLII (10 µg/mL) was used as a positive control. Both proteins were detected with an anti-Biotin antibody. Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

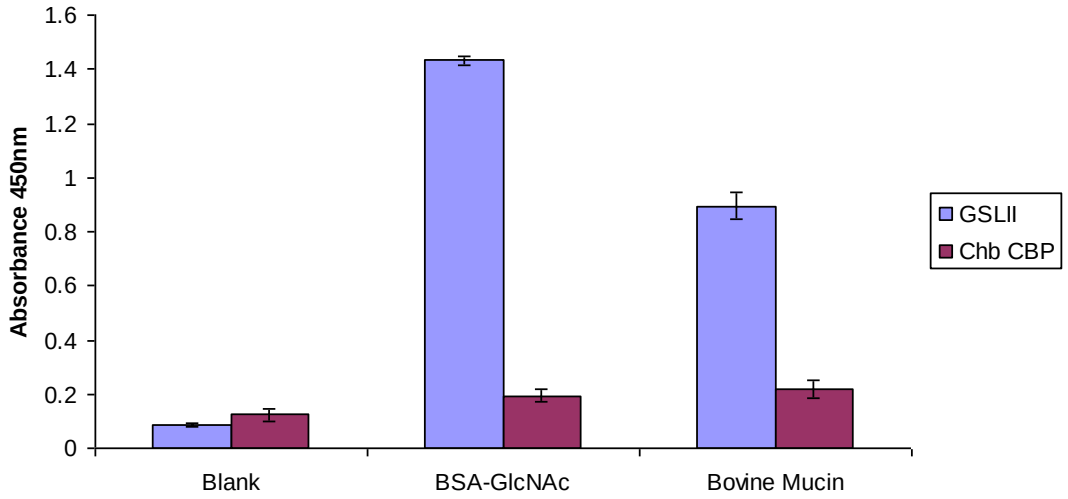


Fig 4.35: ELLA investigating the binding of the biotinylated rChb CBP to BSA-GlcNAc and bovine serum mucin. An ELLA was performed as per Section 2.27. Biotinylated rChb CBPlnk was used at a concentration of 10 µg/mL. GSLII (10 µg/mL) was used as a positive control. Both proteins were detected with an anti-Biotin antibody. Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

Although the problems with direct detection of the rChb CBP protein were overcome through biotinylation there was still no binding detected between rChb CBP and any of the glycoproteins tested nor the BSA-GlcNAc conjugate. The rChb CBPlnk has 57 lysines present in its structure. However, the number of residues that were surface accessible for biotin conjugation was unknown. The location of potentially surface exposed primary amines on the rChb CBP molecule were examined and are highlighted in Figure 4.36. While the lysine residues appear to be distributed evenly over the protein surface there are two possible biotinylation sites located very close to the active site/binding pocket. The presence of the small biotin molecule in these positions would not interfere with direct detection of the rChb CBP protein in the ELLA format but could have rendered the protein incapable of binding to its natural substrate by blocking access to the active/binding site. Therefore binding of the rChb CBP to GlcNAc residues was still not conclusively ruled out.

a) **b)**

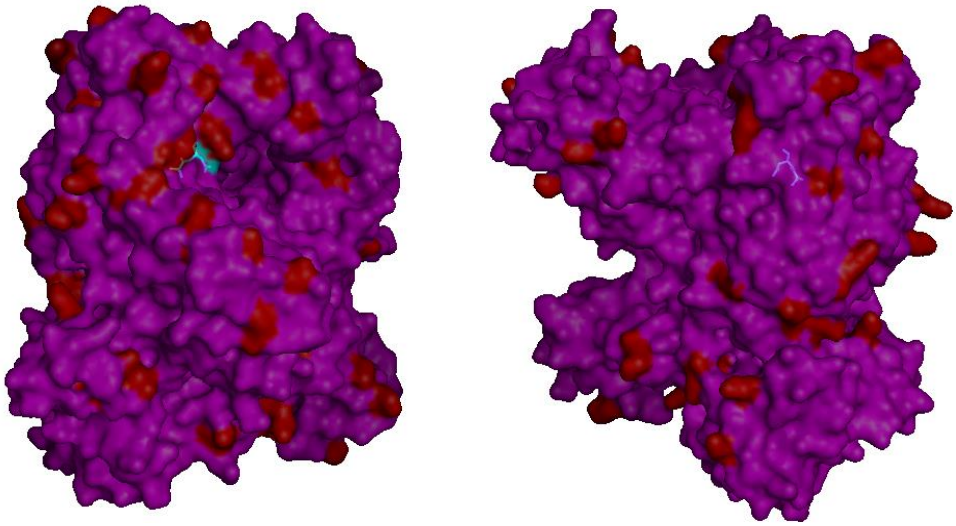


Fig 4.36: Possible biotinylation sites on the Chb CBPlnk a) Front view and b) Back view. The lysine residues are highlighted in red. The active site residues are highlighted in blue. The image was visualised using Pymol software (available from http://www.pymol.org/).

4.11 Immobilised rChb CBP

An alternative method for detecting an interaction between the rChb CBP and GlcNAc residues was employed. The rChb CBP was immobilised on cyanogen bromide-activated sepharose as previously described in Section 3.9. The final concentration of the immobilised rChb CBP was 1.5 mg/mL. The resin was then diluted to 10 µg/mL for the purpose of this experiment. This resin was mixed with BSA-GlcNAc in a column format and the unbound fractions were collected. The resin was washed with five CV of PBS. The resin was then boiled for 10 minutes in 1X protein sample buffer to remove any bound BSA-GlcNAc. The unbound, washed and bound fractions were analysed by SDS-PAGE. It was evident from Figure 4.37 that the rChb CBP resin did not bind to any BSA-GlcNAc, all BSA-GlcNAc passed over the column. As this chemistry was the same as the biotinylation – i.e. primary amines may have immobilised via the same residues surrounding the active site, these may have inhibited binding.

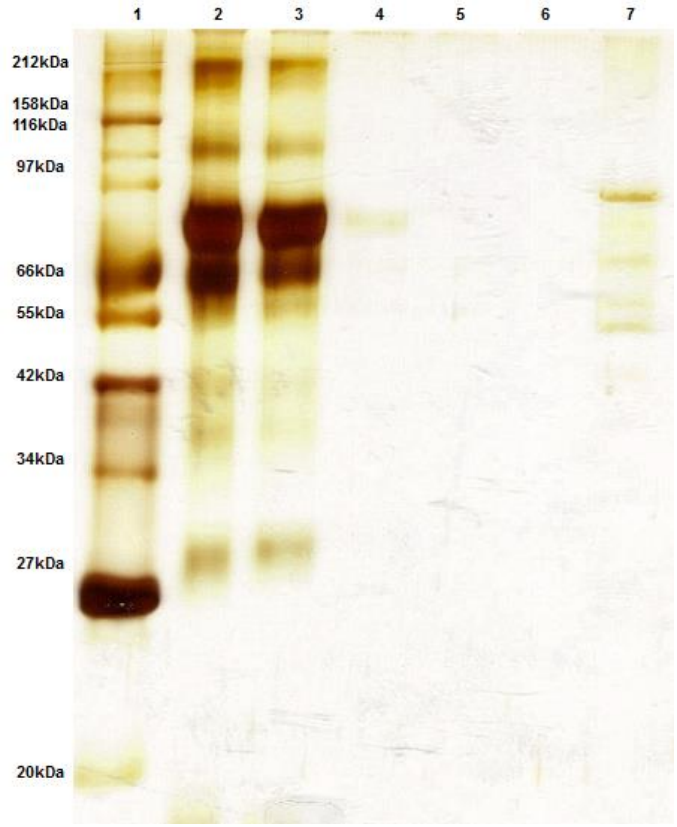


Fig 4.37: Examination of the binding between immobilised rChb CBP and BSA-GlcNAc. Lane 1, Protein Ladder; Lane 2, BSA-GlcNAc sample; Lane 3, Flow through; Lane 4, Wash; Lane 5, Elution 1; Lane 6, Elution 2; Lane 7, Bound Fraction.

4.12 Activity of the ChbL enzyme on glycoproteins

The ability of the rChbL enzyme to cleave GlcNAc residues from glycoproteins was investigated (see Figure 4.38). This was to test the cleavage ability of the rChbL enzyme against GlcNAc residues on glycoproteins, to ensure that the mutations of the catalytic residues (D539A and E540D) to remove the catalytic activity did not affect the proteins ability to interact with its substrate. Various asialo agalacto glycoproteins were treated with either rChbL or a commercial N-acetylglucosamindase. The efficiency of the removal of GlcNAc residues was measured via the use of the ELLA and GSLII. The less signal observed for bound GSLII the more efficient the enzyme was at removing GlcNAc residues.

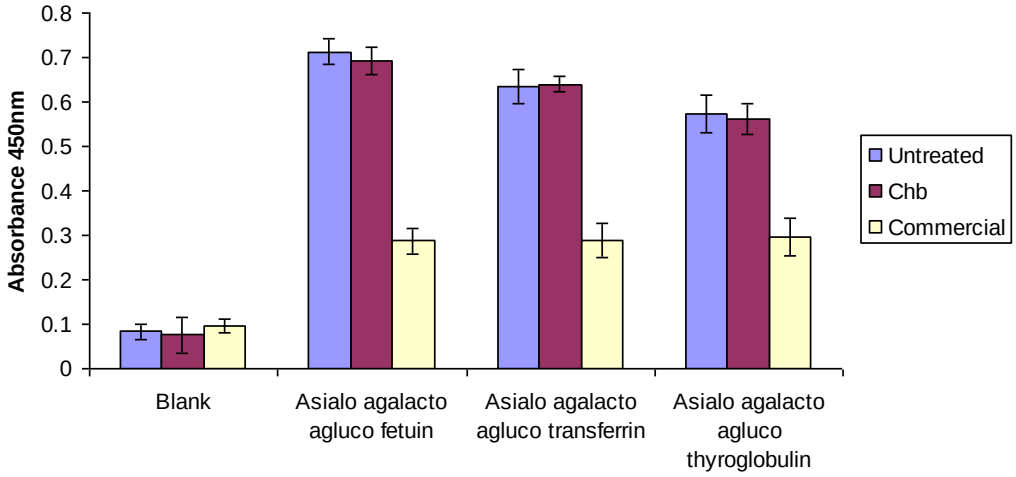


Fig 4.38: Comparison of rChbL and commercial N-acetylglucosaminidase treatment of glycoproteins to remove GlcNAc residues. Asialo agalacto fetuin, asialo agalacto transferrin and asialo agalacto thyroglobulin were treated with either 10 μg/mL of rChbL or a commercial glycosidase (Sigma). The resultant glycoproteins were probed with GSLII which binds to GlcNAc residues. Binding was detected with an anti-biotin antibody. Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

A high signal was observed when the rChbL treated glycoproteins were probed with GSLII. This meant that the rChbL enzymatic treatment was not successful. The terminal GlcNAc residues were not removed from the glycoprotein. A much lower signal was observed when GSLII was used to probe the commercial N-acetylglucosaminidase treated glycoproteins, meaning it effectively cleaved the terminal GlcNAc residues from the glycoproteins tested.

4.13 Discussion

Family 20 glycosyl hydrolases are ubiquitously found in plants, animals, fungi and bacteria. Chitobiases are assigned to Family 20 based on amino acid similarity. Family 20 chitobiase protein sequences from many different bacterial species were aligned and were found to have large regions of conservation. High levels of homology were observed in the residues surrounding the catalytic site. The amino acids critical for catalytic activity were observed to be conserved across all chitobiases examined. The least conservation between sequences was observed at the beginning of the sequence in the signal peptide sequence region. This was to be expected as each species of bacteria has their own recognizable signal peptide sequence characteristics. Low levels of homology were also observed between residues 200 and 300. This region of sequence was described as Domain II by Prag *et al.,* (1996) however no function has been assigned to it. The active site and critical catalytic residues were conserved in the ChbL protein sequence as compared with the Chb from *S. marcescen*s.

Homology modelling allowed a possible structure of rChbL to be assigned. This model was used to great advantage allowing for the study of the structural implications of amino acid mutations surrounding the binding pocket. It was envisioned that a successful mutagenesis strategy of residues present in the catalytic site would generate a catalytically inactive chitobiase similar to the *S. marcescens* Chb generated by Prag *et al.,* (2000). While there is no report of further examination into the binding capabilities of this inactive enzyme, other studies have described the mutagenesis of enzymes and the generation of novel carbohydrate binding proteins with functional applications. Laine *et al.,* (2005) utilized catalytically inactive endoglycosidases for the detection of fungal and bacterial species. The endoglycosidases bound specifically and with high affinity to chitin present in their structure/cell wall. They generated a fusion protein composed of the chitin-binding domain from an inactive chitinase from *Bacillus circulans* linked to GFP, which is now commercially available and used in human diagnostics for fungal testing. Jokilammi *et al.,* (2005) isolated inactive endosialidases from bacteriophages. These enzymes were unable to carry out their normal function of cleaving polysialic acid

present on the surface of *E. coli* K1 however, they still retained their polysialic acid binding capabilities. They exploited this polysialic acid binding specificity for a variety of applications; through fusion to GFP this mutant enzyme was used for the detection of bacterial pathogens with surface polysialic acid such as *Neisseria meningitidis* and *Moraxella nonliquefaciens*, the GFP-endosialidase fusion protein was also used as a diagnostic tool for the detection of neuroblastoma tumors as they have high levels of polysialic acid expression.

The catalytically important residues, D539 and E540 are present in all bacterial homologues examined in this study. Chitobiases have human homologues, know as β-hexosaminidases. These are lysosomal enzymes that cleave terminal GlcNAc from glycoconjugates using the retaining acid hydrolysis mechanism. The acidic residues employed in this mechanism were elucidated by Fernandes *et al.,* (1997). They proposed a glutamate at position 323 as a catalytically important residue, as mutation of this residue dramatically reduced catalytic activity. Through predictive modelling of the hexosamindase protein structure they identified the glutamate at position 540 in the *S. marcescens* Chb was the residue corresponding to the glutamate at position 323 in the hexosaminidase enzyme thus cementing the importance of this acidic residue in the catalytic mechanism of certain chitobiase enzymes.

The generation of an inactive enzyme from the *P. luminescens* Chb required two mutations, D539A (aspartate to alanine) and E540D (glutamate to aspartate). The study of the *S. marcescens* Chb demonstrated that one of the functions of the aspartate at position 539 was to hold the glutamate and N-acetyl group of the sugar in place for the duration of catalysis. Prag *et al*., (2000) observed that when this residue was mutated to the neutral amino acid alanine, the position on the sugar substrate remained the same. However, the N-acetyl group of the sugar twisted 175^{o}. It inhabited the space that had previously been occupied by the side chain of the aspartate. The mutation at residue 540 acted directly on the catalytic activity of the *S. marcescens* Chb. This study also found a mutation to residue 540 of the *P. luminescens* Chb acted directly on catalytic activity. Glutamate and aspartate are amino acids with similar properties; both are acidic with a carboxylic acid side-chain. The major difference is the presence of an extra $-CH_2$ group in the side chain of glutamate. Mutating the glutamate at position 540 to aspartate (a residue with a

slightly smaller side-chain) increased the distance between the amino acid and the non-reducing end of the sugar. While the nature of the side chain remained unchanged the increased distance between the carboxylic acid and the N-acetyl group of the sugar prevented any proton donation occurring and therefore, removed the catalytic activity of the rChbL enzyme.

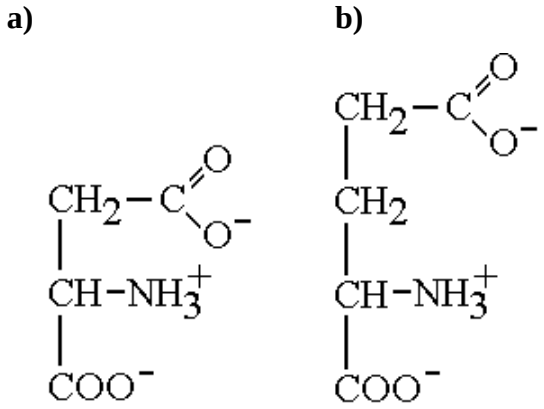


Fig 4.39: (a) Aspartate and (b) Glutamate structures. The similarity between the two acidic amino acids is clear.

Upon the creation of the catalytically inactive rChbL protein rChb CBP the next step was to examine its binding abilities. An Enzyme Linked Lectin Assay (ELLA) was used (as described in Section 2.27). In 1994, Duk *et al.,* first described an enzyme-linked lectin assay in the format that was used in this study. They used biotinylated lectins to detect Tf and Tn receptors from glycophorin in a microwell plate format. The treated glycophorin was immobilised on a microwell plate and probed with biotinylated lectins such as peanut agglutinin (PNA) and jacalin (AIA). The wells were incubated with ExtrAvidin alkaline phosphatase solution after which the p-nitrophenyl phosphate substrate was added, the absorbance of which could be read at 405 nm. This study also immobilised the treated glycoprotein on the surface of a microwell plate. Polyvinyl alcohol (PVA) was used as a blocking solution to prevent any non-specific interactions. PVA was found to be an optimum blocker in ELLAs in a study by Thompson *et al.,* (2011). Bovine serum albumin (BSA) and carbo-free blocker were also tested, however PVA generated the results with the lowest background signal and highest positive signal. The commercial lectins used were biotinylated and detected with an anti-Biotin antibody, however, the recombinant Chb CBP was His_6 tagged and detected with an anti-His antibody. The chromogenic

substrate used in the ELLAs in this study was TMB, which was broken down in the presence of horseradish peroxidise conjugated to the anti-Biotin and anti-His antibodies. The absorbance of the results was also read, however at 450 nm.

The first attempts at examining the binding abilities of the rChb CBP in an ELLA format were unsuccessful. It was thought that the lack of signal was due to a difficulty in detecting the His_6 tag rather than no binding occurring between rChb CBP and it's target glycoprotein. Modeling of the protein structure revealed that the His_6 tag was probably hidden by the rest of the protein structure. In order to overcome this problem an amino acid linker sequence was genetically engineered between the His_6 tag and the protein to allow the tag to protrude from the protein structure and therefore facilitate detection through the use of an anti-His antibody. The linker improved the purification of the rChb CBP resulting in higher yields of pure protein. It also allowed the rChb CBP to be detected on a western blot. The linker did not, unfortunately, improve detection of the protein in the ELLA format. In this instance the His_6 tag was present on the C-terminus of the protein. His_6 tags can also be engineered at the N-terminus of a protein. However, as this protein contained a signal peptide that was cleaved the His_6 tag would also have been removed. The addition of internal His_6 tags are possible however this may have disrupted the structure of the protein. Since ELLA detection of the protein was not feasible via the His_6 a different method of detection had to be devised.

The rChb CBP contains a large number of primary amine containing residues (lysine), distributed evenly throughout the proteins surface. Therefore biotinylation was chosen as the next labeling method of choice. The biotinylated rChb CBP was detectable in the ELLA format, however there was no binding detected between the rChb CBP and the glycoproteins. However, it was noted that there were lysine residues present in the binding pocket of the rChb CBP. These may have hindered binding between the protein and GlcNAc residues.

Indirect ELLA methods were then employed to test the ability of rChb CBP to bind to asialo agalacto glycoproteins but the only signals observed were due to interactions between the GSLII and rChb CBP. GSLII is a glycoprotein itself (Goldstein, 1983). However it was unknown if the interaction was occurring due to the rChb CBP

binding to sugar residues on the surface of GSLII or due to non-specific interactions between the two proteins. The rChb CBP was also tested against a biotinylated chitobiose substrate. A low but definite signal was observed so it was concluded that there was some interaction taking place. The low signal may have been due to the fact that the rChb CBP was immobilized non-specifically on the plate. The orientation of the protein is unknown and the active site/binding pocket may have been inaccessible. While obtaining an affirmative result for an interaction between rChb CBP and chitobiose was positive, terminal chitobiose structures are not found on glycoproteins. The problem with the detection of terminal GlcNAc by rChb CBP on glycoproteins still remained.

Finally the rChb CBP was immobilised on cyanogen bromide activated resin and tested for binding to BSA-GlcNAc; again, no binding was detected. There are other types of affinity tags that can be engineered in a protein to allow purification and detection of the protein such as GST tags, Arg tags or MBP tags when problems with His_6 tags are encountered (Walls and Loughran, 2011). However at this point it had been concluded with relative certainty that the rChb CBP could not bind to any glycan residues on glycoproteins.

The rChbL enzymes ability to cleave GlcNAc residues from the surface of glycoproteins was also examined to ensure that the mutations to remove catalytic activity did not also remove the rChbLs binding ability. The rChbL enzyme was not able to cleave GlcNAc residues from the glycoproteins. This suggests that the mutations did not affect the natural binding abilities of the rChb CBP as both the enzyme and binding protein were unable to interact with the glycoprotein targets.

The structure of the rChb CBP was examined to explain why binding might not be occurring. The natural substrate of the Chb is chitobiose which is a small sugar composed of two N-acetylglucosamine residues. The rChb CBP was shown to bind to this substrate in the ELLA format. The glycoproteins used for analysis in this study are large bulky structures. The active site of the ChbL is relatively deep in the protein and has many over-hanging residues. It was hypothesized that steric hindrance was occurring; the protein was unable to access the GlcNAc residues due to both its bulky structure and the bulky structure of the glycoproteins. As the rChb CBP was generated

for the final application of detecting GlcNAc residues on glycosylated biopharmaceuticals, methods to overcome this steric hindrance issue had to be examined. The active site of the rChb CBP was examined in depth in Chapter 5 and site-specific mutagenesis was carried out in an attempt to open up the binding pocket and allow binding to glycoproteins to take place.

Chapter 5

Site-specific mutagenesis of the rChb CBP and rChbL enzyme

5.1 Overview

Homology modelling of the rChb CBP structure showed an active site deep in the structure of the protein with many over-hanging residues. It was thought that these residues may be interfering with binding between the rChb CBP and glycoproteins. Tews *et al.,* (1996) previously examined the structure of the Chb protein from *S. marcescens* and the amino acids involved in the catalytic mechanism of the Chb enzyme. A series of residues were proposed to interact directly with the chitobiose sugar. It was discovered that a collection of tryptophans in the pocket of the active site interacted with the hydrophobic surface of the chitobiose sugar, holding it in place during catalysis, namely Trp639, Trp685 and Trp737. An arginine at position 349 formed hydrogen bonds with the non-reducing end of the sugar, keeping it stabilised during catalysis. This arginine was held in position through polar interactions with other residues in the active site, specifically Asp346, Asp378, Asp 379 and Glu380.
Figure 5.1 shows the interaction between these residues and the chitobiose sugar as modelled by Tews *et al*., (1996). Figure 5.2 shows the position of the analogous residues in the modelled active site of the rChb from *P. luminescens*. These residues were conserved in the *P. luminescens* Chb and are considered essential for binding and therefore were not targeted for mutagenesis.

a) **b)**

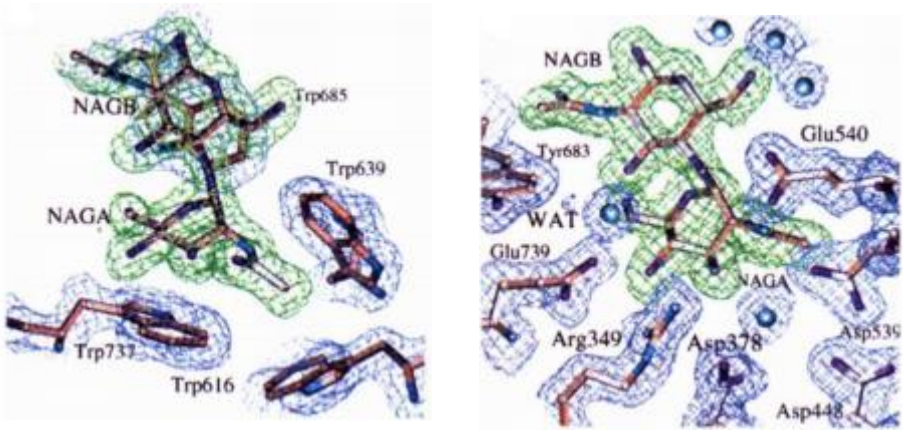


Fig 5.1: Important residues involved in binding of the chitobiose by the Chb from *S. marcescens*. Image (a) highlights the tryptophan residues involved in holding the chitobiose in place while image (b) shows the arginine at position 349 also involved in holding the chitobiose in place. Image obtained from Tews *et al.,* (1996).

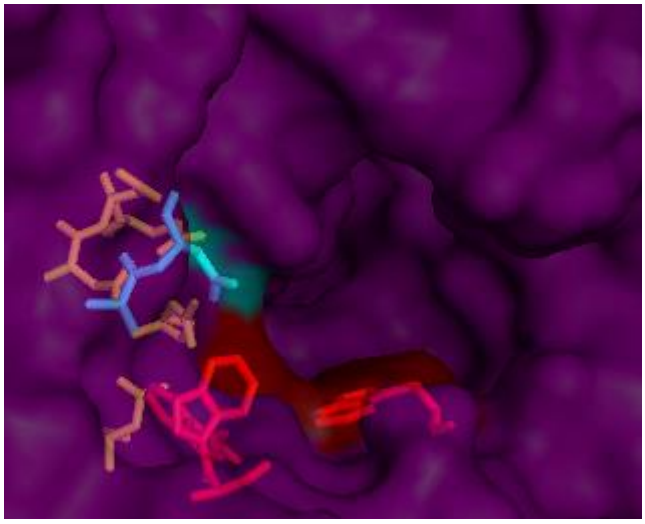


Fig 5.2: Important residues involved in binding of the chitobiose by the Chb from *P. luminescens*. The catalytic residues are highlighted in blue, the important tryptophans are highlighted in red while the arginine and associated residues are highlighted in yellow. The image was generated using Pymol (available from http://www.pymol.org/).

5.2 Site-specific mutagenesis of rChb CBP

It was believed that the inability of rChb CBP to bind to the glycoprotein structures tested was an issue with steric hindrance due to the large bulky nature of the glycoproteins. The active site of the *P. luminescens* Chb was closely examined and five residues were chosen for site-specific mutagenesis in an effort to allow rChb CBP to bind to GlcNAc residues on the surface of glycoproteins. W63, D648, K545 and D551 were all chosen as they were projecting out over the active site of the protein (see Figure 5.3). Alanine scanning mutagenesis was carried out on the four residues, W63, D648, K545 and D551, resulting in four mutants named W63A, D648A, K545A and D551A. The predicted structural modifications are highlighted in Figures 5.4, 5.5, 5.6, 5.7 and 5.9. The primers used are shown in Figures 5.10 to 5.13. A tryptophan at position 689 was also mutated. As tryptophan is such a large amino acid it was thought that mutagenesis of this residue may open up the active site. However, as previously discussed this tryptophan has an important role in maintaining the chitobiose position during catalysis. It was believed that mutagenesis of this position to alanine would result in an extensive change to an important amino acid that would possibly result in a complete loss of function. Therefore the tryptophan at position 689

was mutated to a tyrosine which has a similar ring structure, generating a mutant named W689Y. The predicted structure of this mutant is shown in Figure 5.8. The primers used are shown in Figure 5.14. The new molecular weight, extinction coefficient and pI for each new rChb CBP is shown in Table 5.1.

5.2.1 Structural modelling of the rChb CBP site-specific mutants

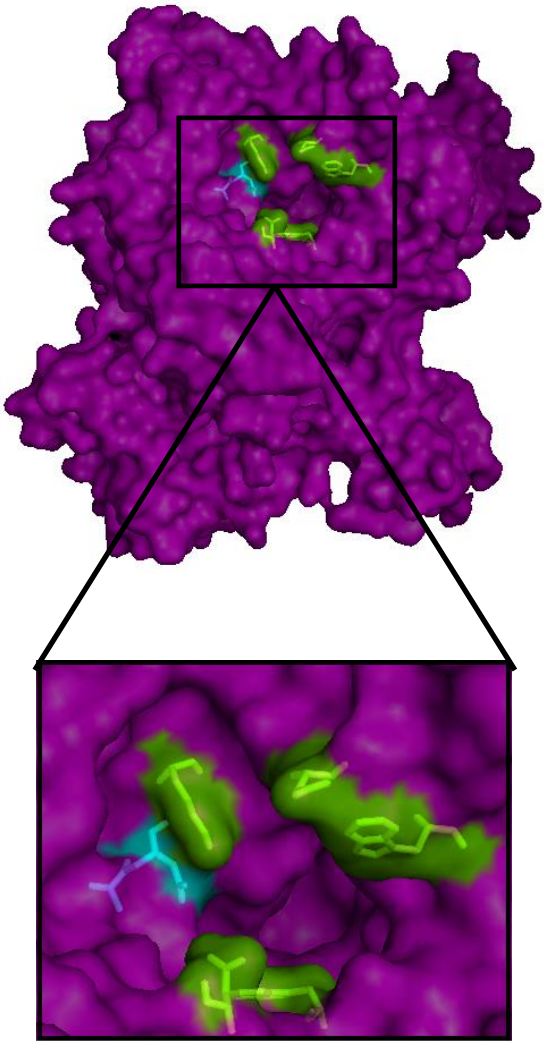


Fig 5.3: Residues chosen for site-specific mutagenesis in the *P. luminescens* rChb CBP active site. The mutagenesis was performed and visualised in Pymol. The mutated catalytic residues (D539A and E540D) are highlighted in blue. The other residues chosen for mutagenesis are highlighted in green, W63, K545, D551, D648 and W689. Images were generated using Pymol (available from http://www.pymol.org/).

a) **b)**

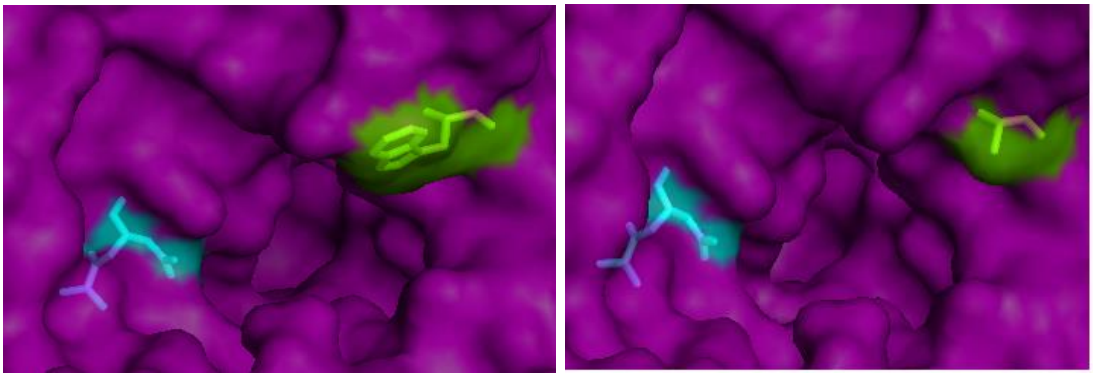


Fig 5.4: Active site of a) Chb CBP and b) Chb CBP W63A. The catalytic residues, D539A and E540D, are highlighted in blue and residue 63 is highlighted in green. The image was generated using Pymol software. Mutations were also carried out and visualised in Pymol (available from http://www.pymol.org/).

a) **b)**

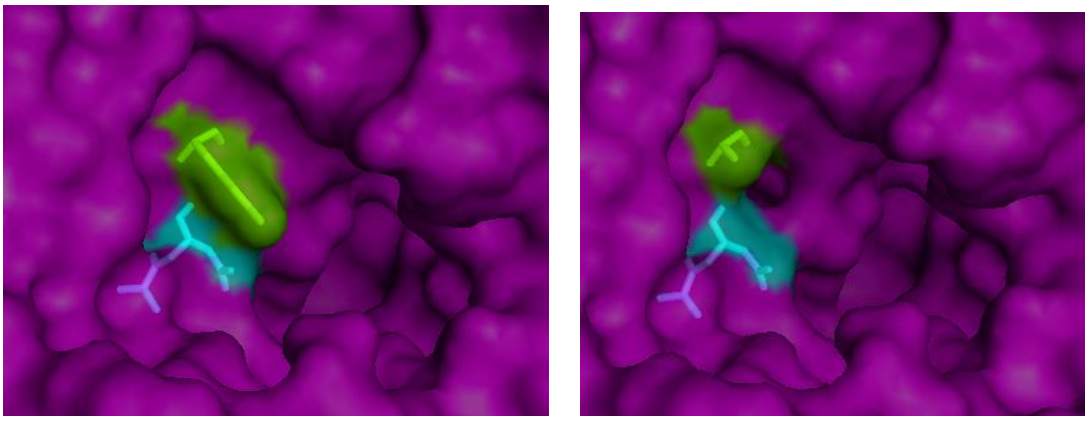


Fig 5.5: Active site of a) Chb CBP and b) Chb CBP K545A. The catalytic residues, D539A and E540D, are highlighted in blue and residue 545 is highlighted in green. The image was generated using Pymol software. Mutations were also carried out and visualised in Pymol (available from http://www.pymol.org/).

a) **b)**

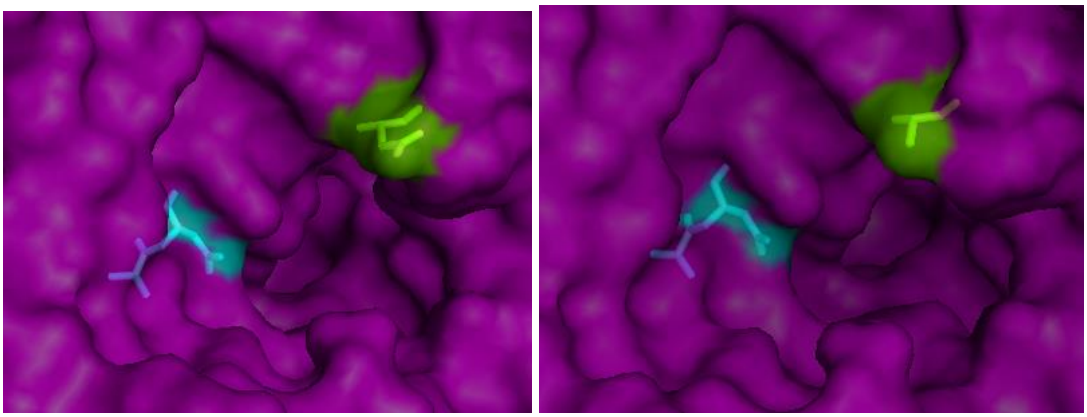


Fig 5.6: Active site of a) Chb CBP and b) Chb CBP D551A. The catalytic residues, D539A and E540D, are highlighted in blue and residue 551 is highlighted in green. The image was generated using Pymol software. Mutations were also carried out and visualised in Pymol (available from http://www.pymol.org/).

a) **b)**

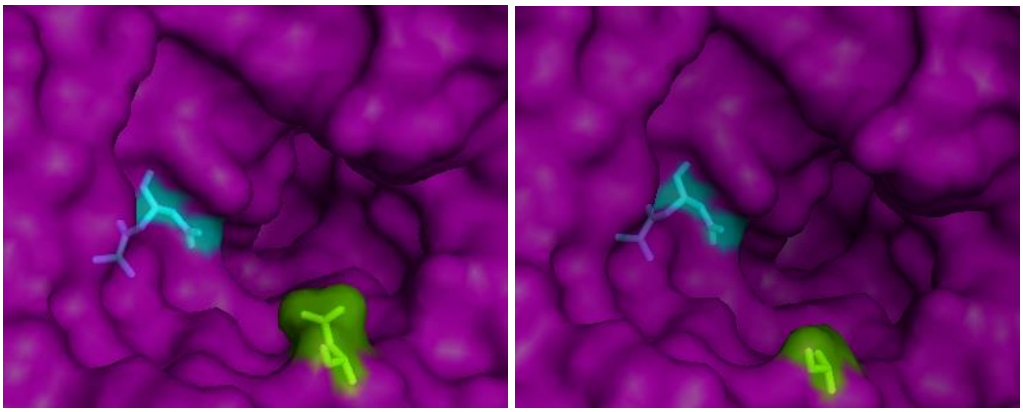


Fig 5.7: Active site of a) Chb CBP and b) Chb CBP D648A. The catalytic residues, D539A and E540D, are highlighted in blue and residue 648 is highlighted in green. The image was generated using Pymol software. Mutations were also carried out and visualised in Pymol (available from http://www.pymol.org/).

a) **b)**

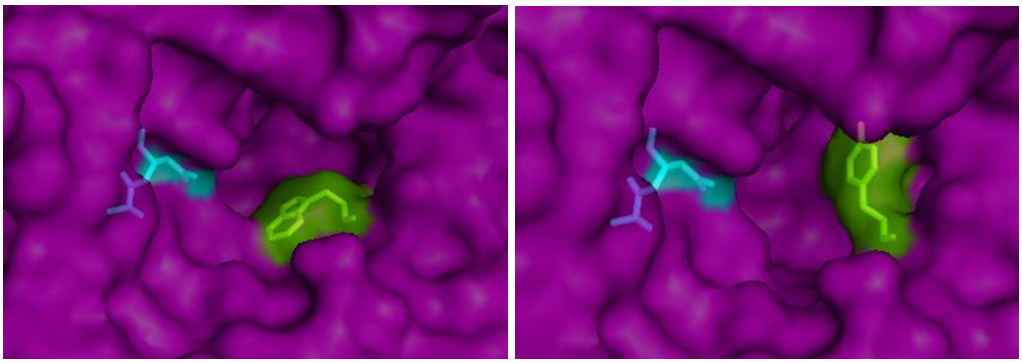


Fig 5.8: Active site of a) Chb CBP and b) Chb CBP W689Y. The catalytic residues, D539A and E540D, are highlighted in blue and residue 689 is highlighted in green. The image was generated using Pymol software. Mutations were also carried out and visualised in Pymol (available from http://www.pymol.org/).

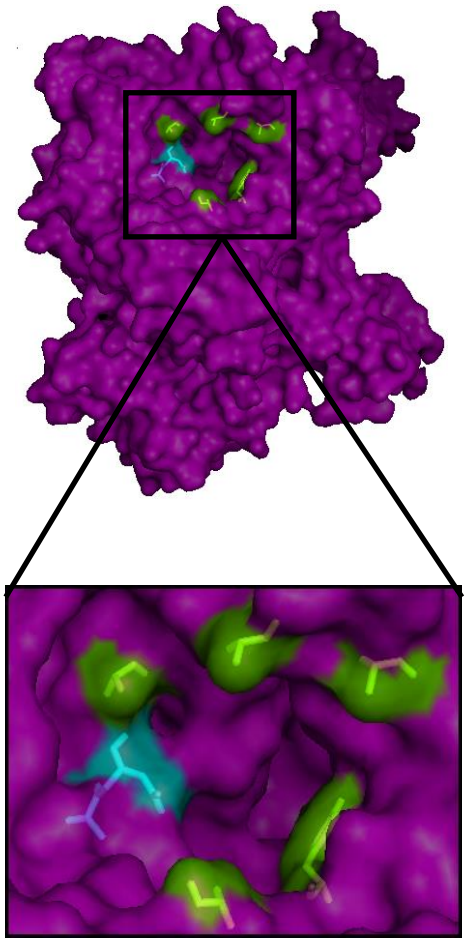


Fig 5.9: Mutated residues chosen in the active site of rChb CBP. The mutated catalytic residues (D539A and E540D) are highlighted in blue. The other mutated residues are highlighted in green, W63A, K545A, D551A, D648A and W689Y. The image was generated and the mutations were performed using Pymol (available from http://www.pymol.org/).

5.2.2 Site-specific mutagenesis of the rChb CBP mutants

Whole vector amplification was used to introduce the specific point mutations into the molecules DNA sequence. This was carried out as outlined in Section 2.13.

GTG CCG CAA

GTT AAT TAT AAA GTT GTT GAT AAT CAG GCA GGA AGT CAC GGC GTT

TTG ACA CGC CAG GCA GGG GCA GAT **GCG** GCT AAA TGT AAT CG

AAC TGT GCG GTC CGT GGG GCA GAT **TGG** GCT AAA TGT AAT CGA GTG

Fig 5.10: Primers used for the construction of rChb CBP mutant W63A. Outline of the position of the primers used for the mutation of the rChb CBP at residue 63. The forward primer 63A_F is shown in red and the reverse primer 63A_R is shown in green. The mutated residue is underlined and in bold.

GT GGT AAC CTC TGA

GGT GAA ATT GCT ATG CAT AAA GTA GCA GGT TCA CCA TTG GAG ACT

ACC GTA AAA CCC GGA GCT GAC GCG **GCG** AAT ATC AGT TTG

TGG CAT TTT GGG GGA GCT GAC GCG **AAG** AAT ATC AGT TTG GGA GAC

Fig 5.11: Primers used for the construction of rChb CBP mutant K545A. Outline of the position of the primers used for the mutation of the rChb CBP at residue 545. The forward primer 545A_F is shown in red and the reverse primer 545A_R is shown in green. The mutated residue is underlined and in bold.

GTA AAA CCC CCT CGA CTG CGC TTC TTA ATC CGT TTG GGA

ACT TGG CAT TTT GGG GGA GCT GAC GCG AAG AAT ATC CGT TTG GGA

GCG GGA TTC CAG G

GAC GGA TTC CAG GAT AAA AAT GGT CCT ATC GTA CCG GGT AAG

Fig 5.12: Primers used for the construction of rChb CBP mutant D551A. Outline of the position of the primers used for the mutation of the rChb CBP at residue 551. The forward primer 551A_F is shown in red and the reverse primer 551A_R is shown in green. The mutated residue is underlined and in bold.

CGG

TGG CAG GAT GGT TTA AAA CAT GCT AGC AGT GCT AAA GAT TTT GCC

TGG TTT GCG CAA CCC CAC TTA AAA AGG GAC ACT CTC TAC **GCG** GGC

ACC AAA CGC GTT GGG GTG AAT TTT TGG GAC ACT CTC TAC **GAT** GGC

GGC

GGC TTT GAT TCA GTC AAT GAT TGG GCG AAT AAG GGC TAT GAC GTG

Fig 5.13: Primers used for the construction of rChb CBP mutant D648A. Outline of the position of the primers used for the mutation of the rChb CBP at residue 648. The forward primer 648A_F is shown in red and the reverse primer 648A_R is shown in green. The mutated residue is underlined and in bold.

TA TAC GGT ATA CTT CAA GTA GGT

AGC CCA GAT TAT ATT TAT CTG GAT ATG CCA TAT GAA GTT CAT CCA

GTT CTT CGT GGC TAT TAT **TGT** GCG GCT CGT TTT

CAA GAA CGT GGC TAT TAT **TGG** GCG GCT CGT TTT AAC GAT GAA GCT

Fig 5.14: Primers used for the construction of rChb CBP mutant W689Y. Outline of the position of the primers used for the mutation of the rChb CBP at residue 689. The forward primer is shown in red and the reverse primer is shown in green. The mutated residue is underlined and in bold.

Table 5.1: Properties of the rChb CBP mutants.

	Molecular Weight (kDa)	Extinction Coefficient	pI
Chb CBP	99.91	1.602	6.66
W63A	99.79	1.548	6.66
K545A	99.85	1.602	6.53
D551A	99.86	1.602	6.81
D648A	99.86	1.602	6.81
W689Y	99.88	1.562	6.66

5.3 Expression and purification of the rChb CBP mutants

The rChb CBP mutants were transformed into *E. coli* JM109 cells for stable storage of the plasmid DNA and *E. coli* KRX cells for the expression of the mutant proteins. The optimised culture conditions for the expression of the rChbL enzyme (see Section 3.6) were also employed here for the expression of the rChb CBP mutants. The rChb CBP mutants were expressed in *E. coli* KRX, in TB, at 30°C and induced with 100 µM IPTG. The expression of the rChb CBP mutant proteins in the cellular fraction and to the culture media was examined (see Figures 5.16 and 5.17). The rChb CBP mutant proteins were collected from the culture media and purified via IMAC (see Figure 5.18) as previously described for the rChbL enzyme (see Section 3.7).

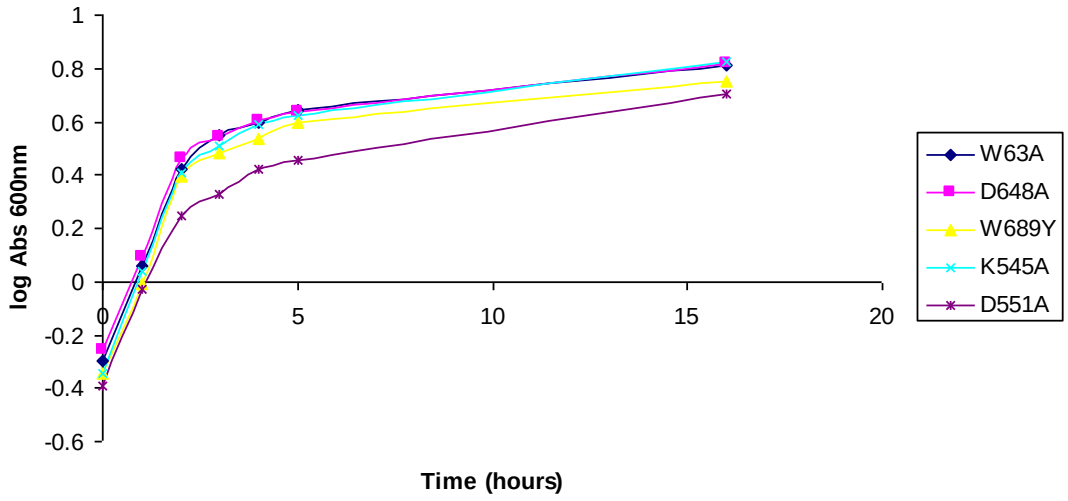


Fig 5.15: Comparison between the growth rates of the rChb CBP mutants. Recombinant protein expression was induced at OD_{600nm} = 0.5. 1 mL samples were taken from the cultures every hour for five hours, following induction and an overnight sample was taken 16 hours post induction (see Section 2.15). Absorbance readings were taken at the different time-points. The log_{Abs} was plotted against time.

The growth rates of each of the rChb CBP mutant cultures were similar to the rChb CBP growth rate (see Figure 4.7). The growth rates of all the rChb CBP mutant cultures were quite similar to each other. The observed cell density of the W63A, D648A and K545A cultures was similar at all time-points, suggesting that these site-specific mutations did not affect the growth rate of the cells. The growth rate of the

W689Y mutant culture was marginally slower while the cell density of the D551A culture at all time-points was lower than all the other mutants.

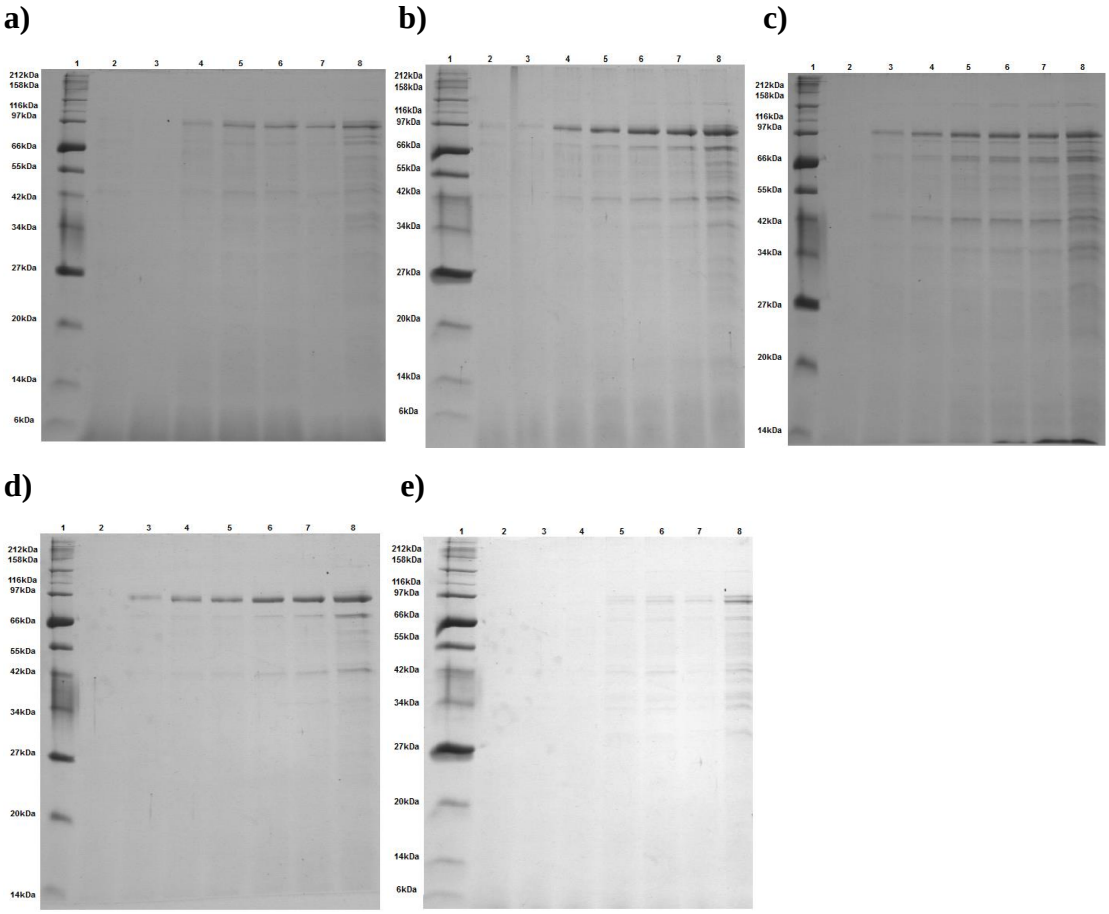


Fig 5.16: Time-course analysis of culture media samples from rChb CBP mutant expression cultures a) W63A b) K545A c) D551A d) D648A and e) W689Y. Recombinant protein expression was induced at OD_{600nm} = 0.5. 5 mL samples were taken from the culture media every hour for five hours, following induction and an overnight sample was taken 16 hours post-induction (see Section 2.17). The samples were then analysed by 12.5% SDS-PAGE. Lane 1, Protein ladder; Lane 2, T=0; Lane 3, T=1 hour; Lane 4, T=2 hours; Lane 5, T=3 hours; Lane 6, T=4 hours; Lane 7, T=5 hours; Lane 8, Overnight (T=16 hours).

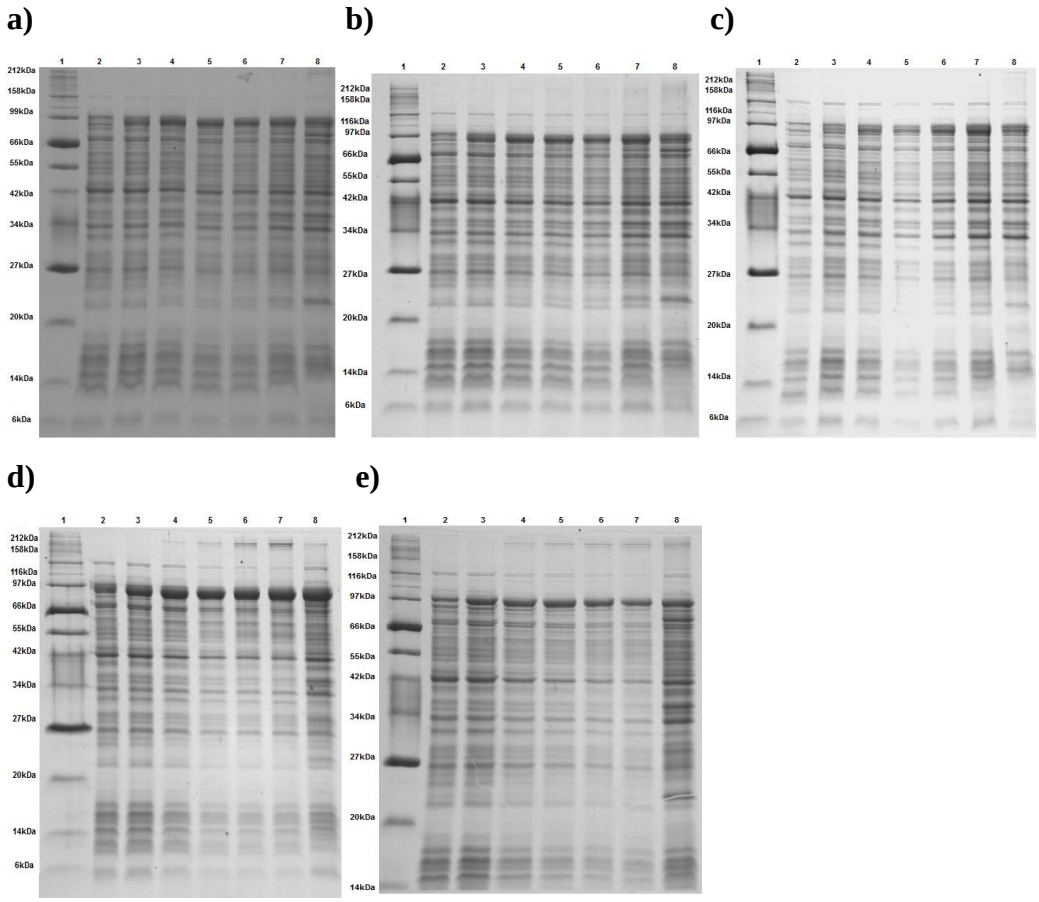


Fig 5.17: Time-course expression analysis of the cellular fractions from rChb CBP mutant cultures a) W63A b) K545A c) D551A d) D648A and e) W689Y. Once the cultures had reached an OD_{600} of 0.5 and had been induced, 1mL samples were taken every hour for 5 hours followed by an overnight sample. Samples from the cellular fractions were obtained (see Section 2.16) and then analysed by 12.5% SDS-PAGE. Lane 1, Protein ladder; Lane 2, T=0 hours; Lane 3, T=1 hour; Lane 4, T= 2 hours; Lane 5, T=3 hours; Lane 6, T=4 hours; Lane 7, T=5 hours; Lane 8, Overnight (T=16 hours).

All of the rChb CBP mutant proteins expressed. Expression of the rChb CBP mutants was similar with that of the native rChbL protein and rChb CBP. Again protein was found in the insoluble cellular fraction and the culture media. As with rChbL and rChb CBP, the rChb CBP mutant proteins began to appear in the culture media two hours post induction. Protein production was optimal five hours after induction. The rChb CBP mutant proteins were therefore collected five hours post-induction. There was a similar level of protein expression observed for all rChb CBP mutants, except W689Y. This mutant protein had a much lower level of expression and the most

protein expression was observed after an overnight incubation (16 hours). This protein was therefore harvested following an overnight incubation.

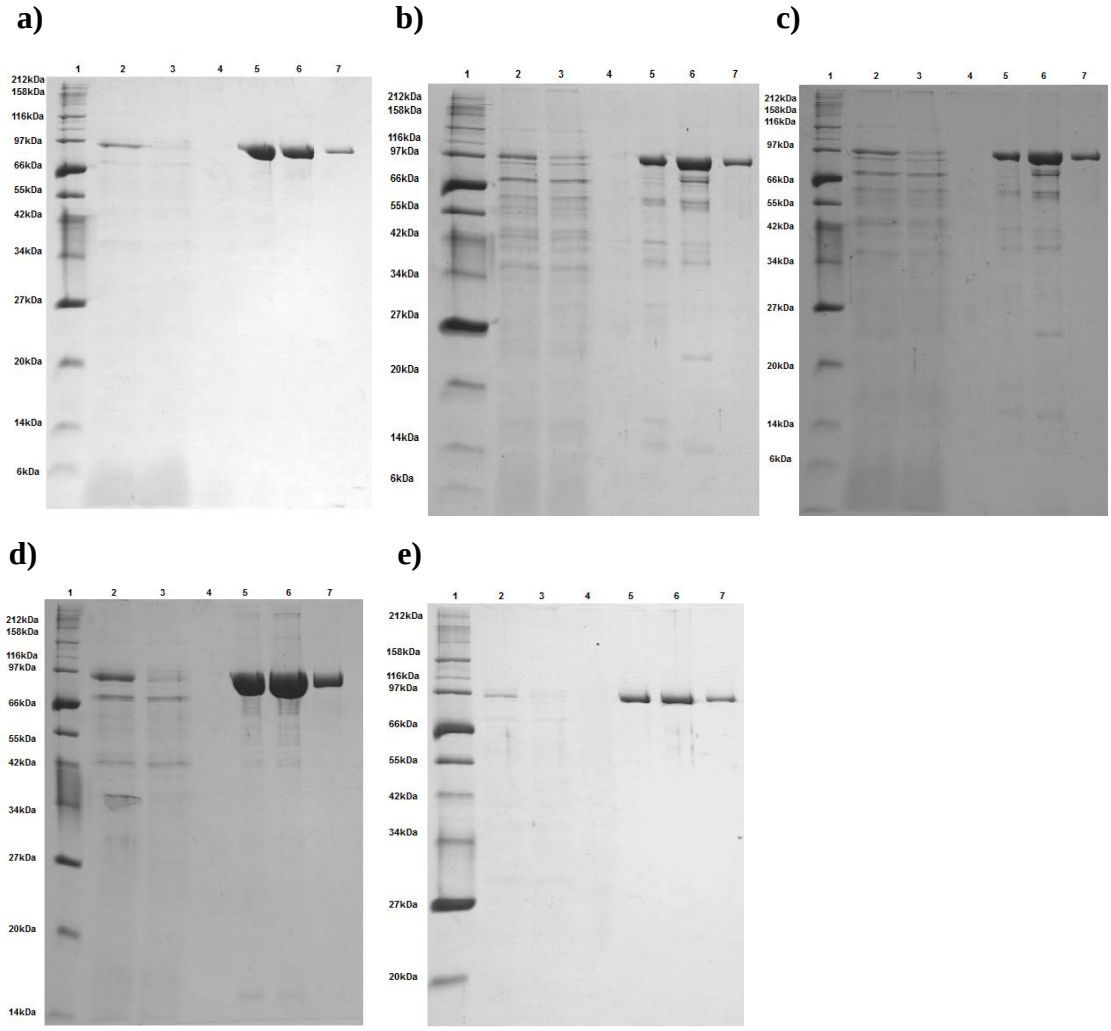


Fig 5.18: Purification of the rChb CBP mutants from the culture media fraction a) W63A b) K545A c) D551A d) D648A and e) W689Y. The culture media was collected after cell harvest from the mutant rChb CBP expression cultures. These were then run over separate IMAC columns, washed with 40 mM imidazole and eluted with 250 mM imidazole. Fractions were run on 12.5% SDS gels.
Lane 1, Protein ladder; Lane 2, Lysate; Lane 3, Flow through; Lane 4, Wash; Lane 5, Elution one; Lane 6, Elution two; Lane 7, Buffer-exchanged sample.

Purification of the rChb CBP mutant proteins was similar to that of the native rChbL enzyme and rChb CBP. Protein was still eluted from the resin at low imidazole concentrations during the washes (40 mM). Yields of protein comparable with the rChbL enzyme and rChb CBP were achieved (approximately 5 mg/250 mL). The

most protein was obtained from the D648A culture while the least protein was obtained from the W689Y culture.

5.4 Characterisation of the rChb CBP mutants

Initially, the ability to detect the rChb CBP mutants in the ELLA format was tested. Access to the His_6 tag was initially tested and unsurprisingly detection was not successful (see Figure 5.19). The mutants were then biotinylated, after which detection with an anti-Biotin antibody was successful (see Figure 5.19). The ability of the rChb CBP mutants to bind glycoproteins was then examined using ELLAs as previously described in Section 2.27. The rChb CBP mutants were tested against GlcNAc residues on asialo agalacto fetuin and also against BSA-GlcNAc. The biotinylated rChb CBP mutants were used in all subsequent ELLAs.

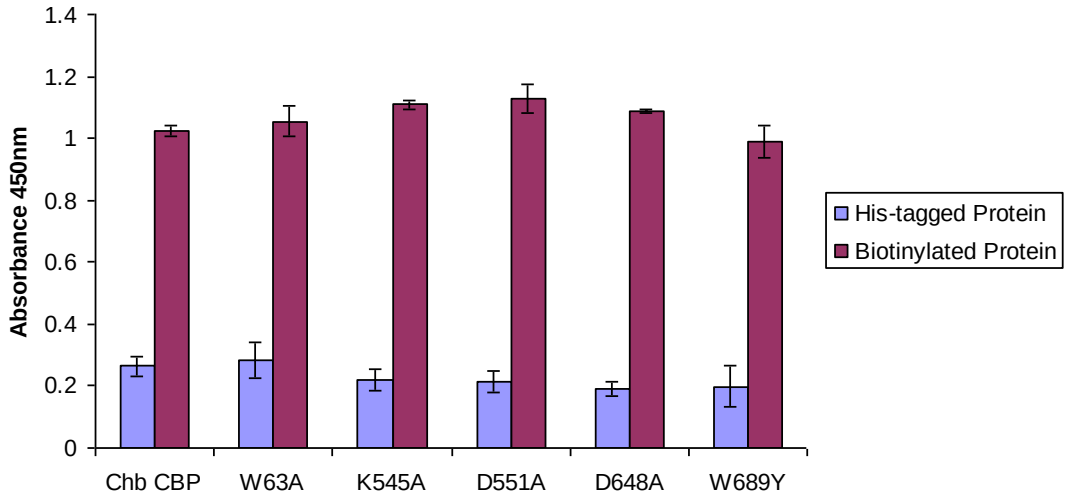


Fig 5.19: Detection of rChb CBP mutants via their His-tag and biotinylation. 50 µL of each mutant protein was laid down on a 96-well plate at 10 µg/mL in triplicate. The plate was blocked as per Section 2.27 and the protein was probed with either an anti-His or anti-Biotin antibody. Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

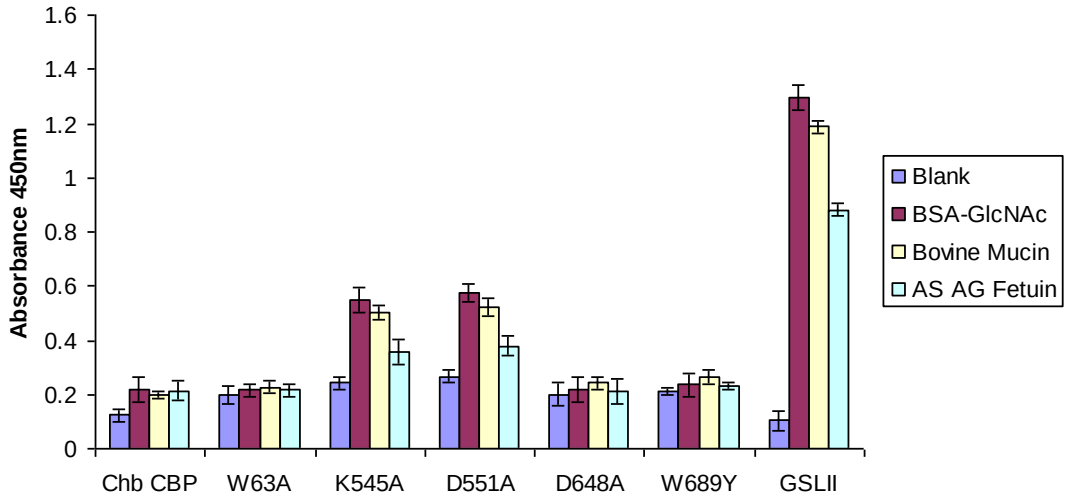


Fig 5.20: ELLA to investigate the binding of the rChb CBP mutants to GlcNAc residues. An ELLA was carried out as per Section 2.27. The rChb CBP mutants were used at a concentration of 10 µg/mL. Commercial GSLII was used as a positive control at a concentration of 10 µg/mL and was detected using an anti-biotin antibody. The rChb CBP mutants were also detected using an anti-Biotin antibody. Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates. An unpaired T-test was used to determine if the increases in binding capabilities were significant in comprison with the rChb CBP (K545A - *$p<0.0115$, D551A - **$p<0.009$).

Differences in the binding abilities of the rChb CBP mutants were observed. Similarly to the rChb CBP, mutants W63A, D648A and W689Y did not bind to GlcNAc residues on asialo agalacto fetuin, BSA-GlcNAc or bovine mucin. However, a significant increase in binding to BSA-GlcNAc and bovine mucin was observed for two of the mutants, K545A and D551A. The signal from the mutants in comparison with rChb CBP was three-fold higher. A slightly higher signal was obtained from D551A than K545A. The signal obtained from GSLII, the positive control, was two-fold higher than the mutants' signal. Binding to asialo agalacto fetuin was also observed, however to a lesser extent. The signal obtained from the mutants was two-fold higher than that of rChb CBP. The binding abilities of K545A and D551A were then tested against increasing concentrations of BSA-GlcNAc. The rChb CBP mutants bound to BSA-GlcNAc in a concentration dependent manner. A slightly higher signal was obtained with D551A (see Figure 5.21). The binding of the rChb CBP mutants

was sugar dependent as inclusion of free GlcNAc sugar (100mM) inhibited binding (see Figure 5.22).

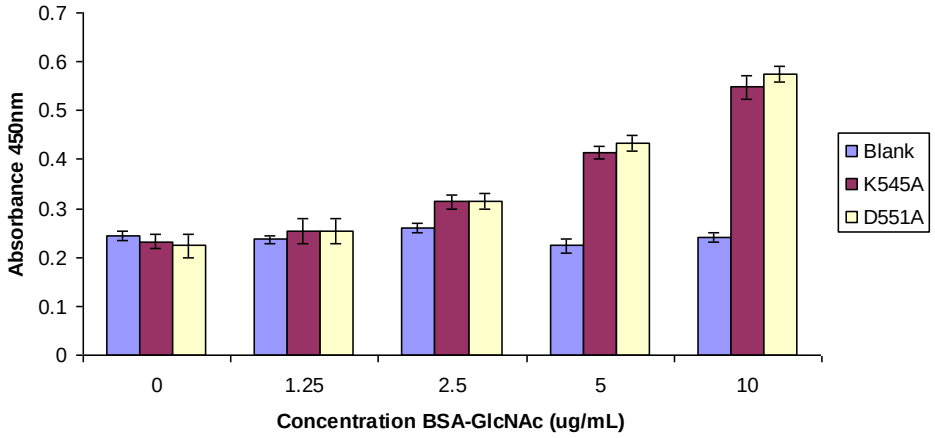


Fig 5.21: ELLA to investigate the binding of the rChb CBP mutants to increasing concentrations of BSA-GlcNAc. An ELLA was carried out as per Section 2.27. BSA-GlcNAc was immobilised at increasing concentrations from 0 – 10 µg/mL. The rChb CBP mutants and commercial GSLII were used at a concentration of 10 µg/mL and detected using an anti-Biotin antibody. Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

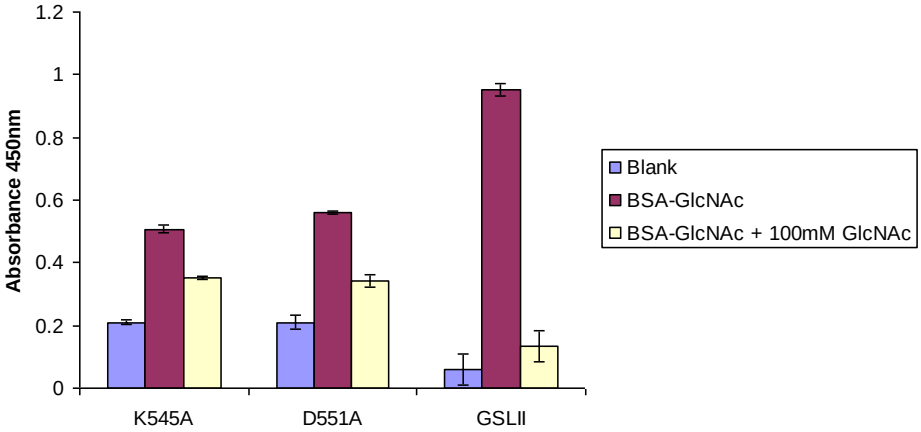


Fig 5.22: Examination of the inhibition of binding to GlcNAc residues with free GlcNAc sugar. An ELLA was carried out as per Section 2.27. BSA-GlcNAc was immobilised at increasing concentrations of 10 µg/mL. The rChb CBP mutants and commercial GSLII were used at a concentration of 10 µg/mL and incubated with 100 mM free GlcNAc for 30 minutes before addition to the plate. They were detected

using an anti-Biotin antibody. Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

5.5 Construction of a double rChb CBP mutant, K545A D551A

As two of the individual site-specific mutations increased the binding abilities of the rChb CBP separately, the next step was to combine these mutants to investigate how the double mutation would affect the binding ability of rChb CBP. While the single mutations may be individually beneficial, the effect of the combination of these mutants is unknown. When mutations are combined the affect on the protein may be synergistic or may be antagonistic (Mildvan *et al*., 1992). The double mutant structure was modelled (see Figure 5.23) and was constructed using the primers depicted in Figure 5.24. The new mutant, titled K545A D551A, was transformed into *E. coli* JM109 cells for stable storage of the plasmid DNA and *E. coli* KRX cells for expression of the protein.

5.5.1 Structural modelling of the rChb CBP double mutant

a) **b)**

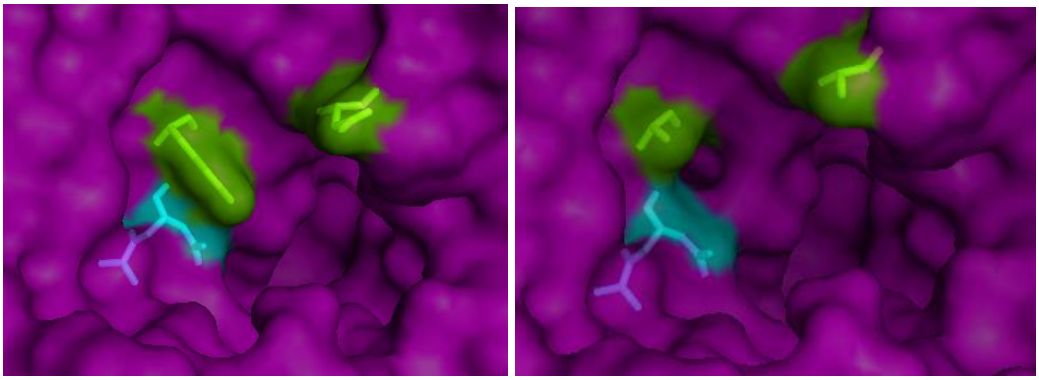


Fig 5.23: Active site of a) rChb CBP and b) rChb CBP double mutant K545A D551A. The catalytic residues, D539A E540D, are highlighted in blue and residues 545 and 551 are highlighted in green. The image was generated using Pymol software (available from http://www.pymol.org/). Mutations were also carried out and visualised in Pymol.

5.5.2 Site-specific mutagenesis of the rChb CBP double mutant

The site-specific mutations were introduced using whole vector amplifications with the desired mutations encoded on the forward primer. PCR was carried out as discussed in Section 2.11.

CCT CGA CTG CGC **GCG** TTA TAG GCA AAC GGA **GCG** GGA TTC CAG
GGA GCT GAC GCG **AAG** AAT ATC CGT TTG GGA **GAC** GGA TTC CAG
GAT AAA AAT GGT CCT ATC
GAT AAA AAT GGT CCT ATC

Fig 5.24: Primers used for the construction of rChb CBP double mutant K545A D551A. Outline of the position of the primers used for the mutation of the rChb CBP at residues 545 and 551. The forward primer 551A_F is shown in red and the reverse primer 551A_R is shown in green. The mutated residue is underlined and in bold.

Table 5.2 Properties of the rChb CBP double mutant K545A D551A

	Molecular Weight (kDa)	Extinction Coefficient	pI
Chb CBP	99.91	1.602	6.66
K545A D551A	99.80	1.603	6.66

5.6 Expression and purification of the rChb CBP double mutant K545A D551A

The rChb CBP double mutant K545A D551A was expressed and purified as previously described for the rChb CBP single mutants. The construct was transformed into *E. coli* JM109 cells for storage of the plasmid DNA and *E. coli* KRX cells for the expression of the mutant protein. The cells were grown in TB, at 30^{o}C and induced with 100 µM IPTG. Both the cellular fraction and culture media of these expression cultures were examined for protein (see Figure 5.25). The protein was purified from the culture media fraction via IMAC (see Figure 5.26) and buffer exchanged into PBS using FPLC. The binding abilities of the double mutant were then examined (see Section 5.7).

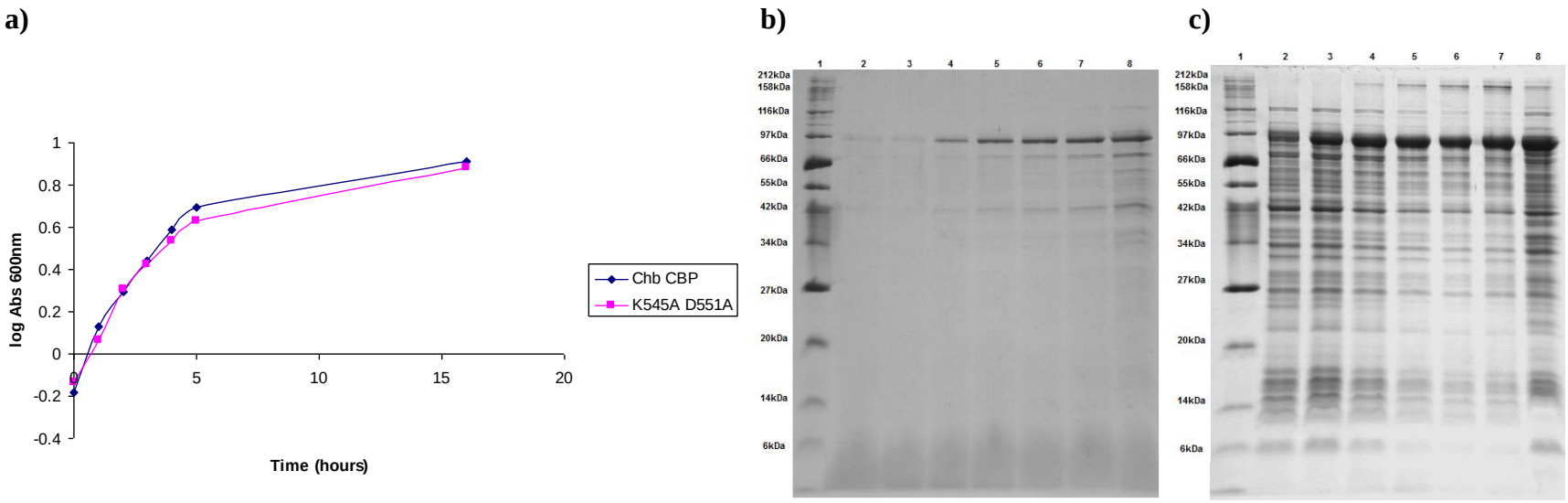


Fig 5.25: a) Analysis of the growth rate of rChb CBP double mutant K545A D551A b) and c) Time-course analysis of rChb CBP double mutant K545A D551A expression culture samples from the culture media (b) and the cellular fraction (c). a) Recombinant protein expression was induced at OD_{600nm} = 0.5. 1 mL samples were taken from the cultures every hour for five hours, following induction and an overnight sample was taken 16 hours post induction. Absorbance readings were taken at the different time-points. The log_{Abs} was plotted against time (hours). b) and c) Recombinant protein expression was induced at OD_{600nm} = 0.5. 5 mL samples were taken from the culture every hour for five hours, following induction and an overnight sample was taken 16 hours post induction. Cellular fraction samples (see Section 2.15) and culture media samples (see Section 2.17) were obtained. The samples were then analysed by 12.5% SDS-PAGE. Lane 1, Protein ladder; Lane 2, T=0; Lane 3, T=1 hour; Lane 4, T=2 hours; Lane 5, T=3 hours; Lane 6, T=4 hours; Lane 7, T=5 hours; Lane 8, Overnight (T=16 hours).

The growth rate of the rChb CBP double mutant culture was slightly lower than the rChb CBP growth rate (see Figure 5.25). The observed cell density of the double mutant culture was lower at all time-points, however this did not affect the expression of the rChb CBP double mutant. Expression of the rChb CBP double mutant was similar with that of the native rChbL protein, rChb CBP and other rChb CBP mutants. Again, protein was found in the cellular fraction and the culture media. As with rChbL, rChb CBP, and the rChb CBP mutant proteins, the rChb CBP double mutant protein began to appear in the culture media two hours post induction. Protein production was optimal after an overnight incubation (16 hours). The rChb CBP double mutant protein were therefore collected 16 hours post-induction. The level of protein produced was similar to the other rChb CBP mutants.

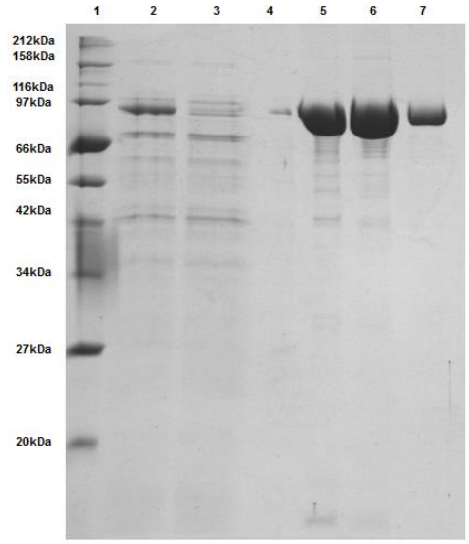


Fig 5.26: Purification of the rChb CBP double mutant K545A D551A from the culture media fraction. The culture media was collected after cell harvest from the rChb CBP double mutant expression cultures. These were then run over separate IMAC columns, washed with 40 mM imidazole and eluted with 250 mM imidazole. Fractions were run on 12.5% SDS gels. Lane 1, Protein ladder; Lane 2, Lysate; Lane 3, Flow through; Lane 4, Wash; Lane 5, Elution one; Lane 6, Elution two; Lane 7, Buffer-exchanged sample.

Purification of the rChb CBP double mutant K545A D551A protein was similar to that of the native rChbL protein, the rChb CBP and the other rChb CBP mutants. Protein was still eluted from the resin at low imidazole concentrations during the washes (40 mM). Relatively high yields of protein comparable with the rChbL protein and rChb CBP were achieved (approximately 6 mg/250 mL).

5.7 Characterisation of the rChb CBP double mutant K545A D551A

The binding ability of the double mutant was examined using ELLAs as described previously (see Section 2.27). The ability to detect the K545A D551A double mutant was examined via the His tag and biotinylation of the protein (see Figure 5.27). The binding of the K545A D551A double mutant was tested against asialo agalacto fetuin, BSA-GlcNAc and Bovine Mucin (see Figure 5.28).

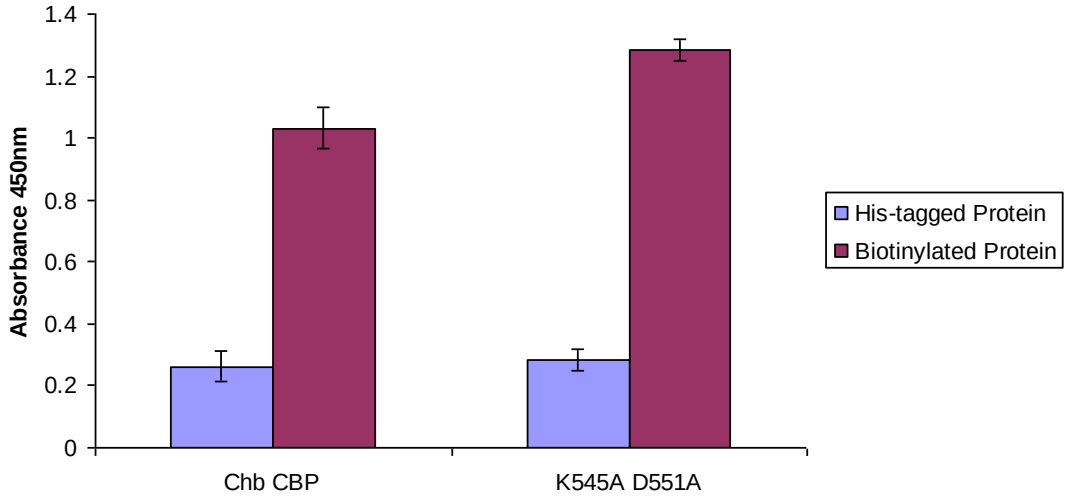


Fig 5.27: Detection of the rChb CBP double mutant K545A D551A via its' His-tag and biotinylation. 50 µL of the double mutant protein was immobilised on a 96-well plate at 10 µg/mL in triplicate. The plate was blocked as per Section 2.27 and the protein was probed with either an anti-His or anti-Biotin antibody. Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

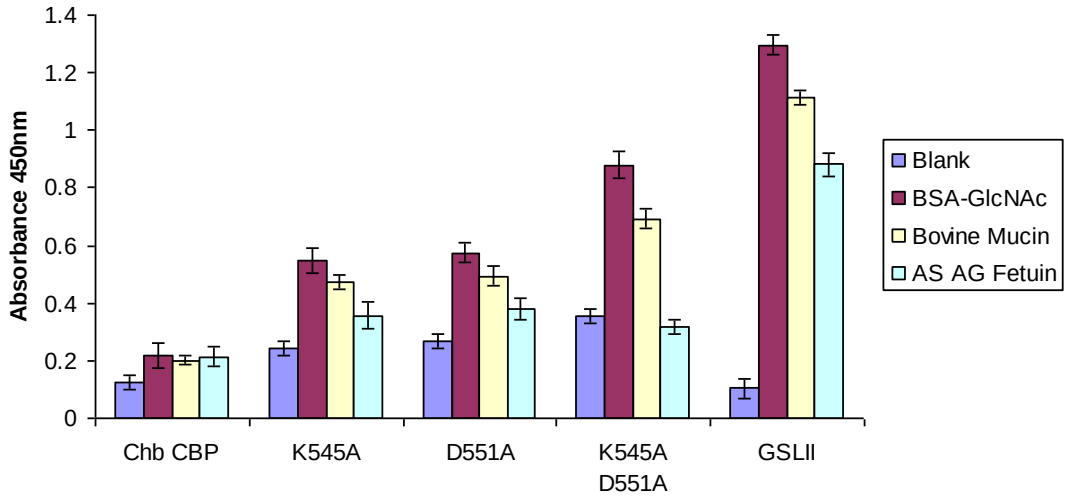


Fig 5.28: ELLA to investigate the binding of the rChb CBP double mutant K545A D551A to GlcNAc residues. An ELLA was carried out as per Section 2.27. Asialo agalacto fetuin, BSA-GlcNAc and Bovine Mucin were immobilised on the plate at 10 µg/mL. The rChb CBP double mutant and commercial GSLII were used at a concentration of 10 µg/mL and were detected using an anti-biotin antibody. Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates. An unpaired T-test was used to determine if the increases in binding capabilities were significant in comprison with the rChb CBP (K545A D551A - ***p<0.0001).

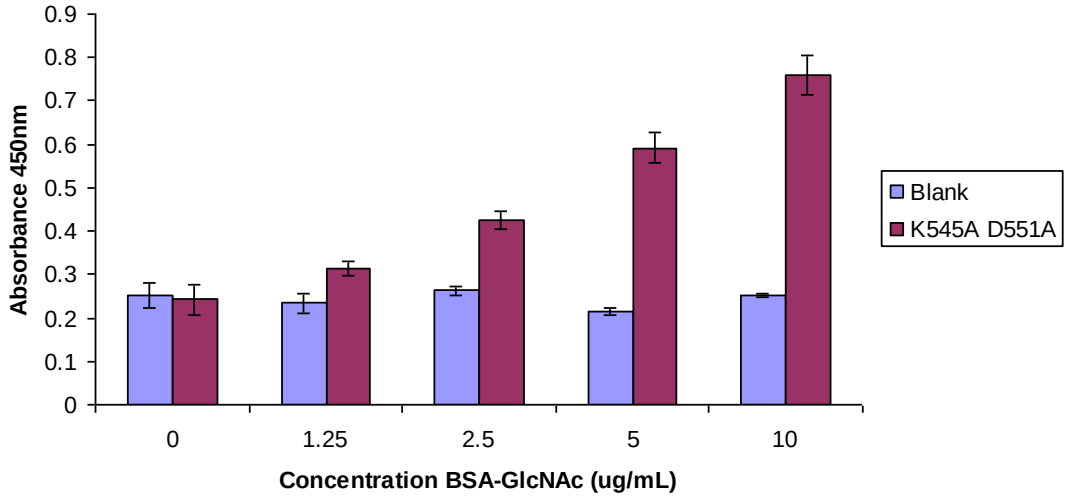


Fig 5.29: ELLA to investigate the binding of the rChb CBP double mutant K545A D551A to increasing concentrations of BSA-GlcNAc. An ELLA was carried out as per Section 2.27. BSA-GlcNAc was immobilised at increasing

concentrations from 0 – 10 μg/mL. The rChb CBP double mutant and commercial GSLII were used at a concentration of 10 μg/mL and were detected using an anti-biotin antibody. Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

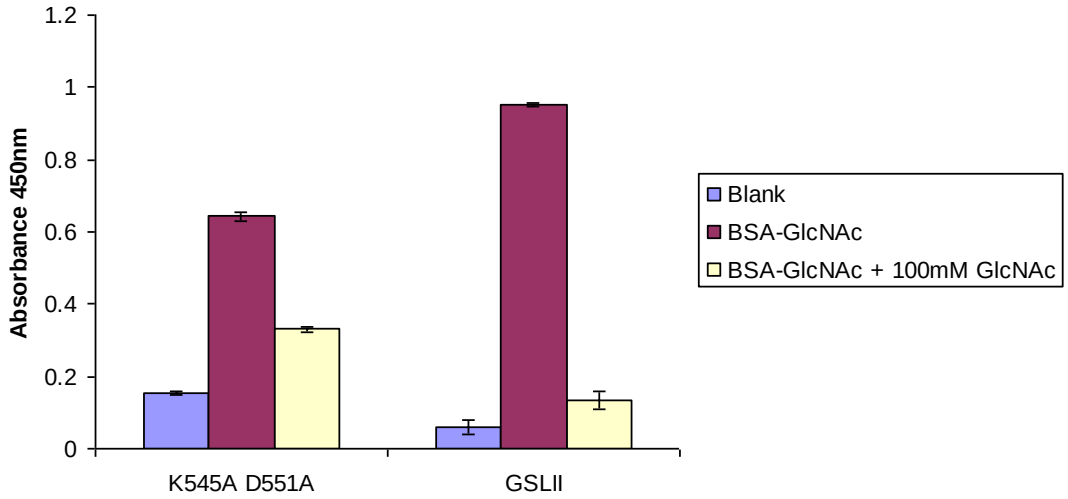


Fig 5.30: Examination of the inhibition of binding of the rChb CBP mutant to GlcNAc residues with free GlcNAc sugar. An ELLA was carried out as per Section 2.27. BSA-GlcNAc was immobilised at increasing concentrations of 10 μg/mL. The rChb CBP double mutant and commercial GSLII were used at a concentration of 10 μg/mL and incubated with 100mM free GlcNAc for 30 minutes before addition to the plate. They were detected using an anti-biotin antibody. Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

The rChb CBP double mutant K545A D551A was not detectable via its' His_6 tag as with the other rChb CBPs, however it was detectable after biotinylation (see Figure 5.27). The biotinylated double mutant protein was therefore used in all subsequent ELLAs. Binding to asialo agalacto fetuin was not noticeably improved. The mutations to the rChb CBP were found to work additively as they significantly increased the binding of the rChb CBP to BSA-GlcNAc and bovine mucin almost two-fold (see Figure 5.28), compared with the single mutations. Binding to asialo agalacto fetuin was not improved. The binding to BSA-GlcNAc was sugar dependent as a high concentration of free GlcNAc sugar inhibited binding (see Figure 5.30).

5.8 Site-specific mutagenesis of the rChbL enzyme

Following the successful mutagenesis of the rChb CBP molecule and the discovery that the K545A and D551A mutations resulted in an increased ability of the protein to bind to GlcNAc epitopes on glycoproteins, the effects of the site-specific mutations on the catalytic activity of the rChbL enzyme was also examined.

The same five site-specific mutations were also carried out on the rChbL enzyme. The residues to be mutated were all present in close proximity to the catalytic site. These residues were mutated to discover more about the catalytic site of the enzyme and to determine which, if any, of these residues had an importance in catalysis.

The site-specific mutations were introduced as previously (see Section 5.8.2); however this time the rChbL enzyme construct pQE60_Chb+ was used as a template. The potential structural effects of the introduced mutations were visualised using Pymol in Figures 5.31 to 5.36.

5.8.1 Structural modelling of the rChbL site-specific mutants

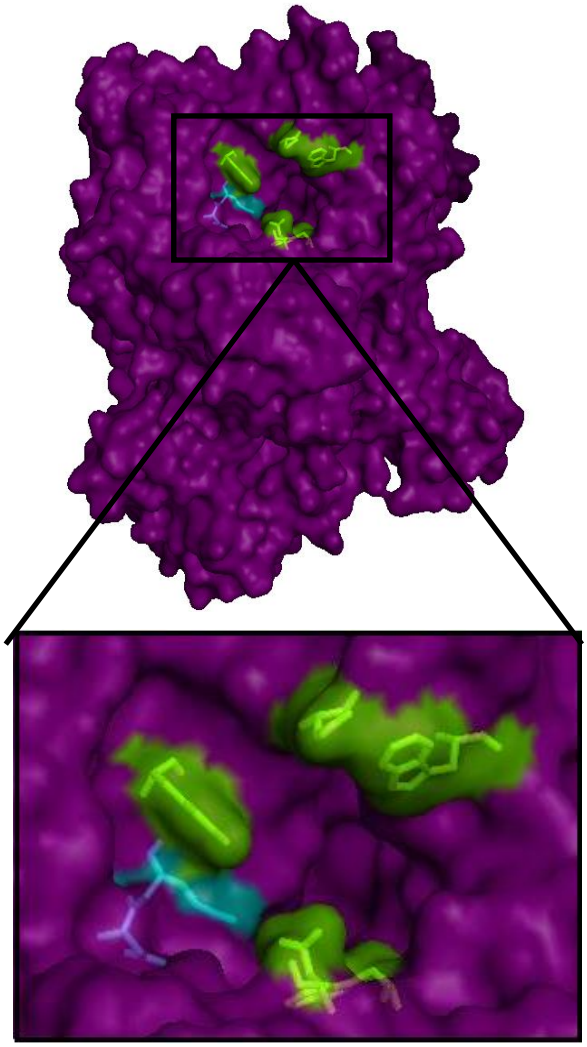


Fig 5.31: Residues chosen for site-specific mutagenesis of the rChbL active site. The catalytic residues, D539 and E540, are highlighted in blue. The residues chosen for mutagenesis, W63, K545, D551, D648 and W689, are highlighted in green. Images were generated using Pymol (available from http://www.pymol.org/). The mutagenesis was performed and visualised in Pymol.

a) **b)**

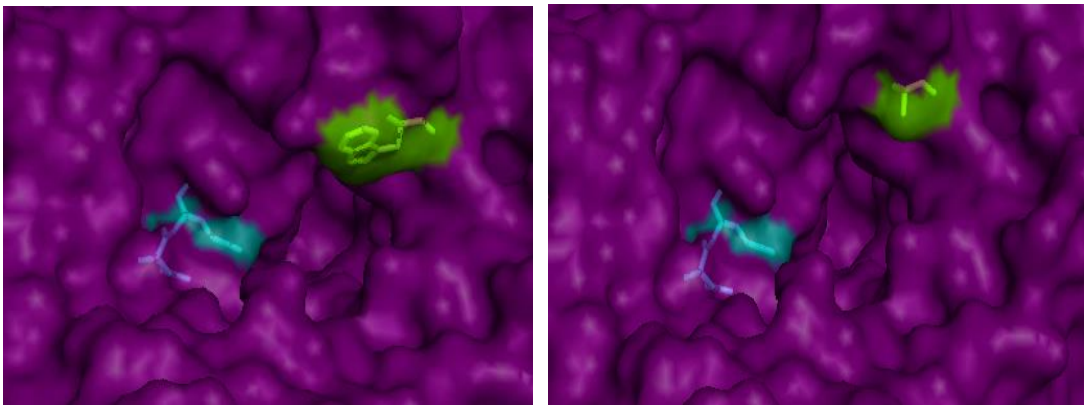


Fig 5.32: Active site of a) rChbL and b) rChbL W63A. The catalytic site residues, D539 and E540, are highlighted in blue and residue 63 is highlighted in green. The image was generated using Pymol software (available from http://www.pymol.org/).

a) **b)**

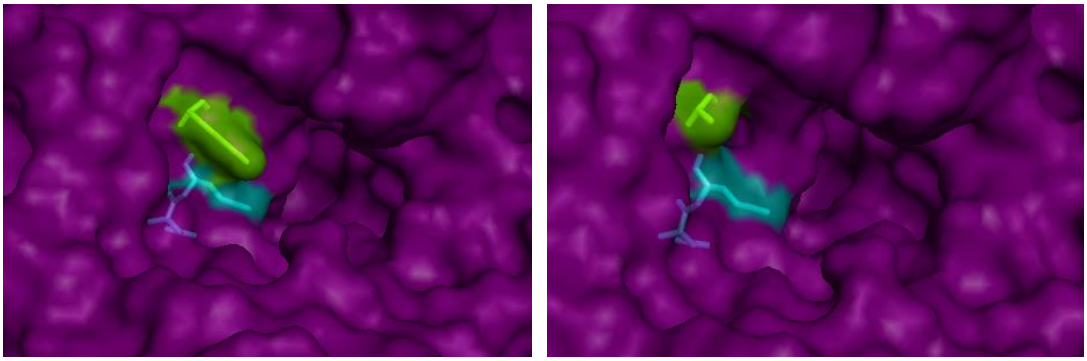


Fig 5.33: Active site of a) rChbL and b) rChbL K545A. The catalytic site residues, D539 and E540, are highlighted in blue and residue 545 is highlighted in green. The image was generated using Pymol software (available from http://www.pymol.org/).

a) **b)**

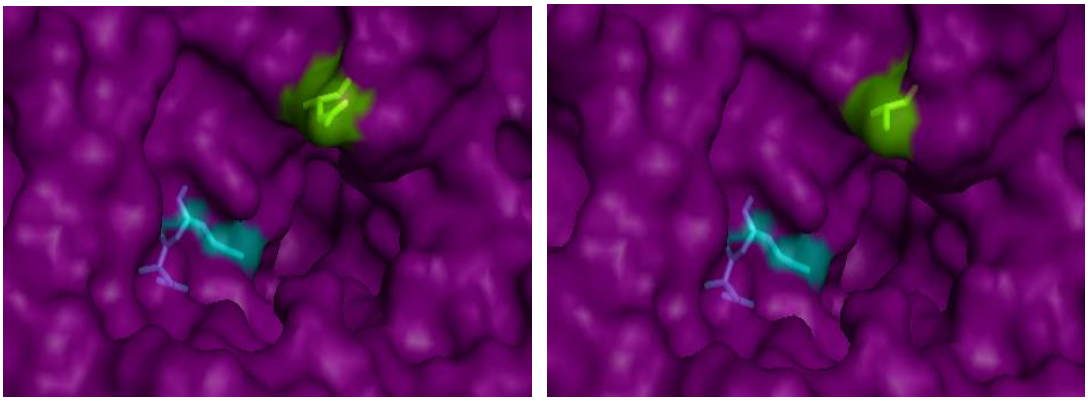


Fig 5.34: Active site of a) rChbL and b) rChbL D551A. The catalytic site residues, D539 and E540, are highlighted in blue and residue 551 is highlighted in green. The image was generated using Pymol software (available from http://www.pymol.org/).

a) **b)**

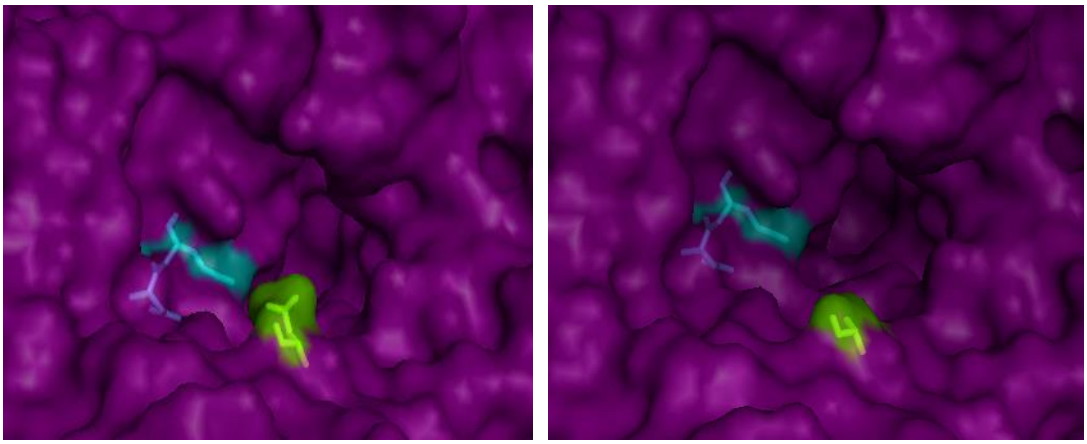


Fig 5.35: Active site of a) rChbL and b) rChbL D648A. The catalytic site residues, D539 and E540, are highlighted in blue and residue 648 is highlighted in green. The image was generated using Pymol software (available from http://www.pymol.org/).

a) **b)**

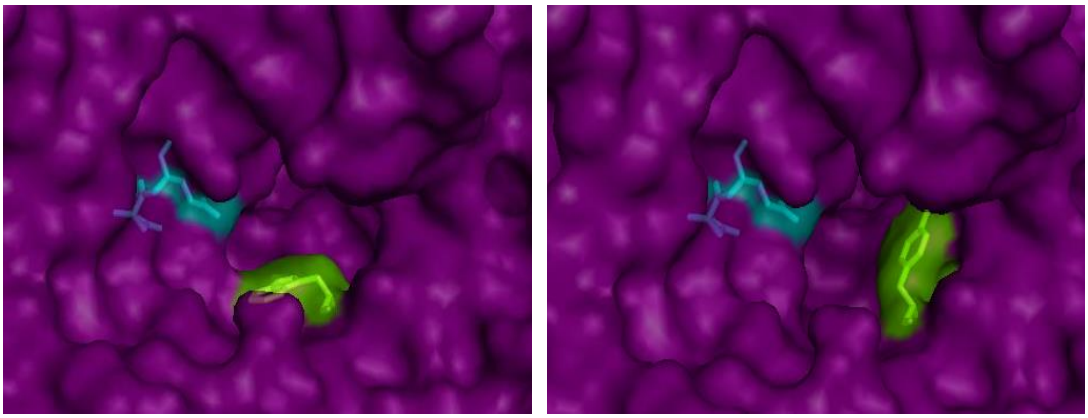


Fig 5.36: Active site of a) rChbL and b) rChbL W689Y. The catalytic site residues, D539 and E540, are highlighted in blue and residue 689 is highlighted in green. The image was generated using Pymol software (available from http://www.pymol.org/).

5.8.2 Site-specific mutagenesis of the rChbL mutants

GTG CCG CAA

GTT AAT TAT AAA GTT GTT GAT AAT CAG GCA GGA AGT CAC GGC GTT

TTG ACA CGC CAG GCA GGG GCA GAT **GCG** GCT AAA TGT AAT CG

AAC TGT GCG GTC CGT GGG GCA GAT **TGG** GCT AAA TGT AAT CGA GTG

Fig 5.37: Primers used for the construction of rChbL mutant W63A. Outline of the position of the primers used for the mutation of the rChbL enzyme at residue 63. The forward primer 63A_F is shown in red and the reverse primer 63A_R is shown in green. The mutated residue is underlined and in bold.

GT GGT AAC CTC TGA

GGT GAA ATT GCT ATG CAT AAA GTA GCA GGT TCA CCA TTG GAG ACT

ACC GTA AAA CCC GGA GAT GAA GCG **GCG** AAT ATC AGT TTG

TGG CAT TTT GGG GGA GAT GAA GCG **AAG** AAT ATC AGT TTG GGA GAC

Fig 5.38: Primers used for the construction of rChbL mutant K545A. Outline of the position of the primers used for the mutation of the rChbL enzyme at residue 545. The forward primer 545A_chb_F is shown in red and the reverse primer 545A_R is shown in green. The mutated residue is underlined and in bold.

GTA AAA CCC CCT CTA CTT CGC TTC TTA ATC CGT TTG GGA

ACT TGG CAT TTT GGG GGA GAT GAA GCG AAG AAT ATC CGT TTG GGA

GCG GGA TTC CAG G

GAC GGA TTC CAG GAT AAA AAT GGT CCT ATC GTA CCG GGT AAG

Fig 5.39: Primers used for the construction of rChbL mutant D551A. Outline of the position of the primers used for the mutation of the rChbL enzyme at residue 551. The forward primer 551A_F is shown in red and the reverse primer 551A_chb_R is shown in green. The mutated residue is underlined and in bold.

CGG

TGG CAG GAT GGT TTA AAA CAT GCT AGC AGT GCT AAA GAT TTT GCC

TGG TTT GCG CAA CCC CAC TTA AAA AGG GAC ACT CTC TAC **GCG** GGC

ACC AAA CGC GTT GGG GTG AAT TTT TGG GAC ACT CTC TAC **GAT** GGC

GGC

GGC TTT GAT TCA GTC AAT GAT TGG GCG AAT AAG GGC TAT GAC GTG

Fig 5.40: Primers used for the construction of rChbL mutant D648A. Outline of the position of the primers used for the mutation of the rChbL enzyme at residue 648. The forward primer 648A_F is shown in red and the reverse primer 648A_R is shown in green. The mutated residue is underlined and in bold.

TA TAC GGT ATA CTT CAA GTA GGT

AGC CCA GAT TAT ATT TAT CTG GAT ATG CCA TAT GAA GTT CAT CCA

TTC TTG CGT GGC TAT TAT **TGT** GCG GCT CGT TTT

CAA GAA CGT GGC TAT TAT **TGG** GCG GCT CGT TTT AAC GAT GAA GCT

Fig 5.41: Primers used for the construction of rChbL mutant W689Y. Outline of the position of the primers used for the mutation of the rChbL enzyme at residue 689. The forward primer is shown in red and the reverse primer is shown in green. The mutated residue is underlined and in bold.

Table 5.3 Properties of the rChbL mutants.

	Molecular Weight (kDa)	Extinction Coefficient	pI
rChbL	99.93	1.601	6.53
W63A	99.84	1.547	6.53
K545A	99.91	1.602	6.42
D551A	99.92	1.601	6.66
D648A	99.92	1.601	6.66
W689Y	99.94	1.561	6.53

5.9 Expression and purification of the rChbL site-specific mutants

After successful cloning of the rChbL site-specific mutant constructs, the plasmids were transformed into *E. coli* JM109 cells for stable storage of the plasmid DNA and *E. coli* KRX cells for expression of the mutant proteins. The optimised culture conditions for the expression of the rChbL enzyme (see Section 3.6) that were used for the expression of the rChb CBP mutants were also employed here for the expression of the rChbL mutant proteins (see Figures 5.42, 5.43 and 5.44). The rChbL mutants were expressed at 30^{o}C in TB, in *E. coli* KRX cells and induced with 100 μM IPTG. The expression of the rChbL mutants in the cellular fraction and the culture media were examined (see Figures 5.42 and 5.43). The mutant proteins were purified via IMAC (see Figure 5.44) and buffer exchanged into PBS using FPLC. The protein was quantified and yields were compared. The catalytic activity of the mutant enzymes was then examined and compared with the rChbL enzyme along with their physical properties.

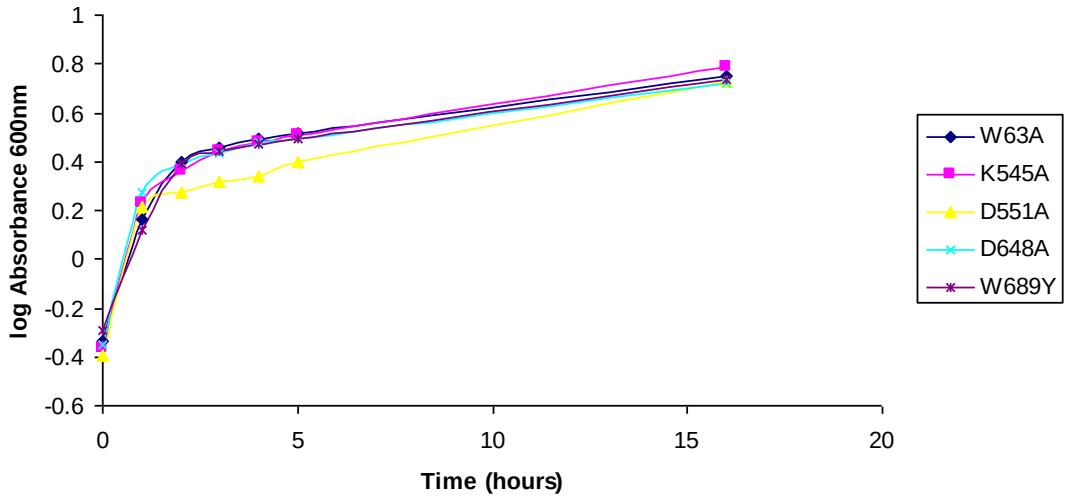


Fig 5.42 Comparison of the growth rates of the rChbL mutants. Recombinant protein expression was induced at OD_{600nm} = 0.5. 1 mL samples were taken from the cultures every hour for five hours, following induction and an overnight sample was taken 16 hours post induction (see Section 2.16). Absorbance readings were taken at the different time-points. The log_{Abs} was plotted against time.

The growth rates of the rChbL mutants were similar to each other (see Figure 5.42) and also to rChbL (see Figure 3.6). The cell densities at each time-point were relatively similar apart from that of D551A. This mutant had a slower rate of growth, with lower cell densities observed from 2 to 5 hours. The final cell density was, however, similar to that of the other rChbL mutants after an overnight incubation.

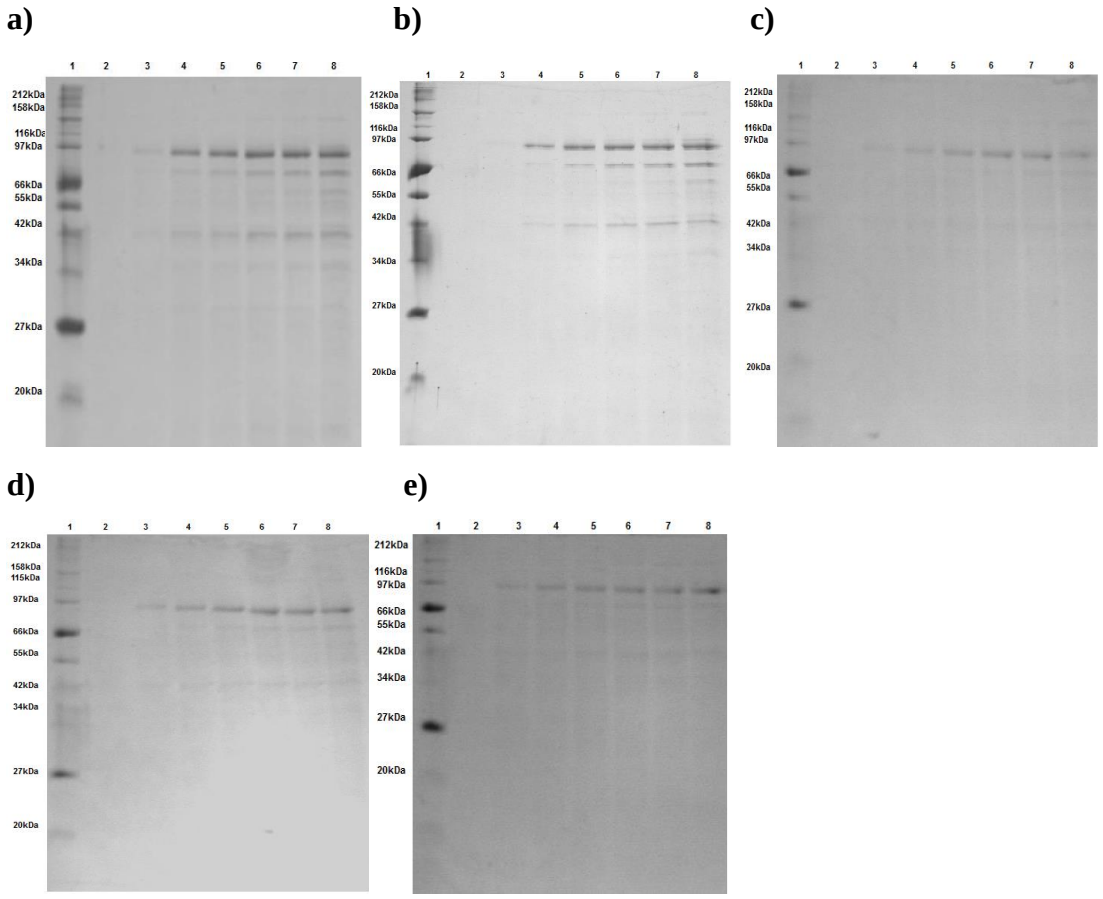


Figure 5.43: Time-course analysis of culture media samples from rChbL mutant expression cultures a) W63A b) K545A c) D551A d) D648A and e) W689Y. Recombinant protein expression was induced at OD_{600nm} = 0.5. 5 mL samples were taken from the culture media every hour for five hours, following induction and an overnight sample was taken 16 hours post-induction (see Section 2.17). The samples were then analysed by 12.5% SDS-PAGE. Lane 1, Protein ladder; Lane 2, T=0; Lane 3, T=1 hour; Lane 4, T=2 hours; Lane 5, T=3 hours; Lane 6, T=4 hours; Lane 7, T=5 hours; Lane 8, Overnight (T=16 hours).

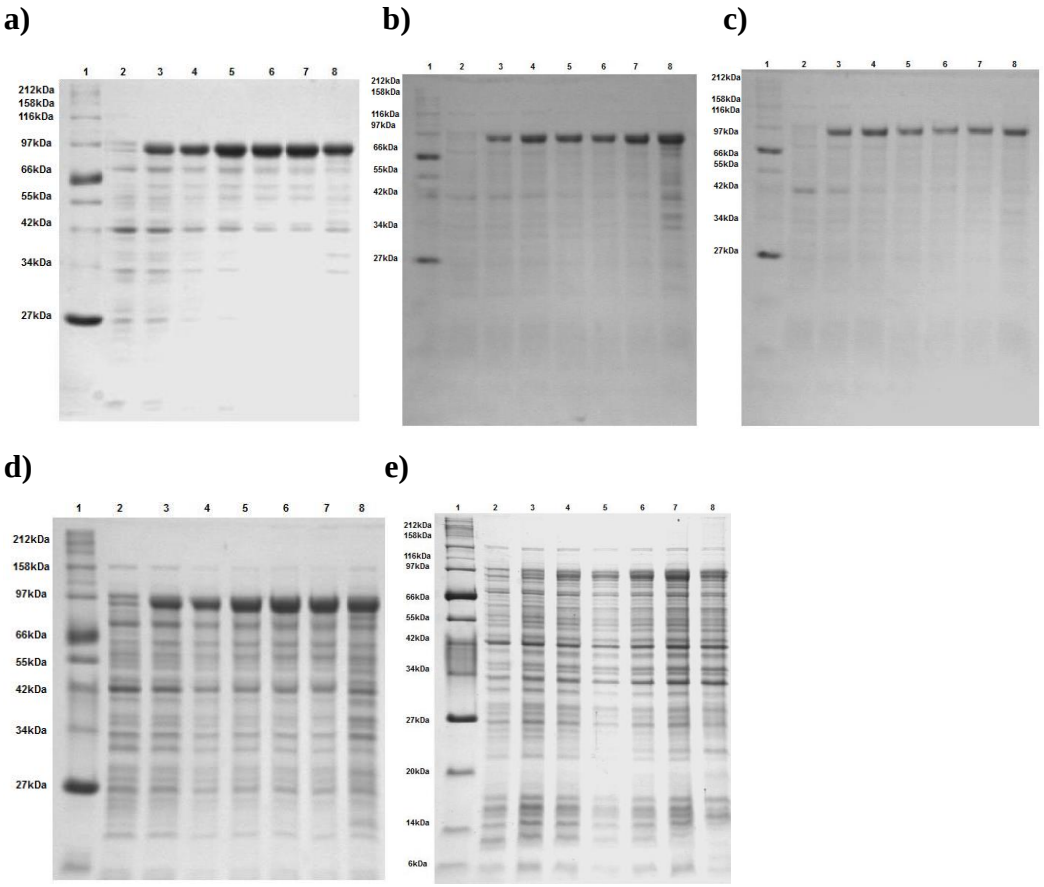


Figure 5.44: Time-course expression analysis of the insoluble cellular fraction from rChbL mutant cultures a) W63A b) K545A c) D551A d) D648A and e) W689Y. Once the cultures had reached an OD_{600} of 0.5 and had been induced, 1 mL samples were taken every hour for 5 hours followed by an overnight sample. Samples from the cellular fractions were obtained (see Section 2.16) and then analysed by 12.5% SDS-PAGE. Lane 1, Protein ladder; Lane 2, T=0 hours; Lane 3, T=1 hour; Lane 4, T= 2 hours; Lane 5, T=3 hours; Lane 6, T=4 hours; Lane 7, T=5 hours; Lane 8, Overnight (T=16 hours).

The rChbL enzyme mutant proteins were found expressed both in the insoluble cellular fraction and in the culture media. The levels of mutant protein expression to the culture media were similar whether the culture was incubated for five hours or incubated overnight. Protein started appearing in the culture media two hours post-induction. Protein was present in the insoluble cellular fraction one hour post-induction. More protein was found in the cellular fraction as was the same with the rChbL, however, protein was again collected from the culture media due to ease of purification and quality of protein obtained.

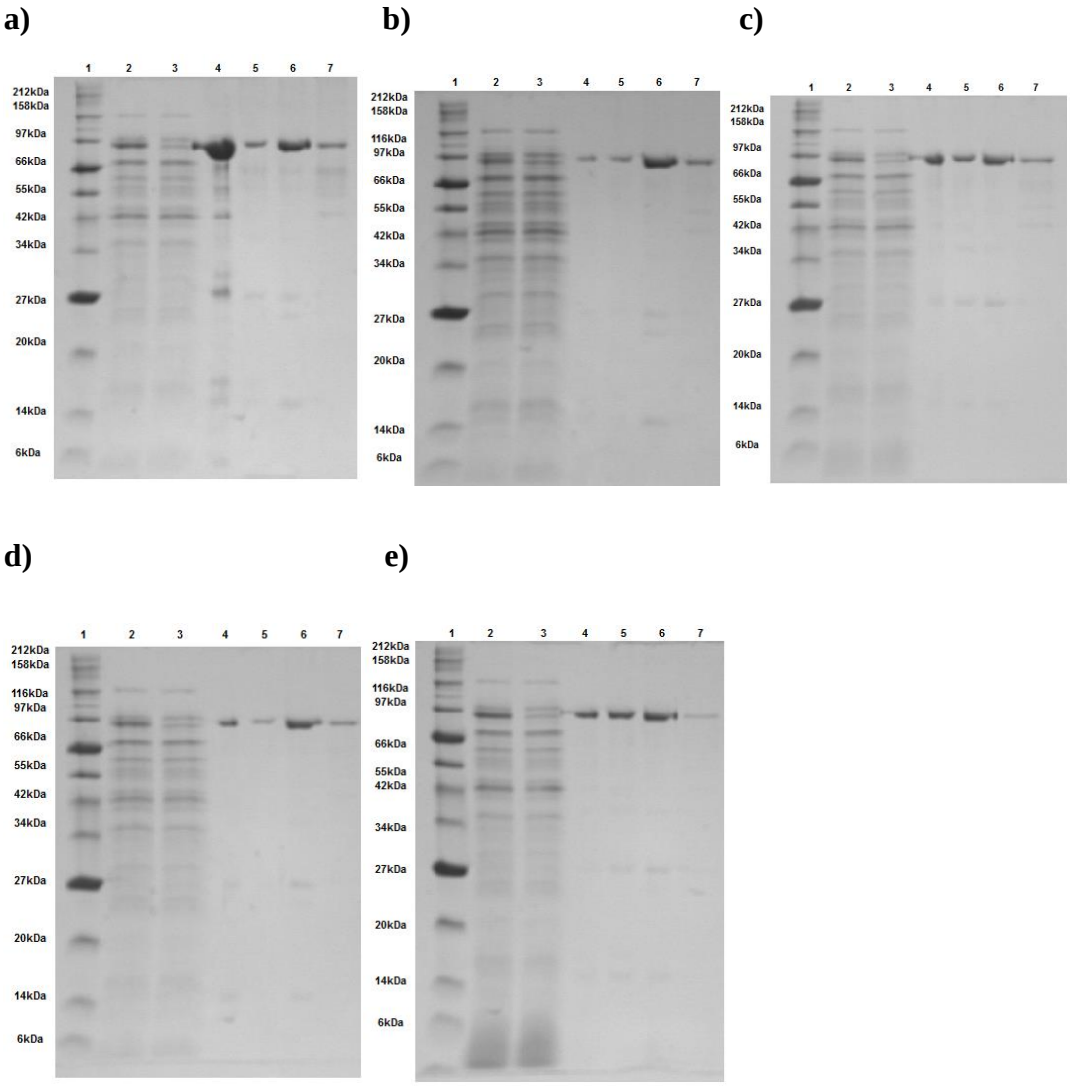


Figure 5.45: Purification of the rChbL mutants from the culture media fraction a) W63A b) K545A c) D551A d) D648A and e) W689Y. The culture media was collected after cell harvest from the mutant rChbL expression cultures. These were then run over separate IMAC columns, washed with 40 mM imidazole and eluted with 250 mM imidazole. Fractions were run on 12.5% SDS gels. Lane 1, Protein ladder; Lane 2, Lysate; Lane 3, Flow through; Lane 4, Wash; Lane 5, Elution one; Lane 6, Elution two; Lane 7, Buffer-exchanged sample

All rChbL mutant proteins were successfully purified and buffer-exchanged. Similar levels of protein were obtained for each mutant. Approximately 6 mg/250 mL culture was obtained for each mutant.

5.10 Characterisation of the rChbL site-specific mutants

The catalytic activity and properties of each of the rChbL mutants was thoroughly examined. The kinetics of each was compared with the rChbL. The specificity of each mutant was examined along with inhibition by various free sugars. Differences in thermodynamic properties were also examined. The Chb assay was carried out as previously described (see Section 2.26) under optimum conditions, 50 mM sodium phosphate pH 8.0, at 37^{o}C for 45 minutes.

5.10.1 Preliminary kinetic characterisation of the rChbL mutants

The kinetics of each mutant was examined using the Michaelis Menten and Hanes-Woolf plots. Each plot was used to calculate the rates of catalytic activity (V_{max}, K_m, k_{cat} and k_{cat}/K_m).

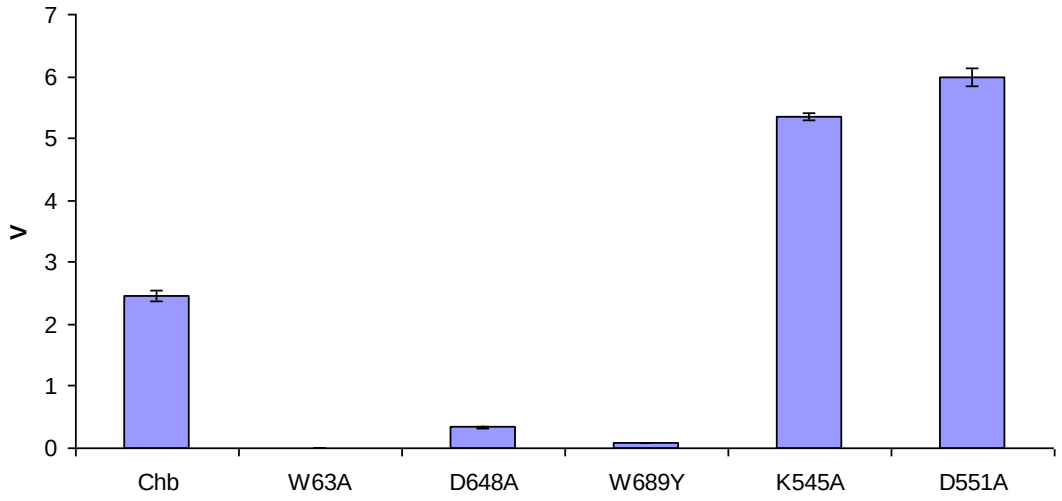


Fig 5.46: Activity assay comparing the rate of catalytic activity between the rChbL and rChbL mutants. The Chb activity assay was performed as per Section 2.26. The rChbL enzyme and rChbL mutant enzymes were used at a concentration of 10 μg/mL. The PNP-GlcNAc substrate was present at 2 mM. The rate, V (mM PNP/min/mL enzyme), of the reaction was plotted. Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates. An unpaired T-test was used to determine if the increases in catalytic activity were significant (K545A - ***$p<0.0001$, D551A - ***$p<0.0001$).

Figure 5.46 demonstrates that the mutations all had very different effects on the catalytic activity of the rChbL enzyme. While mutations W63A, D648A and W689Y reduced the catalytic activity of the rChbL enzyme, mutations K545A and D551A significantly increased the rate of catalytic activity two- fold.

a) **b)**

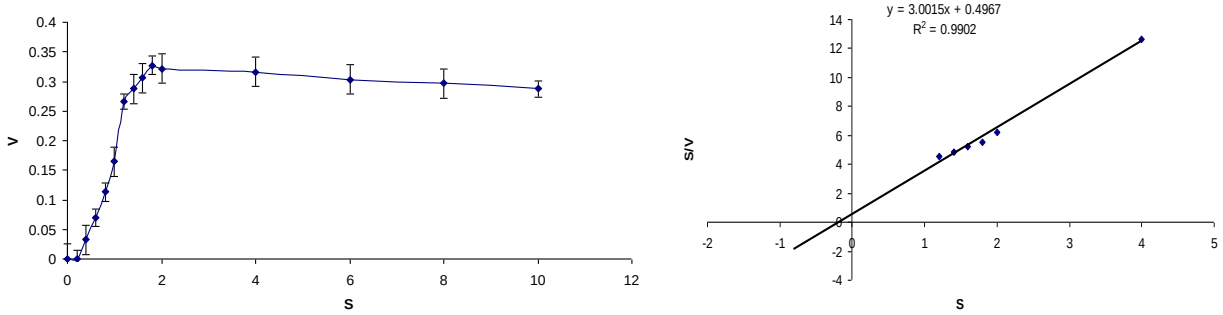


Fig 5.47: Kinetic analysis of the rChbL mutant W63A. The Chb activity assay was performed as per Section 2.26 over a range of substrate concentrations from 0 to 10 mM. The rate, V (mM PNP/min/mL enzyme), was plotted against substrate concentration (S).

a) **b)**

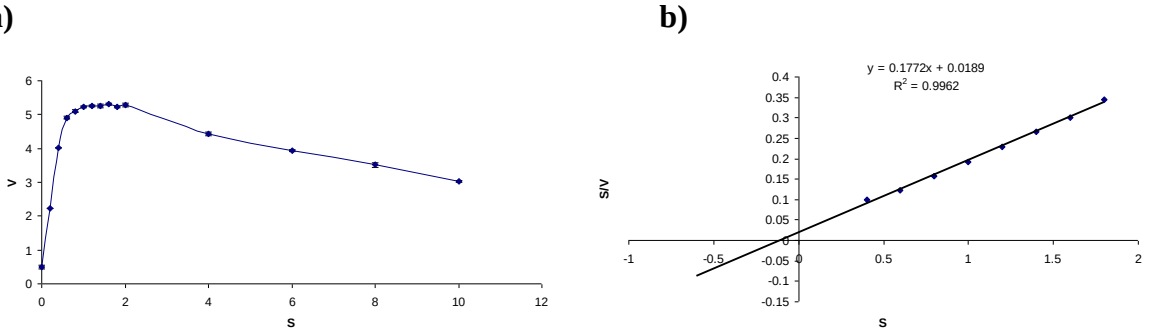


Fig 5.48: Kinetic analysis of the rChbL mutant K545A. The Chb activity assay was performed as per Section 2.26 over a range of substrate concentrations from 0 to 10 mM. The rate, V (mM PNP/min/mL enzyme), was plotted against substrate concentration (S).

a) **b)**

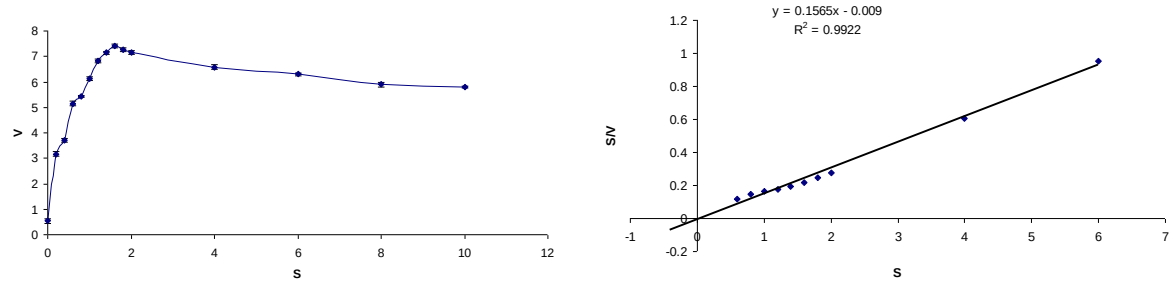


Fig 5.49: Kinetic analysis of the rChbL mutant D551A. The Chb activity assay was performed as per Section 2.26 over a range of substrate concentrations from 0 to 10 mM. The rate, V (mM PNP/min/mL enzyme), was plotted against substrate concentration (S).

a) **b)**

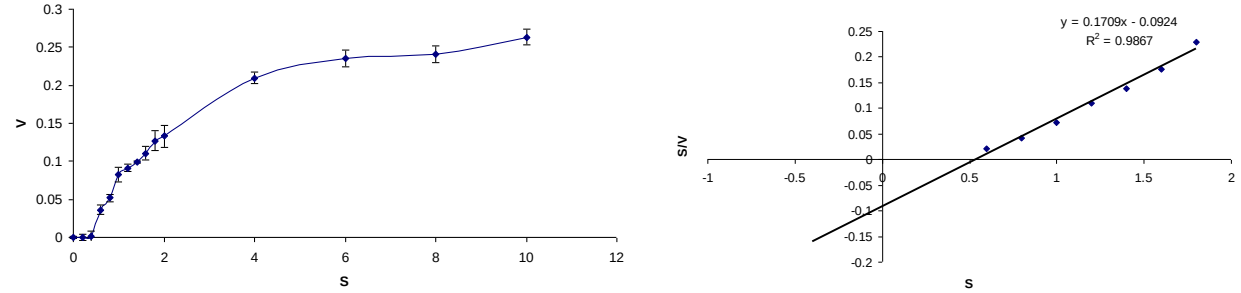


Fig 5.50: Kinetic analysis of the rChbL mutant D648A. The Chb activity assay was performed as per Section 2.26 over a range of substrate concentrations from 0 to 10 mM. The rate, V (mM PNP/min/mL enzyme), was plotted against substrate concentration (S).

a) **b)**

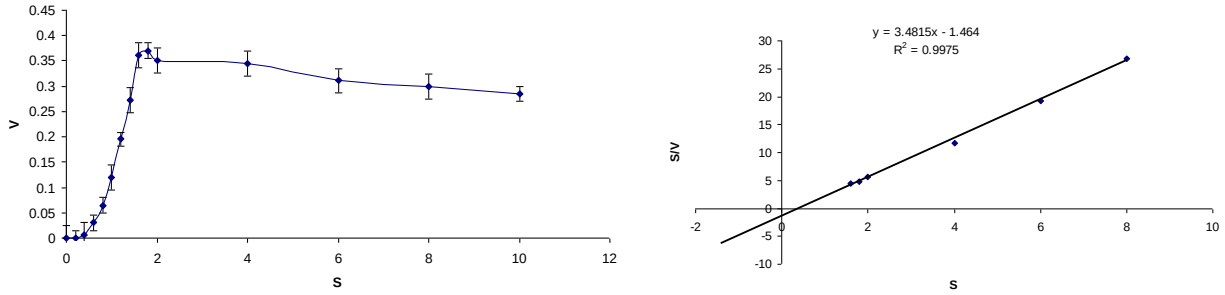


Fig 5.51: Kinetic analysis of the rChbL mutant W689Y. The Chb activity assay was performed as per Section 2.26 over a range of substrate concentrations from 0 to 10 mM. The rate, V (mM PNP/min/mL enzyme), was plotted against substrate concentration (S). Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

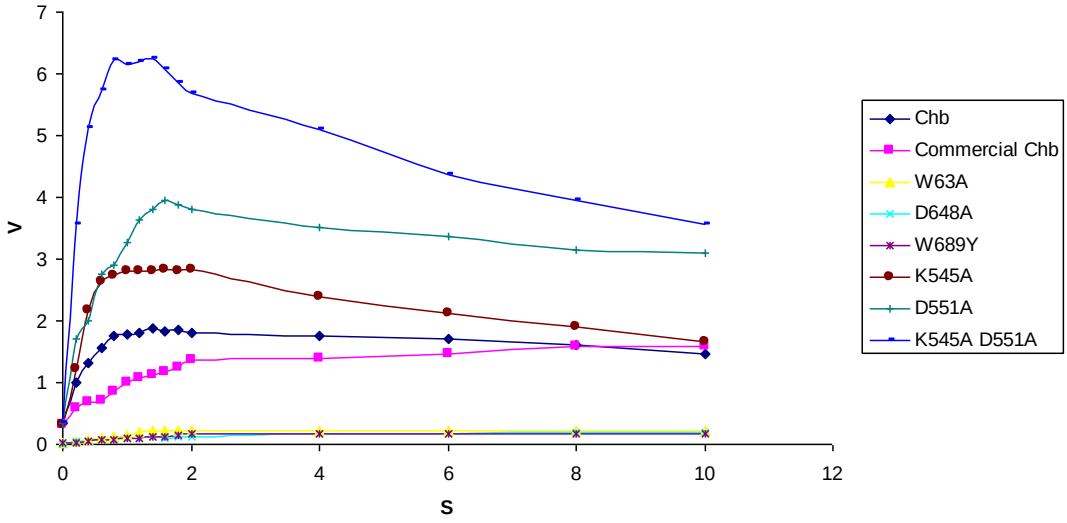


Fig 5.52: Comparison of the catalytic activity of all the rChbL mutants with the rChbL and a commercial Chb. The Chb activity assay was performed as per Section 2.26 over a range of substrate concentrations from 0 to 10 mM. The rate, V (mM PNP/min/mL enzyme), was plotted against substrate concentration (mM). Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

Figures 5.47 to 5.52 demonstrate the differences in catalytic activity between the rChbL enzyme, rChbL mutants and a commercial Chb. The differences in kinetic values are shown in Tables 5.4, 5.5 and 5.6. Mutants W63A, D648A and W689Y reduced catalytic activity in comparison with the native rChbL. Mutants K545A and D551A increased catalytic activity in comparison with the native rChbL. They also had increased catalytic activity in comparison with the commercial Chb.

It can be observed from Table 5.4 that the mutants displaying reduced catalytic activity had a lower V_{max} and higher K_m, calculated using a Michaelis Menten curve. The mutants with increased activity had a higher V_{max} and lower K_m than the rChbL enzyme. The k_{cat} and k_{cat}/K_m of the more active mutants were two orders of magnitude higher than rChbL, while the same values for the less active mutants were slightly lower than the rChbL.

Table 5.4 Michaelis-Menten values for the rChbL mutants.

Michaelis Menten	W63A	K545A	D551A	D648A	W689Y
V_{max} (μM)	327	5300	7420	263	370
K_m (mM)	1	0.22	0.4	2	1.18
k_{cat} (s^{-1})	$3.24x10^5$	$5.25x10^7$	$7.35x10^7$	$2.60x10^5$	$3.66x10^5$
k_{cat}/K_m ($mM^{-1}.s^{-1}$)	$3.24x10^5$	$2.39x10^8$	$1.84x10^8$	$1.30x10^5$	$3.10x10^5$
SD (±)	0.02066	0.04327	0.07082	0.00780	0.02066

Table 5.5 Hanes-Woolf values for the rChbL mutants.

Hanes-Woolf	W63A	K545A	D551A	D648A	W689Y
V_{max} (μM)	333	5649	6925	5851	276
K_m (mM)	0.165	0.106	0.05	0.541	0.51
k_{cat} (s^{-1})	$3.30x10^5$	$5.59x10^7$	$6.86x10^7$	$5.79x10^6$	$2.73x10^5$
k_{cat}/K_m ($mM^{-1}.s^{-1}$)	$1.99x10^6$	$5.25x10^8$	$1.37x10^9$	$1.07x10^7$	$5.36x10^5$
SD (±)	0.02066	0.04327	0.07082	0.00780	0.02066

5.10.2 Characterisation of the rChbL mutants' properties

The properties of the rChbL mutants were examined in the following section. The stability of the mutants at various temperatures was examined in Figure 5.53 while the stability of the mutants at various pH values was examined in Figure 5.54.

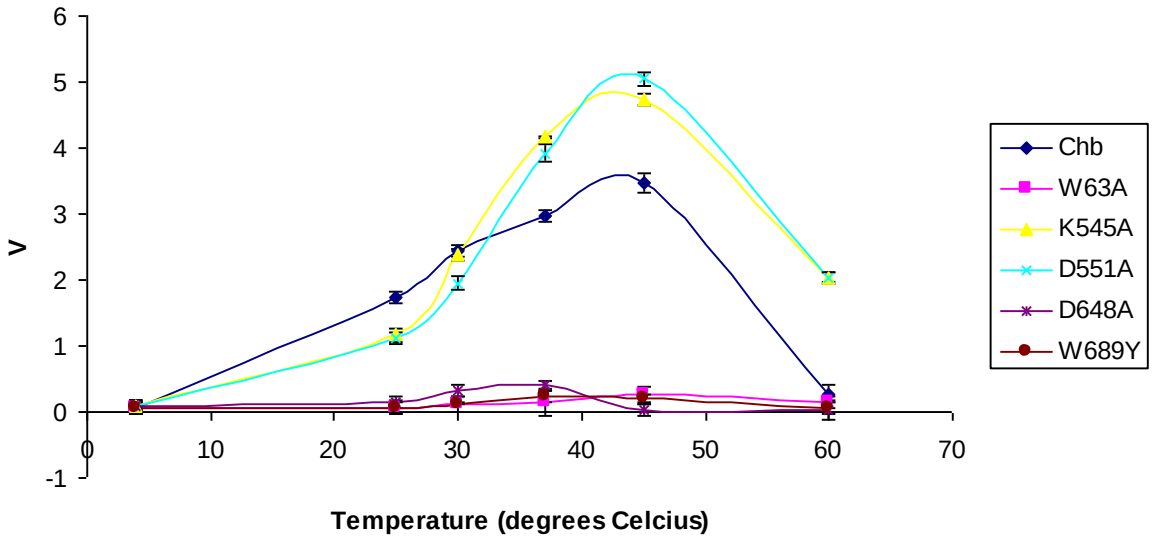


Fig 5.53: The effect of temperature on the catalytic activity of the rChbL mutants. The Chb activity assay was performed as per Section 2.26 over a range of different temperatures, from 4^{o}C to 60^{o}C. The rate, V (mM PNP/min/mL enzyme), was plotted against temperature (^{o}C). Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

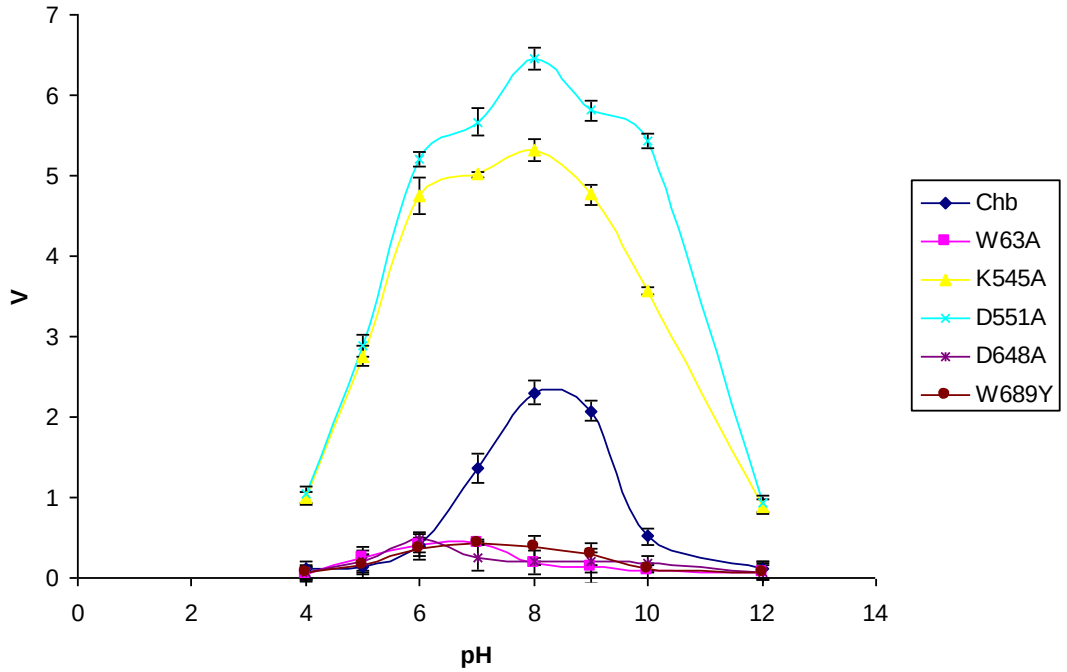


Fig 5.54: The effect of pH on the catalytic activity of the rChbL mutants. The Chb activity assay was performed as per Section 2.26 over a wide pH range, from 4.0 to 12.0. The rate, V (mM PNP/min/mL enzyme), was plotted against pH. Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

The temperature profile of the rChbL mutants was similar to that of the rChbL enzyme. There was relatively little activity at 4^{o}C. As the temperature increased so too did the rate of catalytic activity. The optimum rate of catalytic activity was reached just before 45^{o}C, after which the rate steadily declined. There was very little catalytic activity observed at 60^{o}C. The pH profiles of the mutants were slightly different than that of the rChbL enzyme. While the catalytic activity of the rChbL enzyme peaked at pH 8.0 and declined rapidly after, the mutants with increased catalytic activity (K545A and D551A) were also more stable at higher pH values. They retained the majority of their catalytic activity at pH 9.0 and half of their catalytic activity at pH 10.0. Catalytic activity only sharply declined at pHs higher than 10.0. There was very little catalytic activity observed from the reduced activity mutants, which meant that the optimum pH and temperature for catalytic activity were unable to be determined.

Substrate specificity studies were carried out using a range of PNP-linked sugar substrates including β-GlcNAc, β- and α-glucose, β-galactose and sialic acid (see Figures 5.55 and 5.56). Following these experiments sugar inhibition studies on PNP-GlcNAc cleavage was also carried out using a range of free sugars. The specificity studies and sugar inhibition studies were only carried out on the mutants with observed increases in activity, K545A and D551A, as the rate of catalytic activity of the other mutants was too low.

It can be clearly observed from Figures 5.55 and 5.56 that the specificity of the rChbL mutants had not differed from the rChbL enzyme (see Figure 3.30). The GlcNAc substrate was the only substrate cleaved; there was relatively no cleavage of any of the other substrates examined.

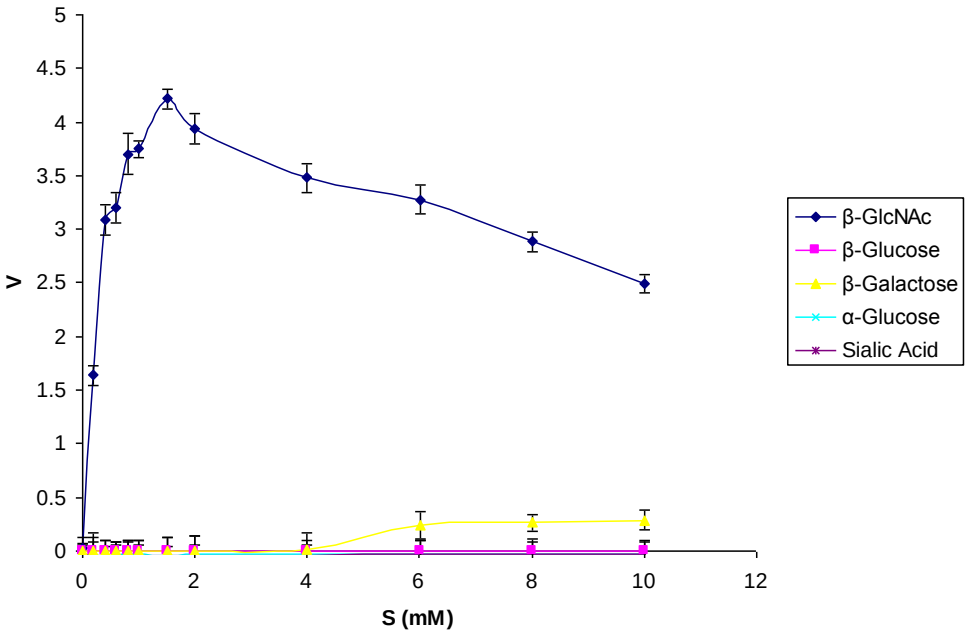


Fig 5.55: Specificity of the rChbL mutant K545A. The Chb activity assay was performed as per Section 2.26 against a wide range of PNP-linked sugar substrates. The rate, V (mM PNP/min/mL enzyme), was plotted substrate concentration (mM). Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

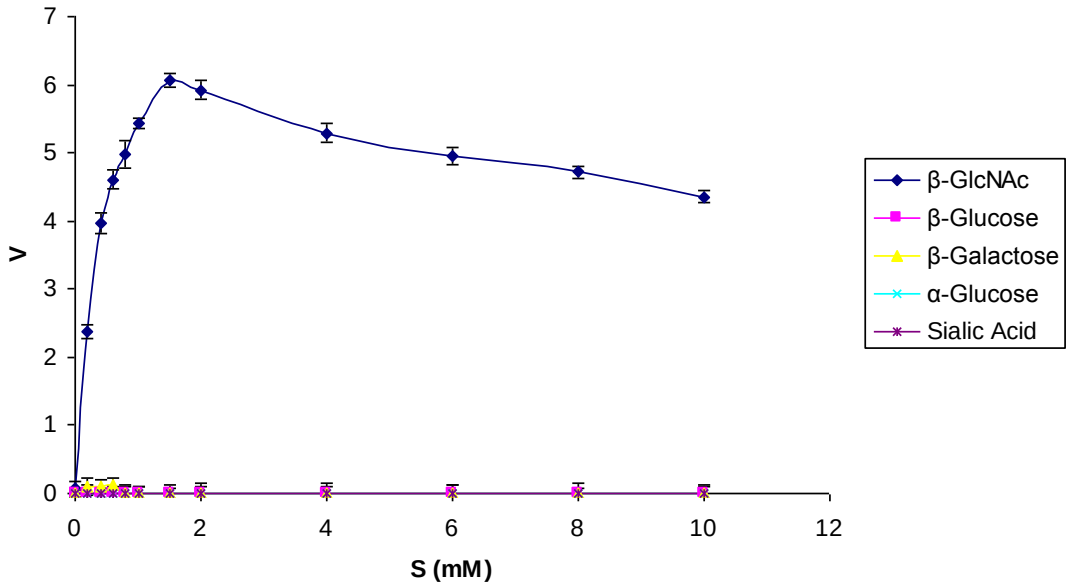


Fig 5.56: Specificity of the rChbL mutant D551A. The Chb activity assay was performed as per Section 2.26 against a wide range of PNP-linked sugar substrates. The rate, V (mM PNP/min/mL enzyme), was plotted substrate concentration (mM). Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

The pattern of inhibition by free sugars was similar between rChbL mutants K545A (see Figure 5.57), D551A (see Figure 5.58) and rChbL itself (see Figure 3.31). While fucose, xylose and galactose caused no inhibition, lactose, mannose, glucose and GlcNAc all caused inhibition to some extent. As with rChbL the most inhibition was observed with GlcNAc, while mannose caused the second highest level of inhibition. As observed for rChbL, lactose was the third best inhibitor for both mutants, K545A and D551A, with galactose the weakest of the inhibiting sugars.

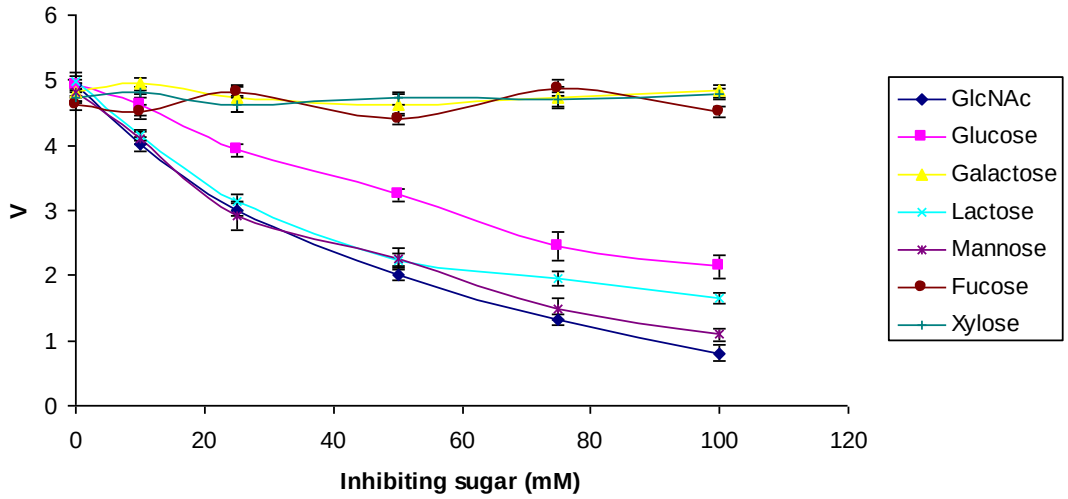


Fig 5.57: The inhibitory effect of free sugars on the rate of rChbL mutant K545A catalytic activity on PNP-GlcNAc. The Chb activity assay was performed as per Section 2.26 with PNP-GlcNAc. The K545A mutant enzyme was used at a concentration of 1 µg/mL and incubated with increasing concentrations of the inhibiting sugar for 30 minutes at 37°C prior to addition to the plate. The rate, V (mM PNP/min/mL enzyme), was plotted substrate concentration (mM). Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

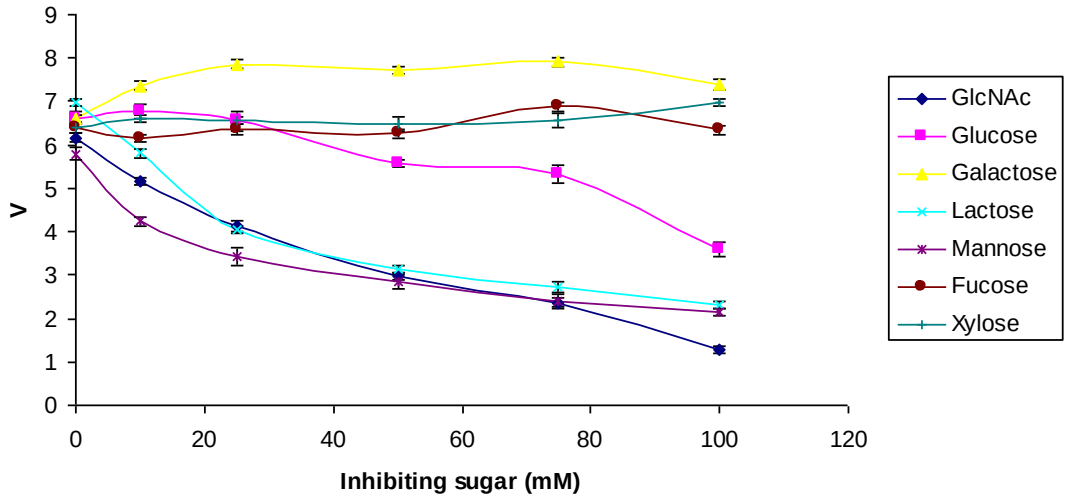


Fig 5.58: The inhibitory effect of free sugars on the rate of rChbL mutant D551A catalytic activity on PNP-GlcNAc. The Chb activity assay was performed as per Section 2.26 with PNP-GlcNAc. The D551A mutant enzyme was used at a concentration of 1 µg/mL and incubated with increasing concentrations of the

inhibiting sugar for 30 minutes at 37^{o}C prior to addition to the plate. The rate, V (mM PNP/min/mL enzyme), was plotted substrate concentration (mM). Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

Lineweaver-Burk plots were generated to describe the inhibition occurring between the rChbL mutants and the free sugars GlcNAc, glucose, mannose and lactose (see Figures 5.59 and 5.60). Tables 5.7 and 5.8 were generated from the results by the calculating 1/V intercept and the slope.

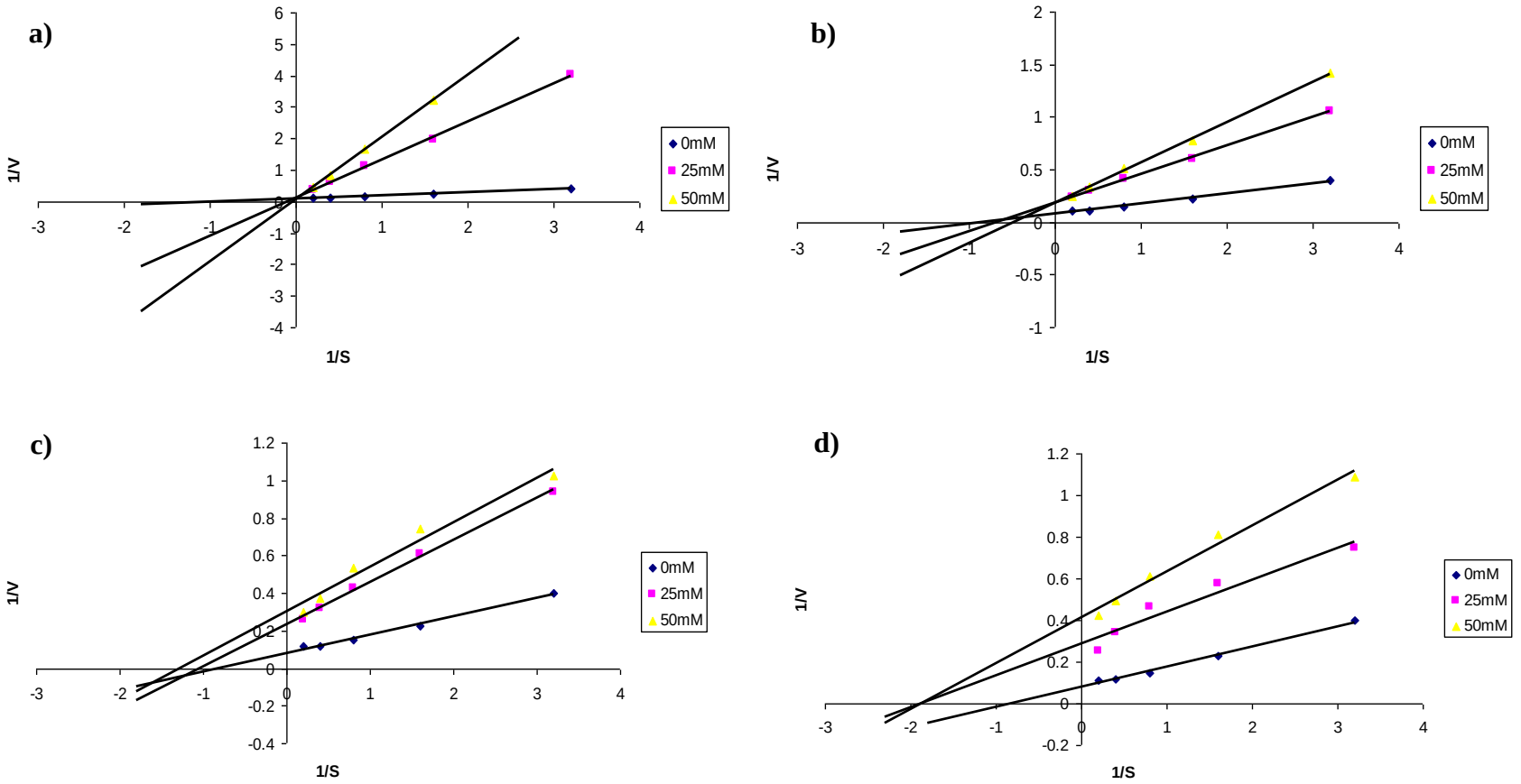


Fig 5.59: Lineweaver-Burk plots of the inhibition occurring with the K545A rChbL mutant and a) GlcNAc, b) glucose, c) mannose and d) lactose. The Chb activity was performed as described in Section 2.26. Increasing concentrations of PNP-GlcNAc were incubated with a constant concentration of enzyme (1 μg/mL) and either 25 mM or 50 mM inhibiting sugar. The reciprocal rate of the reaction was plotted against the reciprocal of the substrate concentration. Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

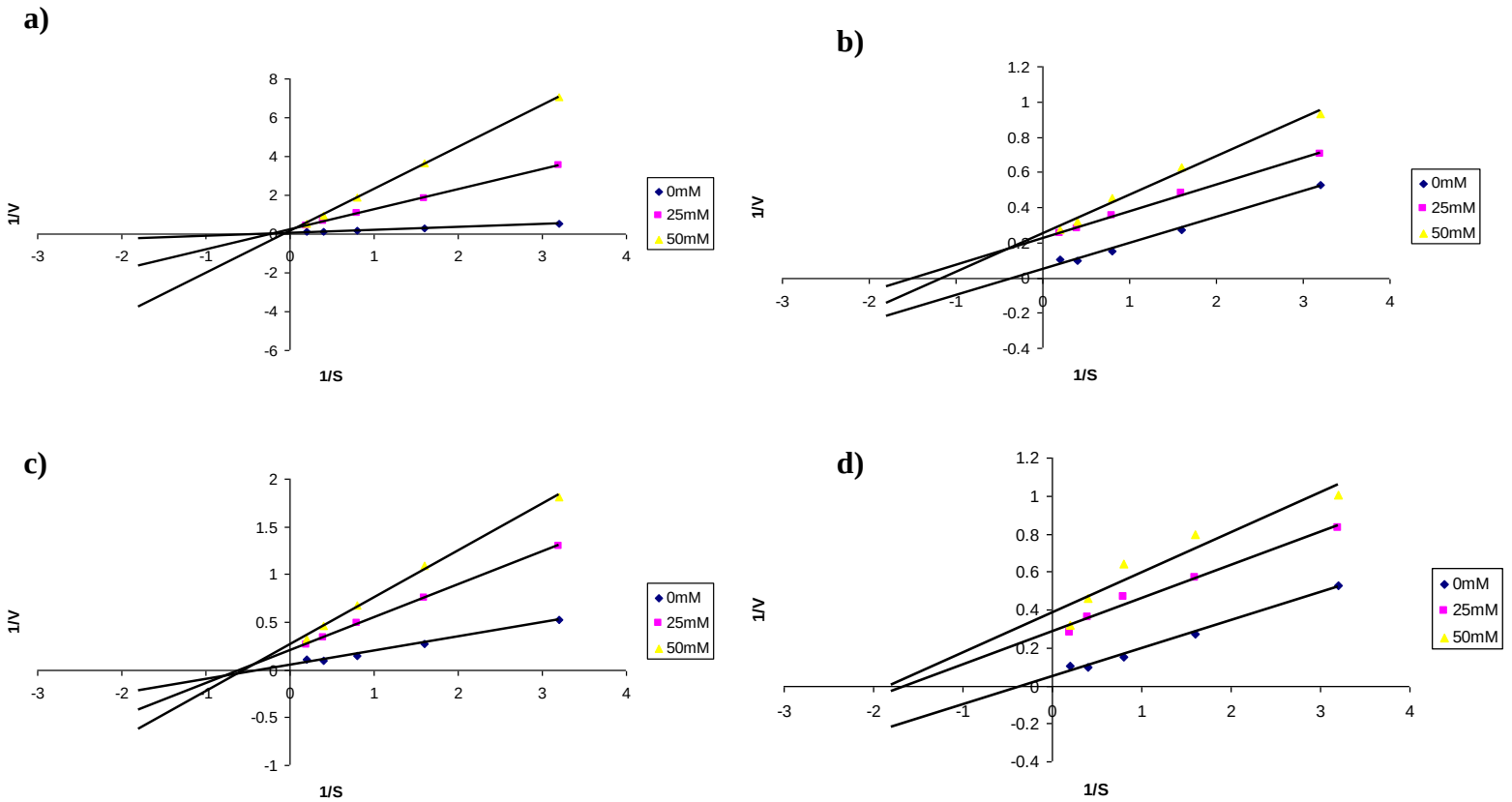


Fig 5.60: Lineweaver-Burk plots of the inhibition occurring with the D551A rChbL mutant and a) GlcNAc, b) glucose, c) mannose and d) lactose. The Chb activity was performed as described in Section 2.26. Increasing concentrations of PNP-GlcNAc were incubated with a constant concentration of enzyme (1 µg/mL) and either 25 mM or 50 mM inhibiting sugar. The reciprocal rate of the reaction was plotted against the reciprocal of the substrate concentration. Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

Table 5.6 Kinetic values describing the inhibition occurring between the rChbL mutant K545A- PNP-GlcNAc reaction and GlcNAc, glucose, mannose and lactose.

	1/V intercept	Slope
GlcNAc (No sugar)	0.0797	0.0973
GlcNAc (25 mM)	0.1223	1.2092
GlcNAc (50 mM)	0.0605	1.9734
Glucose (No sugar)	0.0797	0.0973
Glucose (25 mM)	0.1862	0.2712
Glucose (50 mM)	0.1839	0.384
Mannose (No sugar)	0.0797	0.0973
Mannose (25 mM)	0.2338	0.2237
Mannose (50 mM)	0.3002	0.2374
Lactose (No sugar)	0.0797	0.0973
Lactose (25 mM)	0.2853	0.1533
Lactose (50 mM)	0.4135	0.2195

Table 5.7 Kinetic values describing the inhibition occurring between the rChbL mutant D551A- PNP-GlcNAc reaction and GlcNAc, glucose, mannose and lactose.

	1/V intercept	Slope
GlcNAc (No sugar)	0.0464	0.1477
GlcNAc (25 mM)	0.2128	1.0289
GlcNAc (50 mM)	0.1241	2.1677
Glucose (No sugar)	0.0464	0.1477
Glucose (25 mM)	0.2244	0.1509
Glucose (50 mM)	0.2493	0.2190
Mannose (No sugar)	0.0464	0.1477
Mannose (25 mM)	0.199	0.3454
Mannose (50 mM)	0.264	0.4901
Lactose (No sugar)	0.0464	0.1477
Lactose (25 mM)	0.2854	0.1739
Lactose (50 mM)	0.3851	0.2108

It can be observed from Tables 5.7 and 5.8 that the patterns of inhibition due to the presence of inhibiting sugars are the same for each rChbL mutant, K545A and D551A. However, they differ slightly from those of the original rChbL enzyme. N-acetylglucosamine (GlcNAc) again inhibits both mutants in a competitive fashion, similarly to the rChbL enzyme. However, mannose does not inhibit the mutant enzymes competitively; instead inhibition occurs in a non-competitive manner. As with the rChbL enzyme, glucose and lactose inhibit the catalytic reaction in a non-competitive manner.

5.11 Construction of a double rChbL mutant, K545A D551A

As was the case with the mutations to the rChb CBP, two of the mutations had positive effects on the protein, increasing its ability to bind to GlcNAc epitopes. In this case, both mutations K545A and D551A increased the activity of the enzyme. These mutations were combined in one single mutant, using PCR whole vector amplification, to create the double mutant rChbL K545A D551A. The structural implications of the mutations were modelled using Pymol (see Figure 5.61). The double mutant was constructed using the primers shown in Figure 5.62 and the resultant construct was named pQE60_Chb_545A/551A. The double mutant protein was called rChbL K545A D551A. The properties of the new double mutant are shown in Table 5.9.

a) **b)**

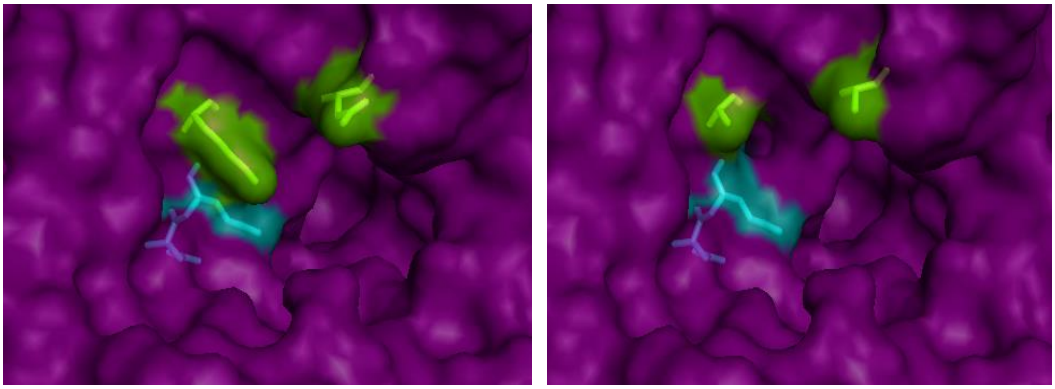


Fig 5.61: Binding pocket of a) wtChbL and b) rChbL K545A D551A. The catalytic site is highlighted in blue, D539 and E540, and residues 545 and 551 are highlighted in green. The image was generated using Pymol software.

CCT TCA CTT CGC **GCG** TTA TAG GCA AAC GGA **GCG** GGA TTC CAG
GGA GAT GAA GCG **AAG** AAT ATC CGT TTG GGA **GAC** GGA TTC CAG
GAT AAA AAT GGT CCT ATC
GAT AAA AAT GGT CCT ATC

Fig 5.62: Primers used for the construction of rChbL double mutant K545A D551A. Outline of the position of the primers used for the mutation of the rChbL enzyme at residues 545 and 551. The forward primer 545A551A_chb_F is shown in

red and the reverse primer 545A551A_chb_R is shown in green. The mutated residues are underlined and in bold.

Table 5.8 Properties of the double mutant rChbL K545A D551A.

	Molecular Weight (kDa)	Extinction Coefficient	pI
rChbL	99.99	1.601	6.53
K545A D551A	99.86	1.602	6.53

5.12 Expression and purification of the double mutant rChbL K545A D551A

Following the successful construction of the double mutant construct rChbL K545A D551A, the plasmid was transformed into *E. coli* JM109 cells for stable storage of the plasmid DNA and *E. coli* KRX cells for expression of the mutant protein. The culture conditions, already optimised for the expression of the rChbL enzyme (see Section 3.6), were used for the expression of the double mutant protein. The rChbL double mutant protein was expressed in *E. coli* KRX cells, in TB, at 30^{o}C and induced with 100 μM IPTG. The expression of the protein to both the insoluble cellular fraction and the culture media were examined (see Figure 5.64). The protein was then purified via IMAC as previously described (see Section 2.19.1) and buffer exchanged using FPLC (see Section 2.20). The catalytic activity of the new enzyme was then examined.

It can be observed from the following figures (5.63 and 5.64) that expression and purification of the rChbL double mutant K545A D551A was similar to that of the rChbL enzyme and the other rChbL mutants previously described (see Section 5.9). There was protein expressed in the cellular fraction and to the culture media. The protein from the cellular lysate was found in the insoluble fraction as before. The protein began appearing in the culture media two hours post-induction. Similar levels of protein were observed in the culture media at five hours and sixteen hours post induction. The protein was collected after five hours incubation. The protein was purified via IMAC as before and was able to withstand a 20 mM imidazole wash, after which it was eluted with 250 mM imidazole. The protein was buffered

exchanged using FPLC and the average protein yield was approximately 6 mg/250 mL culture.

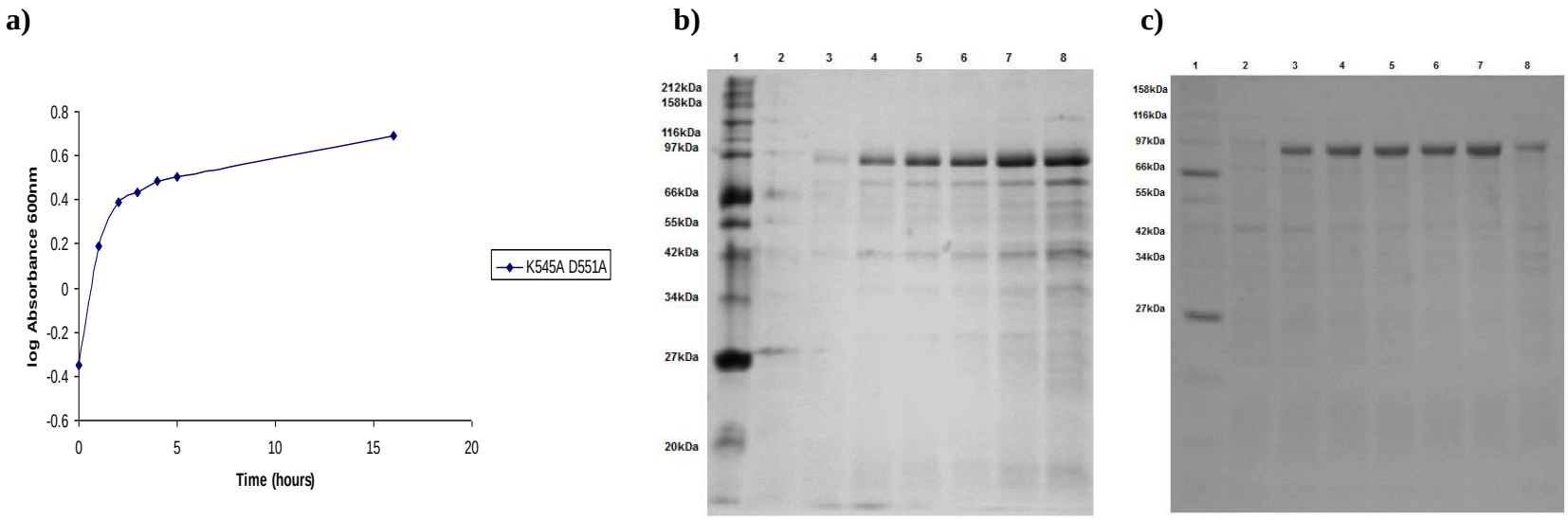


Fig 5.63: a) Analysis of the growth rate of the rChbL double mutant K545A D551A b) and c) Time-course analysis of rChbL double mutant K545A D551A expression culture samples from the culture media (b) and the cellular fraction (c). a) Recombinant protein expression was induced at OD_{600nm} = 0.5. 1 mL samples were taken from the cultures every hour for five hours, following induction and an overnight sample was taken 16 hours post induction. Absorbance readings were taken at the different time-points. The log_{Abs} was plotted against time (hours).

b) and c) Recombinant protein expression was induced at OD_{600nm} = 0.5. 5 mL samples were taken from the culture every hour for five hours, following induction and an overnight sample was taken 16 hours post induction. Cellular fraction samples (see Section 2.15) and culture media samples (see Section 2.16) were obtained. The samples were then analysed by 12.5% SDS-PAGE. Lane 1, Protein ladder; Lane 2, T=0; Lane 3, T=1 hour; Lane 4, T=2 hours; Lane 5, T=3 hours; Lane 6, T=4 hours; Lane 7, T=5 hours; Lane 8, Overnight (T=16 hours).

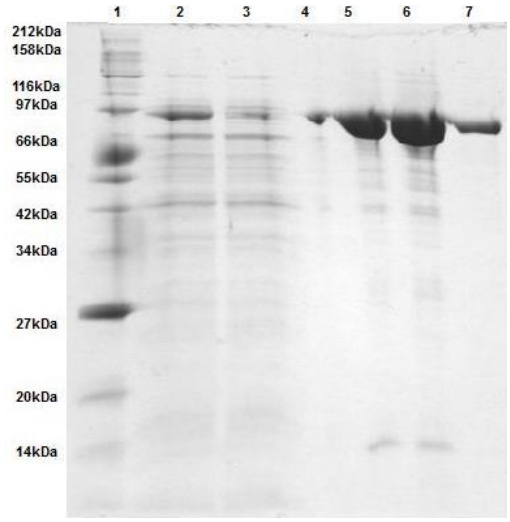


Fig 5.64: Purification of the rChbL double mutant K545A D551A from the culture media fraction. The culture media was collected after cell harvest from the rChbL double mutant expression cultures. These were then run over separate IMAC columns, washed with 40 mM imidazole and eluted with 250 mM imidazole. Fractions were run on 12.5% SDS gels. Lane 1, Protein ladder; Lane 2, Lysate; Lane 3, Flow through; Lane 4, Wash; Lane 5, Elution one; Lane 6, Elution two; Lane 7, Buffer-exchanged sample.

5.13 Characterisation of the rChbL double mutant K545A D551A

The properties of the rChbL double mutant K545A D551A were examined in the following section. The effect of the combination of the mutations on the catalytic activity of the rChbL double mutant was examined (see Figure 5.65) along with their effects on the specificity of the enzyme (see Figure 5.66) and free sugar inhibition (see Figures 5.69 and 5.70). The effect on the physical properties of the enzyme such as its optimum temperature (see Figure 5.67) and its optimum pH (see Figure 5.68) were also examined.

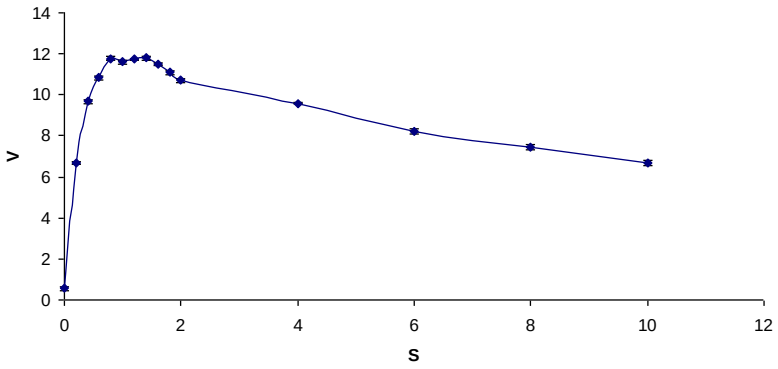
5.13.1 Preliminary kinetic characterisation of the rChbL double mutant K545A D551A

Whether the mutations acted additively to further increase the catalytic rate of the rChbL enzyme or inhibited each other cancelling their positive effects on the catalytic

rate was examined. The catalytic rate of activity was again measured using two plots; Michaelis Menten and Hanes-Woolf. The results are shown in Figure 5.65 and Table 5.10.

The K545A and D551A mutations were found to act additively when combined in the rChbL enzyme. The V_{max} of the rChbL double mutant was almost two-fold greater compared with the single rChbL mutants (see Table 5.10). The K_m was slightly lower than the single rChbL mutants suggesting a greater affinity for the substrate. The k_{cat} and k_{cat}/K_m values were one to two orders of magnitude higher in the rChbL double mutant than the single rChbL mutants.

a)



b)

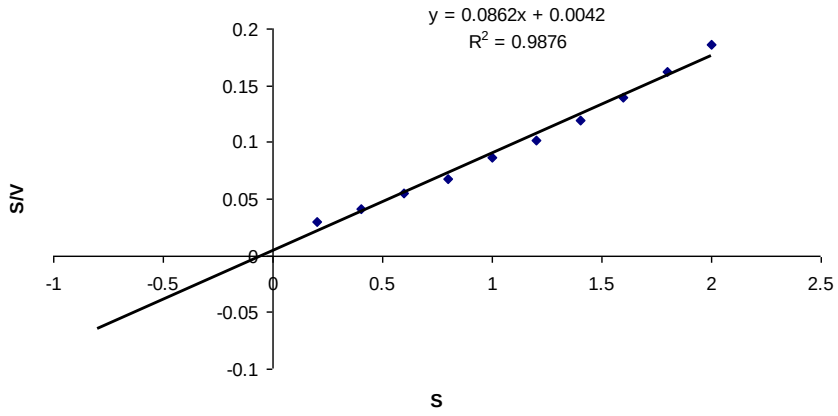


Fig 5.65: Kinetic analysis of the rChbL double mutant K545A D551A. The Chb activity assay was performed as per Section 2.26 over a range of substrate concentrations from 0 to 10 mM. The rate, V (mM PNP/min/mL enzyme), was plotted against substrate concentration (S). Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

Table 5.9 Michaelis Menten and Hanes-Woolf values for the double mutant rChbL K545A D551A.

Rate Constants	Michaelis Menten	Hanes-Woolf
V_{max} (μM)	11778	11601
K_m (mM)	0.18	0.05
k_{cat} (s^{-1})	$1.17x10^8$	$1.15x10^8$
k_{cat}/K_m ($mM^{-1}.s^{-1}$)	$6.48x10^8$	$2.30x10^9$
SD (±)	0.0953	0.0953

5.13.2 Characterisation of the properties of the rChbL double mutant K545A D551A.

The effect of temperature and pH on the catalytic activity of the rChbL double mutant K45A D551A was examined using the Chb activity assay as outlined previously (see Section 2.26) and the results compared with those of the wild type rChbL and the single mutants K45A and D551A.

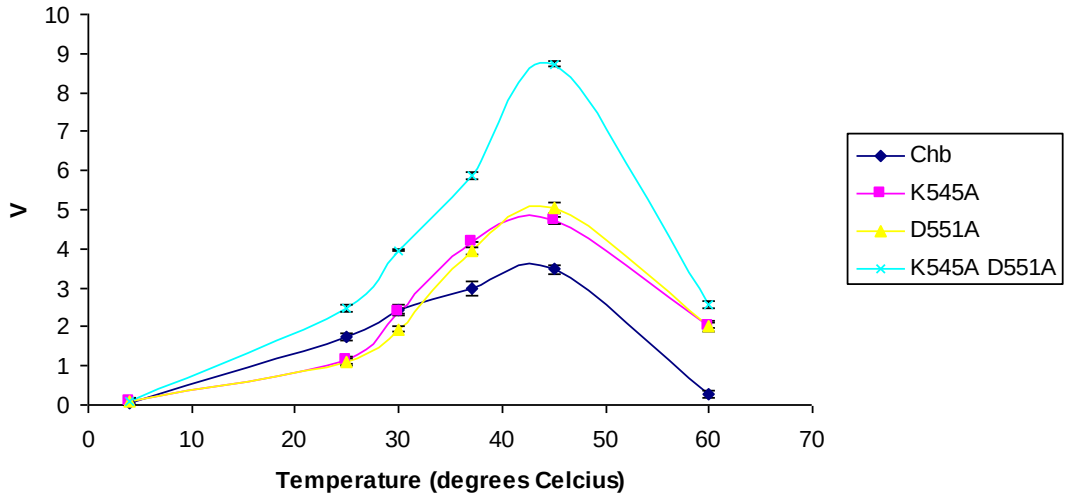


Fig 5.66: The effect of temperature on the catalytic activity of the rChbL double mutant K545A D551A. The Chb activity assay was performed as per Section 2.26 over a range of different temperatures, from 4°C to 60°C. The rate, V (mM PNP/min/mL enzyme), was plotted against temperature (°C). Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

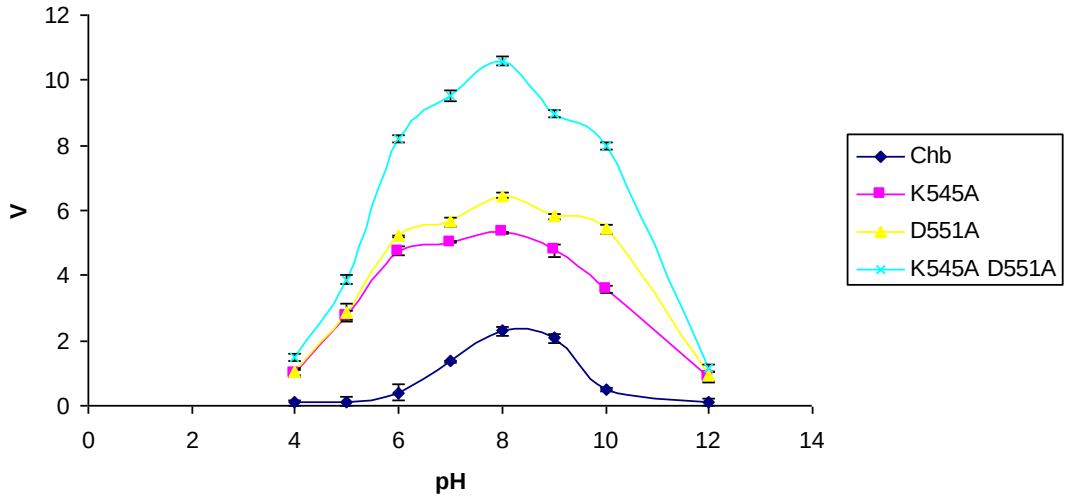


Fig 5.67: The effect of pH on the catalytic activity of the rChbL double mutant K545A D551A compared with the rChbL enzyme. The Chb activity assay was performed as per Section 2.26 over a wide pH range, from 4.0 to 12.0. The rate, V (mM PNP/min/mL enzyme), was plotted against pH. Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

The temperature profile of the rChbL double mutant was similar to that of the rChbL single mutants in that there was no activity observed at 4^{o}C and the rate of activity had greatly declined at 60^{o}C (see Figure 5.66). The optimum rate of activity was observed at approximately 45^{o}C. It can be observed that the double mutants' rate of activity was three times higher than the rChbL at its optimum temperature. The rChbL double mutant was also more stable than the rChbL at alkaline pH. The rate of activity only slightly declined at pH 9.0 and 10.0. The rate of activity was low at pH 12.0. The rChbL double mutants' rate of activity was also lower at acidic pH. The optimum pH was 8.0.

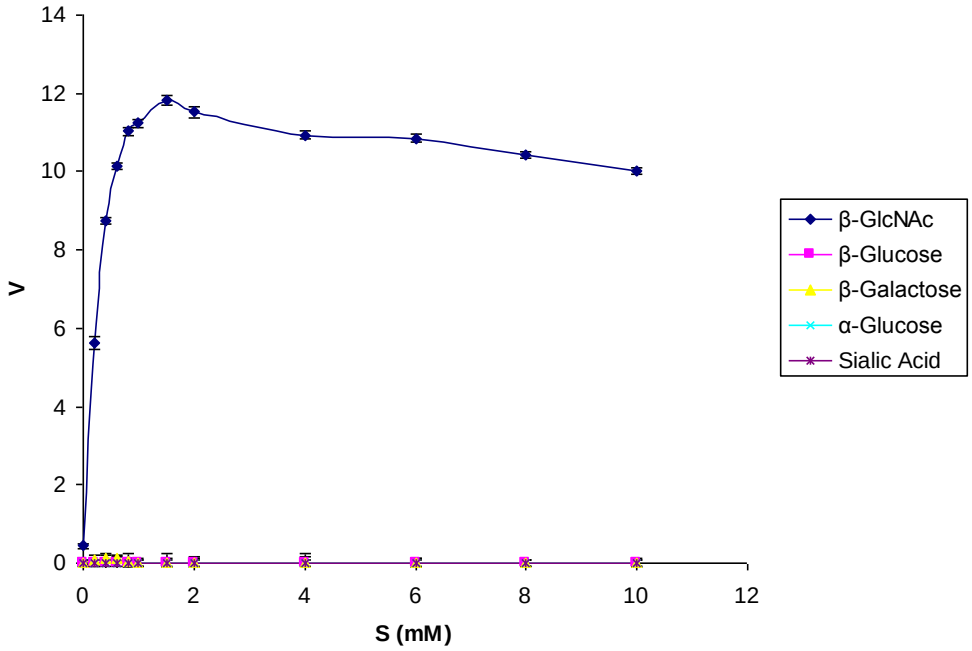


Fig 5.68: Specificity of the rChbL double mutant K545A D551A. The Chb activity assay was performed as per Section 2.26 against a wide range of PNP-linked sugar substrates. The rate, V (mM PNP/min/mL enzyme), was plotted substrate concentration (S). Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

From Figure 5.68 it can be observed that the rChbL double mutant is specific for its substrate. There was only activity observed against β-GlcNAc, no activity was observed against any of the other substrates tested.

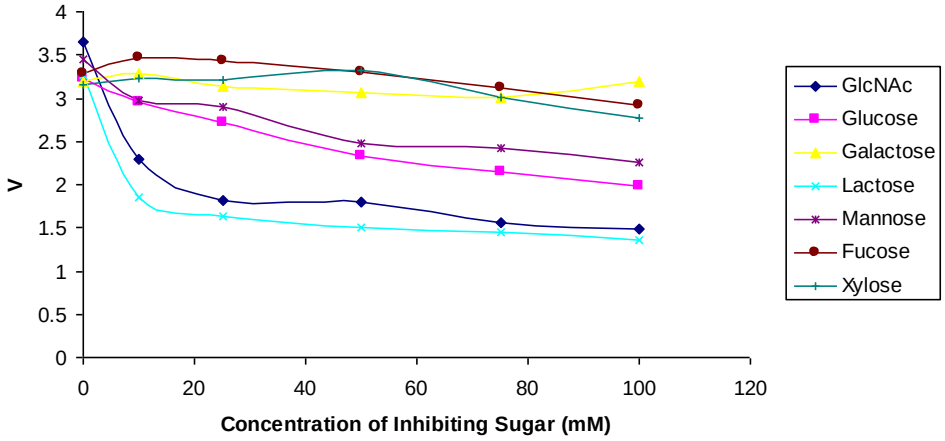


Fig 5.69: The inhibitory effect of free sugars on the rate of rChbL double mutant K545A D551A catalytic activity on PNP-GlcNAc. The Chb activity assay was performed as per Section 2.26 with PNP-GlcNAc. The K545A D551A double mutant enzyme was used at a concentration of 1 µg/mL and incubated with increasing concentrations of the inhibiting sugar for 30 minutes at 37^{o}C prior to addition to the plate. The rate, V (mM PNP/min/mL enzyme), was plotted substrate concentration (S). Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

The rChbL double mutant was inhibited by the same sugars as the rChbL enzyme and rChbL single mutants (see Figure 5.69). GlcNAc and lactose free sugars largely inhibited the enzymes activity on its substrate while glucose and mannose also inhibited activity, however to a lesser degree. There was little to no inhibition observed with galactose, fucose and xylose. How the sugars inhibited catalytic activity was investigated in Figure 5.70.

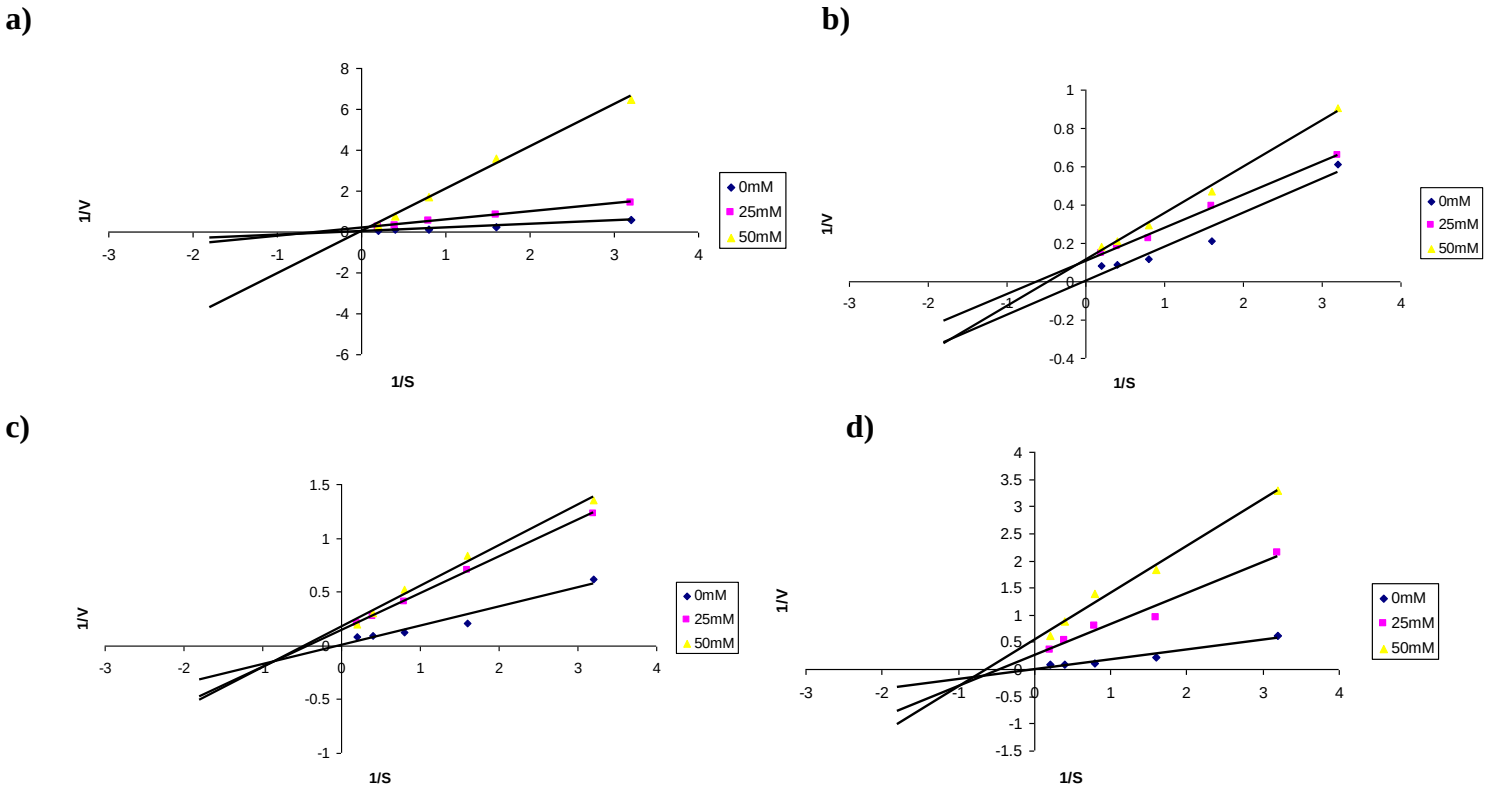


Fig 5.70: Lineweaver-Burk plots of the inhibition occurring with the K545A rChbL mutant and a) GlcNAc, b) glucose, c) mannose and d) lactose. The Chb activity was performed as described in Section 2.26. Increasing concentrations of PNP-GlcNAc were incubated with a constant concentration of enzyme (1 µg/mL) and either 25 mM or 50 mM inhibiting sugar. The reciprocal rate of the reaction was plotted against the reciprocal of the substrate concentration. Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

Table 5.10 Kinetic values describing the inhibition occurring between the rChbL double mutant K545A D551A-PNP-GlcNAc reaction and GlcNAc, glucose, mannose and lactose.

	1/V intercept	Slope
GlcNAc (No sugar)	0.002	0.1785
GlcNAc (25 mM)	0.1631	0.4007
GlcNAc (50 mM)	0.0194	2.0648
Glucose (No sugar)	0.002	0.1785
Glucose (25 mM)	0.1069	0.1724
Glucose (50 mM)	0.1117	0.2425
Mannose (No sugar)	0.002	0.1785
Mannose (25 mM)	0.1394	0.342
Mannose (50 mM)	0.173	0.3795
Lactose (No sugar)	0.002	0.1785
Lactose (25 mM)	0.2582	0.5681
Lactose (50 mM)	0.536	0.8692

The inhibition pattern produced by the different free sugars for the rChbL double mutant was similar to that of the rChbL single mutants (K545A and D551A). From Table 5.11 it can be observed that GlcNAc inhibited the catalytic reaction in a competitive fashion while the three other sugars (glucose, mannose and lactose) inhibited the reaction in a non-competitive manner. This was different to the original rChbL enzyme that was competitively inhibited by both GlcNac and mannose.

5.14 Discussion

This chapter aimed to improve the binding ability of the rChb CBP thereby allowing it to successfully bind glycoproteins in an ELLA format. This aim was achieved through site-directed mutagenesis of residues around the binding pocket of the rChb CBP. Three of the mutations, W63A, D648A and W689Y, inhibited binding, suggesting an important structural role of these residues. Two of these mutations, K545A and D551A, improved the binding ability of the rChb CBP and when combined they worked additively to increase binding of rChb CBP to GlcNAc residues. Alongside this work, these mutations were also carried out on the rChbL enzyme, to investigate the role of these residues in the catalytic activity of the enzyme. While three of these mutations, W63A, D648A and W689Y, destroyed the catalytic activity from the rChbL enzyme, highlighting their importance in catalysis, two of the mutations, K545A D551A, increased the catalytic activity of the rChbL enzyme. When combined these mutations acted synergistically and again increased the catalytic activity of the rChbL enzyme.

Previous studies (Tews *et al.,* 1996) have investigated the structure of the *S. marcescens* chitobiase and proposed functions for certain residues proximal to the binding pocket and active site. While these residues were discovered to play a role in substrate binding and the problem being addressed in this chapter was the inability of the rChb CBP to bind glycoproteins, the issue with the rChb CBP was believed to be due to steric hindrance rather than a lack of substrate binding in the active site. The carbohydrate binding protein constructed in this study was initially an enzyme. Enzymes are extremely specific for their substrate, with the active site usually occupying a deep cleft in the protein structure (Mathews and van Holde, 1991). The position of the active site would not be an issue for chitobiose, the natural substrate of the chitobiase enzyme, as it is a small sugar composed of two N-acetylglucosamine residues. The difficulties arose however, when examining the binding of the rChb CBP to N-acetylglucosamine residues on glycoproteins. While the small chitobiose sugar is able to enter the deep crater, the N-acetylglucosamine residues are attached to a much larger glycoprotein which presumably inhibits binding of the GlcNAc residues.

Structural modelling of the active site/binding pocket suggested several over-hanging residues that may have been inhibiting binding to the N-acetylglucosamine residues on the glycoproteins. Alanine scanning was carried out on these residues. Alanine scanning is the most common method of mutagenesis to decipher the functional relevance of protein residues (Morrison and Weiss, 2001). The substitution of alanine in place of the residue of interest voids the side chain and thus the role of that particular residue can be deduced. Alanine substitutions allow the protein to retain its conformation and structural integrity unlike substitutions with another simple amino acid such as glycine, which could introduce flexibility into the structure (Morrison and Weiss, 2001; Cunningham and Wells, 1989). Cunningham and Wells (1989) carried out alanine scanning when looking at the binding between human growth hormone and its receptors on the human liver. Residues contained on portions of the protein proposed to be involved in receptor recognition were mutated to alanine. When a lower binding affinity was observed, the residues in question were identified as having a role in receptor recognition. A study investigating the functional residues in a protein kinase by Gibbs and Zoller (1991) also used the alanine scanning approach. They identified several charged amino acids in the catalytic site of the enzyme and subsequently mutated to them to alanine. The activity of the resultant mutants was tested and residues were proposed that may have a role in catalysis and substrate binding.

This present study also used alanine scanning mutagenesis, however, rather than using alanine substitution to study the functional relevance of the residues in question, alanine was substituted as it is a small amino acid that allows the protein to retain its structure (Morrison and Weiss, 2001). As previously mentioned in the introduction, the structure and function of an enzyme are intrinsically linked. The mutagenesis of any residue has the risk of altering the structure of the enzyme and rendering it inactive. All mutations were carried out on residues proximal to the active site. Alanine was chosen to minimise structural alteration and preserve the active site.

The expression and purification of both the rChbL enzyme mutants and the rChb CBP mutants did not greatly vary from that of the native rChbL enzyme. All mutant proteins were still expressed to the culture media, appearing two hours post-induction. Purification was carried out in the same manner as the rChbL. The mutant proteins

were also buffer-exchanged using FPLC rather than centrifugal concentrators (vivaspins). This was because of the low protein yields obtained for the native rChbL, due to its binding to the cellulose membrane. Similar yields of recombinant protein were obtained for all mutants in comparison with the native rChbL.

The initial aim of this chapter was the mutagenesis of the rChb CBP to generate a protein capable of binding glycoproteins in the ELLA format. Through site-specific mutagenesis the following mutations were made: W63A, K545A, D551A, D648A and W689Y. From Figure 5.9 it can be seen that these mutations successfully opened up the active site, however there were a variety of consequences. Mutants W63A, D648A and W689Y did not increase the binding between the rChb CBP and GlcNAc, whether it was on a glycoprotein or in the BSA-conjugated format. However, while mutants K545A and D551A only showed a slight increase in their ability to bind GlcNAc on glycoproteins, there was double the signal observed when they were tested against BSA-GlcNAc. An increased signal was also observed when they were tested against bovine mucin. When these mutations were combined in the double mutant K545A D551A they worked in an additive fashion and the signals observed from the rChb CBP when tested against BSA-GlcNAc and bovine mucin were increased again, this time the signal received against BSA-GlcNAc was three times higher than that observed for the rChb CBP. Disappointingly, binding to GlcNAc on glycoproteins was not greatly enhanced. It was thought that the high level display of GlcNAc on the BSA-conjugate boosted the signal from the rChb CBP mutants. The glycans present on bovine mucin are O-linked. While N-linked glycans tend to have branched structures, such as bi-antennary and tri-antennary structures, the terminal ends of O-linked glycans are primarily non-branched. This may be the reason the rChb CBP mutants were capable of binding to the bovine mucin and not the N-linked glycoproteins. The bi-antennary structure of the asialo agalacto fetuin glycoprotein may have inhibited binding. The high signal obtained from GSLII is because it is a dimeric protein compared to the monomeric rChb CBP. Increasing the valency of the rChbL may improve its affinity for GlcNAc substrates. Green fluorescent protein (GFP) is a dimeric protein. Fusion to this protein may increase the binding of the rChb CBP to GlcNAc residues by allowing it to work as a dimer. However, the mutations have shown however that there is the possibility of turning this rChb CBP into a lectin for glycoprotein analysis with further mutagenesis.

Mutant K545A removed a lysine from the binding pocket of the rChb CBP. This lysine molecule is susceptible to the addition of a biotin molecule during biotinylation. The presence of a biotin in this position was proposed to cause interference with binding. Removal of this lysine meant removal of the potential for the acceptance of a biotin molecule in this position within the binding pocket. It was hypothesized that mutation of K545 also increased the space in the binding pocket, allowing the glycoprotein more room. D551 is directly above K545 and mutation of this residue was also hypothesized to increase the space in the active site. As this residue was protruding over the binding pocket it may have been hindering binding of the glycoprotein. Mutagenesis of this residue removed this interference. While the mutations improved the binding abilities of the rChb CBP they were still insufficient in producing a carbohydrate binding protein for the analysis of glycoproteins. However, this study shows the potential of the generation of new carbohydrate binding proteins through the mutagenesis of enzymes.

In a review on enzyme design by Toscano *et al.*, (2007), they state that "enzyme active sites provide highly optimized microenvironments for the catalysis of biologically useful chemical transformations. Consequently, changes at these centres can have large effects on enzyme activity." The mutations carried out on the rChbL enzyme were all proximal to the active site and confirmed this statement. The mutations caused relatively large changes in the catalytic activity of the enzyme to be observed. Mutations W63A, D648A and W689Y caused a considerable decrease in catalytic activity. These mutations succeeded in opening the active site of the enzyme. This was the aim for the carbohydrate binding protein but obviously this was detrimental for the catalytic activity of the rChbL mutant enzymes. It was hypothesized that mutation of residues, W63A and D648A, disrupted the active site conformation in some way causing the enzyme to lose activity. The tryptophan at position 689 functions in holding the chitobiose substrate in position during catalysis. Mutation of this residue to a tyrosine did not allow it to carry out its function. While it was hypothesized that mutation of this residue to tyrosine may open up the binding pocket while still providing the same support to the substrate due to the tyrosine ring structure, this was not the case. These mutations reduced catalytic activity by decreasing the velocity (V_{max}) of the rChbL enzyme while also decreasing its affinity

for its substrate. As the rate of catalytic activity was so low it was difficult to assess their specificity, temperature and pH tolerance.

While mutations to residues W63A, D648A and W689Y on the rChbL enzyme had detrimental effects on catalytic activity, the opposite was the case for mutations to residues K545A and D551A. Their changes to the structure of the active site were obviously not deleterious. The lysine at position 545 covers the catalytically important residues, D539 and E540. By substituting the lysine for an alanine the catalytic residues may have been more accessible for the substrate. This was confirmed as the K_m of the rChbL mutant was decreased compared with the native rChbL. The mutant had a higher affinity for its substrate. The velocity, turnover number and specificity were also increased.

The aspartic acid at position 551 was protruding over the active site. It was directly above residue K545. By mutating this residue to a smaller alanine residue it was again believed that it was easier for the substrate to reach the catalytic residues. This was confirmed by the reduction in K_m and increase in V_{max}. The velocity of the mutant rChbL was increased, while the affinity for the substrate was also increased. The turnover number and specificity of the mutant enzyme were also increased. When combining advantageous mutations the outcome is not necessarily also advantageous. Situations can arise where the advantageous effects are inhibited; if the mutations are in direct or indirect contact with one another or if one of the mutations causes a change in the mechanism of the reaction (Wells, 1990). However, in this case the combination of mutations K545A and D551A had an additive effect on the rate of catalytic activity of the rChbL enzyme. The velocity of the double mutant was almost double that of the single mutants while the affinity for its substrate was also increased. The turnover number and specificity were also increased.

In 2005, Morley and Kazlauskas examined the relationship between the distance of a mutation to the active site of an enzyme and the improvement of certain properties including selectivity, catalytic activity and thermal stability. They revealed that mutations closer to the active site of the enzyme were often more advantageous in increasing the specificity of the enzyme than those further away. Otten *et al*., (2004) discovered that an aspargine at position 266, which was 7.1Å from the active site, was

critical for substrate specificity. Similarly, Cheon *et al.*, (2004) discovered that the site-specific mutation of a phenylalanine at position 159 (11.3Å from the active site) increased substrate specificity 200-fold. This was also the case for the rChbL enzyme mutants, K545A, D551A and the double mutant K545A D551A. The specificity of the enzyme is measured by k_{cat}/K_m (Hedstrom, 2001). This figure was increased from a magnitude of 10^5 to 10^7-10^8. When examining the relationship between the distance of the mutation from the active site and properties such as an increased catalytic rate and increased thermal stability, Morley and Kazlauskas (2005) concluded that there was no relationship between the distance of the mutation from the active site and an improvement in these characteristics. Many studies reported increases in activity when residues close to the active site were mutated such as Ducros *et al.,* (2003) who recorded a 35-fold increase in catalytic activity when they mutated a glutamic acid 9Å from the active site. Alternatively however, many other studies reported increases in catalytic activity from mutations at a much greater distance from the active site such as Whittle and Shanklin (2001) who mutated a histidine 19Å from the active site and recorded a 31-fold increase in activity. This study reported an increased catalytic rate when residues close to the active site were mutated.

Interestingly, the mutations that increased catalytic activity (K545A and D551A) also increased the pH stability of the mutant rChbL enzyme in comparison with the wildtype. The single mutants and double mutant were more stable at pH 9.0 and 10.0 than the native rChbL, only losing activity at pH 12.0. The advantageous mutations were made to residues in the active site. Two charged residues, a lysine (545) and an aspartic acid (551) were mutated to the uncharged amino acid, alanine. This may have altered the charge of the catalytic environment, increasing the positive charge in the active site. This difference in charge may have affected the breakdown of the PNP-GlcNAc substrate in some way. The increase in activity at more alkaline pH may also be due to the general increase in catalysis.

The mutants K545A, D551A and the double mutant K545A D551A were all inhibited by the same free sugars as the original rChbL. N-acetylglucosamine (GlcNAc), glucose, mannose and lactose all inhibited the catalytic reaction between the rChbL mutants and the PNP-GlcNAc substrate. However the mechanism in which they inhibited the reaction was different in certain cases. Again, GlcNAc inhibited all

reactions in a competitive manner, the GlcNAc sugar bound in the active site occupying the place of the PNP-GlcNAc substrate. Lactose also inhibited the catalytic reaction in the same non-competitive manner. However, mannose switched from a competitive inhibitor to a non-competitive inhibitor. It was believed that the changes to the binding pocket did not allow the binding of mannose. It no longer accepted the structural variation of the axial –OH at C-2 that mannose contains rather than the equatorial N-acetyl group that N-acetylglucosamine contains. Glucose was found to inhibit the rChbL enzyme in a non-competitive manner rather than an un-competitive manner. Previous studies have shown mutagenesis of glycosidases changing the specificity of the enzyme. In a study by Corbett *et al*., (2001) it was reported that site-directed mutagenesis altered the binding specificity of a β-glycosidase from *Sulfolobus solfataricus*. A single mutation (M439C) led to enhanced activity on xylose substrates while activity on fucose substrates was decreased. The activity on glucose substrates was not impaired by this mutation. Mutagenesis of the rChbL enzyme did not alter the specificity of the enzyme; it still only catalysed the cleavage of PNP-GlcNAc. However, the inhibition study implied that the rChbL mutants were more specific than the original rChbL due to the fact that only GlcNAc was able to competitively inhibit the catalytic reaction. No other sugars were able to enter the active site. This result was supported by the fact (previously mentioned) that the mutants had an increased specificity for the PNP-GlcNAc substrate as highlighted by an increase in k_{cat}/K_m values.

The work described in this chapter was successful in creating new rChbL enzyme mutants and rChb CBP mutants. The changes to the binding abilities of the rChb CBP mutants were insufficient for them to be used as a carbohydrate binding protein for glycoanalysis, however they did show potential for further development of these rChb CBP mutants. Perhaps further mutagenesis would increase the binding potential. The increases in catalytic activity of the rChbL enzyme mutants and increased stability at higher pH make them promising candidates as commercial biocatalysts. This will be discussed further in the final discussion.

Chapter 6

Interactions with Chitin

6.1 Overview

Hydrolysis of the synthetic substrate PNP-GlcNAc allowed the study of the kinetics of the rChbL enzyme and its mutants. However, this substrate does not reflect the activity of the rChbL in its natural environment. The natural substrate of rChbL is chitobiose, a dimer of GlcNAc, which is a major product of chitin degradation. There are two main forms of chitin found in nature: α- and β-chitin (Rinaudo, 2006). Deacetylation of chitin results in the polymer chitosan. Various forms of chitosan exist that simply differ in the degree of deacetylation. Cellulose is a similar polysaccharide consisting of β1-4 linked D-glucose units. These polysaccharides differ at the C-2 position where chitin has an N-acetyl group and cellulose does not (Dutta *et al.,* 2004). Insoluble substrate assays were utilized to examine the interactions between the rChbL enzyme, CBP and these polysaccharides.

The second aim of this chapter was to investigate the interaction of various length chito-oligosaccharides with the rChbL enzyme. Firstly, the ability of each chito-oligosaccharide to inhibit the rChbL-PNP-GlcNAc reaction was examined to determine the preference of rChbL for each chito-oligosaccharide (see Section 6.4). The breakdown of these chito-oligosaccharides to GlcNAc by rChbL was then investigated using High Performance Liquid Chromatography (HPLC) and will be further explained in Section 6.5.

6.2 rChbL Insoluble Substrate Assays

The rChbL insoluble substrate assays were carried out at 4^{o}C to minimise the occurrence of catalytic activity (see Figure 3.23). The aim was to study the interaction between the enzyme and various insoluble substrates, not its catalytic activity on the insoluble substrates. The assays were carried out as described in Section 2.29.1, analysed by SDS-PAGE (see Section 2.21) and silver stained to visualise the results (see Table 2.6). Briefly once rChbL was mixed with each substrate the levels of unbound protein in the culture media were monitored. The rChbL enzyme and all rChbL enzyme mutants were tested against α-chitin, β-chitin, chitosan and cellulose.

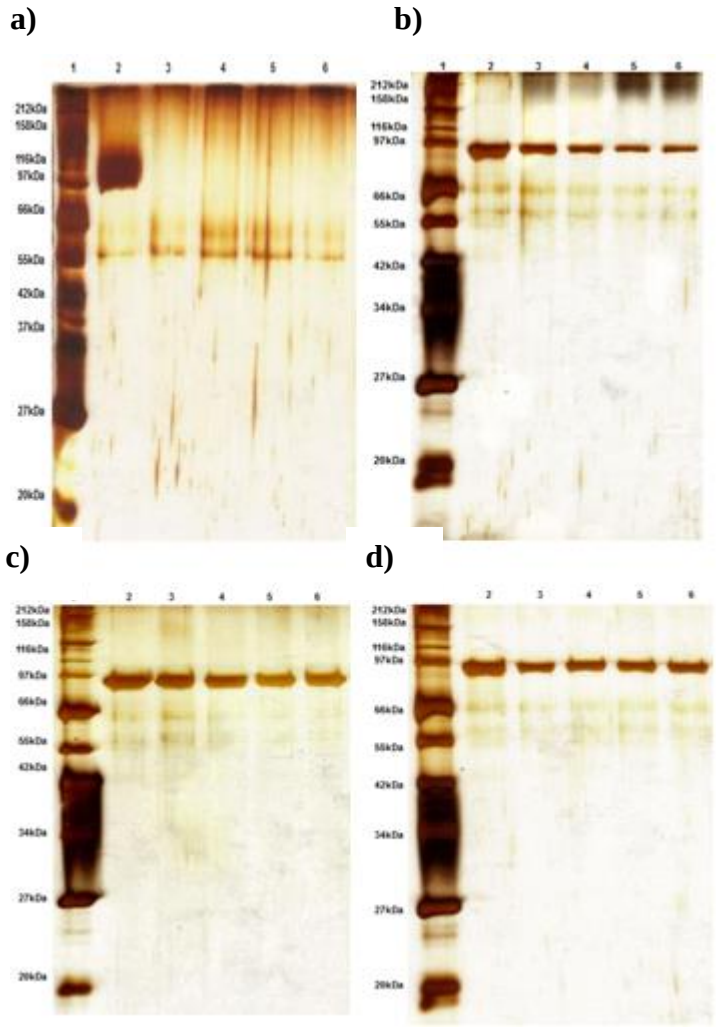


Fig 6.1: Insoluble substrate assay using the rChbL enzyme with a) α-chitin, b) β-chitin, c) chitosan and d) cellulose. The rChbL enzyme sample was continuously mixed with the substrate and samples were taken at various time-points as described in Section 2.29.1. The samples were analysed on SDS-PAGE (see Section 2.21) and visualised using silver nitrate staining (see Table 2.6). Lane 1, Protein Ladder; Lane 2, rChbL sample; Lane 3, 5 minutes incubation; Lane 4, 15 minutes incubation; Lane 5, 30 minutes incubation; Lane 6, 1 hour incubation.

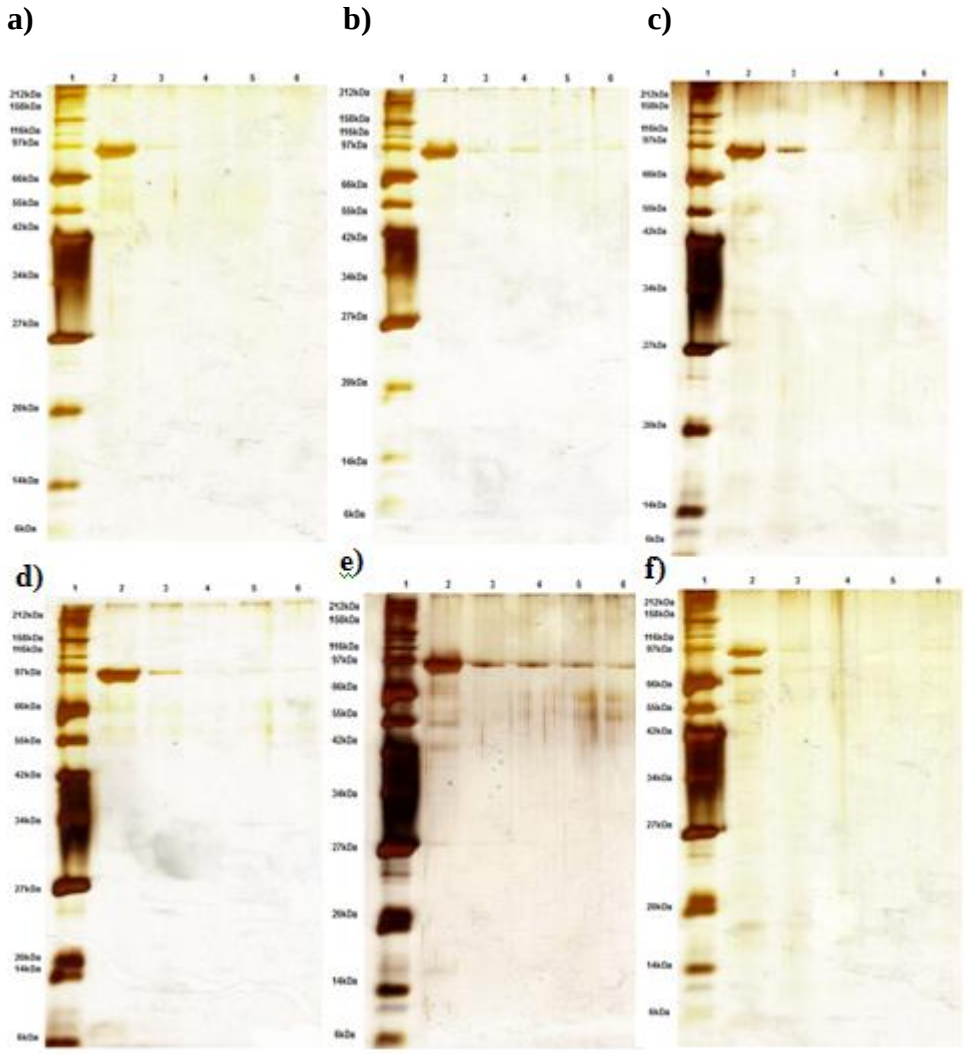


Fig 6.2: Interactions between the rChbL mutants and α-chitin; a) W63A, b) K545A, c) D551A, d) D648A, e) W689Y and f) K545A D551A. The rChbL mutant enzyme samples were continuously mixed with the α-chitin substrate and samples were taken at various time-points as described in Section 2.29.1. The samples were analysed by SDS-PAGE (see Section 2.21) and visualised using silver nitrate staining (see Table 2.6). Lane 1, Protein Ladder; Lane 2, rChbL sample; Lane 3, 5 minutes incubation; Lane 4, 15 minutes incubation; Lane 5, 30 minutes incubation; Lane 6, 1 hour incubation.

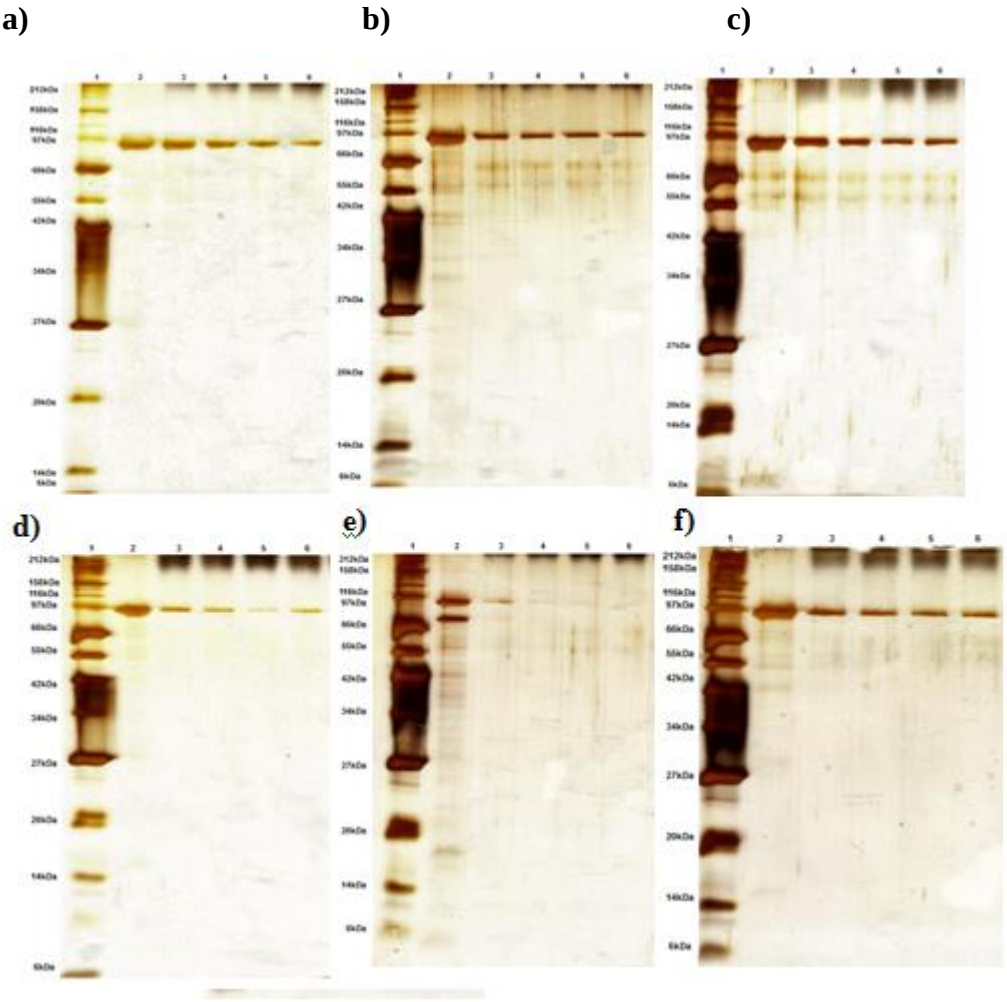


Fig 6.3: Interactions between the rChbL mutants and β-chitin; a) W63A, b) K545A, c) D551A, d) D648A, e) W689Y and f) K545A D551A. The rChbL mutant enzyme samples were continuously mixed with the β-chitin substrate and samples were taken at various time-points as described in Section 2.29.1. The samples were analysed by SDS-PAGE (see Section 2.21) and visualised using silver nitrate staining (see Table 2.6). Lane 1, Protein Ladder; Lane 2, rChbL sample; Lane 3, 5 minutes incubation; Lane 4, 15 minutes incubation; Lane 5, 30 minutes incubation; Lane 6, 1 hour incubation.

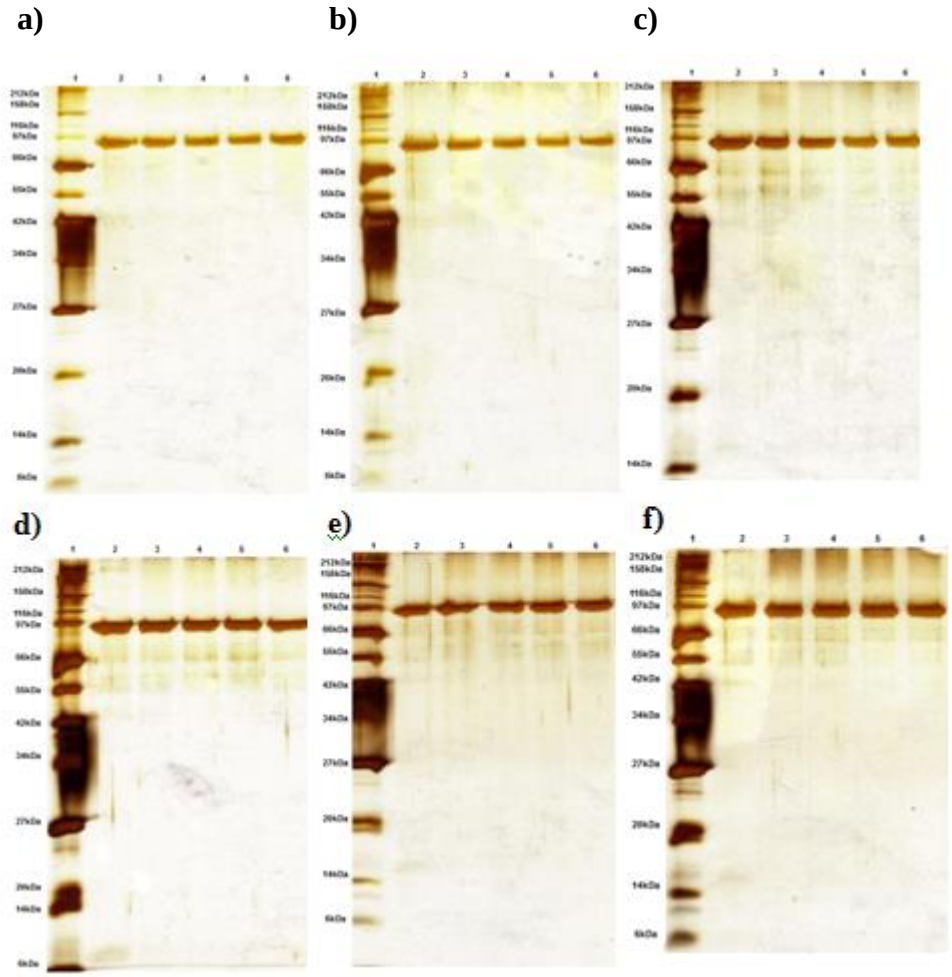


Fig 6.4: Interactions between the rChbL mutants and chitosan; a) W63A, b) K545A, c) D551A, d) D648A, e) W689Y and f) K545A D551A. The rChbL mutant enzyme samples were continuously mixed with the chitosan substrate and samples were taken at various time-points as described in Section 2.29.1. The samples were analysed by SDS-PAGE (see Section 2.21) and visualised using silver nitrate staining (see Table 2.6). Lane 1, Protein Ladder; Lane 2, rChbL sample; Lane 3, 5 minutes incubation; Lane 4, 15 minutes incubation; Lane 5, 30 minutes incubation; Lane 6, 1 hour incubation.

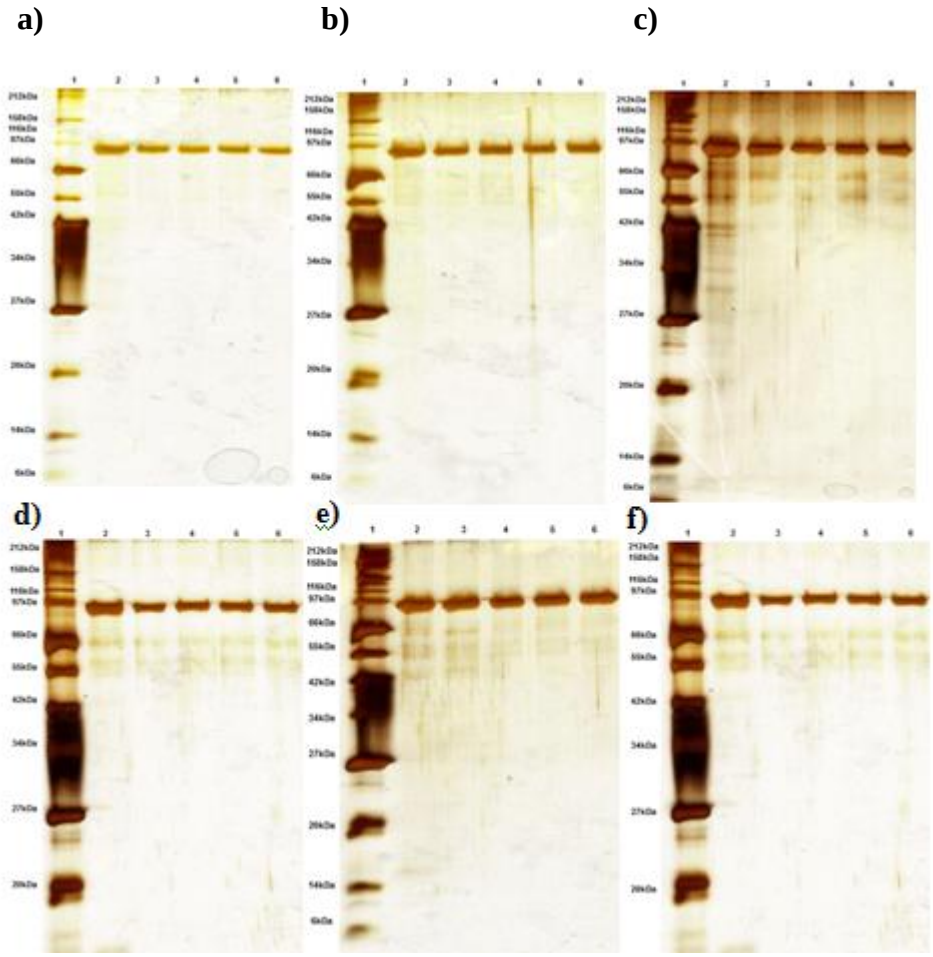


Fig 6.5: Interactions between the rChbL mutants and cellulose; a) W63A, b) K545A, c) D551A, d) D648A, e) W689Y and f) K545A D551A. The rChbL mutant enzyme samples were continuously mixed with the cellulose substrate and samples were taken at various time-points as described in Section 2.29.1. The samples were analysed by SDS-PAGE (see Section 2.21) and visualised using silver nitrate staining (see Table 2.6). Lane 1, Protein Ladder; Lane 2, rChbL sample; Lane 3, 5 minutes incubation; Lane 4, 15 minutes incubation; Lane 5, 30 minutes incubation; Lane 6, 1 hour incubation.

Table 6.1 Summary of the interactions between the rChbL enzyme and rChbL enzyme mutants with α-chitin, β-chitin, chitosan and cellulose.

	α-Chitin	β-Chitin	Chitosan	Cellulose
rChbL	√	x	x	x
W63A	√	x	x	x
K545A	√	x	x	x
D551A	√	x	x	x
D648A	√	x	x	x
W689Y	√	√	x	x
K545A D551A	√	x	x	x

The rChbL enzyme bound to α-chitin with a high affinity (see Figure 6.1). All rChbL was bound to α-chitin after five minutes. There was possibly binding to β-chitin occurring, although to a very small degree compared with α-chitin. Binding to chitosan by the rChbL was not observed. Binding to cellulose was also not observed.

Binding to α-chitin was observed for all rChbL mutants (see Figure 6.2). The lowest affinity was observed for the W689Y mutant. The affinity of the rest of the mutants for α-chitin was similar to the native rChbL enzyme. The rChbL mutant W689Y bound to β-chitin with a relatively high affinity (see Figure 6.3). The majority of the W689Y mutant enzyme was bound after 5 minutes. There was possibly a small amount of binding to β-chitin observed for all the other rChbL mutants, but to a much lesser extent, similar to that of the native rChbL. It can be observed from Figures 6.4 and 6.5 that there was no binding to chitosan or cellulose occurring for any of the rChbL mutants. The amount of rChbL mutant enzyme present was similar at all time-points. This was similar to the native rChbL enzyme.

6.3 rChb CBP Insoluble Substrate Assays

The insoluble substrate assays were repeated using the catalytically inactive rChb CBP. In this instance the reactions were allowed to proceed at the optimum temperature of 45°C (see Figure 3.23). It was expected that the enzyme, rChbL, and carbohydrate binding protein, rChb CBP, should have the same binding pattern and

the temperature difference between both sets of experiments could be used to examine binding at both high and low temperatures, as well as the speed of the reactions. Low affinity binding would occur more slowly at 4°C and may be overlooked during assays carried out at the optimum temperature of 45°C. The assays were carried out as described in Section 2.29.2, and unbound protein samples was analysed by SDS-PAGE (see Section 2.21). Silver staining was used to visualise the results (see Table 2.6). rChb CBP and all of the rChb CBP mutants were tested against the various insoluble substrates, as above.

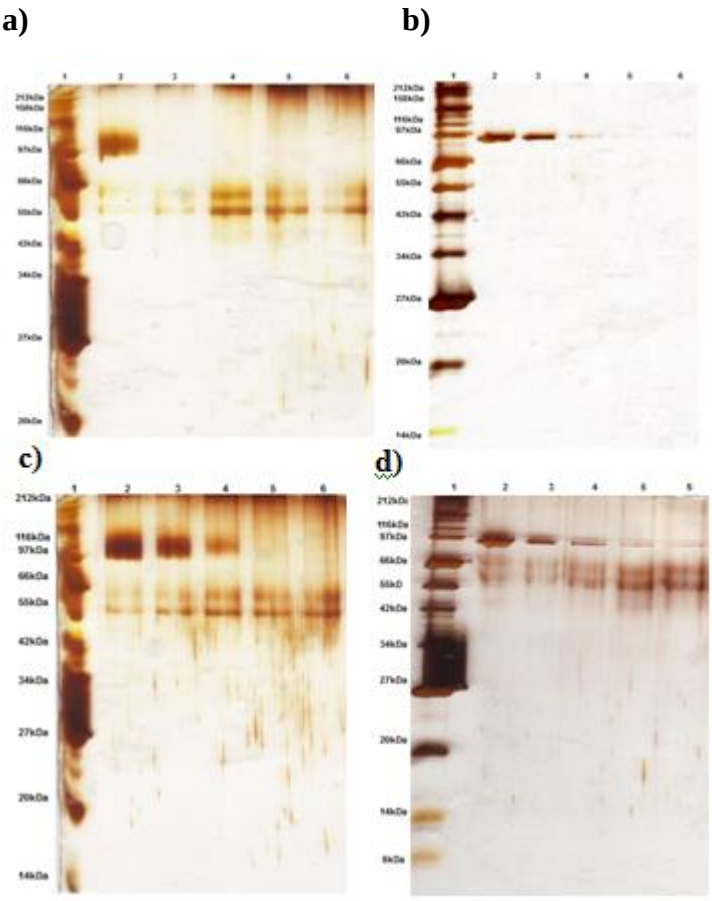


Fig 6.6: Interactions between the rChb CBP with a) α-chitin, b) β-chitin, c) chitosan and d) cellulose. The rChb CBP sample was continuously mixed with the substrate and samples were taken at various time-points as described in Section 2.29.2. The samples were analysed by SDS-PAGE (see Section 2.21) and visualised using silver nitrate staining (see Table 2.6). Lane 1, Protein Ladder; Lane 2, rChbL sample; Lane 3, 5 minutes incubation; Lane 4, 15 minutes incubation; Lane 5, 30 minutes incubation; Lane 6, 1 hour incubation.

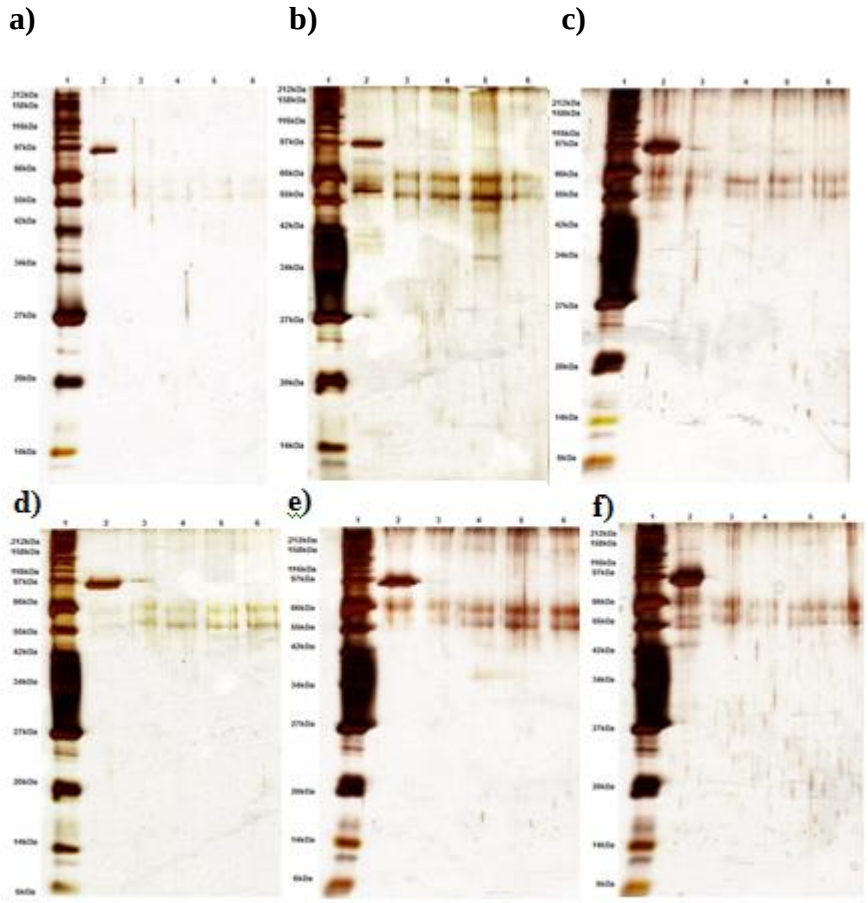


Fig 6.7: Interactions between the rChb CBP mutants and α-chitin; a) W63A, b) K545A, c) D551A, d) D648A, e) W689Y and f) K545A D551A. The rChbL CBP mutant samples were continuously mixed with the α-chitin substrate and samples were taken at various time-points as described in Section 2.29.1. The samples were analysed by SDS-PAGE (see Section 2.21) and visualised using silver nitrate staining (see Table 2.6). Lane 1, Protein Ladder; Lane 2, rChbL sample; Lane 3, 5 minutes incubation; Lane 4, 15 minutes incubation; Lane 5, 30 minutes incubation; Lane 6, 1 hour incubation.

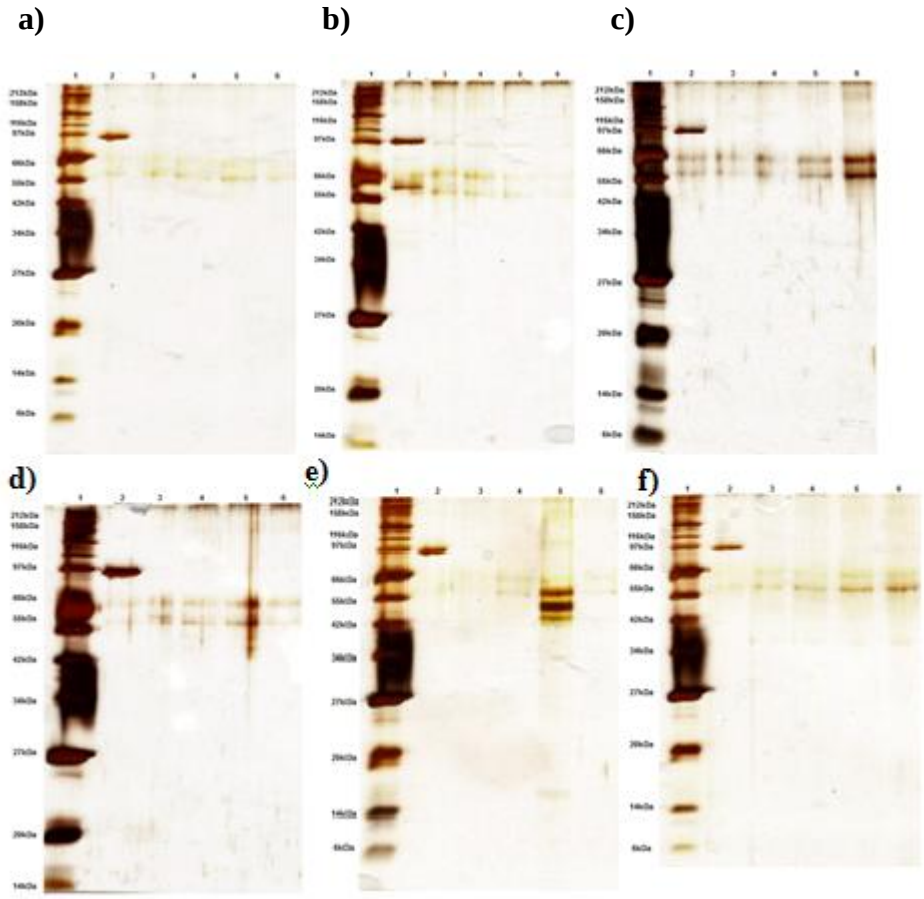


Fig 6.8: Interactions between the rChb CBP mutants and β-chitin; a) W63A, b) K545A, c) D551A, d) D648A, e) W689Y and f) K545A D551A. The rChbL CBP mutant samples were continuously mixed with the β-chitin substrate and samples were taken at various time-points as described in Section 2.29.1. The samples were analysed on SDS-PAGE (see Section 2.21) and visualised using silver nitrate staining (see Table 2.6). Lane 1, Protein Ladder; Lane 2, rChbL sample; Lane 3, 5 minutes incubation; Lane 4, 15 minutes incubation; Lane 5, 30 minutes incubation; Lane 6, 1 hour incubation.

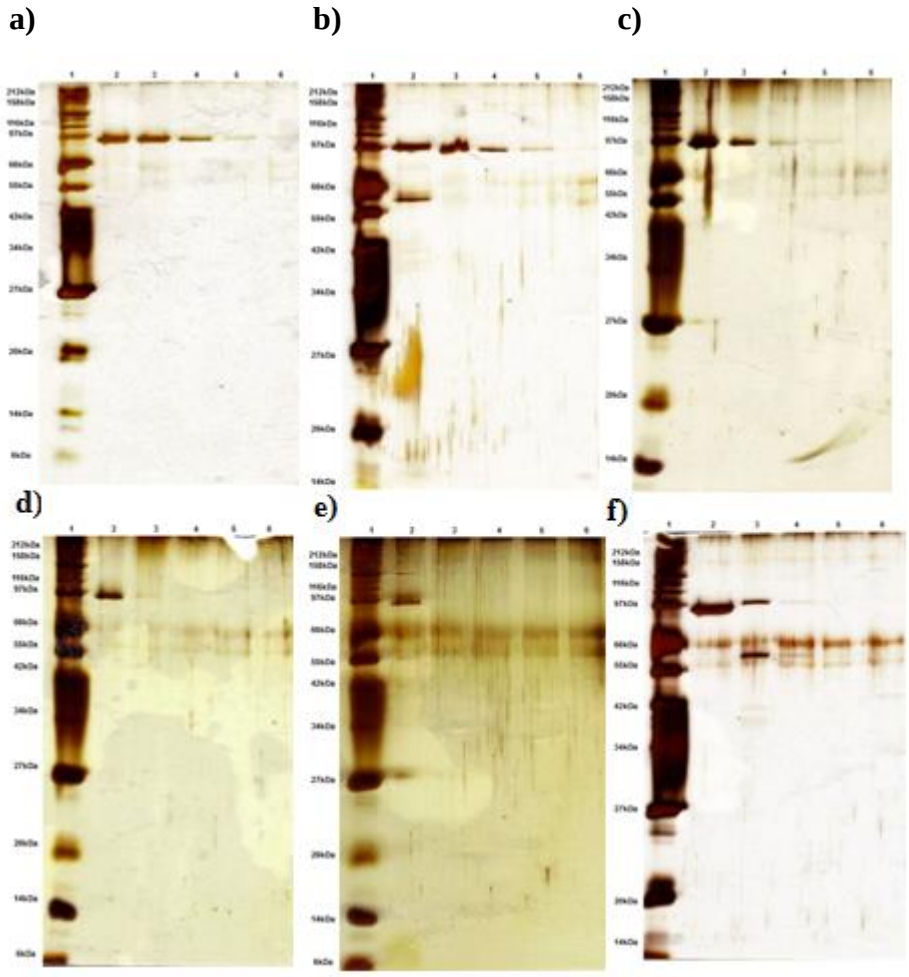


Fig 6.9: Interactions between the rChb CBP mutants and chitosan; a) W63A, b) K545A, c) D551A, d) D648A, e) W689Y and f) K545A D551A. The rChbL CBP mutant samples were continuously mixed with the chitosan substrate and samples were taken at various time-points as described in Section 2.29.1. The samples were analysed by SDS-PAGE (see Section 2.21) and visualised using silver nitrate staining (see Table 2.6). Lane 1, Protein Ladder; Lane 2, rChb CBP sample; Lane 3, 5 minutes incubation; Lane 4, 15 minutes incubation; Lane 5, 30 minutes incubation; Lane 6, 1 hour incubation.

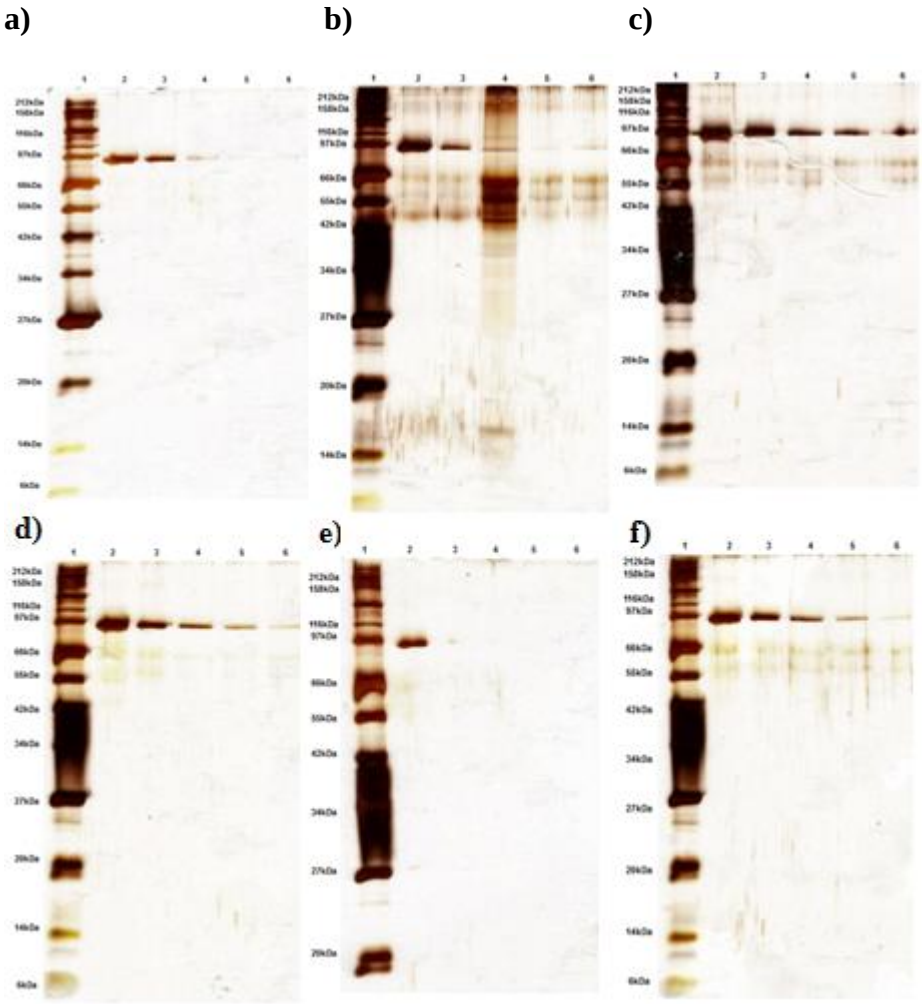


Fig 6.10: Interactions between the rChb CBP mutants and cellulose; a) W63A, b) K545A, c) D551A, d) D648A, e) W689Y and f) K545A D551A. The rChbL CBP mutant samples were continuously mixed with the cellulose substrate and samples were taken at various time-points as described in Section 2.29.1. The samples were analysed on SDS-PAGE (see Section 2.21) and visualised using silver nitrate staining (see Table 2.6). Lane 1, Protein Ladder; Lane 2, rChbL sample; Lane 3, 5 minutes incubation; Lane 4, 15 minutes incubation; Lane 5, 30 minutes incubation; Lane 6, 1 hour incubation.

Table 6.2 Summary of the interactions between the rChb CBP and rChb CBP mutants with α-chitin, β-chitin, chitosan and cellulose.

	α-Chitin	β-Chitin	Chitosan	Cellulose
rChb CBP	√	√	√	√
W63A	√	√	√	√
K545A	√	√	√	√
D551A	√	√	√	√
D648A	√	√	√	√
W689Y	√	√	√	√
K545A D551A	√	√	√	√

Figure 6.6 demonstrates the effect temperature played on the binding capabilities of the rChbL protein. Here the rChb CBP was tested at 45^{o}C and the binding patterns against the insoluble substrates observed were slightly different to those of the rChbL enzyme (see Figure 6.1). The rChb CBP was again observed to bind to α-chitin with high specificity, all the protein was bound to the insoluble substrate after 5 minutes. Definite binding to β-chitin was observed for the rChb CBP. Binding to chitosan and cellulose was observed, in contrast with the rChbL enzyme.

Binding to α-chitin was observed for all the rChb CBP mutants (see Figure 6.7). It was observed that each rChb CBP mutant bound to α-chitin with a similar affinity. All CBP was bound to the insoluble substrate after 5 minutes.

Figure 6.8 demonstrates that the rChb CBP mutants bound to β-chitin with a high affinity, there was no protein remaining in the culture media following 5 minutes of incubation. Binding to chitosan was observed for all the rChb CBP mutants, although to different extents. The highest affinity was observed for mutants D648A and W689Y, followed by K54A D551A. The remaining mutants had similar levels of affinity with almost no unbound protein detectable following 1 hour of incubation (see Figure 6.9). Figure 6.10 demonstrates the ability of the rChb CBP to also bind cellulose, in contrast with the native rChbL and the mutant enzymes. All rChb CBPs had the ability to bind cellulose, W689Y with the highest affinity. The W689Y

mutant had no unbound protein detectable following 5 mins of incubation at 37°C. D551A appeared to have the lowest affinity for cellulose.

6.4 Investigation into the inhibitory effects of the various length chito-oligosaccharides on rChbL catalytic activity

Section 6.3 examined the specificity of the chitobiase against the bulky chitin, chitosan and cellulose substrates and concluded that the rChbL had the highest affinity for α-chitin. This section looks more closely at the specificity of the rChbL enzyme and the active rChbL mutants by examining the effect of shorter chitin chains, known as chito-oligosaccharides, on rChbL activity. Chito-oligosaccharides produced by the depolymerisation of chitin were used in this study. Various length chito-oligosaccharides were used; chitobiose (2 GlcNAc residues), chitotriose (3), chitotetraose (4), chitopentaose (5), chitohexaose (6) and chitoheptaose (7). The inhibitory effect of each chito-oligosaccharide on the native rChbL enzyme and rChbL mutant enzymes catalytic activity was examined (see Figure 6.11).

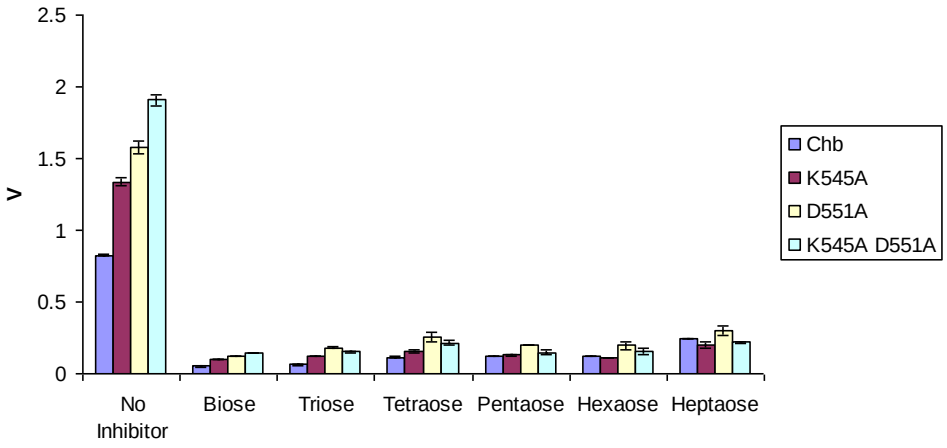


Fig 6.11: Inhibition of the rChbL and rChbL mutants' catalytic activity on PNP GlcNAc by various length chito-oligosaccharides. The Chb activity assay was carried out as described in Section 2.26. 1 µg/mL of enzyme was incubated with 10 mM of various length chito-oligosaccharides at 37^{o}C for 30 minutes before addition to the plate. The rate, V (mM PNP released/min/mL enzyme), of the reaction was plotted. Experiment is representative of three individual experiments. Error bars represent the standard deviation of three replicates.

Figure 6.11 demonstrates that all chito-oligosaccharides examined were able to significantly inhibit rChbL catalytic activity. The inhibition observed for each of the chito-oligosaccharides was relatively similar to one another. The native rChbL was inhibited by all chito-oligosaccharides tested. The highest level of inhibition was observed with chitobiose and the lowest level of inhibition was observed with chitoheptaose, however these values were relatively similar. The rChbL mutant enzymes followed a similar inhibition pattern. The highest level of inhibition for the K545A, D551A and K545A D551A mutants was observed with chitobiose, similarly to the native rChbL. The least inhibition for all the mutants was, again, observed with chitoheptaose. However, all signals were relatively similar indicating that the rChbL mutants were also inhibited by all the various length chito-oligosaccharides examined. These chito-oligosaccharides were potent inhibitors of rChbL activity; even when used at the low concentration of 10 mM, they reduced catalytic activity by over 90%.

6.5 Analysis of the breakdown of various length chito-oligosaccharides using High Performance Liquid Chromatography (HPLC)

The hydrolysis of various length chito-oligosaccharides by the rChbL enzyme was analyzed using High Performance Liquid Chromatography (HPLC). A C-18 silica reversed phase monolith column was used. C-18 is an octadecyl carbon chain bonded silica which makes the stationary phase hydrophobic. The mobile phase used (ammonium acetate) was polar and therefore the hydrophobic stationary phase allowed hydrophobic molecules to adsorb to the column. While sugars are generally hydrophilic, they do possess hydrophobic residues such as valine, leucine and proline. The ring structures of tryptophan, phenylalanine and tyrosine are also hydrophobic. These residues can give the sugar hydrophobic areas which presumably bound to the column. The sugar molecules were eluted by increasing the concentration of organic non-polar solvent (acetonitrile) in the mobile phase. Standards used were N-acetylglucosamine, chitobiose and chitohexaose. The hydrolysis of the chito-oligosaccharides was carried out as described in Section 2.30.2. Sugars are difficult to detect and need to be derivatised to add UV absorptive groups to their structure. Anthranillic acid was chosen for this purpose. It labels the sugars with secondary

amines (UV-absorbing groups) through reductive amination. All samples and standards had to be derivatised prior to addition to the column. They were labelled with the use of anthranillic acid as described in Section 2.30.3. HPLC analysis was carried out as described in Section 2.30.1.

6.5.1 Optimisation of HPLC separation conditions

Conditions for the separation of the various length sugars had to be optimised. The gradient of polar mobile phase to organic solvent was optimised to ensure a good resolution of each sugar (see Figure 6.12). The separation was run multiple times to ensure uniform elution of the sugars (see Table 6.3). The wavelength at which the highest signal for each sugar was obtained was determined (see Figure 6.13). The optimal volume of sugar sample to inject on the column was also ascertained (see Figure 6.14). Finally a standard curve of GlcNAc concentration was obtained for the quantification of chito-oligosaccharide hydrolysis to GlcNAc by the rChbL enzyme (see Figure 6.15).

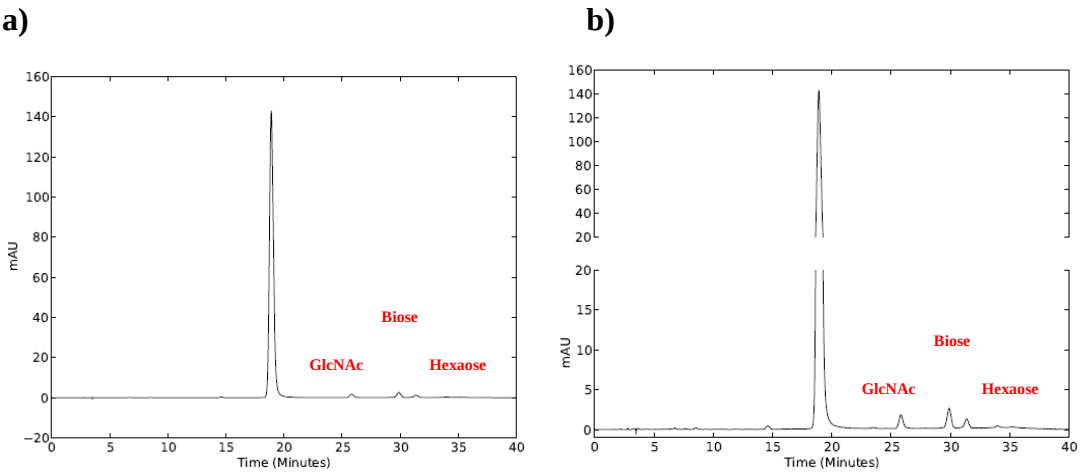


Fig 6.12: Optimised gradient for separation of GlcNAc, chitobiose and chitohexaose. The separation was run as described in Section 2.30. The gradient of polar mobile phase to organic solvent was adjusted. Various run-times and concentrations were examined.

Various run lengths and concentration gradients of mobile phases were used to ensure optimal separation of the sugars. Mobile phase A was sodium acetate while mobile phase B was the organic solvent, acetonitrile. GlcNAc was always resolved from the other sugars, however it was difficult to resolve the chitobiose and longer length chito-oligosaccharides. Figure 6.12 shows the separation of the GlcNAc, chitobiose and chitohexaose sugars. The separation of chitobiose and all the other length sugars was almost identical. Timing did not permit a separation to be optimised that separated out each sugar. The gradient program chosen was: 0-12 minutes (2%

Mobile phase B), 12-27 minutes (2-10% Mobile phase B), 27-35 minutes (10-2% Mobile phase B) and 35-40 minutes (2% Mobile phase B). It showed the optimum resolution between the three sugars.

Table 6.3 Retention times of the various sugars on multiple runs showing consistent separations.

Separation	Deriv. Agent	GlcNAc	Chitobiose	Chitohexaose
1	19.336	26.449	30.507	31.827
2	19.333	26.470	30.580	31.907
3	19.432	26.553	30.616	31.922
4	19.399	26.542	30.611	31.922
5	19.342	26.547	30.488	31.656
6	19.200	26.158	30.195	31.486
7	19.179	25.883	29.938	31.221
8	19.234	25.948	29.967	31.189
9	19.176	25.773	29.721	30.956
10	19.158	25.642	29.562	30.785

The retention times of the sugars were shown to be uniform over multiple runs. The standard deviation of the GlcNAc retention time was calculated to be 0.358424, while the standard deviation of chitobiose was 0.397155 and chitohexaose was 0.481977.

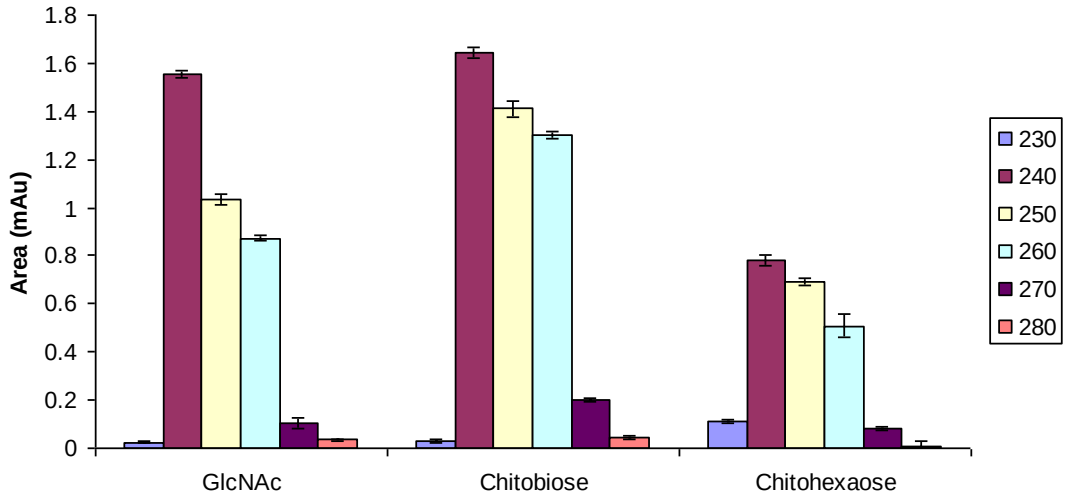


Fig 6.13: Optimisation of the wavelength used for the separation of sugars. The separation was run as described in Section 2.30 at various wavelengths; 230 nm, 240 nm, 250 nm, 260 nm, 270 nm and 280 nm.

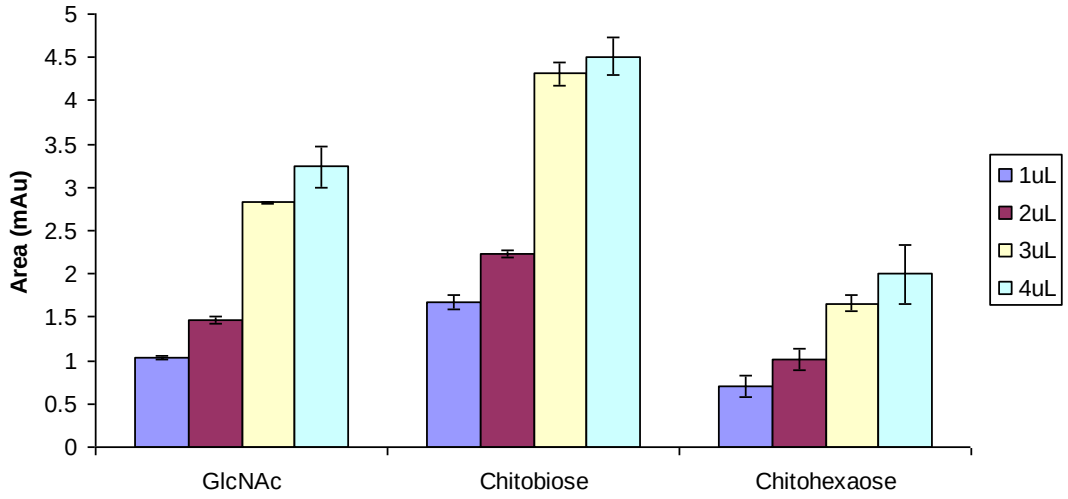


Fig 6.14: Optimisation of the injection volume used for the separation of sugars. The separation was run as described in Section 2.30 Increasing injection volumes were applied; 1µL, 2µL, 3µL and 4µL.

The optimal wavelength for detection of the sugars was determined to be 240nm (see Figure 6.14) while detection of the sugars after a 3 µL and 4 µL injection were relatively similar. A 3 µL injection was chosen as 4 µL caused the peaks to slightly shift.

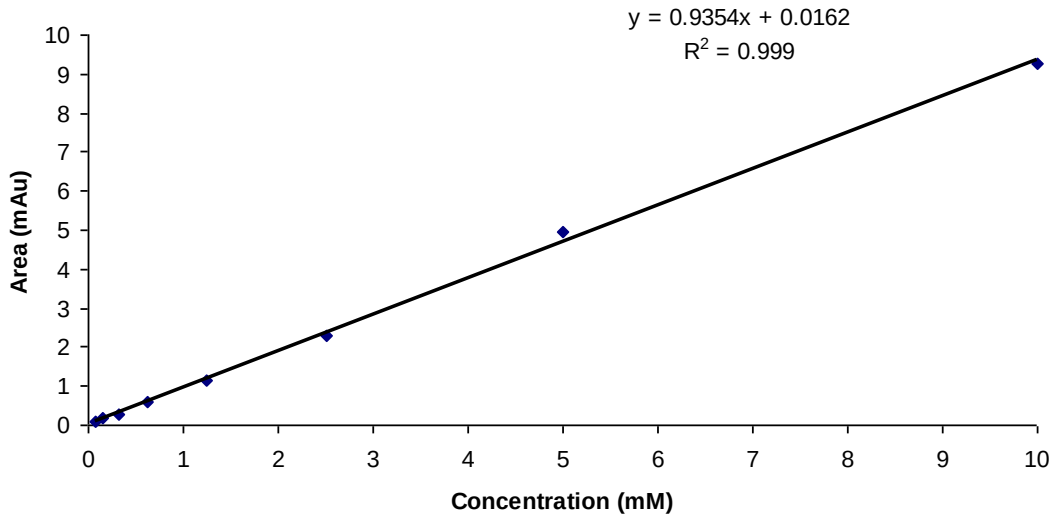


Fig 6.15: GlcNAc standard curve. Increasing concentrations of GlcNAc (from 0 to 10 mM) were derivatised as described in Section 2.30.3 and then analysed as described in Section 2.30.

6.5.2 Examination of the hydrolysis of chito-oligosaccharides by the rChbL enzyme

As the insoluble substrate binding and chito-oligosaccharide inhibition profiles of the rChbL enzyme and rChbL enzyme mutants were found to be similar (see Section 6.2), only the study of the hydrolysis of chito-oligosaccharides by the native rChbL enzyme was deemed necessary. The hydrolysis of these chito-oligosaccharides to GlcNAc was examined. The hydrolysis of chitobiose, chitotriose, chitotetraose, chitopentaose, chitohexaose and chitoheptaose by the rChbL enzyme over various time-points was carried out as described in Section 2.30.2. The samples were then derivatised as described in Section 2.30.3 and were separated via HPLC as described in Section 2.30. The results are shown in Figures 6.16 to 6.21 and Table 6.4.

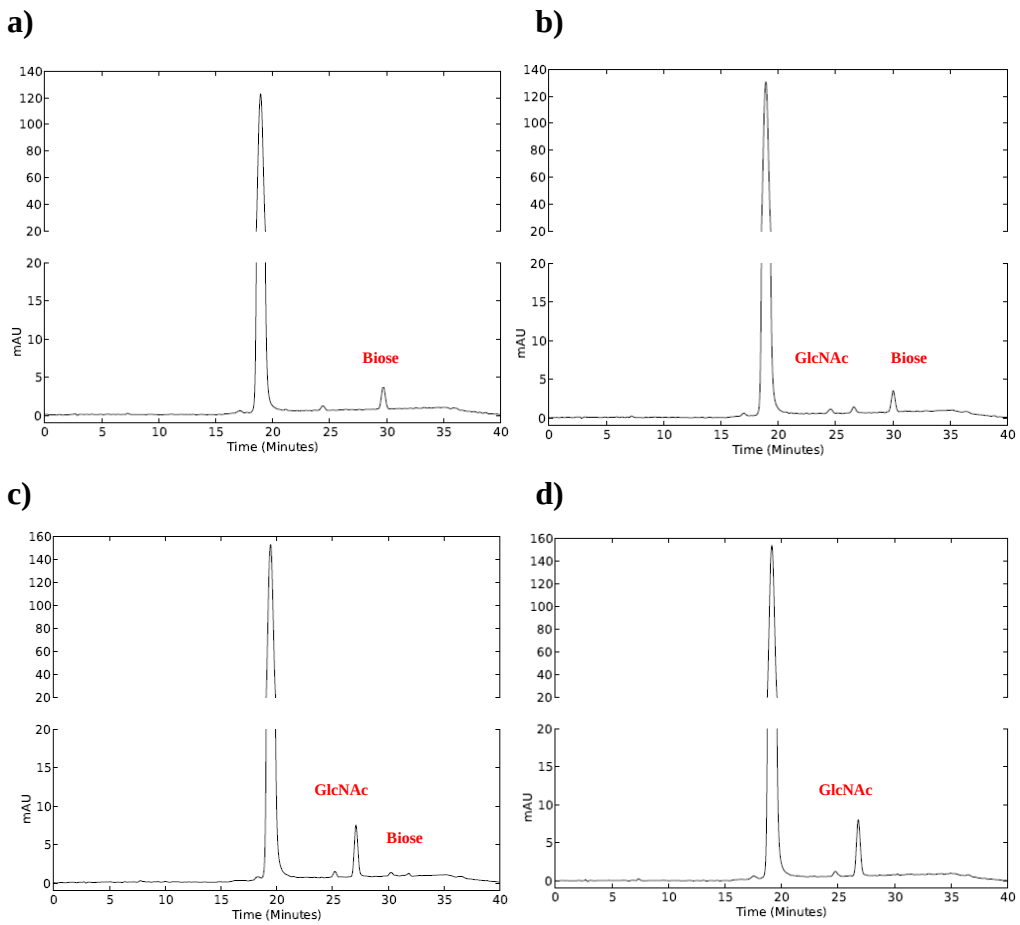


Fig 6.16: HPLC analysis of the hydrolysis of chitobiose by the rChbL enzyme over a time-course. a) No rChbL, b) 1 hour incubation with rChbL, c) 3 hours incubation with rChbL and d) 24 hours incubation with rChbL. The hydrolysis of chitobiose was carried out as described in Section 2.30.2 and the resultant samples were derivatised as described in Section 2.30.3. The samples were then run as described in Section 2.30. The peaks were identified by the separation shown in Figure 6.12 and the amount of GlcNAc released was calculated using the standard curve generated in Figure 6.15.

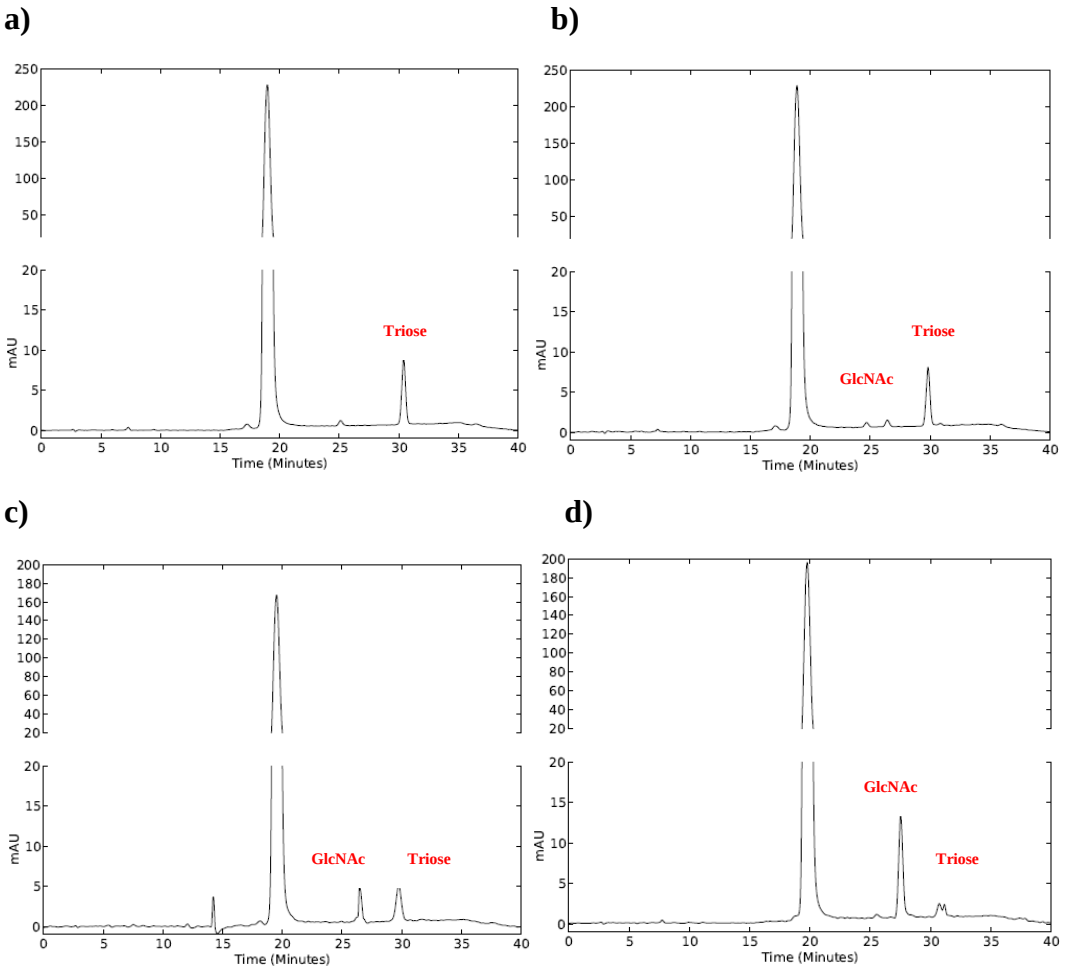


Fig 6.17: HPLC analysis of the hydrolysis of chitotriose by the rChbL enzyme over a time-course. a) No rChbL, b) 1 hour incubation with rChbL, c) 3 hours incubation with rChbL and d) 24 hours incubation with rChbL. The hydrolysis of chitotriose was carried out as described in Section 2.30.2 and the resultant samples were derivatised as described in Section 2.30.3. The samples were then run as described in Section 2.30. The peaks were identified by the separation shown in Figure 6.12 and the amount of GlcNAc released was calculated using the standard curve generated in Figure 6.15.

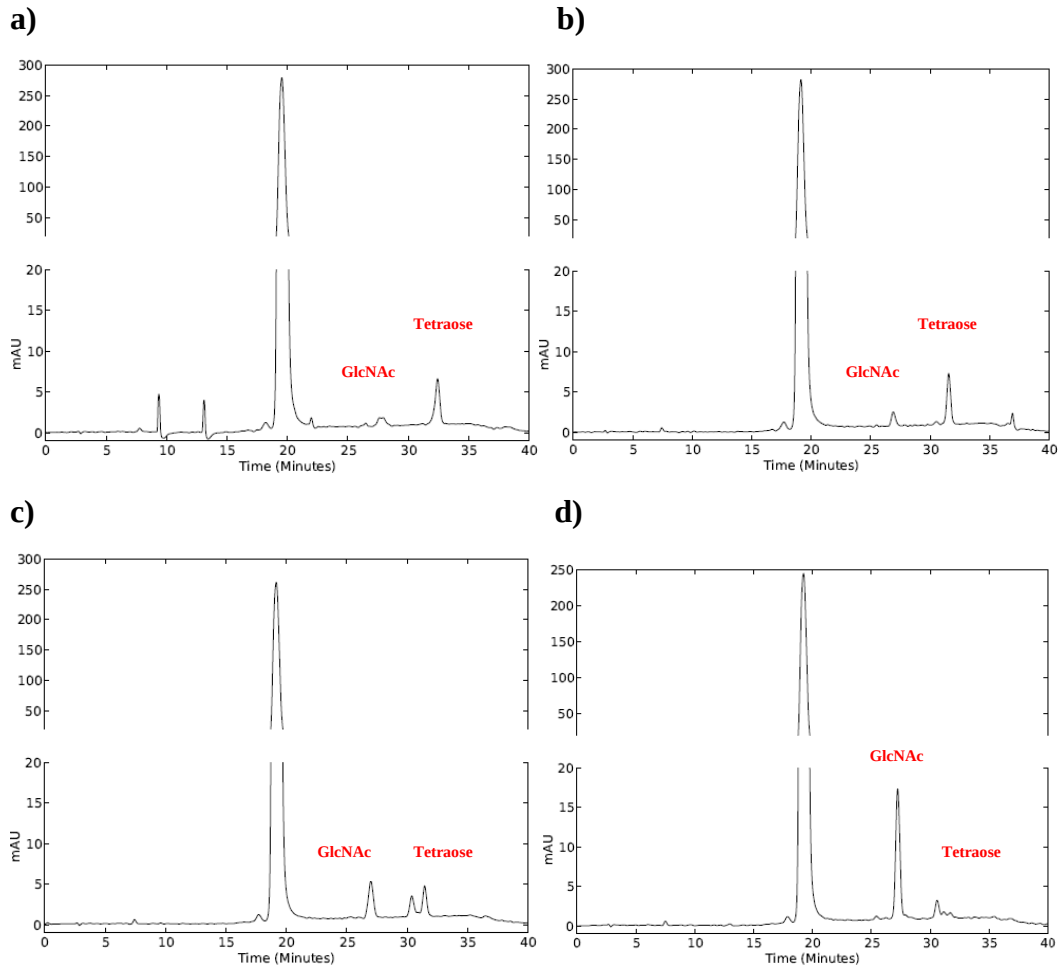


Fig 6.18: HPLC analysis of the hydrolysis of chitotetraose by the rChbL enzyme over a time-course. a) No rChbL, b) 1 hour incubation with rChbL, c) 3 hours incubation with rChbL and d) 24 hours incubation with rChbL. The hydrolysis of chitotetraose was carried out as described in Section 2.30.2 and the resultant samples were derivatised as described in Section 2.30.3. The samples were then run as described in Section 2.30. The peaks were identified by the separation shown in Figure 6.12 and the amount of GlcNAc released was calculated using the standard curve generated in Figure 6.15.

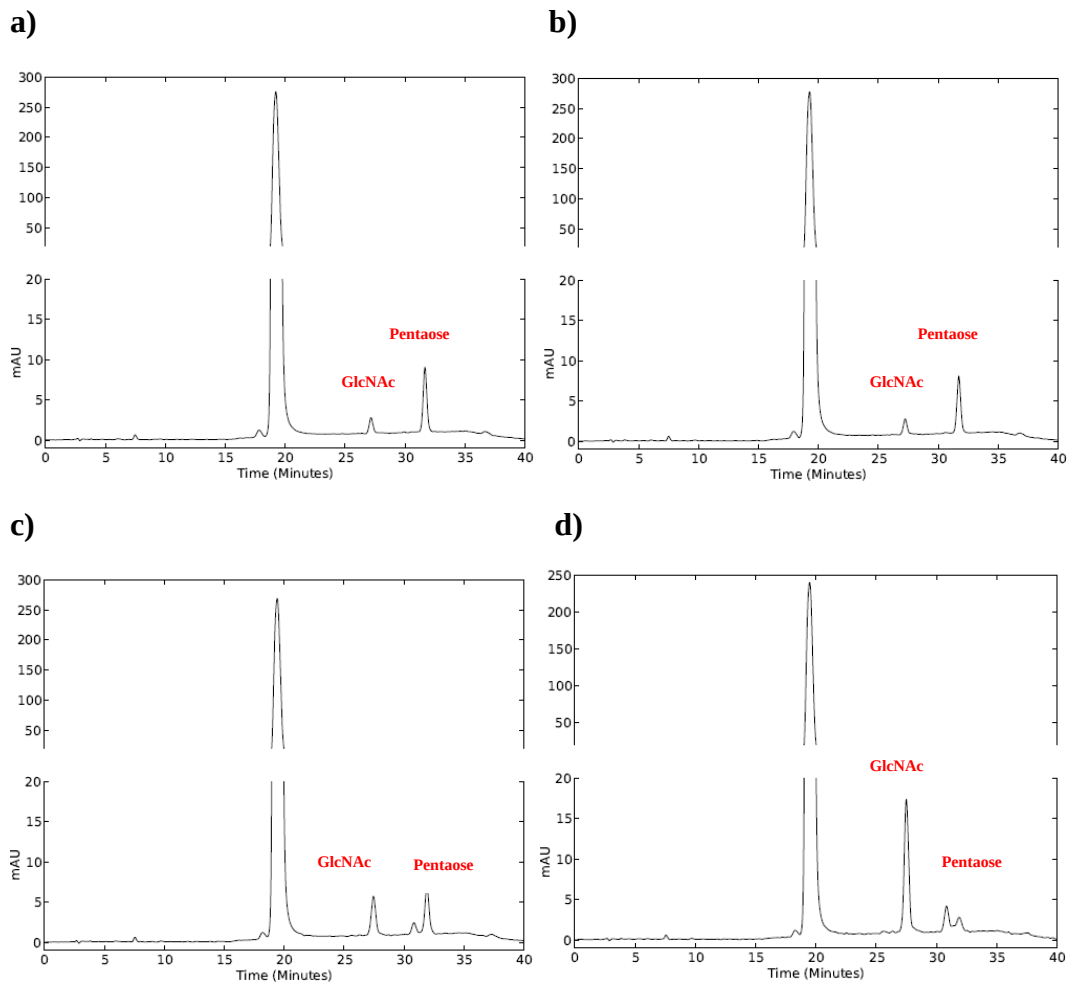


Fig 6.19: HPLC analysis of the hydrolysis of chitopentaose by the rChbL enzyme over a time-course. a) No rChbL, b) 1 hour incubation with rChbL, c) 3 hours incubation with rChbL and d) 24 hours incubation with rChbL. The hydrolysis of chitopentaose was carried out as described in Section 2.30.2 and the resultant samples were derivatised as described in Section 2.30.3. The samples were then run as described in Section 2.30. The peaks were identified by the separation shown in Figure 6.12 and the amount of GlcNAc released was calculated using the standard curve generated in Figure 6.15.

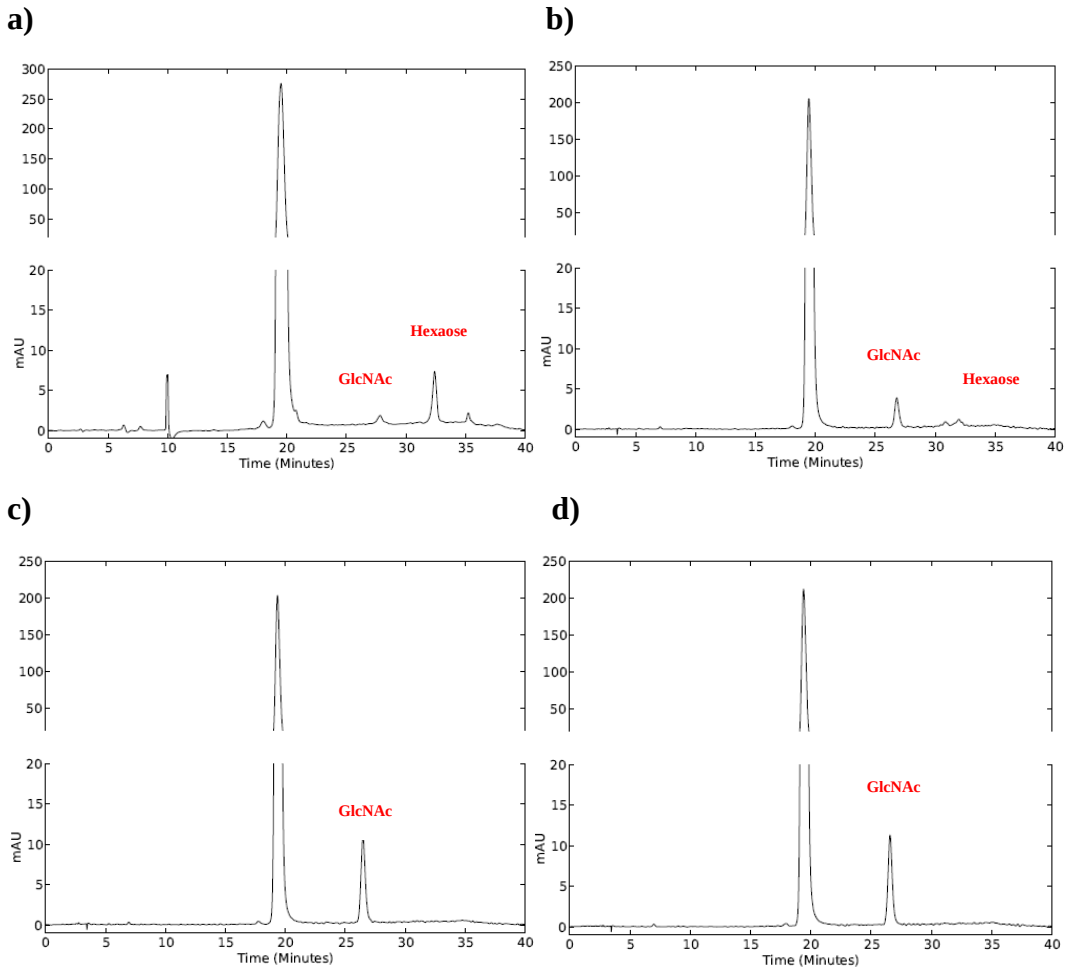


Fig 6.20: HPLC analysis of the hydrolysis of chitohexaose by the rChbL enzyme over a time-course. a) No rChbL, b) 1 hour incubation with rChbL, c) 3 hours incubation with rChbL and d) 24 hours incubation with rChbL. The hydrolysis of chitohexaose was carried out as described in Section 2.30.2 and the resultant samples were derivatised as described in Section 2.30.3. The samples were then run as described in Section 2.30. The peaks were identified by the separation shown in Figure 6.12 and the amount of GlcNAc released was calculated using the standard curve generated in Figure 6.15.

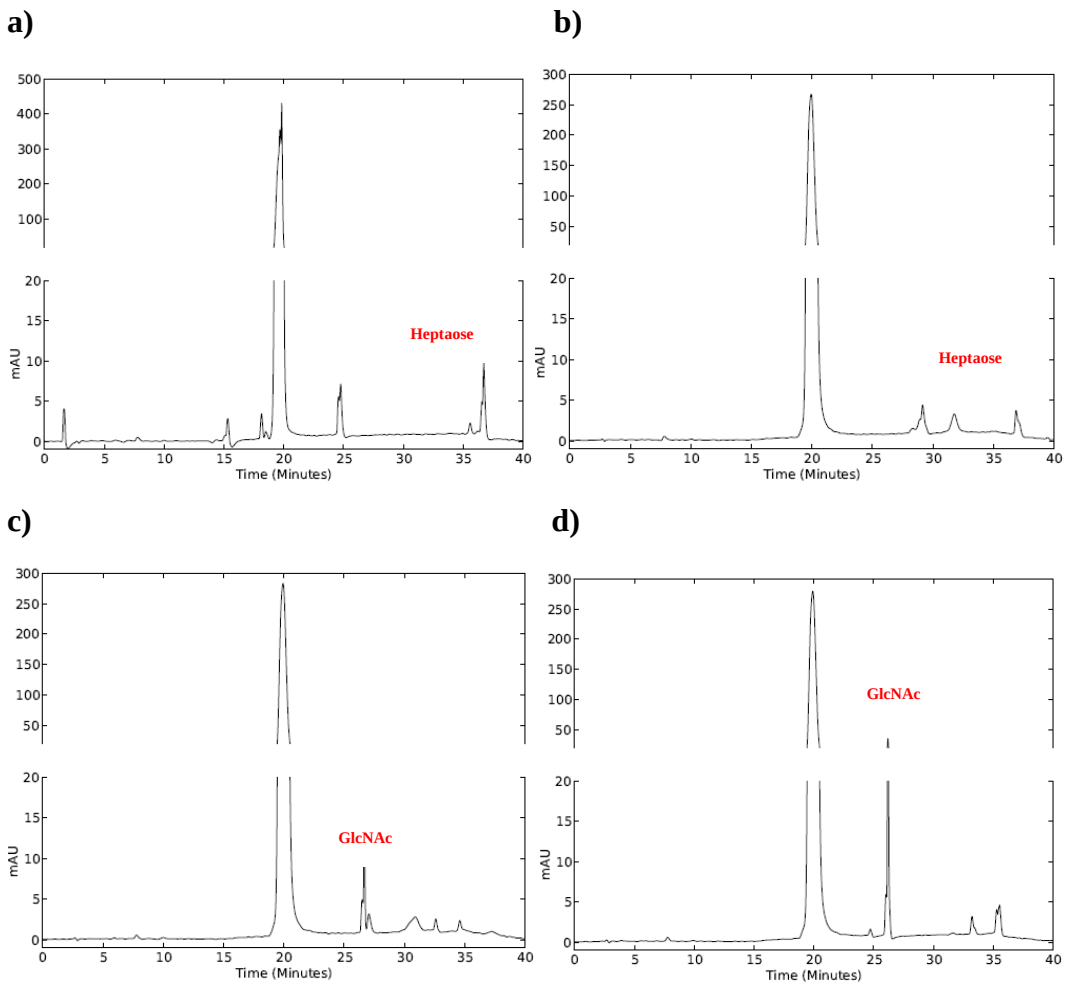


Fig 6.21: HPLC analysis of the hydrolysis of chitoheptaose by the rChbL enzyme over a time-course. a) No rChbL, b) 1 hour incubation with rChbL, c) 3 hours incubation with rChbL and d) 24 hours incubation with rChbL. The hydrolysis of chitoheptaose was carried out as described in Section 2.30.2 and the resultant samples were derivatised as described in Section 2.30.3. The samples were then run as described in Section 2.30. The peaks were identified by the separation shown in Figure 6.12 and the amount of GlcNAc released was calculated using the standard curve generated in Figure 6.15.

Table 6.4 The millimolar concentration of GlcNAc released from the hydrolysis of the chito-oligosaccharides by rChbL.

	0 hour	1 hour	3 hour	24 hour
Chitobiose	0.237	0.282	2.721	3.102
Chitotriose	0	0.078	1.466	5.450
Chitotetraose	0.670	0.808	2.223	7.244
Chitopentaose	0.633	0.707	2.244	7.684
Chitohexaose	0.505	1.707	4.848	4.811
Chitoheptaose	0	0	2.218	5.596

The hydrolysis of all chito-oligosaccharides by the rChbL enzyme released GlcNAc, each to varying degrees (see Figures 6.16 to 6.21). The large peak observed around 20 minutes was the excess derivatisation agent present in the sample. The highest levels of GlcNAc were generated through the hydrolysis of chitotetraose and chitopentaose. Less GlcNAc was released from the longer length chito-oligosaccharide, chitoheptaose. The hydrolysis of chitobiose proceeded at the fastest rate, with the majority of chitobiose hydrolysed to GlcNAc after 3 hours. The hydrolysis of chitohexaose also proceeded quickly, again with the majority of chitohexaose hydrolysed to GlcNAc after 3 hours. The hydrolysis of chitotetraose, chitopentaose and chitoheptaose proceeded at a slower rate.

Chitobiose was completely hydrolysed to GlcNAc after 24 hours. No other products were observed. Not all the chitotriose was hydrolysed to GlcNAc after the 24 hours. The presence of a double peak after 24 hours suggests that the chitotriose was broken down into smaller length chito-oligosaccharide chains, as well as GlcNAc. This was presumably chitobiose, the only chito-oligosaccharide smaller than chitotriose but larger than GlcNAc.

The majority of chitotetraose was hydrolysed to GlcNAc. Some intermediate hydrolytic products were observed after three hours, presumably chitobiose or chitotriose. The hydrolysis of chitopentaose proceeded at a relatively slow rate. Again, the majority of chitopentaose was hydrolysed to GlcNAc, with a small amount of intermediate product observed. The chitohexaose was completely hydrolysed to yield only GlcNAc after 24 hours. Intermediate hydrolytic products were present at

earlier time-points, presumably the lower length chito-oligosaccharides. The majority of chitoheptaose was hydrolysed to GlcNAc after 24 hours. Again, lower length chito-oligosaccharides were present.

6.6 Discussion

A chitin-binding assay protocol described by Suginta *et al.*, (2007) was used to study the binding patterns of rChbL to various insoluble substrates. This assay has been used previously to examine the effect of certain mutations on the specificity of a chitinase from *Vibrio carchariae*. Although the optimum temperature for the catalytic activity of rChbL has been demonstrated to be 45°C (see Figure 3.23) this assay was carried out at 4°C in order to minimise any catalytic activity. Quantification of unbound protein in the culture media by spectroscopy was impossible due to signal disruption caused by particulate insoluble matter. As chitin and its derivatives are insoluble substances it is necessary to visualise unbound and bound protein using SDS-PAGE analysis. The binding patterns of the rChbL enzyme and rChbL enzyme mutants yielded some interesting results.

The binding patterns of rChbL and its mutants were very similar with only one exception, mutant W689Y. All of the enzymes bound to α-chitin with high affinity. No unbound enzyme was detected in the culture media following five minutes of incubation and the bound enzyme was not removed even with lengthy boiling. On the other hand, there was very little, if any, binding observed to β-chitin for all but one mutant. Mutant W689Y was found to bind to β-chitin with a relatively high affinity compared with the native rChbL enzyme. The point mutation may have slightly changed the structure of the active site allowing the binding of β-chitin.

The predominant form of chitin found in nature is α-chitin. β-chitin is much rarer as it is only found in squid pens. *P. luminescens* is an insect pathogen. Although it is not known what role chitin-degrading enzymes may play in this process, high affinity binding of rChbL to α-chitin was expected. The difference in the structures of α-chitin and β-chitin may also explain the difference in binding and affinity. The structures of α-chitin and β-chitin are described thoroughly in the introduction. The chains of N-

acetylglucosamine residues are organised in a parallel fashion in α-chitin while they are organised in an anti-parallel fashion in β-chitin. The parallel sheets of α-chitin are held together by strong hydrogen bonds while the anti-parallel sheets of β-chitin are only held together by weaker hydrogen bonds (Jang *et al*., 2004). Perhaps the rChbL enzyme has a high affinity for the tightly organised structure of α-chitin rather the more loosely packed together β-chitin. The distribution of charge may also play a role here. The active site of the rChbL is found in a deep cleft in the enzyme structure. The organised α–chitin structure may enter the active site with more ease than the loosely packed β–chitin.

Binding of rChbL and its mutants to chitosan and cellulose was not observed following incubation at 4^{o}C for one hour. While chitosan is a deacetylated form of chitin, cellulose differs from chitin due to the absence of an N-acetyl group at the C-2 position. These variations in structure obviously play a role in the ability of rChbL to bind to these substrates with a comparable affinity to α-chitin binding.

Although it was necessary to perform the rChbL insoluble substrate binding assays at 4^{o}C to prevent catalytic activity the low temperature may also have inhibited binding to the substrates. The results demonstrated however that α-chitin was the preferred substrate of the rChbL enzyme, while the W689Y mutant was capable of binding to both α- and β-chitin.

The advantage of using the catalytically inactive rChb CBP was that insoluble substrate assays could be carried out at the proteins optimum temperature without the worry of catalytic activity occurring. In comparison at 45°C, the rChb protein was not as selective as rChbL; it bound to all insoluble substrates tested. High affinity binding to α-chitin and β-chitin was observed, while binding to chitosan and cellulose was observed, albeit with a lower affinity. It was concluded that while rChb CBP bound to all of the insoluble substrates tested, it did have a preference for α-chitin. Similar results may also have been observed for the catalytically active rChb had the reaction been monitored for a longer period of time at 4°C. The binding of chitinases to chitin was examined by Suginta *et al.,* (2007). They also mutated various residues in the active site of their chitinase and reported no difference in the binding affinities towards chitin.

While analysis of the binding affinities of the rChbL enzyme was useful and gave insight into its affinity for the different forms of chitin, it gave no indication as to the catalytic action of the enzyme on chitin. As stated previously rChbL is analogous to the *S. marcescens* chitobiase Chb. Chitobiases such as Chb are capable of cleaving chitobiose into GlcNAc monomers and do not break down chitin into smaller chito-oligosaccharides. In order to determine the hydrolytic action of rChbL HPLC analysis was employed to study the break-down of varying length chito-oligosaccharides by rChb. First, the inhibitory effect of the chito-oligosaccharides on activity was examined using the PNP-GlcNAc assay. This experiment revealed that all of the chito-oligosaccharides were potent inhibitors, inhibiting the reaction by 90%. There was very little difference in the levels of inhibition by each chito-oligosaccharide observed. This result demonstrated that rChbL was able to bind to each chito-oligosaccharide. Again, however, binding a substrate was not the same as hydrolysing it.

There are many studies that describe the hydrolysis of chito-oligosaccharides by various chitinases and chitobiases and the resulting analysis of the reaction products. The methods typically used for this analysis are either TLC or HPLC. The examination of chito-oligosaccharides by HPLC has been extensively described in the literature and was the method of choice in this study. Van Eikeren and McLaughlin (1977) described the use of a carbohydrate column and an eluent gradient of 70:30 acetonitrile:water. More recently Chang and Fu (2000) optimised this acetonitrile:water gradient going from 80:20 to 60:40. The derivatisation and separation protocols used in this work were taken and adapted from work by Alwael *et al.*, (2011). Alwael *et al.* devised a HPLC method for the determination of monosaccharides in complex samples originating from the production of biopharmaceuticals. Samples such as cell-culture media and fermentation broths were tested for the presence of sugars such as glucose, glucosamine, galactose, galactosamine, mannose, mannosamine, fucose and sialic acid. Monosaccharides must be labelled for detection on a HPLC system due to the absence of UV absorbing groups in their structure. The monosaccharides were derivatised using anthranillic acid before injection onto the column. This added a secondary amine group through reductive amination to the sugar which is a UV absorbing group. However, the presence of N-acetyl groups on sugars can cause lower rates of reductive amination as

their reducing ends are known to be less reactive and therefore lower yields of secondary amines and lower UV signals are observed (Rustighi *et al.*, 2009). Alwael *et al.*, (2011) did not include any sugars with N-acetyl groups in their study. Rustighi *et al.*, (2009) compared the derivatisation efficiency of various sugars with and without N-acetyl groups using three different chemicals dyes; 4-aminobenzonitrile, 2-aminopyridine and 2-aminobenzoic acid (anthranillic acid). The resultant samples were analysed by capillary electrophoresis (CE) with UV detection. They found that the optimal dye for monosaccharide detection was anthranillic acid; this dye allowed for the most sensitive detection of the sugars containing N-acetyl groups.

Suginta *et al.*, (2010) studied the hydrolysis of various length chito-oligosaccharides by an N-acetylglucosaminidase from *Vibrio harveyi 650* using HPLC. They discovered that this N-acetylglucosamindase hydrolysed all chito-oligosaccharides tested but had a slight preference for chitotetraose, followed by chitotriose, while its least favourite chito-oligosaccharide was chitobiose. TLC analysis showed that the N-acetylglucosamindase was exo-acting, cleaving GlcNAc from the end of each chito-oligosaccharide, the final product being only GlcNAc. Hoell *et al.*, (2005) used HPLC to analyse chito-oligosaccharide hydrolysis by a novel chitinase from *Trichoderma atroviride*. This chitinase was found to be endo-acting. For example, the hydrolysis of chitohexaose yielded chitotetraose, chitotriose and chitobiose, but not GlcNAc.

Analysis in this study revealed the rChbL enzyme was primarily exo-acting. The main hydrolytic product at all time-points was GlcNAc. It was hypothesized that the rChbL enzyme primarily cleaved GlcNAc from the non-reducing ends of the chito-oligosaccharides. Its preferential substrate was deemed to be chitobiose as this was the reaction that proceeded the fastest. It also favoured chitotriose and chitohexaose over the other substrates. The hydrolysis of chitoheptaose proceeded at the slowest rate and this substrate was therefore deemed to be the least favoured by the rChbL enzyme. In the end it was concluded that rChbL was an exo-acting *N*-acetylglucosaminindase (chitobiase), capable of cleaving single GlcNAc units from various length chito-oligosaccharides.

This chapter succeeded in determining the binding specificities of rChbL and rChb CBP for various insoluble substrates as well as the hydrolytic capabilities of the N-

acetylglucosaminidase rChbL. These results highlight the promise of rChbL as a commercial enzyme for the production of GlcNAc from chitin substrates.

Chapter 7

Final Discussion, Conclusions and Recommendations

Chitin is a readily available bioresource, with between 10^{10}-10^{12} metric tons produced annually. Chitin breakdown is a commercially lucrative process due to the economic value of its' breakdown products such as chito-oligosaccharides and N-acetylglucosamine (Gortari and Hours, 2013). These products have uses in areas such as the food industry, pharmaceutical industry, cosmetic industry and textile industry. As previously discussed chemical methods have been widely used for chitin breakdown, however these have many disadvantages including non-specific products and the generation of toxic waste. There is a general movement towards the biological hydrolysis of chitin biowaste, this can be carried out through microbial fermentation or enzymatic hydrolysis (Gortari and Hours, 2013). This study is interested in the enzymatic hydrolysis of chitin with the use of enzymes such as chitinases and chitobiases. The hydrolysis of chitin by chitinolytic enzymes from various sources is difficult to compare due to the different processes and parameters employed. The source of chitin differs, along with how the waste is pre-treated, the temperature/pH of the process and how the supplements added. Enzymes from various sources such as *Serratia marcescens*, *Streptococcus pneumoniae* and *Vibrio harveyi* have all been employed for chitin breakdown (Gortari and Hours, 2013; Suginta *et al.,* 2010; Tews *et al*. 1996).

This study described the cloning, expression and purification of a chitobiase enzyme from *P. luminescens* with the final aim of a creating an enzyme for commercial chitin breakdown. The enzyme was cloned from the *chb* gene from *P. luminescens* and expressed in a bacterial expression system (*E. coli*). The addition of a His_6 tag to the recombinant enzyme allowed the use of affinity purification. This expression/purification system enabled the fast production of relatively high yields of pure enzyme (rChbL). The purified enzyme was then characterised, its kinetic properties were investigated and compared with a commercial chitobiase using a modified assay. The major issue encountered was the identification of the best fraction from which to harvest the recombinant protein. The protein was initially found in the insoluble cellular fraction, namely the periplasmic fraction. On further investigation it was noted that a portion of the protein was being exported to the culture media. The recombinant protein was expressed with its' signal peptide however this is a signal peptide recognised by *P. luminescens*. Efforts to increase the amount of secreted recombinant protein could include the addition of a signal peptide

recognised by *E. coli* such as pelB. This signal peptide directs the protein to the periplasmic space where it is cleaved and secreted. This would possibly increase the recombinant protein yields obtained.

A recombinant chitobiase enzyme was generated that had comparable activity to an already commercially available recombinant chitobiase enzyme from *Xanthomonas manihotis* available from New England Biolabs. Sigma Aldrich also advertises a recombinant chitobiase from a microbial source, from *Streptococcus pneumoniae*. This highlights that there is a market for these products. The rChbL from this study could be less costly to produce due to the engineered recombinant purification tag which leads to less downstream processing and therefore less cost. The fact that it can also be obtained directly from the culture media reduces production costs.

A catalytically inactive chitobiase protein (rChb CBP) was generated to establish whether the homologous residues in the *S. marcescens* enzyme were also the catalytic residues in the *P. luminescens* enzyme. The catalytically active residues were mutated and the catalytic activity was shown to be removed from the enzyme. The carbohydrate-binding properties of the new carbohydrate-binding protein were examined against terminal GlcNAc residues on a variety of glycoproteins as well as a synthetic BSA-GlcNAc conjugate and an O-linked glycan, bovine serum mucin. Binding was examined using various ELLA formats, however no interaction between the rChb CBP and GlcNAc was observed. The immobilisation of the rChb CBP to a sepharose matrix was completed successfully, however binding to GlcNAc was still not observed. Detection of binding between the rChb CBP and GlcNAc was the major issue encountered in this chapter. While the His_6 tag allowed for the purification of the recombinant protein it did not allow detection in an ELLA format. The addition of a linker to extend the His_6 tag from the protein did not enable detection in the ELLA format but did improve purification of the protein. Biotinylation of the protein enabled detection but it was noted that there were potential biotinylation sites in the binding pocket of the protein that could interfere with potential binding of GlcNAc residues. Further work could be carried out to enable better detection of the rChb CBP in the ELLA format. A possible way to do this would be the addition of a His patch. While the His tag was at the C-terminus of the protein and was possibly hidden by the tertiary protein structure a His patch may be engineered so that it would be present on the protein surface. However, this may cause issues with protein folding and structure

and ultimately affect its binding. Alternatively, the His_6 tag may be engineered at the N-terminus of the protein. While not ideal in this case due to the cleavage of the signal peptide and consequent removal of the His_6 tag, it would be possible to engineer the tag after the signal peptide sequence.

The next objective was to improve the binding ability of the rChb CBP through site-directed mutagenesis. Residues were chosen around the active site that were proposed to be interfering with the binding of the bulky glycoprotein structure. Five single mutants were successfully cloned, expressed and purified. Two mutations (K545A and D551) resulted in binding to a BSA-GlcNAc conjugate being observed, while the combination of these mutations in a single mutant increased the signal observed again. This reinforced the hypothesis that the rChb CBP was unable to bind to the GlcNAc residues on glycoproteins due to its binding pocket being located deep inside the enzyme and denying access to the large glycoprotein. The binding of the BSA-GlcNAc conjugate was a positive result, and the slight increase in binding to bovine serum mucin and asialo agalacto fetuin demonstrated that there was potential to generate a carbohydrate binding protein for glycoanalysis. However the overall aim of creating a carbohydrate-binding protein capable of binding GlcNAc residues on glycoproteins was not achieved. Further mutagenesis may be carried out on the binding pocket of the rChb CBP to make it more accessible for glycoprotein binding. The rChb CBP double mutant (K545A D551A) may be used as a starting point and further mutagenesis of the surrounding residues to alanine carried out as this technique was successful in this study. Random mutagenesis of these and surrounding residues could also generate interesting results. Currently the only N-acetylglucosamine binding protein available on the market is GSLII, available from Vector Labs. As N-acetylglucosamine is part of the core N-glycan structure and present on almost all glycoproteins there is a large gap in the glycoanalysis market for a protein with N-acetylglucosamine binding capabilities. The issue with GSLII is that it is obtained from plant origin so there are issues encountered with purity, batch-batch-to batch variation and supply. The generation of a recombinant N-acetylglucosamine binding protein would be commercially valuable as there would be a constant supply with no batch-to-batch variation. Carbohydrate-binding proteins are critical for the biopharmaceutical industry as they are used for glycoanalysis to

examine the various glycoforms of its protein therapeutics. It is a critical quality control measure.

We also investigated what effect the site specific mutations had on the rChbL enzyme due to their proximity to the active site. Again, five single mutants were successfully cloned, expressed and purified. Interestingly, the mutations that increased the binding abilities of the rChb CBP (K545A and D551A) also increased the catalytic activity of the rChbL enzyme. The combination of these mutations in a single mutant acted additively and significantly increased the catalytic activity of the rChbL double mutant threefold in comparison with the native rChbL. The mutants with increased catalytic activity also demonstrated an increased stability over a wider pH range and higher temperatures, all properties that are of interest in the generation of an enzyme for commercial/industrial use. The increased activity of these mutants highlighted that they could be a more cost effective product than those already available, due to their higher levels of activity. Also activity at wider pH ranges and temperatures is important due to the fact that chitin hydrolysis can involve the use of acids/bases and high temperatures.

Finally the activity of rChbL against natural substrates was investigated. At 4^{o}C it was shown to bind both α–chitin and β–chitin, but not cellulose and chitosan. The rChb CBP was used to examine its interactions with these insoluble substrates at a higher temperature (45^{o}C), as 4^{o}C is not very industrially relevant. Binding to all insoluble substrates was observed at this higher temperature. The use of the lower temperature allowed the relative affinity of the rChbL for the insoluble substrates to be determined. The rChbL had a higher affinity for α–chitin and β–chitin than chitosan and cellulose. The study of the interaction of the rChbL with different length chito-oligosaccharides generated some interesting results. The enzyme was inhibited by all length chito-oligosaccharides. It also hydrolysed all length chito-oligosaccharides to GlcNAc after 24 hour incubation. The study of the ability of the rChbL site specific mutants to hydrolyse the different length chito-oligosaccharides would be an interesting future study. The development of a separation of all the different length sugars would also be worthwhile. This would enable the various breakdown products to be identified and provide a more thorough insight into the preferences of the rChbL enzyme and its mechanism of action. This work so far has allowed us to conclude that

the enzyme is probably an exo-acting N-acetylglucosaminidase cleaving GlcNAc from the terminal end of chito-oligosaccharides.

While the rChb CBP is not a viable option for use in glycoanalysis, this work has provided a potential principle for the generation of carbohydrate-binding protein for glycoanalysis. Enzymes are extremely specific in the substrates they bind which can be exploited for the field of glycoanalysis. Enzymes of a substrate specificity of interest may be catalytically inactivated and used as novel carbohydrate-binding proteins. There is a lack of commercial GlcNAc binders available for glycoanalysis, with only GSLII available. Other chitobiases or chitin hydrolysing enzymes could be catalytically inactivated in the search for a GlcNAc binding protein that could be produced recombinantly to ensuring a constant, consistent supply.

The rChbL enzyme has proven itself as a possible candidate for commercial use. Industrially enzymes need to be able to perform at fast rates, at relatively high temperatures and over a broad pH range to account for fluctuations that occur over the duration of the process. The rChbL double mutant particularly, has many of these characteristics. It catalysed the hydrolysis of PNP-GlcNAc three times faster than a commercially available chitobiase; it has proven to be active at 45^{o}C and has optimal activity over a wide pH range, from pH 6-10. The fact that the rChbL enzyme retained some of its catalytic activity after immobilisation to a sepharose matrix is a major advantage, which could possibly be improved by using different immobilisation techniques. The enzymatic hydrolysis of substances, including chitin, is being carried out in bioreactors using immobilised enzymes. This allows the re-use of the enzyme and the continuous removal of product.

The rChbL has potential uses in the breakdown of chitin and chito-oligosaccharides. Used in conjunction with endo-acting chitinases it would be of possible use in the hydrolysis of chitin and chito-oligosaccharides to N-acetylglucosamine. This is useful for the production of N-acetylglucosamine which has many uses as previously described (see Section 1.1.3.3). It could also be used for the breakdown of chitin biomass.

Chapter 8

References

Aam, B., Heggset, E., Norberg, A., Sorlie, M., Varum, K. and Vincent, E. 2010. Production of chito-oligosaccharides and their potential applications in medicine. *Marine Drugs* 8(5), pp1482-1517.

Alwael, H., Connolly, D. and Paull, B. 2011. Liquid chromatographic profiling of monosaccharide concentrations in complex cell-culture media and fermentation broths. *Analytical Methods* 3, pp62-69.

Ambrosi, M., Cameron, N.R. and Davis, B.G. 2005. Lectins: tools for the molecular understanding of the glycocode. *Organic and Biomolecular Chemistry* 3(9), pp1593-1608.

Anumula, K.R. 1994. Quantitative determination of monosaccharides in glycoproteins by high-performance liquid chromatography with highly sensitive fluorescence detection. *Analytical Biochemistry* 220(2), pp275-283.

Apweiler, R., Hermjakob, H. and Sharon, N. 1999. On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochimica et Biophysica Acta* 1473(1), pp4-8.

Ashry, E. and Aly, M. 2007. Synthesis and biological relevance of N-acetylglucosamine-containing oligosaccharides. *Pure and Applied Chemistry* 79(12), pp2229-2242.

Aub, J.C., Sanford, B.H. and Cote, M.N. 1965. Studies on reactivity of tumor and normal cells to a wheat germ agglutinin. *Procedures of the National Academy of Science* 54(2), pp396-399

Aub, J.C., Tieslau, C. and Lankester, A. 1963. Reactions of normal and tumor cell enzymes. I. Wheat-germ lipase and associated mucopolysaccharides. *Procedures of the National Academy of Science* 50(4), pp613-619.

Baker, J., Celik, E. and DeLisa, M. 2013. Expanding the glycoengineering toolbox: the rise of bacterial N-linked protein glycosylation. *Trends in Biotechnology* 31(5), pp313-323.

Baneyx, F. (1999). Recombinant protein expression in *Escherichia coli. Current Opinion in Biotechnology* 10(5), pp411-421.

Beckham, G.T. and Crowley, M.F. 2011. Examination of the α–chitin strucuture and decrystallization thermodynamics at the nanoscale. *The Journal of Physical Chemistry* 115(15), pp4516-4522.

Bendiak, B., Harris-Brandts, M., Michnick, S., Carver, J. and Cumming, D. 1989. Separation of the complex asparagine-linked oligosaccharides of the glycoprotein fetuin and elucidation of three triantennary structures having sialic acids linked only to galactose residues. *Biochemistry* 28(15), pp6491-6499.

Bhattacharya, D., Nagpure, A. and Gupta, R. 2007. Bacterial chitinases: properties and potential. *Critical Reviews in Biotechnology* 27(1), pp21-28.

Birnboim, H.C. and Doly, J. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acid Research.* 7(6), pp1513-1523.

Bissett, D.L. 2006. Glucosamine: an ingredient with skin and other benefits. *Journal of Cosmetic Dermatology* 5(4), pp309-315.

Bisset, D.L., Robinson, L.R., Raleigh, P.S., Miyamoto, K., Hakozaki, T., Li, J. and Kelm, G.R. 2006. Reduction in the appearance of facial hyperpigmentation by topical N-acetylglucosamine. *Journal of Cosmetic Dermatology* 6(1), pp20-26.

Blake, C., Koenig, D., Mair, G., North, A., Phillips, D. and Sarma, V. 1965. Structure of hen egg white lysozyme. A three dimensional Fourier synthesis at 2 Å resolution. *Nature* 206(4986), pp757-763.

Boyd, W.C. and Shapleigh, E. 1954. Specific precipitating activity of plant agglutinins (lectins). *Science* 119 (3091), pp419.

Brameld, K. and Goddard III, W. 1998. The role of enzyme distortion in the single displacement mechanism of family 19 chitinases. *Procedures of the Natural Academy of Science* 95(8), pp4276-4281.

Brena, B.M and Batista-Viera, F. 2006. Immobilization of Enzymes – A literature Survey. *IN:* Guisan, J.M. *Methods in Biotechnology: Immobilization of Enzymes and Cells*, 2nd Ed. New Jersey: Humana Press pp15-30.

Brurberg, M.B., Nes, I.F. and Eijsink, V.G. 1996. Comparative studies of chitinases A and B from *Serratia marcescens*. *Microbiology* 142(7), pp1581-1589

Brurberg, M., Synstad, B., Klemsdal, S., van Aalten , D., Sundheim, L. and Eijsink, V. 2001. Chitinases from *Serratia marcescens*. *Recent Research Developments in Microbiology* 5(1), pp187-204

Burda and Aebi 1999. The dolichol pathway of N-linked glycosylation. *Biochimica et Biophysica Acta* 1426(2), pp239-257.

Burtan, A. and Freeman, H. 1993. *N-acetylglucosamine as a gastroprotective agent*, WO9323055.

Camby, I., Le Mercier, M., Lefranc, F. and Kiss, R. 2006. Galectin-1: a small protein with major functions. *Glycobiology* 16(11), pp.137-157.

Cantarel, B.L., Countinho, P.M., Rancurel, C., Bernard, T., Lombard, V. and Henrissat, B. 2008. The carbohydrate-active enzymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Research* 37(2), pp233-238.

Chen, G., Zhang, Y., Li, J., Dunphy, G.B., Punja, Z. and Webster, J.M. 1996. Chitinase activity of *Xenorhabdus* and *Photorhabdus* species, bacterial associates of entomopathogenic nematodes. *Journal of Invertebrate Pathology* 68(2), pp101-108.

Chen, J. and Chang, K. 1994. Immobilization of chitinase on a reversibly soluble-insoluble polymer for chitin hydrolysis. *Journal of Chemical technology and Biotechnology* 60(2), pp133-140.

Chen, J., Shen, C. and Liu, C. 2010. N-acetylglucosamine: Production and applications. *Marine Drugs* 8(9), pp2493-2516.

Cheon, Y., Park, H., Kim, J., Kim, Y. and Kim, H. 2004. Manipulation of the active site loops of D-Hydantoinase, a $(\beta/\alpha)_8$-barrel protein, for modulation of the substrate specificity. *Biochemistry* 43(22), pp7413-7420.

Clarke, V., Platt, N. and Butters, T. 1995. Cloning and Expression of the β-N-Acetylglucosaminidase Gene from *Streptococcus pneumoniae*. *The Journal of Biological Chemistry* 270(15), pp8805-8814.

Corbett, K., Fordham-Skelton, A.P., Gatehouse, J.A. and Davis, B.G. 2001. Tailoring the substrate specificity of the β-glycosidase from the thermophilic archaeon *Sulfolobus solfataricus*. *FEBS Letters* 509(3), pp355-360.

Cunningham, B.C., and Wells, J.A. 1989. High-resolution epitope mapping of hGH-Receptor interactions by alanine-scanning mutagenesis. *Science* 244(4908), pp1081-1085.

Dahiya, N., Tewari, R. and Hoondal, G.S. 2005. Biotechnological aspects of chitinolytic enzymes: a review. *Applied Microbiology and Biotechnology* 71(6), pp773-782.

Dalpathado, D.S. and Desaire, H. 2008. Glycopeptide analysis by mass spectrometry. *The Analyst* 133(6), pp731-738.

Davies, G. and Henrissat, B. 1995. Structures and mechanisms of glycosyl hydrolases. *Structure* 3(9), pp853-859.

De Assis, C.F., Araujo, N.K., Pagnoncelli, M.G., Pedrini, M.R., De Macedo, G.R. and Santos, E.S. 2010. Chitooligosaccharides enzymatic production by *Metarhizium anisopliae*. *Bioprocess and Biosystems Engineering* 33(7), pp893-899.

De Marco, J., Valadares-Inglis, M.C. and Felix, C.R. 2004. Purification and characterization of an N-acetylglucosaminidase produced by *Trichoderma harzianum* strain which controls *Crinipellis perniciosa*. *Applied Microbiology and Biotechnology* 64(1), pp70-75.

Delorme, E., Lorenzini, T., Giffin, J., Martin, F., Jacobsen, F., Boone, T. and Elliott, S. 1992. Role of glycosylation on the secretion and biological properties of erythropoietin. *Biochemistry* 31(41), pp9871-9876.

Drouillard, S., Armand, S., Davies, G.J., Vorgias, C.E. and Henrissat , B. 1997. *Serratia marcescens* chitobiase is a retaining glycosidase utilizing substrate acetamido group participation. *Biochemistry Journal* 328(3), pp945-949.

Dube, D.H. and Bertozzi, C.R. 2005. Glycans in cancer and inflammation – potential for therapeutics and diagnostics. *Nature* 4(6), pp477-488.

Ducros, V.M., Tarling, C.A., Zechel, D.L., Brzozowski, A.M., Frandsen, T.P., von Ossowski, I. Schulein, M., Witehrs, S.G. and Davies, G.J. 2003. Anatomy of glycosynthesis: Structure and kinetics of the *Humicola insolens* Cel7B E197A and E197S glycosynthase mutants. *Chemistry and Biology* 10(7), pp619-628.

Duk, M., Lisowska, E., Wu, J.H., and Wu, A.M. 1994. The Biotin/Avidin-Mediated Microtiter Plate Lectin Assay with the Use of Chemically Modified Glycoprotein Ligand. *Analytical Biochemistry* 221(2), pp.266-272.

Dutta, P.K., Dutta, J. and Tripathi, V. 2004. Chitin and chitosan: Chemistry, properties and applications. *Journal of Scientific and Industrial Research* 63(1) pp20-31.

Dwek, R. 1996. Glycobiology: Toward understanding the function of sugars. *Chemical Reviews* 96(2), pp683-720.

Erbayraktar, S. *et al.* 2003. Asialoerythropoieten is a nonerythropoietic cytokine with broad neuroprotective activity in vivo. *Procedures of the National Academy of Science* 100(11), pp6741-6746.

Fersht, A. 1999. *Structure and Mechanism in Protein Science*. New York: W.H Freeman and Company.

Forst, S., Dowdes, B., Boemare, N. and Stackebrandt, E. 1997. Xenorhabdus and Photohabdus Spp.: Bugs that kill bugs. *Annual Reviews Microbiology* 51(1), pp47-52.

Forst, S. and Nealson, K. 1996. Molecular Biology of the Symbiotic-Pathogenic Bacteria *Xenorhabdus* spp., and *Photorhabdus* spp. *Microbiological Reviews* 60(1), pp21-43.

Fuchs, R., McPherson, S. and Drahos, D. 1986. Cloning of a *Serratia marcescens* gene encoding chitinase. *Applied and Environmental Biology* 51(3), pp504-509.

Funkhouser, J. and Aronson Jr., N. 2007. Chitinase Family GH18: evolutionary insights from genomic history of a diverse protein family. *BMC Evolutionary Biology* 7(96) pp1-16.

George, R.A. and Heringa, J. 2003. An analysis of protein domain linkers: their classification and role in protein folding. *Protein Engineering* 15(11), pp871-879.

Geyer, H. and Geyer, R. 2006. Strategies for analysis of glycoprotein glycosylation. *Biochimica et Biophysica Acta* 1764(12), pp1853-1869.

Ghazarian, H., Idoni, B. and Oppenheimer, S. 2011. A glycobiology review: Carbohydrates, lectins and implications in cancer therapeutics. *Acta Histochemica* 113(3), pp236-247.

Ghormade, Z., Kulkarni, S., Doiphode, N., Rajamohanan, P. and Deshpande, M. 2010. Chitin deacetylase: a comprehensive account on its role in nature and its biotechnological applications. *IN:* Mendez-Vilas, A. *Current Research, Technology and Education Topics in Applied Microbiology and Microbial Biotechnology* Spain: Formatex pp1054-1060.

Gibbs, C. and Zoller, M. 1991. Rational scanning mutagenesis of a protein kinase identifies functional regions involved in catalysis and substrate interactions. *The Journal of Biological Chemistry*. 266(15), pp8923-8931.

Goldstein, I. 1983. Lectins from *Griffonia simplicifolia* seeds. *Journal of Biosciences* 5(1), pp65-71.

Gortari, M.C. and Hours, R.A. 2013. Biotechnological processes for chitin recovery out of crustacean waste: A mini-review. *Electronic Journal of Biotechnology* 16(3), pp1-18.

Grimont, F. and Grimont, P. 2006. The genus *Serratia. Prokaryotes* 6(3), pp219-244.

Hanahan, D. 1985. Techniques for transformation of *E. coli*. *IN: DNA Cloning, Vol.1*. *Ireland:* IRL Press Oxford.

Hedstrom, L. 2001. Enzyme specificity and selectivity. *IN: Encyclopedia of Life Science,*. United Kingdom: Macmillan Publishers Limited.

Hejazi, A. and Falkiner, F.R. 1997. *Serratia marcescens. Journal of Medical Microbiology* 46(11), pp903-912.

Henderson, P.J.F. and MacPherson, J.S. 1986. Assay, genetics, proteins, and reconstitution of proton-linked galactose, arabinose, and xylose transport systems of *Escherichia coli*. *Methods in Enzymology* 125(5), pp387-429.

Henrissat, B. 1991. A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochemical Journal* 280(2), pp309-316.

Henrissat, B. and Davies, G.J. 2000. Glycoside hydrolases and glycosyltransferases. Families, modules and implications for genomics. *Plant Physiology* 124(4), pp1515-1519.

Hitchen, P.G. and Dell, A. 2006. Bacterial glycoproteomics. *Microbiology-Sgm* 152(6), pp1575-1580.

Hossler, P., Khattak, S. and Li, Z. 2009. Optimal and consistent protein glycosylation in mammalian cell culture. *Glycobiology* 19(9), pp936-949.

Imai, T., Watanabe, T., Yu, T. and Sugiyama, J. 2003. The directionality of chitin biosynthesis: a revisit. *Biochemical Journal* 374(3), pp755-760.

Imperiali, B. and O'Connor, S.E. 1999. Effect of N-linked glycosylation on glycopeptide and glycoprotein structure. *Current Opinion in Chemical Biology* 3(6), pp643-649.

Intra, J., Pavesi, G. and Horner, D.S. 2008. Phylogenetic analyses suggest multiple changes of substrate specificity within the Glycosyl hydrolase 20 family. *BMC Evolutionary Biology* 8(214) pp1-9

Jaeken, J. and Matthijis, G. 2007. Congenital disorders of glycosylation: a rapidly expanding disease family. *Annual Review of Genomics and Human Genetics* 8(1) pp261-278.

Jang, M., Kong, B., Jeong, Y. Lee, C. and Nah, J. 2004. Physiochemical characterization of α–chitin, β-chitin and γ–chitin separated from natural resources. *Journal of Polymer Science* 42(14), pp3423-3432.

Jeon, Y. Shahidi, F. and Kim, S. 2000. Preparation of chitin and chitosan oligomers and their applications in physiological functional foods. *Food Reviews International* 16(2), pp159-176.

Jokilammi, A., Korja, M., Jakobsson, E. and Finne, J. 2007. Generation of lectins from enzymes: use of inactive endosialidase for polysialic acid detection. *IN:* Nilsson, C.L. *Lectins: Analytical Technologies.* Amsterdam: Elsevier, pp385-396.

Jones, A.J.S., Papac, D.I., Chin, E.H., Keck, R., Baughman, S.A., Lin, Y.S., Kneer, J. and Battersby, J.E. 2007. Selective clearance of glycoforms of a complex glycoprotein pharmaceutical caused by terminal N-acetylglucosamine is similar in humans and cynomolgus monkeys. *Glycobiology* 17(5), pp529-540.

Joshi, M., Sidhu, G., Pot, I., Brayer, G., Withers, S. and McIntosh, L. 2000. Hydrogen bonding and catalysis: A novel explanation for how a single amino acid substitution can change the pH optimum of a glycosidase. *Journal of Molecular Biology* 299(1), pp.255-279.

Juge, N. 2012. Microbial adhesins to gastrointestinal mucus. *Trends in Microbiology* 20(1), pp30-39.

Kamerling, J., Rijkse, I., Maas, A., van Kuik, J.and Vliegenthart, J. 1988. Sulfated N-linked carbohydrate chains in porcine thyroglobulin. *FEBS* 241(1), pp246-250.

Kanninen, K., Goldsteins, G., Auriola, S. 2004. Glycosylation changes in Alzheimer's disease as revealed by a proteomic approach. *Neuroscience letters* 367 (2), pp235-240.

Katchalski-Katzir, E. 1993. Immobilized enzymes: Learning from past successes and failures. *Trends in Biotechnology* 11(11), pp471-478.

Khoushab, F. and Yamabhai, M. 2010. Chitin research revisited. *Marine Drugs* 8(7), pp1988-2012.

Kim, S. and Rajapakse, N. 2005. Enzymatic production and biological activities of chitosan oligosaccharides (COS); A review. *Carbohydrate Polymers* 62(4) pp357-368.

Krokeide, I., Synstad, B., Gaseidnes, S., Horn, S., Eijsink, V. and Sørlie, M. 2007. Natural substrate assay for chitinases using high-performance liquid chromatography: A comparison with existing arrays. *Analytical Biochemistry* 363(1), pp128-134.

Kumar, M. 2000. A review of chitin and chitosan applications. *Reactive and Functional Polymers* 46(1) pp1-27.

Laine, R., Lo, J. Zhu, B. 2007. Catalytically inactive endoglycosidases as microbial diagnostic reagents: chitinases and lysozymes as fungal and bacterial capture/label agents. *IN:* Nilsson, C.L. *Lectins: Analytical Technologies.* Amsterdam: Elsevier, pp373-384.

Laine, R. 1994. A calculation of all possible oligosaccharide isomers both branched and linear yields 1.05 x 10^{12} structures for a reducing hexasaccharide: the Isomer Barrier to development of single-method saccharide sequencing or synthesis systems. *Glycobiology* 4(6), pp759-767.

Laine, R. 1997. Information capacity of the carbohydrate code. *Pure and Applied Chemistry* 69(9), pp1867-1873.

Langer, E. 2012. Biopharmaceutical Industry in 2012: optimism on the rise. *Pharmaceutical Technology*.

Lee, S., Senevirathne, M., Ahn, C. Kim, S. and Je, J. 2009. Factors affecting anti-inflammatory effect of chitooligosaccharides in lipopolysaccharides-induced RAW264.7 macrophage cells. *Bioorganic and Medicinal Chemistry Letters* 19(23) pp6655-6658.

Lenardon, M.D., Munro, C.A. and Gow, N. 2010. Chitin synthesis and fungal pathogenesis. *Current Opinion in Microbiology* 13(13), pp416-423.

Linehan, S., Martinez-Pomares, L. and Gordon, S. 2000. Macrophage lectin in host defence. *Microbes and Infection* 2(3), pp279-288.

Lis, H. and Sharon, L. 1998. Lectins: Carbohydrate-Specific Proteins that Mediate Cellular Recognition. *Chemical Reviews* 98(2), pp637-674.

Liu, T., Zhang, H., Liu, F., Wu, Q., Shen, X. and Yang, Q. 2011. Structural determinants of an insect β-N-acetyl-D-hexosaminidase specialized as a chitonlytic enzyme. *The Journal of Biological Chemistry* 286(6), pp4049-4058.

Maeda, Y. and Kimura, Y. 2004. Anti-tumor effects of various low-molecular-weight chitosans are due to increased natural killer activity of intestinal intraepithelial lymphocytes in sarcoma 180-bearing mice. *Nutrition and Cancer* 134(4), pp945-950.

Makrides, S.C. 1996. Strategies for achieving high-level expression of genes in *Escherichia coli. Microbiological Reviews.* 60(3), pp512-538.

Marek, K., Vijay, I. and Marth, J. 1999. A recessive deletion in the GlcNAc-1-phosphotransferase gene results in peri-implantation embryonic lethality. *Glycobiology* 9(11), pp1263-1271

Marth, J. and Grewal, P. 2008. Mammalian glycosylation in immunity. *Nature Reviews Immunology* 8(11), pp874-887.

Mathews, C.K. and van Holde, K.E. 1990. *Biochemistry*. United States: The Benjamin/Cummings Publishing Company.

Matthews, B. and Remington, S. 1974. The three-dimensional structure of lysozyme from bacteriophage. *Procedures of the National Academy of Science* 71(10), pp4178-4181.

Mazzei, R., Giorno, L., Piacentini, E., Mazzuca, S. and Drioli, E. 2009. Kinetic study of a biocatalytic membrane reactor containing immobilized β-glucosidase for the hydrolysis of oleuropein. *Journal of Membrane Science* 339(1-2), pp.215-233.

Merzendorfer, H. 2006. Insect chitin synthases: a review. *Journal of Comparative Physiology B* 176(1), pp1-15.

Merzendorfer, H. and Zimoch, L. 2003. Chitin metabolism in insects: structure, function and regulation of chitin synthases and chitinases. *The Journal of Experimental Biology* 206(24), pp4393-4412.

Mescher, M. and Strominger, J. 1976. Purification and characterization of a prokaryotic glucoprotein from the cell envelope of *Halobacterium salinarium*. The *Journal of Biological Chemistry* 251(7), pp2005-2014.

Mildvan, A., Weber, D. and Kuliopulos, A. 1992. Quantitative interpretations of double mutations in enzymes. *Archives of Biochemistry and Biophysics* 294(2), pp327-340.

Minke, R. and Blackwell J. 1978. The structure of α–chitin. *Journal of Molecular Biology* 120(2), pp167-181.

Miot, M. and Betton, J.M. 2004. Protein quality control in the bacterial periplasm. *Microbial Cell Factories* 3(1), pp4.

Molloy, C., Cannon, R., Sullivan, P. and Shepherd, M. 1994. Purification and characterization of two forms of N-acetylglucosaminidase from *Candida albicans* showing widely different outer chain glycosylation. Microbiology 140(7), pp1543-1553.

Monreal, J. and Reese, E. 1969. The chitinase of *Serratia marcescens*. *Canadian Journal of Microbiology* 16(7), pp689-696.

Moran, A. Gupta, A. and Joshi L. 2011. Sweet-talk: role of host glycosylation in bacterial pathogenesis of the gastrointestinal tract. *Gut* 60(10), pp1412-1425.

Morely, K. and Kazlauskas, R. 2005. Improving enzyme properties: when are closer mutations better?. *Trends in Biotechnology* 23(5), pp231-237.

Morrison, K. and Weiss, G. 2001. Combinatorial alanine-scanning. *Current Opinion in Chemical Biology* 5(3), pp302-307.
Moskowitz, R. 2001. Osteoarthritis: simple analgesics versus nonsteroidal antiinflamatory drugs. *Journal of Rheumatology* 28(5), pp932-934.

Muntoni, F., Brockington, M., Blake, J. 2002. Defective glycosylation in muscular dystrophy. *Lancet* 360(9343), pp1419-1421.

Nagatomo, H., Matsushita, Y., Sugamoto, K. and Matsui T. 2005. Preparation and properties of a gelatine immobilized β-Glucosidase from *Pyrococcus furiosus*. *Bioscience, Biotechnology and Biochemistry* 69(1), pp. 128-136.

Ngo, D., Lee, S., Kim, M. and Kim, S. 2009. Production of chitin oligosaccharides with different molecular weights and their anti-oxidant effect in RAW 264.7 cells. *Journal of Functional Foods* 1(2), pp188-198.

Nishiyama, Y., Noishiki, Y. and Wada, M. 2011. X-ray structure of anhydrous β–chitin at 1 Å resolution. *Macromolecules* 44(4), pp950-957.

Noguchi, A., Mukuria, C., Suzuki, C. and Naiki, M. 1995.Immunogenicity of N-glycolylneuraminic acid-containing carbohydrate chains of recombinant human erythropoietin expressed in Chinese hamster ovary cells. *Journal of Biochemistry* 117(1), pp59-62.

Nothaft, H. and Szymanski, C.M. 2010. Protein glycosylation in bacteria: sweeter than ever. *Nature Reviews Microbiology* 8(11), pp765-778.

Nowell, P.C. 1960. Phytohemagglutinin: an initiator of mitosis in culture of animal and human leukocytes. *Cancer Research* 20(4), pp462-466.

Ohtsubo, K. and Marth, J. 2006. Glycosylation in cellular mechanisms of health and disease. *Cell* 126(5). pp855-867.

Otten, L.G., Sio, C.F., van der Sloot, A.M., Cool, R.H. and Quax, W.J. 2004. Mutational analysis of a key residue in the substrate specificity of a cephalosporin acylase. *ChemBioChem* 5(6), pp820-825.

Palcic, M.M. 2011. Glycosyltransferases as biocatalysts. *Current Opinion in Chemical Biology* 15(2), pp226-233.

Park, Y., Kim, M., Park, C., Cheong, H., Jang, M., Nah, J. and Hahm, K. 2008. Investigation of the anti-fungal activity and mechanism of action of LMWS-Chitosan. *Journal of Microbiology and Biotechnology* 18(10), pp1729-1734.

Patil, R.S., Ghormade, V.V. and Deshpande, M.V. 2000. Chitinolytic enzymes: an exploration. *Enzyme and Microbial Technology* 26(7), pp473-483.

Perrakis, A., Tews, I., Dauter, Z. 1994. Crystal-Structure of a Bacterial Chitinase at 2.3-Angstrom Resolution. *Structure* 2 (12), pp1169-1180.

Perrakis, A., Tews, I., Dauter, Z., Oppenheim, A., Chet, I., Wilson, K. and Vorgias, C. 1994. Crystal structure of a bacterial chitinase at 2.3Å resolution. *Structure* 2(12), pp1169-1180.

Pilobello, K. and Mahal, L. 2007. Deciphering the glycocode: the complexity and analytical challenges of glycomics. *Current Opinion in Chemical Biology* 11(3), pp300-305.

Plapp, B. 1995. Site-directed mutagenesis: a tool for studying enzyme catalysis. *IN:* Purich, D.L. *Methods in Enzymology 249 Enzyme Kinetics and Mechanisms Part D Developments in Enzyme Dynamics*, Calafornia: Academic Press, pp91-119.

Prag, G., Papanikolau, Y., Tavlas, G., Vorgias, C.E., Petratos, K., Oppenheim, A.B. 2000. Structures of chitobiase mutants complexed with the substrate Di-N-acetyl-D-glucosamine: the catalytic role of the conserved acidic pair, aspartate 539 and glutamate 540. *Journal of Molecular Biology* 300(3), pp611-617.

Prasanth, K. and Tharanathan, R. 2007. Chitin/chitosan: modifications and their unlimited application potential: an overview. *Trends in Food Science and Technology* 18(3), pp117-131.

Ramirez-Coutino, L., Marin-Cervantes, M. Huerta, S., Revah, S. and Shirai, K. 2006. Enzymatic hydrolysis of chitin in the production of oligosaccharides using *Lecanicillium fungicola* chitinases. *Process Biochemistry* 41(5), pp1106-1110.

Reece, R. J. 2004. *Analysis of genes and genomes*. England: Wiley.

Reginster, J., Deroisy, R., Rovati, L., Lee, R., Lejeune, E., Bruyere, O., Giacovelli, G., Henrotin, Y., Dacre, J. and Gossett, C. 2001. Long-term effects of glucosamine sulphate on osteoarthritis progression: a randomised, placebo-controlled clinical trial. The Lancet 357(9252), pp251-256.

Rhodes, J. 1996. Unifying hypothesis for inflammatory bowel disease and associated colon cancer: sticking the pieces together with sugar. *The Lancet* 347(8993), pp40-44.

Rhoades, J., Gibson, G., Formentin, K., Beer, M. and Rastall, R. 2006. Inhibition of the adhesion of enteropathogenic Escherichia coli strains to HT-29 cells in culture by chito-oligosaccharides. *Carbohydrate Polymers* 64(1), pp57-59.

Rinaudo, M. 2006. Chitin and chitosan: Properties and applications. *Progress in Polymer Science* 31(7), pp603-632.

Rovati, L., Casula, P. and Mascherpa, S. *N-acetylglucosamine for treating degenerative afflications of the joints,* 3697632 1972.

Rustighyi, I., Campa, C., Rossi, M., Semeraro, S., Vetere, A. and Gamini, A. 2009. Analysis of N-acetylaminosugars by CE: A comparative derivatization study. *Electrophoresis* 30(15), pp2632-2639

Rye, C.S. and Withers, S.G. 2000. Glycosidase mechanisms. *Current Opinion in Chemical Biology* 4(5), pp573-580.

Sahai, A. and Manocha, M. 1993. Chitinases of fungi and plants: their involvement in morphogenesis and host-parasite interaction. *FEMS Microbiology Reviews* 11(4), pp317-338.

Salvatore, S., Heuschkel, R., Tomlin, S., Davies, S., Edwards, S., Walker-Smith, J., French, I. and Murch, S. 2000. A pilot study of N-acetylglucosamine, a nutritional substrate for glycosaminoglycan synthesis, in paediatric chronic inflammatory bowel disease. *Alimentary Pharmacology and Therapeutics* 14(12), pp1567-1579.

Schwartz, F. and Aebi, M. 2011. Mechanisms and principles of N-linked protein glycosylation. *Current Opinion in Structural Biology* 21(5), pp576-582.

Seo, J. and Lee, K. 2004. Post-translational modifications and their biological functions: proteomic analysis and systemic approaches. *Journal of Biochemistry and Molecular Biology* 37(1), pp35-44.

Sethuraman, N. and Stadheim, T. 2006. Challenges in therapeutic glycoprotein production. *Current Opinion in Biotechnology* 17(4), pp341-346.

Sharon, L. and Lis, H. 2004. History of lectins: from hemagglutinins to biological recognition molecules. *Glycobiology* 14(11), pp53-62.

Shen, K., Chen, M., Chan, H., Jeng, J. and Wang, Y. 2009. Inhibitory effects of chitooligosaccharides on tumor growth and metastasis. *Food and Chemical Toxicology* 47(8), pp1864-1871.

Sheridan, C. 2007. Commercial interest grows in glycan analysis. *Nature Biotechnology* 25(2), pp145-146.

Shikhman, A., Amiel, D., D'Lima, D., Hwang, S., Hu, C., Xu, A., Hashimoto, S., Kobayashi, K., Sasho, T. and Lotz, M. 2005. Chondroprotective activity of N-acetylglucosamine in rabbits with experimental osteoarthritis. *Annals of Rheumatic Diseases* 64(1), pp89-94.

Shikhman, A., Kuhn, K., Alaaeddine, N. and Lotz, M. 2001. N-acetylglucosamine prevents IL1β-Mediated activation of human chondrocytes. *The Journal of Immunology* 166(8), pp5155-5160.

Sikorski, P., Hori, R. and Wada, M. 2009. Revisit of α–chitin structure using high resolution x-ray diffraction data. *Biomacromolecules* 10(5), pp110-1105.

Sinclair, A.M. and Elliott, S. 2004. Glycoengineering: the effect of glycosylation properties on therapeutic proteins. *Journal of Pharmaceutical Sciences* 94(8), pp1626-1635.

Sletyr, U. and Thorne, K. 1976. Chemical characterization of the regularly arranged surface layers of *Clostridium thermosaccharolyticum* and *Clostridium thermohydrosulfuricum*. *Journal of Bacteriology* 126(1), pp377-383.

Sorbotten, A., Horn, S., Eijsink, V. and Varum, K. 2005. Degradation of chitosans with chitinase B from *Serratia marcescens*: Production of chito-oligosaccharides and insight into enzyme processivity. *FEBS Journal* 272(2), pp538-549.

Sorlie, M., Zakariassen, H., Norberg, A. and Eijsink, V. 2012. Processivity and substrate-binding in family 18 chitinases. *Biocatalysis and Biotransformation* 30(3), pp353-365.

Spik, G., Baynard, B., Fournet, B., Strecker, G., Bouquelet, S. and Montreuil, J. 1975. Studies on glycoconjugates. LXIV. Complete structure of two carbohydrates of human serotransferrin. *FEBS Letters* 50(3), pp296-299.

Spik, G., Debruyne, V., Montreuil, J., van Halbeek, H. and Vliegenthart, J. 1985. Primary structure of two sialylated triantennary glycans from human serotransferrin. *FEBS* 183(1), pp65-69.

Suginta, W., Chuenark, D., Mizuhara, M. and Fukamizo, T. 2010. Novel β-N-acetylglucosaminidases from Vibrio Harvey 650: Cloning, expression, enzymatic properties and subsite identification. *BMC Biochemistry* 11(40), pp1-10.

Suginat, W., Songsiriritthigul, C., Kobdaj, A., Opassiri, R. and Svasti, J. 2007. Mutations of Trp275 and Trp397 altered binding selectivity of *Vibrio carchariae* chitinase A. *Biochimica et Biophysica Acta* 1770(8), pp1151-1160

Sugiyama, J., Boisset, C., Hashimoto, M. and Watanabe, T. 1999. Molecular directionality of β-chitin biosynthesis. *Journal of Molecular Biology* 286(1), pp247-255.

Sumner, J.B. and Howell, S.F. 1936. The identification of hemagglutinin of the jack bean with concanavalin A. *Journal of Bacteriology* 32(2), pp227-237.

Suzuki, K., Mikami, T., Okawa, Y., Tokoro, A., Suzuki, S. and Suzuki, M. 1986. Antitumor effect of hexa-N-acetylchitohexaose and chitohexaose. *Carbohydrate Research* 151(1), pp403-408.

Synstad, B., Gaseidnes, S., van Aalten, D., Vriend, G., Nielsen, J. and Eijsink, V. 2004. Mutational and computational analysis of the role of conserved residues in the active site of a family 18 chitinase. *European Journal of Biochemistry* 271(2), pp253-262.

Szymanski, C., Yao, R., Ewing, C., Trust, T. and Guerry P. 1999. Evidence for a system of general protein glycosylation in *Campylobacter jejuni*. *Molecular Microbiology* 32(5), pp1022-1030.

Takenaka, Y., Fukumori, T. and Raz, A. 2004. Galectin-3 and metastasis. *Glycoconjugate Journal* 19(7-9), pp543-549.

Tamai, Y., Miyatake, K., Okamoto, Y., Takamori, Y., Sakamoto, K. and Minami, S. 2003. Enhanced healing of cartilaginous injuries by N-acetyl-D-glucosamine and glucuronic acid. *Carbohydrate Polymers* 54(2), pp251-262.

Tews, I., Perrakis, A., Oppenheim, A., Dauter, Z., Wilson, K. and Vorgias, C. 1996. Bacterial chitobiase structure provides insight into catalytic mechanism and the basis of Tay-Sachs disease. *Nature* 3(7), pp638-648.

Tews, I., Vincentelli, R. and Vorgias, C.E. 1996. N-Acetylglucosaminidase (chitobiase) from *Serratia marcescens*: gene sequence and protein production and purification in *Escherichia coli. Gene* 170(1), pp63-67.

Tharanathan , R. and Kittur, F. 2003. Chitin – The undisputed biomolecule of great potential. *Critical Reviews in Food Science and Nutrition* 43(1), pp61-87.

Toratani, T., Shoji, T., Ikehara, T., Suzuki, K. and Watanabe, T. 2008. The importance of chitobiase and N-acetylglucosamine uptake (GlcNAc) uptake in N,N'-diacetylchitobiose [$(GlcNAc)_2$] utilization by *Serratia marcescens* 2170. Microbiology 154(5), pp1326-1332.

Toscano, M., Woycechowsky, K. and Hilvert, D. 2007. Minimailst active-site redesign: Teaching old enzymes new tricks. *Angewandte Chemie International Edition.* 46(18), pp3212-3236.

Tran, D., Pham, G., Nguyen, X., Vu, D., Nuguyen, N, Tran, V., Mai, T., Nguyen, H., Le, Q., Nguyen, T. and Ba, T. 2011. Some biomedical applications of chitosan-based hybrid nanomaterials. *Advances in Natural Sciences: Nanoscience and Nanotechnology* 2(4), pp1-12.

Tsou, C. 1998. The role of active site flexibility in enzyme catalysis. *Biochemistry (Mosc).* 63(3), pp.253-258.

Ubhayasekera, W. 2011. Structure and function of chitinases from glycoside hydrolase family 19. *Polymer International* 60(6), pp890-896.

Vaaje-Kolstad, G., Horn, S., van Aalten, D., Synstad, B. and Eijsink, V. 2005. The non-catalytic chitin-binding protein CBP21 from *Serratia marcescens* is essential for chitin degradation. *The Journal of Biological Chemistry* 280(31), pp28492-28497.

Vaaje-Kolstad, G., Houston, D., Riemen, A., Eijsink, V. and van Aalten, D. 2005. Crystal structure and binding properties of *Serratia marcescens* chitin-binding protein CBP21. *The Journal of Biological Chemistry* 280(12), pp11313-11310.

Van den Steen, P., Rudd, P., Dwek, R. and Opdenakker, D. 1998. Concepts and principles of O-linked glycosylation. *Critical Reviews in Biochemistry and Molecular Biology* 33(3), pp151-208.

Van Kooyk, Y. and Rabinovich, G. 2008. Protein-glycan interactions in the control of innate and adaptive immune responses. *Nature Immunology* 9(6), pp593-601.

Varki, A., Cummings, R., Esko, J., Freeze, H., Stanley, P., Bertozzi, C., Hart, G. and Etzler, M. (2009). *Essentials of Glycobiology*, 2nd edition, New York: Cold Springs Harbor Laboratory Press.

Vethanayagam, J. and Flower, A. 2005. Decreased expression from T7 promoters may be due to the impaired production of active T7 RNA polymerase. *Microbial Cell Factories* 4(1) pp3.

Vuong, T. and Wilson, D. 2010. Glycoside hydrolases: catalytic base/nucleophile diversity. *Biotechnology and Bioengineering* 107(2), pp195-205.

Walls, D. and Loughran, S. 2011. Tagging recombinant proteins to enhance solubility and aid purification. *Protein Chromatography: Methods and Protocols, Methods in Molecular Biology* 681, p151-175.

Walsh, G. 2006. Biopharmaceutical benchmarks 2006. *Nature Biotechnology* 24(7), pp769-776.

Walsh, G. 2010. Post-translational modifications of protein biopharmaceuticals. *Drug Discovery Today* 15(17), pp773-780.

Waterfield, N. Ciche, T. and Calrke, D. 2009. Photorhabdus and a Host of Hosts. *Annual Reviews Microbiology* 63, pp557-574.

Wells, J. 1990. Additivity of mutational effects in proteins. *Biochemistry* 29(37), pp8509-8517.

Werner, R., Kopp, K. and Schuelter, M. 2007. Glycosylation of therapeutic proteins in different production systems. *Acta Paediatrica* 96(455), pp17-22.

Whittle, E. and Shanklin, J. 2001. Engineering Δ^9-16:0-Acryl Carrier Protein (ACP) desaturase specificity based on combinatorial saturation mutagenesis and logical redesign of the castor Δ^9-18:0-ACP desaturase. *The Journal of Biological Chemistry* 276(24), pp21500-21505.

Williamson, V. and Kaya, H. 2003. Sequence of a symbiont. *Nature* 21(11), pp1294-1295.

Wong-madden, S. and Landry, D. 1995. Purification and characterization of novel glycosidases from the bacterial genus *Xanthomonas*. *Glycobiology* 5(1), pp19-28.

Wu, H., Yao, Z., Bai, X., Du, Y. and Lin, B. 2008. Anti-angiogenic activities of chitooligosaccharides. *Carbohydrate Polymers* 73(1), pp105-110.

Xia, W., Liu, P., Zhang, J. and Chen, J. 2011. Biological activities of chitosan and chitooligosaccharides. *Food Hydrocolloids* 25(2), pp170-179.

Xu, Q., Dou, J., Wei, P., Tan, C., Yun, X., Wu, Y., Bai, X., Ma, X. and Du, Y. 2008. Chitooligosaccharides induce apoptosis of human hepatocellular carcinoma cells via up-regulation of Bax. *Carbohydrate Polymers* 71(4), pp509-514.

Yang, E., Kim, J., Kim, J., Kim, S. Lee, N. and Hyun, C. 2010. Anti-inflammatory effect of chitosan oligosaccharides in RAW 264.7 cells. *Central European Journal of Biology* 5(1), pp95-102.

Young, N., Brisson, J., Kelly, J., Wwatson, D., Tesier, L., Lanthier, P., Jarrell, H., Cadotte, N., Michael, F., Aberg, E. and Szymanski, C. 2002. Structure of the N-linked glycan present on multiple glycoproteins in the gram-negative bacterium, *Campylobacter jejuni*. *The Journal of Biological Chemistry* 277(45), pp42530-42539.

Zechel, D. and Withers, S. 2000. Glycosidase mechanisms: anatomy of a finely tuned catalyst. *Accounts of Chemical Research* 33(1), pp11-18.

Zhao, Y., Park, R. and Muzzarelli, R. 2010. Chitin deacetylases: properties and applications. *Marine Drugs* 8(1), pp24-46.