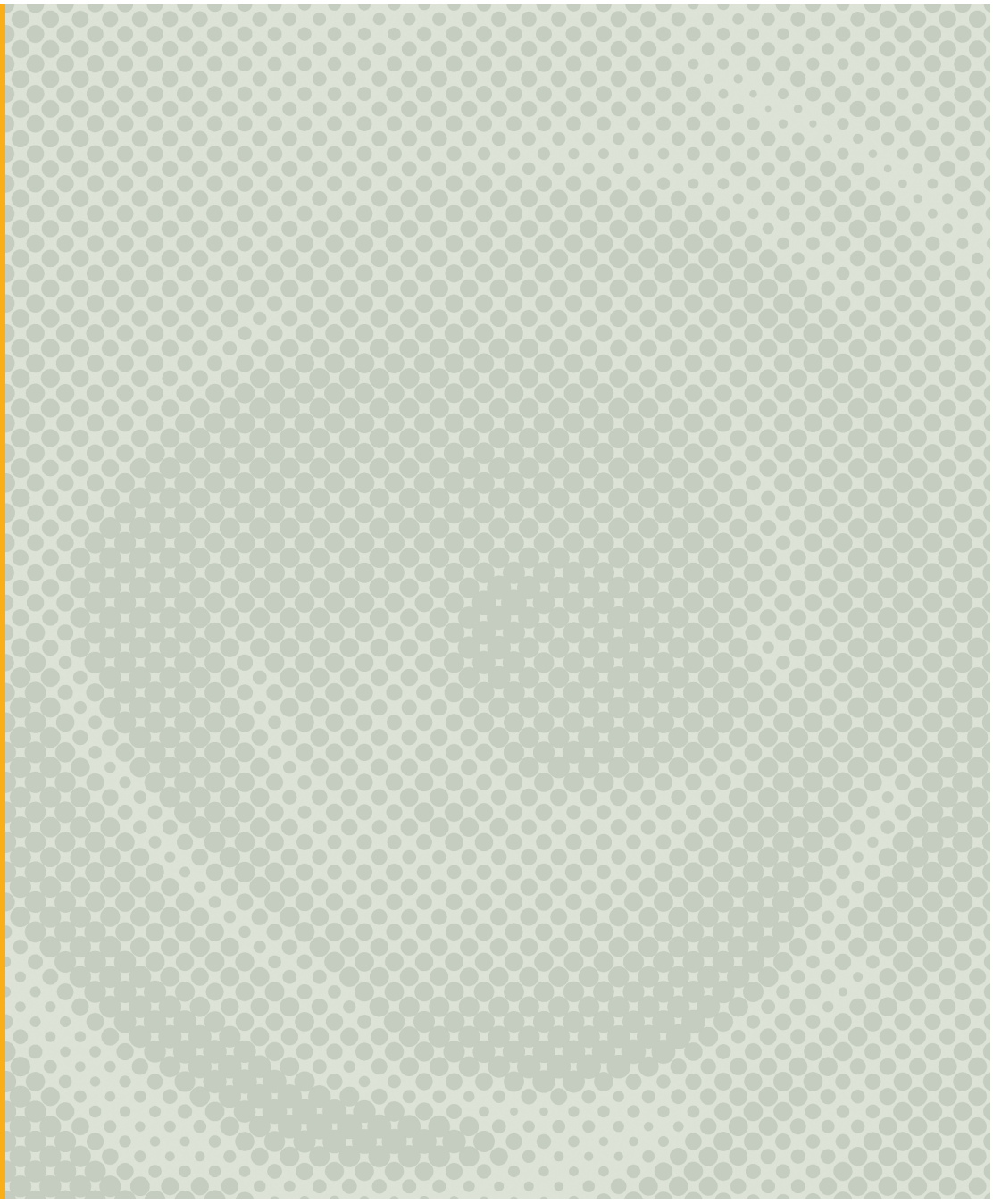
Expression and characterization of CBM33 proteins from *Cellulomonas flavigena* and *Aspergillus terreus*

Trine Isaksen

Norwegian University of Life Sciences
Department of Chemistry, Biotechnology and Food science
Master Thesis 60 credits 2012



ACKNOWLEDGEMENTS

The present work was carried out at the Department of Chemistry, Biotechnology and Food Science at the Norwegian University of Life Sciences with Prof. Vincent Eijsink, Dr. Gustav Vaaje-Kolstad and Ph.D. Zarah Forsberg as supervisors, and partly at the Department of Systems Biology, Technical University of Denmark supervised by Associate Professor Maher Abou Hachem.

First, I would like to thank Vincent Eijsink for the opportunity to write my thesis in his group and Gustav Vaaje-Kolstad who has guided and encouraged me during this exciting year. I am also very grateful to Zarah Forsberg for all your help and always finding time to answer my questions. Thanks to Jane Wittrup_Agger and Alasdair Mackenzie for sharing your expertise with me and to all the members at the PEP-group for your daily positivism and cheerfulness. I am looking forward to four more years!

Special thanks go to Birte Svensson, Maher Abou Hachem and the other group-members at EPC, DTU for including me in your group and taking care of the Norwegian girl. You all made my stay in Denmark special.

I would like to express my deepest gratitude to my family and friends for always pushing me, supporting me and making me smile even on rainy days. Finally, thanks Pappa for proof reading of my thesis, always believing in me and for being my best friend.

Thank you!

Ås, August 15th 2012

Trine Isaksen

ABSTRACT

An essential element of a modern biorefinery is the enzymatic conversion of biomass to soluble sugars. Although the optimization of this process has been pursued by both academia and industry for decades, it still represents a bottleneck in the biorefinery concept. Recent discoveries of a lytic polysaccharide monooxygenase (LPM) activity among members of family 33 of carbohydrate binding modules (CBM33), boosting the degradation of recalcitrant polysaccharides, have given more insight to the degradation of complex polysaccharides in nature that could be adapted for biorefining processes. These findings have also prompted an effort in cloning, expressing and characterizing a wide variety of CBM33s in order to get information on CBM33-family diversity in terms of function and mechanism.

The genome of the Gram negative soil bacterium *Cellulomonas flavigena* encodes four CBM33s (*Cf*CBM33s), all possessing C-terminal CBM2a domains, indicating substrate specificity towards either cellulose or chitin. The bacterium is previously known to metabolize cellulose and xylan. Conversely, results from binding and activity assays performed with two of the *Cf*CBM33s (*Cf*CBM33A-N and *Cf*CBM33B-N) in this study show no specificity towards either of these substrates. However, both show affinity for chitin, which is intriguing as the genome of *C. flavigena* contains no identified chitinases. In addition to the unexpected binding affinity towards chitin, *Cf*CBM33B-N shows lytic chitin monooxygenase activity and is also able to boost the degradation of β-chitin by chitinase A, B and C from *Serratia marcescens*. Moreover, intriguingly, *Cf*CBM33B-N generates partly deacetylated products from the oxidation of β-chitin possibly showing a new function or binding specificity not previously reported for LPMs.

Finally, a CBM33 containing a CBM20 (indicating binding affinity for starch) from the fungus *Aspergillus terreus* was successfully cloned and expressed in *Pichia pastoris.* Preliminary binding experiments using isothermal titration calorimetry indicate that *At*CBM33A binds specifically to starch. Analysis of LPM activity was prevented by the time restrains of this study, but will be the focus of consecutive work. A starch active LPM would be completely new to the field and could be an important finding for the starch processing industry.

SAMMENDRAG

Enzymatisk nedbrytning av biomasse er en viktig prosess innen det moderne bioraffinerikonseptet. På tross av at både den akademiske verden og industrien har jobbet målrettet for å optimalisere denne prosessen i tiår, er den fullstendige nedbrytningen til løselige sukker fremdeles flaskehalsen i bioraffinerien. Nylige oppdagelser viser en lytisk polysakkarid monooksygenase (LPM)-aktivitet blant medlemmer i familie 33 av karbohydratbindende moduler (CBM33) som øker nedbrytningshastigheten av gjenstridige polysakkarider. Disse oppdagelsene har gitt større innsikt i den naturlige nedbrytningen av komplekse polysakkarider som kan benyttes i bioraffineringsprosessen. I tillegg er det nå økt interessen for å klone, utrykke og karakterisere ulike CBM33er for å kartlegge ulike funksjoner og mekanismer innad i familien.

Genomet til den Gram-negative jord bakterien *Cellulomonas flavigena* koder for fire CBM33 (*Cf*CBM33) med C-terminale CBM2a-domener som indikerer substratspesifisitet mot cellulose eller kitin. Bakterien har tidligere vist å metabolisere cellulose og xylan. Resultater fra bindings- og aktivitetsstudier med *Cf*CBM33A-N og *Cf*CBM33B-N viser derimot ingen spesifisitet for disse substratene. Derimot viser de affinitet for kitin, noe som er svært interessant siden genomet til *C. flavigena* ikke inneholder identifiserte kitinaser. I tillegg til denne uventede substratspesifisiteten utøver *Cf*CBM33B-N lytisk kitinmonooxygenase aktivitet og øker nedbrytningshastigheten av β-kitin i synergi med chitinase A, B og C fra *Serratia marcescens.* I oksidasjon av β-kitin genererer *Cf*CBM33B-N delvis deacetylerte produkter, en funksjon som tidligere ikke er beskrevet for LPM.

I tillegg ble en CBM33 fra soppen *Aspergillus terreus* (*At*CBM33A) suksessfullt klonet og uttrykt i *Pichia pastoris.* *At*CBM33A inneholder et CBM20 domene som indikerer bindingsaffinitet for stivelse, som også ble observert ved innledende bindings eksperimenter utført ved hjelp av isotermisk titrerings-kalorimetri. Studiets tidsbegrensing hindret videre analyser av *At*CBM33As LPM-aktivitet, men oppfølgende studier vil fokusere på dette. Stivelseaktive LPMer er ikke tidligere beskrevet i litteraturen og vil være svært betydningsfullt for industrielle prosesser hvor nedbrytning av stivelse er nødvendig.

ABREVIATIONS

aa	Amino acids
*At*CBM33A	A CBM33 protein from Aspergillus terreus
BHI	Brain Heart Infusion
BMGY/ BMMY	Buffered Complex Glycerol/ Methanol Medium
CAZY	Carbohydrate-Active Enzymes
CBM	Carbohydrate Binding Module
CBM33	Carbohydrate Binding Module of family 33
CBP	Chitin Binding Protein
*Cf*CBM33A, -B, -C, -D	CBM33 proteins from *Cellulomonas flavigena*, full-length protein
*Cf*CBM33A-N, -B-N, -C-N, D-N	CBM33 proteins from *Cellulomonas flavigena*, truncated version
ChiA/ ChiB/ ChiC	ng a CBM20 (indicating binding affi
dH_2O	Sterile water (Milli-Q)
dNTP	Deoxynucleoside triphosphate
EC	Enzyme Commission
g	gravity
GH	Glycosyl Hydrolase
GlcN	β-glucosamine
GlcNAc	*N*-acetyl-β-glucosamine
ITC	Isothermal titration calorimetry
IPTG	Isopropyl β-D-1-thiogalactopyranoside
kb	Kilobases
kDa	Kilo Dalton
LB	Luria Bertani
LPM	Lytic Polysaccharide Monooxygenase
MALDI-TOF	Matrix-Assisted Laser Desorption and Ionization Time Of Flight
MS / MS/MS	Mass spectrometry / Tandem mass spectrometry
OD_{600}	Optical density at 600 nanometer
PCR	Polymerase Chain Reaction
rpm	Rotations per minute
SDS-PAGE	Sodium Dodecyl Sulphate PolyacrylAmide Gel Electrophoresis
TB	Terrific Broth
UV	Ultraviolet
v/v	Volume/volume
w/v	Weight/volume
YPDS/ YPD	Yeas Extract Peptone Dextrose with/or without Sorbitol

INNHOLD

1 INTRODUCTION

1.1 Polysaccharides

A carbohydrate is an organic compound with the formula $C_m(H_2O)_n$. Carbohydrates represent one of the most common biomolecules in nature and possess important roles including roles in cell signalling and energy storage to structural roles, e.g. as components of the cell wall. Carbohydrates are divided into four groups: i) monosaccharides, which form the basic carbohydrate unit, ii) disaccharides, which are composed of two monosaccharides, iii) oligosaccharides, typically composed of two to ten monosaccharides, and iv) polysaccharides, that are long carbohydrate polymers with structures ranging from linear to highly branched. Mono-, di-, and oligosaccharides are important in cell signalling and metabolism, while polysaccharides function as structural components (e.g. cellulose, xylan, and chitin) and in energy storage (e.g. glycogen and starch).

In carbohydrate polymers the monosaccharides are covalently joined by glycosidic bonds. A glycosidic linkage is formed between the hemiacetal group in one saccharide and the hydroxyl group of another saccharide. A glycosidic bond can either have α– or β-conformation depending on the axial or equatorial orientation, respectively, of the bond (Figure 1.1). Different conformations of glycosidic linkages combined with the large amount of different monosaccharides enable the enormous molecular diversity seen among polysaccharides (Vaaje-Kolstad 2005).

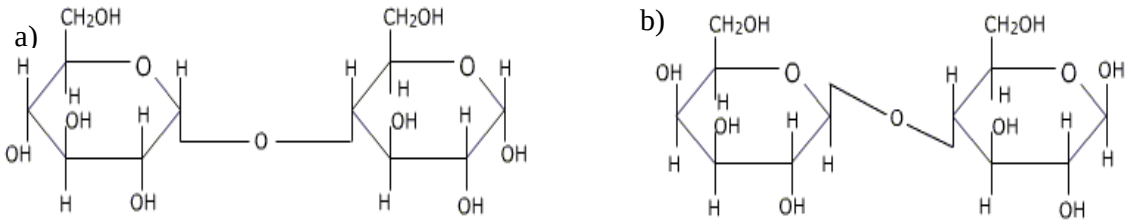


Figure 1.1. Glycosidic linkages. Glycosidic linkages can either be in α-conformation as in maltose (a), or in β-conformation, as in lactose (b).
Figure source: http://chem-guide.blogspot.com/2010/04/carbohydrates.html.

The annual production of polysaccharides in nature is dominated by cellulose, hemicellulose (with xylan as main component) and chitin. Their main function is to protect the organism or cell from mechanical and chemical stress, by forming rigid insoluble and sometimes crystalline structures that are insoluble and difficult to degrade chemically.

1.1.1 Cellulose

The cell wall of plant is a complex structure made up of several layers of microfibrils of cellulose in a crosshatched pattern, impregnated with other polysaccharides (hemicellulose, pectin), lignin, and some proteins; alltogether known as lignocellulose (Lynd et al. 2002). The main component, cellulose, is a linear polymer of β-1,4 linked D-glucose with the consecutive monomers rotated 180 degrees to each other(Figure 1.2). Thus, the repeating units of the cellulose chain are the disaccharide cellobiose. Parallel chains of cellulose are connected through hydrogen bonds, forming a microfibril network with great mechanical strength (Viëtor et al. 2000). These microfibrils are crystalline and non-soluble.

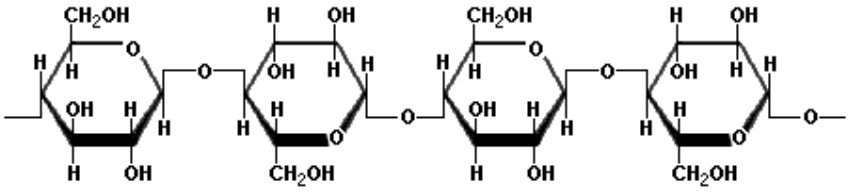


Figure 1.2. The structure of cellulose. Cellulose consists of β-1,4 linked D-glucose. Note that the repeating unit of the cellulose chain is a glucosamine-dimer as every monosaccharide unit is rotated 180° in relation to the next unit. Figure source: http://www.scientificpsychic.com/fitness/carbohydrates1.html.

Crystalline cellulose is found in two forms, Iα and Iβ (Iβ being the major type of cellulose found in plants), while other types (i.e. type II, III and IV) of crystalline cellulose can be gained by pre-treatment (chemically or enzymatically).

1.1.2 Xylan

Besides cellulose, the plant cell wall is composed of other polysaccharides generically termed hemicelluloses, a collectively name for various non-cellulose polysaccharides. The predominant component of hemicellulose in grasses, angiosperms and hardwood is the polysaccharide xylan, which is also found in the cell walls of green algae. Xylan has a backbone of β-1,4 linked D-xylopyranosyl residues (1.3) which are pentose sugars (whereas glucose is a hexose). This backbone often has side chain substitutions and may be decorated with α-L-D-glucuronopyranosyl, 4-O-methyl-α-D-glucuronopyranosyl, α-L-arabinofuranosyl, O-acetyl, feruloyl or coumaroyl residues.

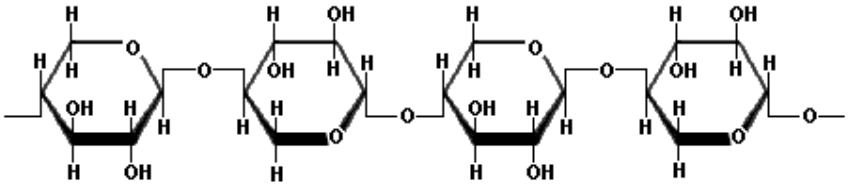


Figure 1.3. The structure of xylan. The backbone of xylan is made up of D-xylopyranosyl residues connected by β-1,4 linkages. Figure source: http://www.scientificpsychic.com/fitness/carbohydrates1.html.

1.1.3 Starch

Starch is present in all green plants as energy storage of glucose derived from carbon dioxide through photosynthesis. Starch is stored in semi-crystalline granules the shape and size of which vary between plant organs and species (Smith 2001). The two major polysaccharides in starch are amylose and amylopectin, with a relative weight percentage ranging from 72% to 82% for amylopectin and from 18% to 28% for amylose (Buléon et al. 1998). The most abundant compound in starch granules, amylopectin (Figure 1.4), is a large molecule with a main chain composed of α-1,4 linked D-glucopyranosyl residues which is heavily branched through α-1,6 linkages.

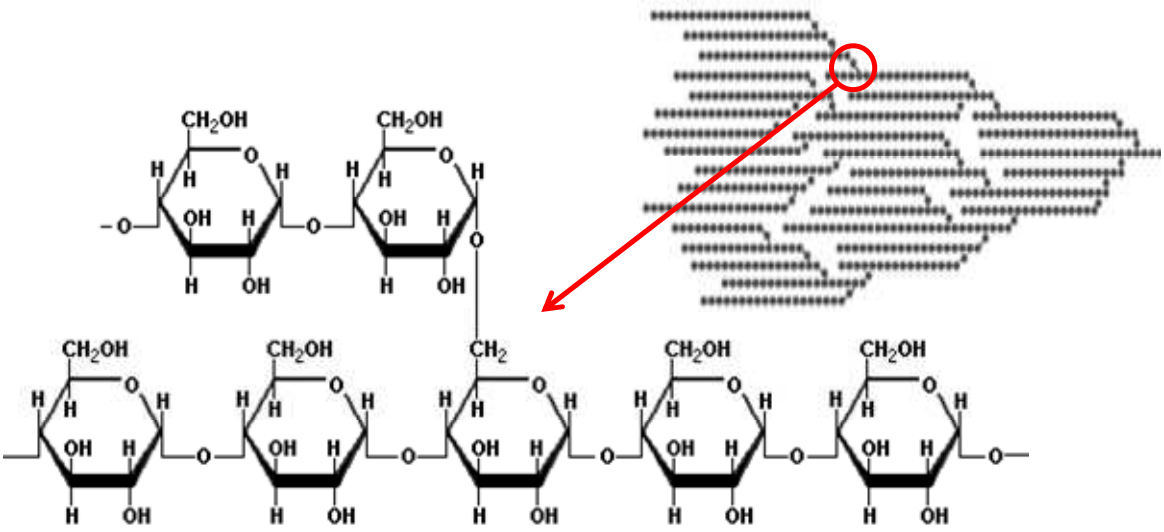


Figure 1.4. The branched nature of amylopectin. Amylopectin has a main-chain of 1,4-linked α-D-glucopyranosyl residues and is heavily branched with α-1,6 linkages.
Figure source: http://www.scientificpsychic.com/fitness/carbohydrates1.html.

Amylose is a relatively long and linear chain of α-D-glucopyranosyl units connected through α-1,4 linkages (Figure 1.5) with few α-1,6 linkages (Buléon et al. 1998).

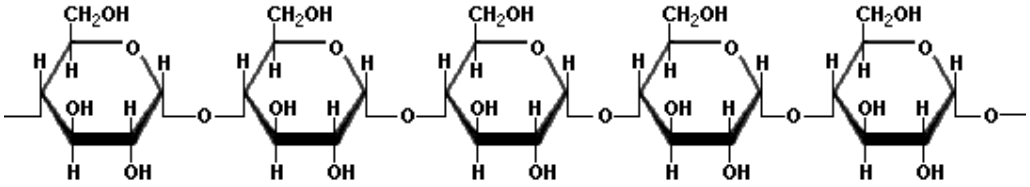


Figure 1.5. The structure of amylose. Amylose consists of α-1,4 linked D-glucopyranosyl units.
Figure source: http://www.scientificpsychic.com/fitness/carbohydrates1.html.

Both the amylose chains and some exterior chains of amylopectin can form double helices which, in turn, may associate to form crystalline domains. Starch is biosynthesized as semi-crystalline granules, and the varying content of amylopectin and amylose in the granules gives varying degrees of crystallinity (Buléon et al. 1998).

1.1.4 Chitin

Chitin is an important structural biopolymer that functions as the main component of the cell walls of fungi, the exoskeletons of arthropods such as insects and crustaceans (e.g. crabs and shrimps) (Rinaudo 2006) as well as in molluscs (Peters 1972) and algae(Jeuniaux 1972). Chitin is a linear insoluble homopolymer of β-1,4 linked *N*-acetyl-β-D-glucosamine (GlcNAc) (Blackwell 1988) as shown in Figure 1.6. The repeating disaccharide units of the unbranched chitin chain contain two GlcNAcs that are rotated 180^{o} relative to each other.

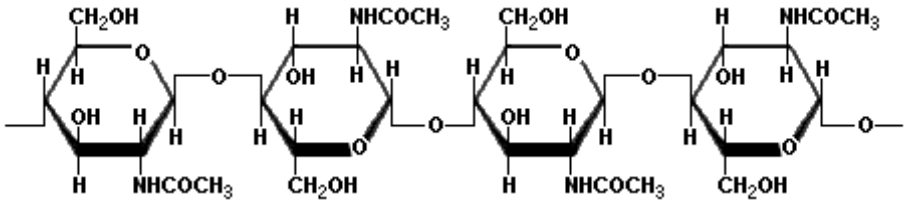


Figure 1.6. The structure of chitin. The repeating units in chitin are disaccharides with the subsequent monomers of β-1,4 linked N-acetylglucosamine (GlcNAc) rotated 180oC to each other.
Figure source: http://www.scientificpsychic.com/fitness/carbohydrates1.html.

Chitin is similar to cellulose, only differing by the acetamido group on C2 in GlcNAc being replaced by a hydroxyl group in cellulose. Like cellulose, chitin chains also form rigid crystal structures by intermolecular bonding. However, because of the presence of the acetamido group, the structure of chitin is somewhat more open than the structure of cellulose (Eijsink et al. 2008) In nature, crystalline chitin is found in three different polymorphic forms dependent on the orientation of the GlcNAc-chains; α-chitin has an anti-parallel arrangement of the individual chains (Minke & Blackwell 1978) while the chains in β-chitin have a parallel arrangement

(Gardner & Blackwell 1975). The least frequent form of chitin, γ-chitin, has a repeating structure of two parallel and one single anti-parallel chain.

The α-form is the most rigid and abundant chitin polymer and the anti-parallel packing of the chains is stabilized by a high number of intra- and intermolecular hydrogen bonds (Carlström 1957). This arrangement contributes to the mechanical strength and stability seen in α-chitin containing organisms such as crustaceans and insects. The crystal structure of β-chitin shows looser packing of the chains and fewer interchain hydrogen bonds, which is reflected in the ability to accommodate water molecules and results in a more flexible chitinous structure (Merzendorfer & Zimoch 2003). β-chitin is frequently found in squid pen, a rigid internal body part found in some squids, and in certain phytoplankton (Blackwell 1988).

Figure 1.6 above shows the fully acetylated form of chitin, however, deacetylation by concentrated NaOH or enzymatic hydrolysis yields a water-soluble copolymer of GlcNAc and D-glucosamine (GlcN) referred to as chitosan (Horn et al. 2006a). Chitosans may differ in their degree of acetylation and solubility.

The annual production of chitin in nature is estimated to be 10^{10}- 10^{11} tons (Gooday 1990). In principle, chitin is available in considerable amounts as underutilized waste product from e.g. the shrimp and crab industry. Currently, only a small part of this biopolymer is utilized for the production of glucosamine and chitosan for use in e.g. water treatment, agriculture, food and paper industry, and personal-care products (Ravi Kumar 2000). There are several other potential applications for chitin or chitosan based materials with both antibacterial and antifungal effects (Hoell et al. 2010). Alternatively, the polymers can be converted to bioactive oligomeric compounds or building blocks for bioactive glycol-conjugates (Aam et al. 2010). However, lack of efficient enzyme technology is a hinder for the optimal degradation and utilization of chitin (Horn et al. 2006a).

1.2 Microbial degradation of polysaccharides

As mentioned, the annual production of polysaccharides in nature is dominated by cellulose, hemicellulose and chitin. Their main use is to protect the organism or cell from mechanical and chemical stress, thus forming rigid crystalline structures that are insoluble and difficult to chemically degrade. In nature, however, there is no long term accumulation of these abundant polysaccharides, indicating efficient break down by microorganisms. Bacterial and fungal degradation of complex polysaccharides is performed by extracellular enzymes which are either released in the surroundings or remain associated with the cell surface. Secreted enzymes act synergistically to hydrolyse crystalline substrates into soluble mono-and oligosaccharides (Merino & Cherry 2007) which can be taken up by the organism and further processed.

1.2.1 Degradation of cellulose

Cellulose is degraded by bacteria and fungi which secrete cellulolytic enzymes (cellulases). Cellulases (EC 3.2.1.4) are found in family 5-10, 12, 18, 19, 26 44, 45, 48, 51, 61, 74 and 124 of glycoside hydrolases (GH; for details on the classification of carbohydrate active enzymes see section 1.3.1). They are divided into exoglucanases/ cellobiohydrolases and endoglucanases. Synergistic cooperation of these enzymes is required for efficient cellulose degradation (Kurašin & Väljamäe 2011). Endoglucanases are non-processive enzymes and cleave at random positions along the crystalline cellulose, thereby generating new chain ends for the processive cellobiohydrolases to attach to and continue degradation, releasing cellobiose units (Kurašin & Väljamäe 2011). β-glucosidases (EC 3.2.1.21, found in GH family 1, 3, 5, 9, 30, 116) finally hydrolyse cellobiose into single glucose units.

In recent years (Forsberg et al., 2011; Quinlan et al., 2011) it has become clear that microbial degradation of cellulose involves at least one additional class of (oxidative) enzymes known as lytic polysaccharide monooxygenases (LPMs); this is discussed in more detail in section 1.3.4.

1.2.2 Starch degradation

The hydrolysis of starch is performed by α-amylases (EC 3.2.1.1), belonging to GH families 13, 14, 57 and 119. Various physical and structural features influence the degradation of starch granules. The main parameter determining degradation efficiency is granule crystallinity, depending on the extent of helix formation in amylose and amylopectin. However, also granular

size, phosphorous contents, and complex formation between amylose and lipids affect the enzymatic degradation (Asare et al. 2011).

1.2.3 Chitin degradation

Two major are known. The first pathway encompasses the deacetylation of chitin to chitosan by chitin deacetylases (EC 3.5.1.41, belonging to family 4 of carbohydrate esterase), followed by hydrolysis of the β-1,4 glycosidic bond in chitosan by chitosanases (EC 3.2.1.132, belonging to GH families 5, 7, 8, 46, 75 and 80) (Hoell et al. 2010). The other pathway, the chitinolytic pathway, involves initial hydrolysis of the β-1,4 glycosidic linkages in chitin by chitinases (EC 3.2.1.14) belonging to GH families 18 and 19 (Henrissat 1991; Henrissat & Davies 1997). This hydrolysis results in production of mainly dimers of GlcNAc that are subsequently degraded to GlcNAc monomers by family 20 glycoside hydrolases called chitobiases (EC 3.2.1.29). Like for cellulose, microbial degradation of chitin also involves LPMs (see section 1.3.4) (Vaaje-Kolstad et al. 2010).

Chitinolytic degradation is achieved in a synergistic manner by mixtures of hydrolytic exo- and endo-acting enzymes (Horn et al. 2006b), similar to what is seen for cellulose degradation. These enzymatic machineries have the potential to convert the biomass to fermentable sugars, and are therefore of great interest for industrial purpose Enzymes involved in degradation of crystalline polysaccharides have to be able to associate with the insoluble substrate, disrupt the polymer packing, and guide a single polymer chain into the catalytic centre (Eijsink et al. 2008).

1.2.4 Biomass degrading microorganisms relevant for this study

1.2.4.1 *Serratia marcescens*

One of the most efficient chitin degraders described is the Gram-negative enterobacterium *S. marcescens.* The bacterium is commonly found in soil, water and plants.

The chitinolytic machinery of *S. marcescens* is induced by chitin (Brurberg et al. 1995; Monreal & Reese 1969) and consists of three family 18 chitinases, two processive enzymes working in opposite directions called ChiA and ChiB, and a non-processive endochitinase called ChiC, an LPM known as CBP21 (CBP for chitin binding protein; see section1.3.4 for more details), and a family 20 chitobiase (Brurberg et al. 1994; Brurberg et al. 1995; Fuchs et al. 1986; Horn et al. 2006b).

1.2.4.2 *Cellulomonas flavigena*

C. flavigena (Bergey et al. 1923) is an aerobic Gram-positive non-motile bacterium with snapping division (a post fission movement after rapid horizontal division of the cell resulting in an angular (V-) arrangement of the daughter cells before complete separation (Krulwich & Pate 1971)). *C. flavigena's* preferred habitats are cellulose enriched environments such as soil, bark, wood, and sugar fields. The bacterium secretes amylases in addition to multiple enzymes for utilization of different lignocellulosic substrates (Sanchez-Herrera et al. 2007). Members of the genus *Cellulomonas* are known to degrade cellulose, xylan and starch (Abt et al. 2010). The genome of *C. flavigena* encodes four family 33 carbohydrate-binding modules (CBM33s) all connected to a CBM2 domain (see section 1.3.5.1) and these proteins' substrate specificity and function as putative lytic polysaccharide monooxygenases (LPMs, described further in section 1.3.4) were the prior focus of this study.

1.2.4.3 *Aspergillus terreus*

The *Aspergillus* species are adapted for the degradation of complex plant polymers by secreting acids and enzymes into the surrounding environment, and are thus found in various terrestrial habitats. *A. terreus* is a filamentous fungus that produces statins, clinically relevant secondary metabolites used in cholesterol-lowering drugs, and toxins associated with aspergillosis of the lungs and/ or disseminated aspergillosis (Bennett 2010). The genome of *A. terreus* encodes one CBM33 connected to a CBM20 domain (see section 1.3.6.3.1) ant the substrate specificity and putative LPM-activity of this protein were the second focus of this study.

1.3 Enzymes and binding modules involved in the degradation of polysaccharides

1.3.1 Classification of carbohydrate active enzymes

All enzymes are provided with an Enzyme Commission (EC) number by the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology representing the reaction catalyzed and the substrate specificity of the enzyme. However, this coding fails to identify enzymes with multiple substrate specificities and does not account for evolutionary relationships provided through sequence and structure data. In 1991, Henrissat et al introduced a

new classification system to improve the classification of glycoside hydrolases This system is based on amino acid sequence similarities (Henrissat 1991; Henrissat & Davies 1997) and formed the basis of the Carbohydrate-Active Enzymes (CAZY) database. Today, an extended version of CAZY describes different families of structurally related enzymes that degrade, modify or create glycosidic bonds, and their carbohydrate-binding modules (Cantarel et al. 2009). As of June 2012 the, the database includes 130 families of glycoside hydrolases, 94 families of glycosyl transferases, 22 families of polysaccharide lysases, 16 families of carbohydrate esterases and 64 families of carbohydrate-binding modules.

1.3.2 Glycoside hydrolases

Glycoside hydrolases (GHs, EC 3.2.1.-) are enzymes catalyzing the hydrolysis of glycosidic bonds in di-, oligo- or polysaccharides. GHs acting on polysaccharides can be either exo-acting, cleaving the substrate from one end or endo-acting, only cleaving interior bonds in the substrate chain (Dilokpimol 2010). GHs have a high structural diversity but the shape of the active site is more conserved and can be divided into three general classes: (i) pocket, (ii) cleft, and (iii) tunnel (Davies & Henrissat 1995).

A deep cleft or a tunnel topology allows the enzyme to remain attached to the substrate during saccharification and degrade the substrate in a processive manner. Processivity is a common feature of glycosidases that degrade crystalline polysaccharides such as cellulose and chitin. As the process of gaining access to a single polymer chain is energetically unfavorable for an enzyme, a processive mechanism is thought to be beneficial for degradation of crystalline substrates. However, the same properties that make an enzyme processive, may contribute to a reduced enzyme efficiency for certain substrates as processive enzymes tends to “get stuck” on their substrates.

1.3.3 A closer view at GH18 chitinases and the chitinolytic machinery of *S. marcescens*

One of the best-studied enzyme systems for the degradation of recalcitrant polysaccharides is the chitinolytic system from *S. marcescens* (Eijsink et al., 2008). The lytic polysaccharide monooxygenases, which are the main topic of the present study (see below), where originally discovered in this bacterium (Vaaje-kolstad et al., 2005a,b, 2010).

As stated in section 1.2.4.1, *S. marcescens* produces three GH18 chitinases when grown on chitin; ChiA, ChiB, and ChiC. All these chitinases employ a substrate-assisted double displacement catalytic mechanism and contain the diagnostic sequence motif D-X-X-D-X-D-X-E (D corresponds to aspartic acid, E to glutamic acid and X can be any amino acid), where the final glutamate in the motif acts as the catalytic acid (Durand et al. 2005). Crystal structures of both ChiA (Papanikolau et al. 2003; Perrakis et al. 1994) and ChiB (van Aalten et al. 2000) reveal deep "tunnel-like" substrate-binding clefts having a path of aromatic residues extending over the surface of the substrate-binding domain (Figure 1.7) (Perrakis et al. 1994; van Aalten et al. 2000). In ChiB, this substrate-binding domain, a carbohydrate-binding module family 5 (CBM5)-domain, extends the cleft on the side where the reducing end of the substrate binds, where in ChiA another type of substrate-binding domain, a fibronectin III-domain (see section 1.3.6.2), extends the cleft towards where the non-reducing end of the substrate binds. The deep substrate-binding clefts in both ChiA and ChiB underpin the notion that these enzymes work in a processive manner.

Glycosidases acting processively on non-soluble polysaccharides share a common feature of having aromatic residues, particularly tryptophans, lining the substrate-binding cleft. Together with hydrogen bonding, aromatic residues are the dominant interaction in protein-carbohydrate complexes (Sørlie et al. 2012). The impact of aromatic residues near the catalytic centres of ChiA and ChiB was demonstrated by Zakariassen et al. (2009) and Horn et al. (2006a), respectively. In the former study, it was shown that by mutating two tryptophan residues near the catalytic centre (Trp^{167} → Ala and Trp^{275} → Ala, see Figure 1.7) ChiA showed a decrease in the processive degradation of crystalline β-chitin. However, the loss of processivity resulted in increased activity towards chitosan, indicating that the rate-limiting step of the reaction depends on the solubility of the polymeric substrate (Zakariassen et al. 2009, (Zakariassen et al. 2009; Zakariassen et al. 2010). Similar observations have been made for ChiB (Horn et al. 2006a).

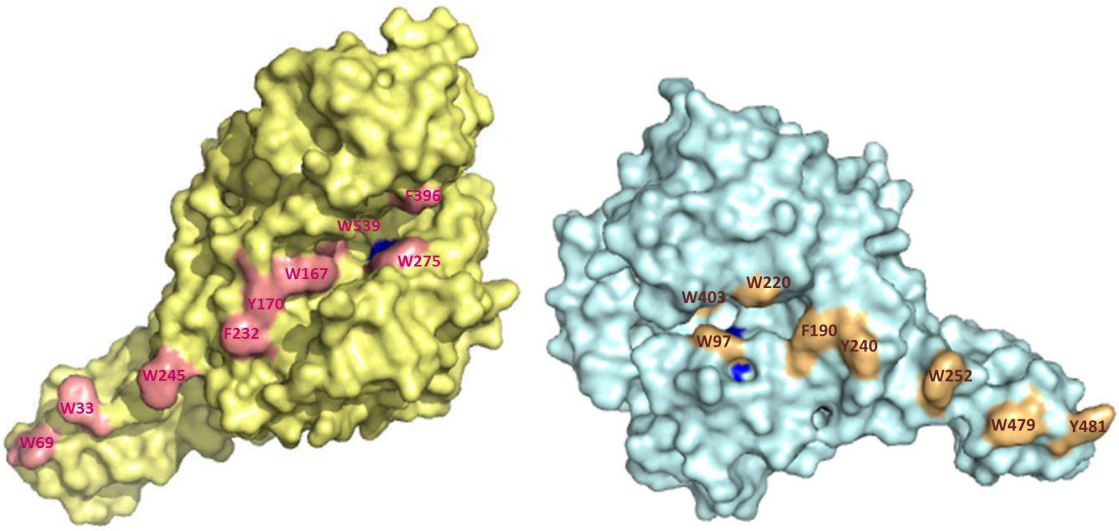


Figure 1.7. Crystal structures of ChiA (left, PDB: 1ctn) and ChiB (right, PDB: 1e15). The catalytic acids, ChiA-Glu[315] and ChiB-Glu[144], are colored blue and the aromatic residues involved in substrate binding are highlighted and numbered. Note that the catalytic domains are oriented in a similar manner and that they are extended by CBMs in opposite directions (to the "non-reducing side" in ChiA, and to the "reducing side" in ChiB. Protein structure graphics were made using PyMol (DeLano 2002).

While both ChiA and ChiB are processive, exo-acting enzymes, the third chitinase expressed in *S. marcescens*, ChiC, is a non-processive endo-chitinase. Judged from sequence alignments with ChiA and ChiB, ChiC has a much more open substrate binding groove (Suzuki et al. 1999). ChiC randomly hydrolyzes the polymer and yields longer chitooligosaccharides as products (Baban et al. 2010; Horn et al. 2006b). It has two extra domains; one fibronectin III-like domain and one CBM12 domain, both with chitin-binding properties. ChiC often occurs in two forms in cultures of *S. marcescens*; the complete protein, referred to as ChiC1, and a proteolytically truncated variant, called ChiC2, lacking the two chitin-binding domains (Horn et al. 2006b; Suzuki et al. 1999).

ChiA and ChiB degrade chitin chains from opposite ends, ChiA from the reducing end and ChiB from the non-reducing end (Baban et al. 2010; Horn et al. 2006b). When including the non-processive endo-acting ChiC, the three *S. marcescens* chitinases show strongly synergistic effects in degradation of different chitinous substrates (Brurberg et al. 1996; Hult et al. 2005; Suzuki et al. 2002). The available data convincingly show that the three chitinases have different and complementary activities and directionalities, which explains the synergism (Horn et al. 2006b). Still, the picture is not completely clear, as ChiB and ChiC show little synergy in the hydrolysis of powdered chitin, while ChiA in combination with either ChiB or ChiC, show a clear synergy on the same substrate (Suzuki et al. 2002). Recent findings on the role of CBP21 from *S. marcescens*, which disrupts and depolymerizes the structure of crystalline chitin by an

oxidative mechanism, and thereby potentiates the chitinase activity (Vaaje-Kolstad et al. 2005a; Vaaje-Kolstad et al. 2010) suggest other explanations for the synergistic effects as discussed in section 1.3.4).

1.3.4 Lytic Polysaccharide Monooxygenases

It has recently been discovered that enzymes now called lytic polysaccharide monooxygenases (LPMs) contribute to the depolymerization of polysaccharides. These enzymes are currently represented by two families in CAZY, carbohydrate-binding module family 33 (CBM33) and GH family 61 (GH61); both these families await re-classification). The first structure of LPMs to be solved was of a CBM33, the chitin-binding protein CBP21 from *S. marcescens* (Figure 1.8) (Vaaje-Kolstad et al. 2005b). In 2005 Vaaje-Kolstad et al (2005a; 2005b) discovered that CBP21 increases chitinase efficiency in degradation of certain forms of crystalline chitin. At the time it was thought that the binding of CBP21 to crystalline chitin somehow led to changes in the substrate structure and increased substrate accessibility. It was shown that CBP21 promoted hydrolysis of β-chitin by ChiA and ChiC, while it was essential for full degradation by ChiB. In 2010, Vaaje-Kolstad et al. revealed that CBP21 is a metal-dependent enzyme which generates chain breaks and oxidizes chain ends on the surface of crystalline chitin, thereby depolymerizing the substrate and making it more available for degradation by chitinases.

Similar activities have been observed for members of the GH family 61 (GH61) (Phillips et al. 2011; Quinlan et al. 2011; Westereng et al. 2011). The GH61 enzymes share structural similarity with CBM33s (Fig. 1.8), but share a sequence identity of less than 10%. The GH61 family consists of fungal proteins up regulated during growth on cellulose and other polysaccharides from plant biomass (Hori et al. 2011). They were originally classified as glycoside hydrolases in the CAZy database on the basis of weak endo-1,4-β-D-glucanase activity detected for one of the family members. However, Harris et al. (2010) concluded that GH61 proteins are unlikely to be classical glycoside hydrolases based on the lack of the characteristics of a glycosidase hydrolase. Like CBM33s, GH61s have a remarkably flat substrate-binding surface and lack catalytic residues that are typical for glycoside hydrolases. Furthermore, unlike GHs, CBM33s and GH61 require divalent metal ions for optimal activity. Harris et al. (2010) also described synergistic effects of GH61 and cellulases in hydrolyzing lignocellulose, indicating that GH61 activity corresponds to the activity of CBP21 (Harris et al. 2010). In their paper on the discovery of the lytic oxidative properties of CBP21, Vaaje-Kolstad et al. (2010) pointed out that GH61s were likely to carry out the same chemistry, as suggested by conserved

active site architectures (Fig 1.8) and as later confirmed experimentally (Phillips et al. 2011; Quinlan et al. 2011; Westereng et al. 2011).

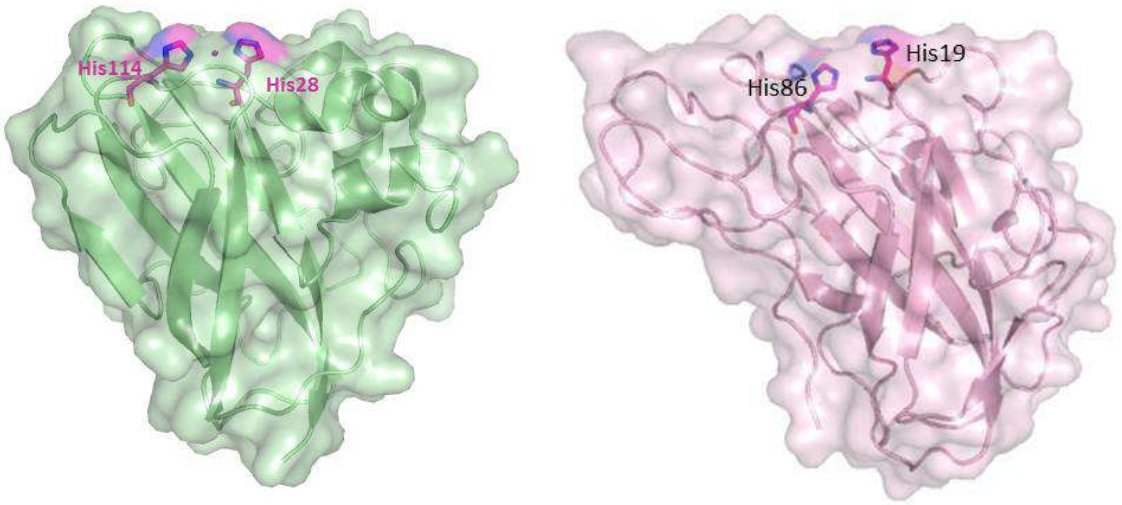


Figure 1.8. Crystal structure of CBP21 (left, PDB ID: 2BEM) and GH61E from *Thielavia terrestris* (right, PDB ID: 3EII). The side chains of conserved histidine residues in the active site involved in binding of divalent copper are coloured in magenta in both models. Note that His28 and His19 in the left and right panels, respectively, are the N-terminal residues of correctly processed secreted proteins. Protein structure graphics were made using PyMol (DeLano 2002).

In 2011, the oxidative properties and direct degradation of cellulose by a GH61 from *Thermoascus aurantiacus* (*Ta*GH61) were described (Quinlan et al. 2011). *Ta*GH61's oxidative activity was proven to be copper-dependent. The copper binds to a type II copper site involving the methylated N-terminal histidine of the enzyme that acts as a metal-coordinating residue (Quinlan et al. 2011). A similar activities were detected for *Pc*GH61D, a GH61 from *Phanerochaete chrysosporium* (Westereng et al. 2011) and for Cdh-1, from *Neuropora crassa* (Phillips et al. 2011). The GH61s cleave cellulose at the glycosidic bond and oxidize one of the new chain ends. This is analogous to what has been described for the chitin-active CBM33s CBP21 (Vaaje-Kolstad et al. 2010) and CBM33A from *Enterococcus faecalis* (Vaaje-Kolstad et al. 2012) and the cellulose-active CBM33 CelS2 (Forsberg et al. 2011) Importantly, while data indicate that CBM33s and *Pc*GH61D exclusively oxidize C1 in the scissile C1-C4 bond, it seems that other GH61 oxidize C4 in the scissile bond and perhaps even C6 at the newly generated non-reducing end.

Available data show that the activity of CBM33s and GH61s is increased by the presence of external electron donors such as ascorbic acid (Quinlan et al. 2011; Vaaje-Kolstad et al. 2010; Westereng et al. 2011), while Langston et al. (2011)and Phillips et al. (2011) showed the same using an electron donating enzyme, cellobiose dehydrogenase (enzymes oxidizing cellobiose to

the corresponding 1-5-δ-lactone). Furthermore, after some initial confusion (Harris et al. 2010; Vaaje-Kolstad et al. 2010), several studies have convincingly shown that CBM33s and GH61s are copper dependent oxidases (Aachmann et al. 2011; Phillips et al. 2011; Quinlan et al. 2011; Westereng et al. 2011). Details of the catalytic mechanism remian to be unravelled but reaction mechanisms and names for these novel enzymes have been proposed. One proposed reaction mechanism proposed by Phillips et al. (2011) is shown in

Figure 1.**9**. Possible names for these enzymes include polysaccharide monooxygenase (PMO) or the name adapted in this report, lytic polysaccharide monooxygenase (LPM). In their 2011 paper presenting a possible mechanism for GH61s, Phillips et al. (2011) only consider C1- or C4- oxidation generating a lactone or a ketoaldose, respectively.

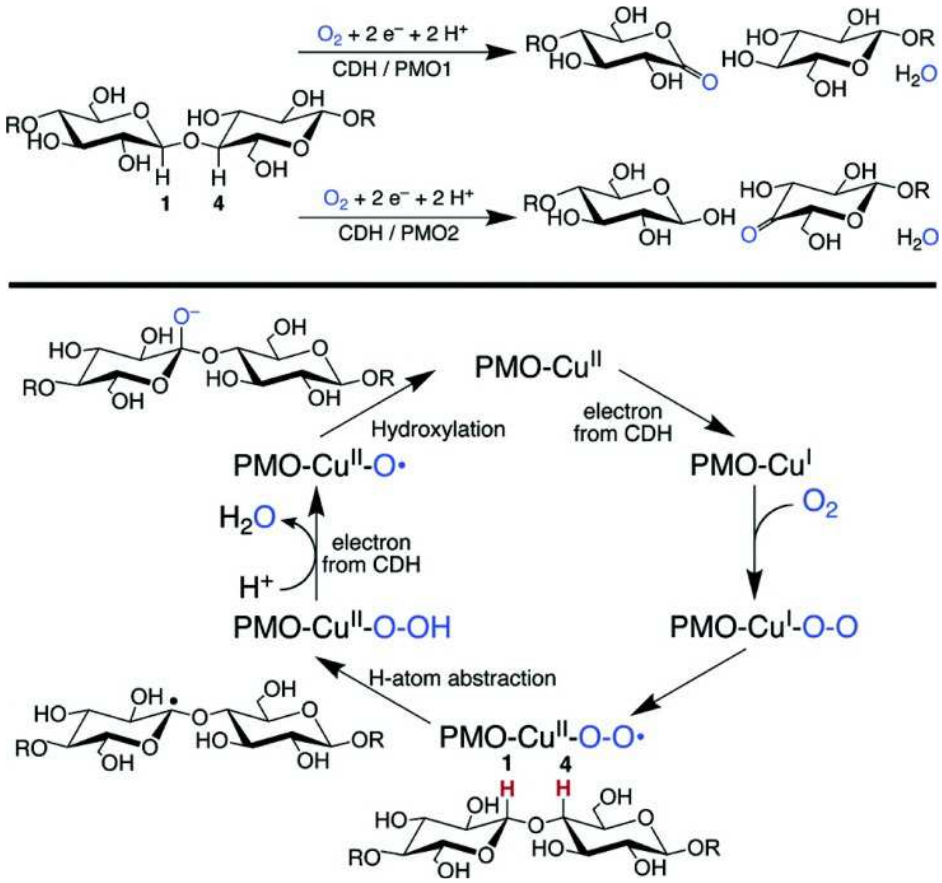


Figure 1.9. **PMO/LPM reactions and proposed mechanism. (Top**) Type 1 PMOs abstract a hydrogen atom from carbon 1 leading to the formation of a sugar lactones, which will be spontaneously hydrolyzed to aldonic acids, depending on pH. Type 2 PMOs catalyze hydrogen atom abstraction from carbon 4 leading to formation of ketoaldoses. (**Bottom**) PMO mechanism: an electron, in this case from a heme domain of the cellobiose dehydrogenase, reduces the PMO Cu(II) to Cu(I) so O_2 can bind. By internal electron transfer, a copper superoxo intermediate is formed, which then abstracts an $H^{\bullet}$ from C1 or C4 on the carbohydrate. A second electron from the cellobiose dehydrogenase leads to a cleavage of the Cu-bound hydroperoxide. The copper oxo-species ($Cu–O^{\bullet}$) then couples with the substrate radical, hydroxylating the substrate. Addition of the oxygen atom destabilizes the glycosidic bond and leads to elimination of the adjacent glucan (Phillips et al. 2011). The Figure is from Phillips et al. (2011).

CBM33s and GH61s are not active on soluble cello- or chito-oligosaccharides (Vaaje-Kolstad et al., 2010; Westereng et al., 2011) and their flat active site surfaces (Fig. 1.8), suggest that these enzymes are optimized for interacting with ordered substrate surfaces as in crystalline substrates. Interestingly, the chromatographic profiles of products generated by oxidation of various cellulose or chitin substrates (PASC, Avicel, cellulose nanofibrils) differ from each other in terms of the dominance of even-numbered products. This might reflect different binding modes towards different substrates with varying degree of crystallinity (Vaaje-Kolstad et al. 2010; Westereng et al. 2011). Westereng et al. (2011) point out that putative functional difference between cellulolytic GH61s and CBM33s may yield synergistic effects when combining several of these proteins working on the same substrate. This is interesting, as the genomes of most biomass-degrading microorganisms contain multiple genes encoding CBM33s or GH61s.

Recently, Medie et al. (2012) introduced these enzymes as lytic oxidases, a name that is broad as oxidases covers all oxidizing enzymes and more informative than "PMO" as it indicates bond cleavage ("lytic"). However, as monooxygen is incorporated into the product of the CBM33/GH61 catalyzed reaction (convincing experimental evidence from both Vaaje-Kolstad et al. (2010) and Beeson et al (2012)), the optimum name that would cover all aspects of the mechanism would be lytic polysaccharide monooxygenase (LPM), which is the name adopted in this report. An overview of the degradation of cellulose by the synergistic action of LPMs, endoglucanases and cellobiohydrolases is shown in Figure 1.10.

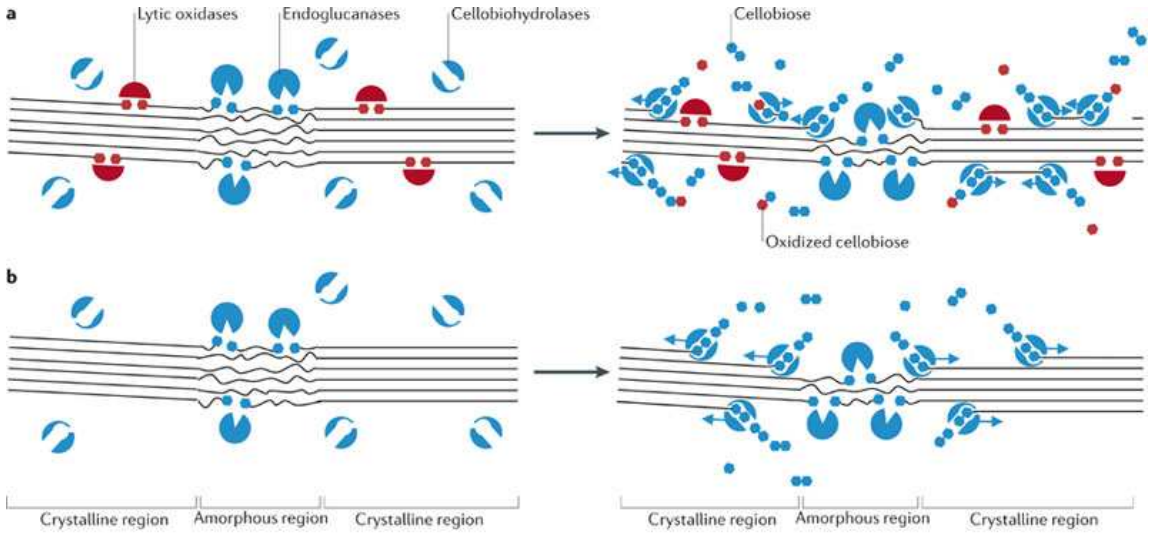


Figure 1.10. An overview of the enzymatic degradation of cellulose in presence and absence of LPMs (referred to as "lytic oxidizes" in the Figure). **a)** Endoglucanases catalyze hydrolytic chain cleavage, resulting in non-oxidized chain ends (blue hexagons). LPMs introduce chain cleavages on the crystalline substrate, resulting in oxidized chain ends (red hexagons). Cellobiohydrolases are processive enzymes, and hydrolyze the cellulose into cellobiose. β-D-glucosidases can then further cleave the cellobiose to glucose monomers. **b)** In the absence of oxidative cleavage, endoglucanases and processive cellobiodydrolases degrade the cellulose, but the rate limiting step is the numbers of chain ends (Medie et al. 2012). The Figure is from Medie et al. (2012).

1.3.5 Carbohydrate-binding modules

Carbohydrate-binding modules (CBMs) are non-catalytic domains with a carbohydrate-binding activity that enhance the activity of many enzymes acting on complex carbohydrates (Cantarel et al. 2009). CBMs are usually small and occur as clearly distinguishable domains separate from the catalytic domains of hydrolytic enzymes (Sorimachi et al. 1997). They are thought to enhance the catalytic efficiency of the enzyme towards insoluble substrates by bringing the catalytic module in intimate contact with the substrate (Hashimoto et al. 2000), thus increasing the local concentration of substrate. There are indications that certain CBMs may have a substrate-disrupting effect (Eijsink et al. 2008). The CBM33s discussed above are no true CBMs but enzymes and will be re-classified in the near future.

1.3.6 Carbohydrate-binding modules

Carbohydrate-binding modules (CBMs) are non-catalytic domains with a carbohydrate-binding activity that enhance the activity of many enzymes acting on complex carbohydrates (Cantarel et al. 2009). CBMs are usually small and physically separate from the catalytic domains of hydrolytic enzymes (Sorimachi et al. 1997). They are thought to enhance the catalytic efficiency

of the enzyme towards insoluble substrates by bringing the catalytic module in intimate contact with the substrate (Hashimoto et al. 2000), thus increasing the local concentration of substrate.

1.3.6.1 CBM2

CBM family 2 (CBM2) contains members that primarily bind cellulose (CBM2a) and xylan (CBM2b), and in some instances chitin(Fujii & Miyashita 1993). The subdivision is based on an 8-residue loop in CBM2a that is absent from CBM2b (Simpson et al. 2000). Family 2a has three surface-exposed tryptophan residues, forming a planar binding surface (Xu et al. 1995) (Figure 1.11) ideal for binding to the planar surface of crystalline cellulose. One of these tryptophans (W17 in Figure 1.11, left) is located within the 8-residue loop region. CBM2b contains only two conserved surface-exposed tryptophan residues, and one of these (W259 in Figure 1.11, right) is rotated by 90^{o} compared to its corresponding position in CBM2a (W17 in Figure 1.11, left) (Simpson et al. 2000). In crystalline xylan, the subsequent pyranose rings are rotated 120° to each other and the orientation of the surface exposed trytophans in CBM2b (W291 and W259 in Figure 1.11, right) favours binding via stacking interactions between the tryptophans and every second pyranose ring of xylan.

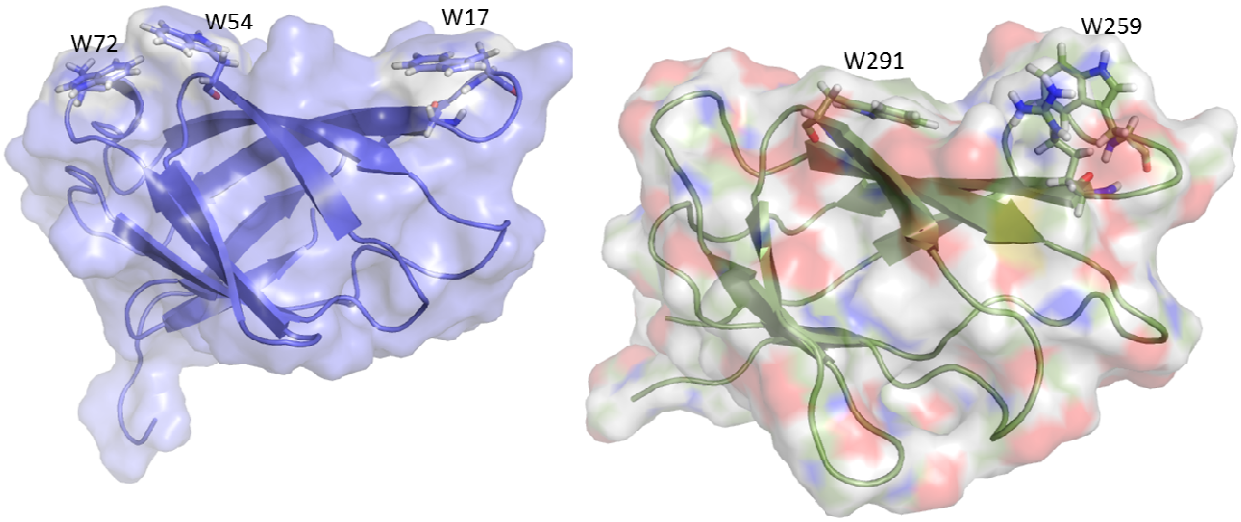


Figure 1.11. Crystal structure of the CBM2a of xy1anase B (*Cfi*xyn10A, PDB ID: 1EXG, left) and the CBM2b domain of xylanase D (*Cfi*xyn11A, PDB ID: 1XBD, right) both found in the genome of *Cellulumonas flavigena*. The surface exposed tryptophans involved in substrate binding are shown in sticks and labelled. The substrate specificity-determining residues, Glycine in Xyn10A and Arginine in Xyn11A (see text), are both shown in sticks, but not labeled. Protein structure graphics were made using PyMol (DeLano 2002).

Regarding the substrates, the only difference between glucose and xylose (see Figure 1.2 and Figure 1.3) is the presence of a CH_2OH group attached to carbon 5 in glucose. Simpson et al. (2000) describes how a mutation of an arginine near the rotated tryptophan in Xyn11A-CBM2b (Uniprot ID: P54865) to a glycine resulted in loss of affinity for xylan and a gain of cellulose-binding properties. They conclude that the arginine holds the tryptophan in an orientation favorable for xylan-binding (Figure 1.12). The structure of the R262G mutant, determined by nuclear magnetic resonance, supports this, as Trp259 had a planar orientation in the mutated protein (Simpson et al. 200).

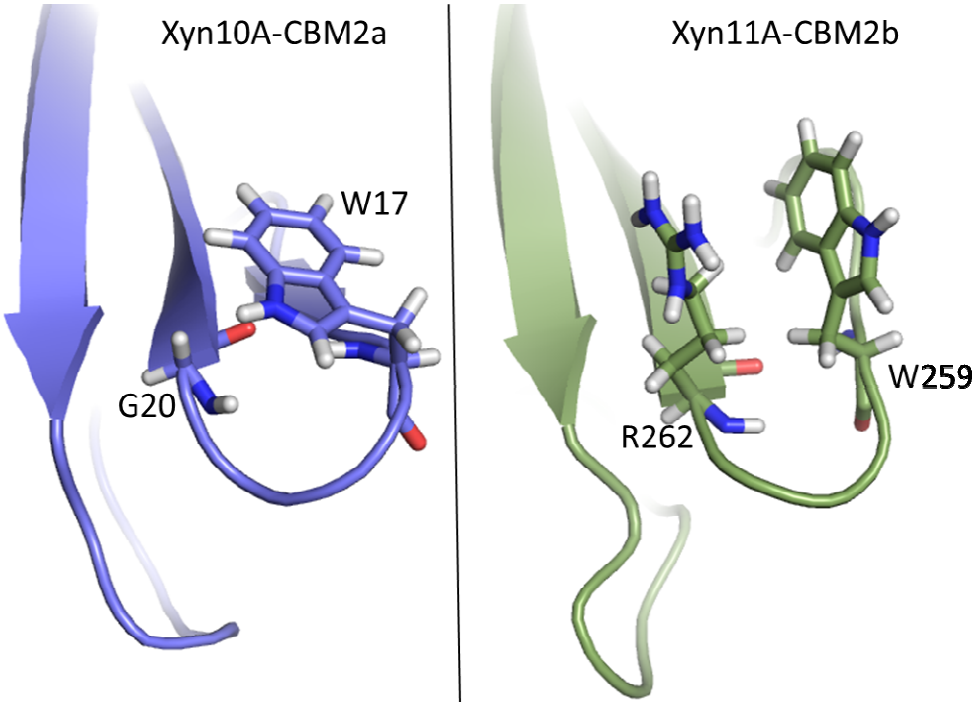


Figure 1.12. Cartoon and stick representation of the major functional difference between CBM families 2a and 2. The representation shows a zoom-in showing the orientation of the key surface tryptophan in *Cfi*xyn10A **(left)** and *Cfi*xyn11A **(right)** and the residue determining the conformation of this tryptophan (glycine in CBM2a and arginine in CBM2b). Protein structure graphics were made using PyMol (DeLano 2002).

1.3.6.2 Fibronectin III-like domains and immunoglobulin-like folds

The immunoglobuline fold consists of a beta-sandwich of seven or more strands divided into two sheets (http://www.ebi.ac.uk/interpro/IEntry?ac=IPR014756). It is one of the most common protein modules found in animals, but modules with structural similarity have also been found in bacterial genomes (Perrakis et al. 1997). Some of these immunoglobulin-folds are structurally similar to fibronectin type III domains.

Both ChiA and ChiC contain an N-terminal domain that has an immunoglobulin-like fold, both structurally similar to fibronectin III domains. However, the N-terminal chitinase domains share no sequence similarity with fibronectin III domains (Perrakis et al. 1997). Perrakis et al. (1997) proposed that such fibronectin III-like domains in chitinases are involved during catalysis by forming interactions with the chitin chain. The role of these domains in chitin-hydrolysis has been confirmed by several experiments (Watanabe et al. 1994).

1.3.6.3 Starch binding domains

Starch binding domains (SBDs) are present in approximately 10% of amylases and are usually located at the C-terminal end of the enzyme. Starch binding domains are classified in CBM families 20, 21, 25, 26, 34, 41, 45, 48 and 53 in the CAZy database (Christiansen et al. 2009b). These domains mainly act on the surface of crystalline starch, allowing the enzyme to degrade whole starch granules (Morris et al. 2005). Morris et al. (2005) also showed that proteolytic removal of the starch binding domain from an *Aspergillus niger* glucoamylase led to a dramatic reduction in activity towards granular starch, whereas this modification did not affect the activity towards soluble starch and oligosaccharides.

1.3.6.3.1 CBM20

Among the SBDs, the CBM family 20 (CBM20) is the best-studied family. CBM20s are found in archaea, bacteria and eukaryotes, mainly linked to amylolytic enzymes (Christiansen et al. 2009a). The structure of the CBM20 of glucoamylase 1 (1,4-α-D-glucan glucohydrolase, EC 3.2.1.3) from *Aspergillus niger* has been determined both when free in solution (Sorimachi et al. 1996) (right in Figure 1.13) and when bound to β-cyclodextrin, an analogue of starch (Sorimachi et al. 1997) (left in Figure 1.13).

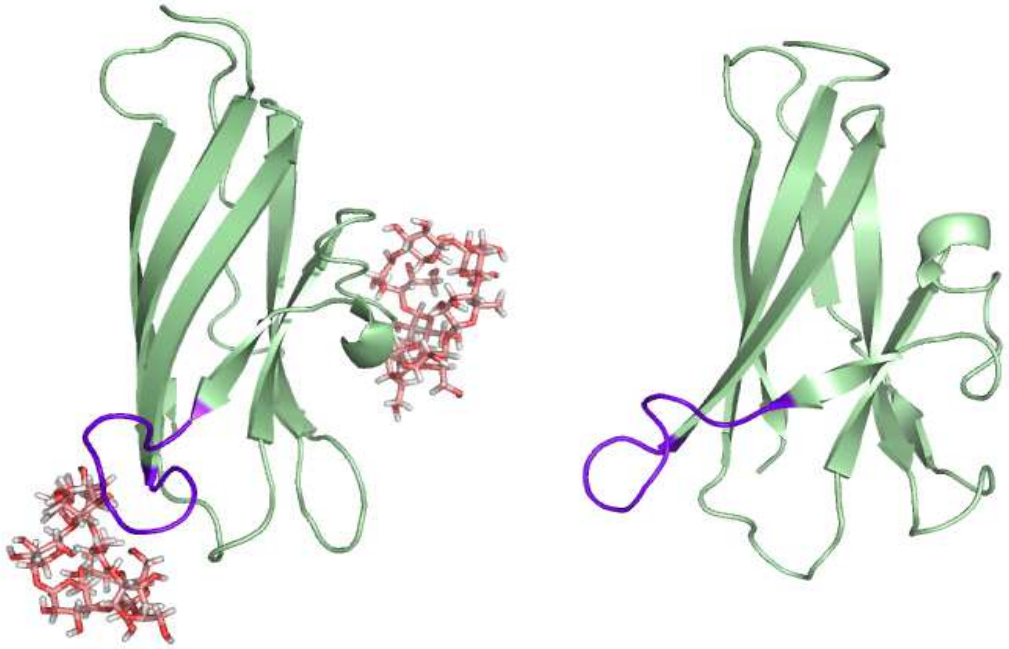


Figure 1.13. Cartoon representation of the CBM20 domain from *A.* niger glucoamylase 1 when bound to β-cyclodextrin (left, PDB ID: 1AC0) and with no ligand bound (right, PDB ID: 1KUL). The flexible loop associated with binding site 2 and showing the largest conformational change upon substrate binding is colored purple. Protein structure graphics were made using PyMol (DeLano 2002).

A. niger's glucoamylase 1 is a multi-domain enzyme containing an N-terminal hydrolase (GH15) domain and a C-terminal CBM20 joined by a highly *O*-glycosylated linker. The enzyme has exo-acting activity, releasing β-D-glucose from the non-reducing end of starch and related oligo- and polysaccharides (Morris et al. 2005). The structural fold is conserved among SBDs, all having a β-sandwich fold with an immunoglobulin-like fold (Christiansen et al. 2009b). The best charachterized SBD is the CBM20 from glucoamylase 1 (Figure 1.13) consisting of seven beta-stands (one parallel and six antiparallel pairs) forming an open-sided distorted beta-barrel (Sorimachi et al. 1996). This well-defined β-sheet structure is maintained when the SBD/CBM20 is bound to substrate (Sorimachi et al. 1997).

Most CBM20s have bivalent interactions with the substrate via two carbohydrate-binding sites, each involving two or three conserved solvent accessible aromatic residues (Penninga et al. 1996; Sorimachi et al. 1997). Site 1 is considered to be the site where the binding step commences (Morris et al. 2005). It is shallower and more solvent exposed than site 2 and site 1 shows only minor structural changes upon binding to a substrate. Site 2, compared to site 1, is more extended and undergoes substantial structural changes upon substrate binding (Sorimachi et al. 1997). This conformational change, combined with a stronger binding in binding site 2, is

thought to lock the substrate into position for catalysis. A flexible loop associated with binding site 2 (Fig. 1.13) allows binding of the SBD to starch strands in a variety of orientations (Sorimachi et al. 1997).

CBM20 domains are proposed to play two roles in starch degradation. Firstly, they help enzymes to attach to granular starch and to increase the enzyme concentration at the substrate surface (Morris et al. 2005). Secondly, they "unwind" α-glucan helices on the granule surface, enhancing the cleavage by the catalytic domain (Southall et al. 1999).

1.4 Goal of this study

The aim of this study was to provide more insight into the roles of different putative lytic polysaccharide monooxygenases (LPMs) from the family 33 of carbohydrate binding modules. These LPMs are of great interest both for the understanding of microbial polysaccharide degradation as well as for industrial purposes, to facilitate the degradation of crystalline substrates. The main objective has been to clone, express, purify and characterize CBM33s from two different organisms, the bacterium *C. flavigena*, the primary target being CBM33s with activity on chitin or cellulose, and the fungus *A. terreus*, with the primary target being CBM33s with activity on starch.

The first issue to be addressed was the characterization of the four CBM33s in the genome of *C. flavigena* (*Cf*CBM33A, *Cf*CBM33B, *Cf*CBM33C and *Cf*CBM33D; UniProt IDs are listed in Table 3.1). The intention was to clone, express and purify all four *Cf*CBM33s to study whether the *Cf*CBM33s have different substrate specificities, if they work synergistically in substrate degradation, and whether or not the C-terminal CBM2 present in all four *CfCBM33*s is required for substrate specificity and degradation.

The second part of this study was focused on a putative CBM33 from *A. terreus* (*At*CBM33A; UniProt ID in Table 3.1). This CBM33 is of great interest as it is linked to a CBM20 domain, suggesting specificity towards starch. So far, nothing is known about the starch degrading properties of CBM33s. Since LPMs have proven to boost the degradation of chitin and cellulose, two recalcitrant crystalline structures, it is conceivable that CBM33s from a starch degrading fungus such as *A. terreus* work on starch. Notably, there are no published data on CBM33s from fungi.

2 MATERIALS AND METHODS

2.1 Materials

2.1.1 Chemicals

Chemical	Supplier
2,5-Dihydroxybenzoic acid (DHB)	Bruker Daltonics
Acetic acid 99.8 %	VWR
Acetonitrile (CH_3CN)	Fulltime
Agar bacteriological (Agar No. 1)	Oxoid
Agarose, SeaKem®	Lonza
Albumin, bovine serum (BSA), Fraction V	Sigma-Aldrich
Ammonium bicarbonate (NH_4HCO_3)	Sigma-Aldrich
Ammonium chloride (NH_4Cl)	Fluka/ Sigma-Aldrich
Ammonium sulphate ($NH_4)_2SO_4$	Merck
Ampicillin	Sigma-Aldrich
Ascorbic acid	Sigma-Aldrich
Bacto™ Peptone	Becton, Dickinson and Company
Bacto™ yeast extract	Becton, Dickinson and Company
Bacto™Tryptone	Becton, Dickinson and Company
Biotin	Sigma-Aldrich
Bis-Tris ($C_8H_{19}NO_5$)	Sigma-Aldrich
Brain heart infusion (BHI)	Oxoid
Calcium chloride ($CaCl_2$)	Sigma-Aldrich
Calcium sulphate ($CaSO_4$)	Sigma-Aldrich
Coomassie Brilliant Blue R250	Merck
D(+)-Glucose monohydrate	VWR
DL-Dithiothreitol (DTT)	Sigma-Aldrich
D-Sorbitol	Sigma-Aldrich
Ethanol 96 % (v/v)	Arcus
Ethidium bomide, ultrapure Bioreagent	J.T. Baker
Ethylenediaminetetraacetic acid (EDTA)	Sigma-Aldrich
Ethylenediaminetetraacetic acid disodium salt dihydrate (EDTA-Na_2)	Sigma-Aldrich
Gallic acid	Sigma-Aldrich
Glycerol 85 % (w/v)	Merck
Hepes ($C_8H_{18}N_2O_4S$)	Sigma-Aldrich
Hydrochloric acid (HCl)	Merck
Imidazole	Sigma-Aldrich

Iodoacetamide	Sigma-Aldrich
Isopropyl β-D-1-thiogalactopyranoside	Sigma-Aldrich
L-Gluthation, reduced	Sigma-Aldrich
Magnesium chloride ($MgCl_2$)	Qiagen
Magnesium sulphate ($MgSO_4$)	Merck
Magnesium sulphate heptahydrate ($MgSO_4$) x $7H_2O$	Merck
MES ($C_6H_{13}NO_4S$) hydrate	Sigma-Aldrich
Methanol, HPLC grade	LAB-SCAN
N-Acetyl-D-glucosamine	Sigma-Aldrich
Phenylmethanesulfonylfluoride (PMSF)	Sigma-Aldrich
Phosphoric acid (KOH) 85 % (w/v)	Merck
Potassium chloride (KCl)	Merck
Potassium dihydrogen phosphate (KH_2PO_4)	Merck
Potassium phosphate dibasic (K_2HPO_4)	Sigma-Aldrich
Potassium sulphate (K_2SO_4)	Sigma-Aldrich
Simply Blue SafeStain	Invitrogen
Sodium acetate ($C_2H_3NaO_2$)	Sigma-Aldrich
Sodium acetate (CH_3COONa)	Sigma-Aldrich
Sodium chloride (NaCl)	Sigma-Aldrich
Sodium hydroxide (NaOH) 50 % (w/v)	J.T. Baker
Sodium sulphate (Na_2SO_4)	Sigma-Aldrich
Sodium phosphate dibasic heptahydrate (Na_2HPO_4 x $7H_2O$)	Sigma-Aldrich
Sodiumdodecylsulfate (SDS)	Bie & Berntsen
Sulphuric acid (H_2SO_4)	Sigma-Aldrich
Thiamine	Sigma-Aldrich
Trifluoro acetic acid (TFA)	Sigma-Aldrich
Tris(hydroxymethyl)aminomethan (Tris-HCl)	Sigma-Aldrich
Yeast extract	Remel
Yeast Nitrogen Base (YNB)	Fluka
Zeocin™	Invitrogen
α-cyano-hydroxy-cinnamic acid (CHCA), 97 % (w/v)	Aldrich
B-lactoglobulin from Bovine Milk, 90 % (w/v)	Sigma-Aldrich

2.1.2 Carbohydrate substrates

Table 2.1. Polysaccharide substrates used for cultivation experiments, binding assays and degradation experiments. Note that only some of the substrate specifications are known.

Substrate	Origin	Specifications provided by supplier	Supplier
α-chitin[A]	Shrimp shell	Dried and milled	Hov Bio, Tromsø, Norway
α-chitin[B]	Crab shell	Dried and milled	Y. Nakagawa, Japan
α-chitin[C]	Crab shell	Milled by cutter mill 60 sec, 88 % crystallinity	Y. Nakagawa, Japan
α-chitin[D]	Crab shell	Ball mill 30 min, 30% crystallinity	Y. Nakagawa, Japan
α-chitin[E]	Nanofiber		S. Ifuku, Japan
β-chitin[A]	Squid pen	dried and milled	France chitin, France
β-chitin[B]	Squid pen	dried and milled	Y. Nakagawa, Japan
β-chitin[C]	Squid pen	Cutter mill 60 sec, 78 % crystallinity	Y. Nakagawa, Japan
β-chitin[D]	Squid pen	Ball mill 60 min, 32 % crystallinity	Y. Nakagawa, Japan
β-chitin[E]	Nanofiber		S. Ifuku, Japan
Cellulose[A]	Avicel	~50 µm particle size	PH101Sigma-Aldrich
Cellulose[B]	PASC	Phosphoric acid swollen cellulose	K. Igarashi
Cellulose[C]	Filter paper	0.5 mm	Whatman no. 1
Xylan	Birchwood xylan		Roth, Karlsruhe, Germany

2.1.3 Primers

Table 2.2. Primers by name and sequence.

Primer name	Primer sequence (5' → 3')
InFCf33AF	CGCAACAGGCGAATGCTCACGGCTCCGTCACCGACCC
InFCf33AR1	CAGCCGGATCAAGCTTTCAGGCCGCGGTGCAGGT
InFCf33AR2	CAGCCGGATCAAGCTTTTACGCGGTCCCGTTGATGTT
InFCf33BF	CGCAACAGGCGAATGCTCACGGTGCCGTGTCCGACC
InFCf33BR1	CAGCCGGATCAAGCTTTCAGGCGGCTGCGCAGGT
InFCf33BR2	CAGCCGGATCAAGCTTTTAGCCACCGAAGGTCAC
InFCf33CF	CGCAACAGGCGAATGCTCACGGGTGGATCTCCGACC
InFCf33CR1	CAGCCGGATCAAGCTTTCAGGCGGCCGCGCAGGT
InFCf33CR2	CAGCCGGATCAAGCTTTTAGCCACCGGGCGTGATGT
InFCf33DF	CGCAACAGGCGAATGCTCACGGCGGTCTGACGAACCC
InFCf33DR1	CAGCCGGATCAAGCTTTCAGCGGACGATGCAGGGCT
InFCf33DR2	CAGCCGGATCAAGCTTTTAACCACCGAAGTCCACGT
Seq1Cf33A	CATCAACGGGACCGCGCCGACGCAG
Seq2Cf33A	CCGCCTGGTAGAGACCGGTCGTGCC
Seq1Cf33B	TGACCTTCGGTGGCGGTGGGACGCC
Seq2Cf33B	TGCTTCGTCACGTAGATCTTCAGGT
Seq1Cf33C	CACCGCGACGTTCAAGACGAACAAC
Seq2Cf33C	CTCGGGCAGGTTGTTGATCGTGTGC
Seq1Cf33D	CGTTCTACAACTGCGTCGACGTGGA
Seq2Cf33D	CTTGCCGCTCTTGTTCGACGGCAGC
3320_XhoI_SP	ATAACTCGAGAAAAGAGAGGCTGAAGCTCACGGCTACTTGACTATTC
3320_XbaI_ASP	GCGTCTAGATCAACGCCAAGAAGCGGCTG
pRSET-B SeqF	GATCTCGATCCCGCGAAATT
pRSET-B SeqR	TGTTAGCAGCCGGATCAAGC

Table 2.3. Primers by name and description.

Primer name	Primer description
InFCf33AF	*Cfcbm33A*, forward cloning primer
InFCf33AR1	*Cfcbm33A*, reverse cloning primer, full-length protein
InFCf33AR2	*Cfcbm33A-N*, reverse cloning primer, N-terminal domain
InFCf33BF	*Cfcbm33C*, forward cloning primer
InFCf33BR1	*Cfcbm33C*, reverse cloning primer, full-length protein
InFCf33BR2	*Cfcbm33C-N*, reverse cloning primer, N-terminal domain
InFCf33CF	*Cfcbm33D*, forward cloning primer
InFCf33CR1	*Cfcbm33D*, reverse cloning primer, full-length protein
InFCf33CR2	*Cfcbm33C-D*, reverse cloning primer, N-terminal domain
InFCf33DF	*Cfcbm33B*, forward cloning primer
InFCf33DR1	*Cfcbm33B*, reverse cloning primer, full-length protein
InFCf33DR2	*Cfcbm33B-N*, reverse cloning primer, N-terminal domain
Seq1Cf33A	*Cfcbm33A*, forward sequencing primer (annealing in the middle of the gene)
Seq2Cf33A	*Cfcbm33A*, reverse sequencing primer (annealing in the middle of the gene)
Seq1Cf33B	*Cfcbm33C*, forward sequencing primer (annealing in the middle of the gene)
Seq2Cf33B	*Cfcbm33C*, reverse sequencing primer (annealing in the middle of the gene)
Seq1Cf33C	*Cfcbm33D*, forward sequencing primer (annealing in the middle of the gene)
Seq2Cf33C	*Cfcbm33D*, reverse sequencing primer (annealing in the middle of the gene)
Seq1Cf33D	*Cfcbm33B*, forward sequencing primer (annealing in the middle of the gene)
Seq2Cf33D	*Cfcbm33B*, reverse sequencing primer (annealing in the middle of the gene)
3320_XhoI_SP	*Atcbm33A*, forward cloning primer
3320_XbaI_ASP	*Atcbm33A*, reverse cloning primer
pRSET-B SeqF	pRSET-B, forward sequencing primer
pRSET-B SeqR	pRSET-B, reverse sequencing primer

2.1.4 Bacterial strains and plasmids

Table 2.4. Bacterial strains

Strain	Source
Escherichia coli DH5α	Invitrogen
Escherichia coli TOP10	Invitrogen
Cellulomonas flavigena (DSM 20109)	ATCC (http://www.lgcstandards-atcc.org/)
Pichia pastoris X-33	Invitrogen

Table 2.5. Plasmids

Plasmid	Source or reference
pPICZα-A	Invitrogen (vector map is shown in Appendix D, Fig D.1)
pRSET-B/*cbp21*	(Vaaje-Kolstad et al. 2012) (pRSET-B vector map retrived from Invitrogen is shown in Appendix D, Fig D.2)

2.2 Cultivation of microorganisms

2.2.1 Agars, cultivation media and substrates

All medium and agar solutions were autoclaved on liquid cycle at 15 psi (1 bar) and 121°C for 20 minutes.

2.2.1.1 Luria Bertani (LB)

Liquid medium

10 g Bacto Trypton

5 g Bacto yeast extract

10 g NaCl

After carefully mixing and dissolving the ingredients, the volume was adjusted to 1 litre with dH_2O and the medium was autoclaved. If the medium was used to select for positive transformants, 50 µg/ml ampicillin was added prior to cultivation.

Agar plates

15 g/liter agar was added to the LB medium before autoclaving. After cooling to ~55°C and, if selecting for transformants, addition of 50 µg/ml ampicillin; the medium (1 liter) was distributed on 20 Petri dishes. After cooling for 20 minutes the plates were stored wrapped in plastic at +4°C.

2.2.1.2 Low salt LB

Liquid medium

10 g Bacto Tryptone

5 g Bacto yeast extract

5 g NaCl

After carefully mixing and dissolving the ingredients with 900 ml dH_2O the pH was adjusted to 7.5 using 1 M NaOH. The volume was adjusted to 1 liter with dH_2O before autoclaving. Prior to use, Zeocin™ was added to 25 µg/ml final concentration.

Agar plates

15 g/liter agar was added to the low salt LB medium before autoclaving. After cooling to 55°C, Zeocin™ was added to 25 µg/ml final concentration and the medium (1 liter) was distributed on 20 petri dishes. After cooling for 20 minutes, the plates were wrapped in foil and stored at 4°C.

2.2.1.3 Terrific broth (TB)

10 x TB salts:

23.12 g KH_2PO_4

125.41 g K_2HPO_4

The chemicals were carefully mixed, and dissolved in dH_2O to a final volume of 1 liter and autoclaved.

Liquid medium

12 g Bacto Trypton

24 g Bacto yeast extract

4 g glycerol

All ingredients were carefully mixed and dissolved in dH_2O to a final volume of 900 ml before autoclavation. After autoclavation the medium was added 100 ml 10 x TB salts.

2.2.1.4 Brain Heart Infusion (BHI)

Liquid medium

37 g BHI

BHI was completely dissolved in dH_2O to a final volume of 1 liter and autoclaved.

2.2.1.5 Yeast extract Peptone Dextrose (YPD)

Liquid medium

10 g yeast extract

20 g peptone

100 ml 20% (w/v) glucose, filter-sterilized

Yeast extract and peptone were dissolved in dH_2O to a total volume of 900 ml and autoclaved; after cooling down glucose was added. When selecting for transformants of *P. pastoris* Zeocin™ was added to the medium to a final concentration of 1µg/ml right before use.

Agar plates

20 g agar was added per 1 liter final volume of medium before autoclaving. After cooling to below 60°C, glucose and 1 µg/ml Zeocin™ were added before spreading on plates. Plates containing Zeocin™ were stored in the dark at 4°C.

2.2.1.6 Yeast extract Peptone Dextrose medium with Sorbitol (YPDS)

Liquid medium

10 g yeast extract

20 g peptone

182.2 g sorbitol

100 ml 20% (w/v) glucose, filter-sterilized

Yeast extract, peptone and sorbitol were dissolved in dH_2O to a total volume of 900 ml and autoclaved; after cooling down glucose was added. When selecting for transformants of *P. pastoris,* Zeocin™ was added to the medium to a final concentration of 1 µg/ml right before use.

Agar plates

For agar plates, 20 g agar was added per 1 liter final volume of medium before autoclaving. After cooling to below 60°C, glucose and 1 µg/ml Zeocine™ were added before spreading on plates. Plates containing Zeocin™ were stored in the dark at 4°C.

2.2.1.7 Buffered complex media

1M potassium phosphate buffer, pH 6.0:

66 ml 1M K_2HPO_4

434 ml 1M KH_2PO_4

The solutions were mixed together and the pH assured to be 6.0 before autoclaving.

2.2.1.8 Buffered complex glycerol medium (BMGY):

10 g yeast extract

20 g peptone

100 ml 1M potassium phosphate buffer, pH 6.0

100 ml 13.4 % (w/v) YNB

2 ml 0.02 % (w/v) Biotin

100 ml 10 % (w/v) Glycerol

Yeast extract and peptone were dissolved in 700 ml dH_2O and autoclaved. After cooling to room temperature, potassium phosphate buffer and filter-sterilized YNB, biotin and glycerol were added.

2.2.1.9 Buffered complex methanol medium (BMMY)

The recipe for BMGY was followed but instead of glycerol 100 ml 5 % (v/v) methanol was added as carbon source.

2.2.2 Antibiotics used for selective growth of microorganisms

Zeocin™

Zeocin™ (Invitrogen) is an antibiotic isolated from the bacterium *Streptomyces verticillus.* It shows strong toxicity against bacteria, fungi (including yeast), plants and mammalian cell lines. It binds to DNA and introduces nucleotide-specific DNA cleavage (Bostock et al. 2003). Organisms with the *Shble* gene (*Streptoalloteichus hindustanus* bleomycin gene) incorporated in their genome are resistant to Zeocine™. Bleomycine-protein binds stoichiometrically to Zeocin™ and inhibits its activity. In this study, Zeocin™ was used in selecting for positively transformed *E. coli* and *P. pastoris*, carrying the pPICZα-A vector containing the *Shble*-gene and the gene of interest.

Ampicillin

Ampicillin (Sigma-Aldrich) is a beta-lactam antibiotic active against both Gram-negative and Gram-positive bacteria. Ampicillin is commonly used as a selective marker in routine biotechnology. The β-lactamase (*bla*) promoter in pRSET-B allows expression of the ampicillin resistance gene, β-lactamase, which, in turn, allows selection of the plasmid in *Escherichia coli.*

2.2.3 Cultivation of bacterial strains

All reagents and media used for cultivation were sterilized either by autoclaving or sterile-filtration (0.45 μm pore size). All culturing work was performed in sterile conditions. To start a new culture, a single colony from an agar plate or a piece of glycerol stock (see section 3.3) was inoculated in 5 ml medium of choice in culture tubes and incubated overnight at 30°C or 37°C with shaking at 220 rpm as described in further detail below. For media recipes, see section 2.2.1. Growth of microorganisms was monitored by measuring optical density at a wavelength of 600 nm (OD_{600}).

C. flavigena strain DSM 20109 (TaxID: 446466) was cultivated in LB-medium, either on solid agar plates or in liquid medium, both without antibiotics, at 30°C in a heating cabinet overnight. Liquid cultures were incubated with shaking at 220 rpm.

E. coli is the most used model organism in prokaryotic genetics, physiology and biochemistry, and the most used bacterium in production of heterologous proteins. The Gram-negative, rod-shaped bacterium is typically present in the lower intestine of humans. Most *E. coli* strains have a generation time of about 20 minutes.

E. coli strains without antibiotic resistance were cultivated on LB- agar plates or in liquid LB-medium at 37°C in a heating cabinet overnight. Cultures in liquid LB were incubated with shaking at 220 rpm.

E. coli TOP10 cells (Invitrogen) carrying plasmids encoding ampicillin resistance were cultivated in liquid LB-medium supplemented with 50 µg/ml ampicillin or on LB- agar plates supplemented with 100 µg/ml ampicillin.

E. coli Bl21 Star™ (DE3) cells (Invitrogen) carrying plasmids encoding ampicillin resistance were cultivated in LB- medium supplemented with 50 µg/ml ampicillin or on LB-agar plates supplemented with 100 µg/ml ampicillin. For optimal expression of *Cf*CBM33 (see below), cells were grown in TB-medium supplemented with 50 µg/ml Ampicillin at 25°C for 3 days.

E. coli XL gold cells (Stratagene) carrying plasmids encoding Zeocin™ resistance were cultivated in Low Salt LB- media or on Low Salt LB- agar plates, both containing 25 µg/ml Zeocin™. It is necessary to use a low salt medium with neutral pH because both high ionic strength and acidity or basicity inhibits the antibiotic activity of Zeocin™.

2.2.4 Cultivation of *Pichia pastoris*

The yeast *P. pastoris* is frequently used as a protein expression system. It combines important properties of higher eukaryotic systems related to protein processing, protein folding and post translational modifications, with being almost as easy to grow and manipulate as *E. coli* or *Saccharomyces cerevisiae.*

P. pastoris strain X-33 was cultivated in yeast extract peptone (YPD)-medium or on YPD-agar plates. To start a new culture, a single colony from an agar plate was inoculated in 5 ml YPD-medium in culture tubes and grown at 30°C overnight with shaking at 220 rpm. *P. pastoris* X-33

has a generation time of about 2.7 hours when grown on YPD-medium and an OD_{600} of 1 corresponds to a cell density of 5 x 10^7 cells/ ml (M. Abou Hachem, personal communication).

For selection of positive transformants of *P. pastoris*, the culture was cultivated in YPD-medium (liquid or agar plates) supplied with Zeocin™ accordign to section 2.2.2.

For protein expression, *P. pastoris* was grown in buffered complex methanol medium (BMMY) with methanol as the sole carbon source. Every day methanol was added to a final concentration of 0.5 % (v/v) to compensate for consumption and, to some extent, evaporation.

2.2.5 *C. flavigena* cultivation experiments

As noted in section 1.2.4.2, *C. flavigena* is reported to use both xylan and cellulose as carbon source. Previous studies have shown successfully growth of *C. flavigena* with xylan and cellulose as sole carbon source (Sanchez-Herrera et al. 2007). To examine *C. flavigena*'s metabolic adaptation, the bacterium was cultivated in minimal medium supplied with different substrates with glucose as a control.

Materials

- M9-salts:

  64 g $Na_2HPO_4 \times 7H_2O$

  15 g KH_2PO_4

  2.5 g NaCl

  5.0 g NH_4Cl

  All ingredients were dissolved in dH_2O to a final volume of 1 liter and autoclaved.
- 1M $MgSO_4$
- 1M $CaCl_2$
- Autoclaved dH_2O
- Substrates (see section 2.1.2, Table 2.1):
  - Cellulose[A]
  - Xylan
  - α-chitin[A]
  - β -chitin[A]
  - Glucose
- 0.1 mg/ml Biotin
- 10 mg/ml Thiamine

Procedure

For each culturing flask, minimal medium was made by mixing 30 ml M9-salts, 0.3 ml 1M $MgSO_4$ and 15 µl 1M $CaCl_2$ with 116,7 ml dH_2O. A 5 ml overnight inoculum of *C. flavigena* (prepared according to section 2.2.3) was centrifuged and the pellet resuspended in 1 ml minimal medium. Out of this suspension, 150 µl was transferred to each culturing flask (1 litre) supplemented with 10 µg/L biotin, 1 mg/L thiamine, and 1% (w/v) substrate (all final concentrations). As a control, one culture was set up in minimal medium containing 0.2 % (w/v) glucose. All cultures (total volume 250 ml) were incubated at 30°C with shaking at 220 rpm for a total of 7 days. Each day, 10 ml of each culture was taken out. After removing the cells by centrifugation the culture supernatants were concentrated by the use of Amicon Ultra centrifugal filter units with 10 kDa cut-off (Millipore) and analyzed by SDS-PAGE according to section 2.18.

2.3 Long-term storage of microorganisms

For long term storage of microorganisms, glycerol was added as a cryoprotectant in order to ensure the survival of most cells. Cultures of the different microorganisms, containing different constructs, were preserved by glycerol as follows:

- 1 ml overnight bacterial or yeast culture
- 300 µl glycerol (85 % (w/v), sterile)

After carefully mixing the culture with glycerol in a cryo-tube, the glycerol stock was kept at -80°C. Inoculations of new cultures were made by scraping small amounts of the frozen stock with a toothpick and inoculating tubes containing 5 ml appropriate growth medium including antibiotics when required. Alternatively, the cells were first inoculated on an agar plate for obtaining single colonies.

2.4 Extraction of chromosomal DNA from *C. flavigena*

Materials

- TE-buffer pH 7.0:
  1 ml 1M Tris-HCl (final concentration 10 mM)
  0.2 ml 0.5 M EDTA (final concentration 1 mM)

Tris-HCl and EDTA were completely dissolved in dH_2O to a total volume of 60 ml and the pH regulated to 7.0 with 6 M HCl. Finally, the volume was adjusted to 100 ml with dH_2O.

- o E.Z.N.A.® Bacterial DNA Kit (Omega Bio-Tek)
  - Lysozyme Solution
  - BTL Buffer
  - Proteinase K
  - RNaseA
  - DBL Buffer
  - DNA Wash Buffer
  - Elution Buffer, preheated to 65°C
  - HiBind DNA columns, collection tubes
- o 96% Ethanol

Procedure

Extraction of the chromosomal DNA from *C. flavigena* was performed according to the E.Z.N.A.® Bacterial DNA Kit, spin protocol:

2 ml *C. flavigena* overnight culture was harvested by centrifugation at 4,000 x g for 10 min at room temperature. After completely removing the medium, cells were resuspended in 100 µl TE buffer before 10 µl 50 mg/ml lysozyme solution was added. The sample was incubated at 37°C for 10 minutes for lysozyme digestion of the cell wall. To digest proteins in the mixture, and thereby incapacitate any DNases, 100 µl BTL buffer and 20 µl 2 mg/ml proteinase K were added to the sample. After incubation at 55°C in a water bath for 1 hour, 5 µl RNaseA was added and the tube that was immediately inverted at least 6 times, followed by 5 minute incubation at room temperature. 220 µl DBL buffer was added, and after vortexing the sample was incubated at 65° C for 10 min. To ensure optimal binding of sample DNA to the column, 220 µl 96% ethanol was added and the sample vortexed for 20 seconds (no precipitation was observed). The sample was transferred to a HiBind DNA column placed in a collection tube and centrifuged at 10,000 x g for 1 minute. The flow-through was discarded and two wash steps were performed each with 700 µl DNA wash buffer. All flow-through was discarded. The HiBind DNA Column was then dried by centrifugation at max speed for 2 min and then placed in a nuclease-free 1.5 ml microfuge tube. 50 µl elution buffer preheated to 65°C was then applied to the HiBind matrix followed by a 3 minute incubation room temperature. The DNA was eluted by centrifugation at 10,000 x g for 1 min. For maximum yield, the elution step was repeated once.

2.5 Polymerase Chain Reaction

The Polymerase chain reaction (PCR) is an *in vitro* technique used for amplification of specific nucleotide sequences. The reactions require both sequence specific oligonucleotide primers complementary to the sequence of interest, dNTPs, and a thermo-stable DNA polymerase.

2.5.1 PCR using Phusion™ High-Fidelity DNA Polymerase

Phusion™ High-Fidelity DNA Polymerase (Finnzymes) works fast and generates long DNA molecules with a high accuracy. In this study, the polymerase was used for three purposes: i) Amplification of putative LPM-genes *Cfcbm33A*, *Cfcbm33B*, *Cfcbm33C*, and *Cfcbm33D* from *C. flavigena*, all without the signal sequence and in two variants, one with and one without the C-terminal CBM2-domain. ii)Amplification of the codon optimized and synthetically made putative LPM-gene *Atcbm33A* from *A. terreus* NIH2624 (Taxid: 341663; see www.genescript.com/ for details on codon optimization algorithms and gene synthesis protocol). iii) Verification of transformed *P. pastoris* X-33 cells carrying the *Atcbm33A* gene in the genome.

Materials

- Phusion™ High-Fidelity DNA Polymerase (Finnzymes):
  - dNTP mix, 10 mM
  - 5x Phusion™ GC Buffer
  - Phusion™ DNA polymerase (2 U/µl)
- Primers (see section 2.1.3, Table 2.2 and 2.3)
- Nuclease-free dH_2O

Procedure

PCR reactions (50µl) were set up on ice in 0.2 ml PCR tubes according to table 2.6. Reaction mixes were placed in a Master cycler gradient 120V (Eppendorf) and amplification was carried out using the settings shown in table 2.7.

Table 2.6. Reaction setup for PCR using Phusion™ High-Fidelity DNA Polymerase

Reagents	Volume (Final concentration)
dNTPs 10 mM	1 µl (2 mM)
5x Phusion™ HF / GC buffer*	10 µl
DNA Template	2 µl (approximately 80 ng)
Forward primer	1 µl (0.5 µM)
Reverse primer	1 µl (0.5 µM)
dH_2O	34.5 µl
Phusion™ DNA Polymerase	0.5 µl (1 U)

*5x Phusion HF- buffer was used in PCR reactions with *Atcbm33A* as template, whereas the GC-buffer was used in reactions with *Cfcbm33*'s because of a high GC- content in the *C. flavigena* genome.

Table 2.7. Cycling parameters for PCR using Phusion™ High-Fidelity DNA Polymerase.

Step	Temperature	Time (minutes:seconds)	Number of cycles
Initial denaturation	98°C	1:30	1
Denaturation	98°C	0:10	25
Annealing	60°C	0:30	
Elongation	72°C	3:00	
Final elongation	72°C	5:00	1

After amplification, the PCR products were analyzed by agarose gel electrophoresis, described in section 2.5.3. The appropriate DNA band was cut from the gel and the DNA fragment purified as described in section 2.5.4 .

2.5.2 PCR using Red Taq DNA Polymerase Master Mix

Red Taq DNA Polymerase Master Mix (VWR) contains Taq polymerase, dNTPs and magnesium chloride. This mix was used in PCR reactions to verify transformed products.

Materials

- Red Taq DNA Polymerse Master Mix (VWR)
- Primers (see section 2.1.3, Table 2.2)
- dH_2O

Procedure

PCR reactions were set up on ice in 0.2 ml PCR tubes following table 2.8. The reaction mixes were placed in a Master cycler gradient 120V (Eppendorf) and the amplification was carried out according to table 2.9.

Table 2.8. PCR reactions set up using Red Taq DNA Polymerase Master Mix.

Reaction Component	Volume
Forward sequencing primer for pRSET-B	5 µl (50 pmol)
Reverse sequencing primer for pRSET-B	5 µl (50 pmol)
Template DNA	1/5 of a colony with transformed cells dissolved in 5 µl dH_2O
dH_2O	10 µl
Red Taq DNA Polymerase Master Mix	25 µl

Table 2.9. Program for the thermal cycler when using Red Taq DNA Polymerase Master Mix.

Step	Temperature	Time (minutes:seconds)	Number of cycles
Initial denaturation	64°C	6:00	1
Denaturation	95°C	0:30	25
Annealing	60°C	0:30	
Elongation	72°C	0:30	
Final Elongation	72°C	5:00	1

After amplification, the PCR product was analyze by agarose gel electrophoresis as described in section 2.5.3.

2.5.3 Agarose gel-electrophoresis

The final PCR products were verified by separation on a 1% agarose gel. Linear DNA molecules are charged and, when subject to an electric field in a gel matrix they are separated according to size. For DNA visualization, ethidium bromide was added to the gel. Ethidium bromide intercalates between the stacked nucleotide bases and because of its fluorescent properties, DNA binding ethidium bromide can be visualized under a UV lamp. The size of the DNA molecules is determined by comparison to a DNA ladder comprised of DNA fragments of known size.

Materials

- Agarose
- 1x TAE-buffer:
  - 4.85 g Tris-base
  - 1.14 ml Acetic acid, 99.8% (v/v)
  - 2 ml 0.5 M EDTA, pH 8.0
  - Dissolved and mixed in 1 liter dH_2O, yielding a final pH of 8.5.
- Ethidium bromide, 10 mg/ml
- 10x loading buffer (Takara)
- 1 kb DNA ladder (Fermentas) or 100 bp DNA ladder (New England Biolabs)

Procedure

1 % (w/v) agarose gels were made by dissolving 0.5 g agarose in 50 ml 1xTAE-buffer by careful heating in a microwave-oven. The solution was cooled to below 60°C before 1 µl Ethidium bromide was added. The gel solution was poured into the gel-chamber of a Mini-Sub Cell GT cell with a UV-transparent gel casting tray (Bio-Rad) with a 15-wells spacer, and left to solidify. After 30 minutes the, spacer was removed and the gel transferred to an electrophoresis chamber. 1xTAE- buffer was added to the chamber covering the gel completely. Both the 1 kb DNA ladder and samples were added 0.1 volumes 10x loading dye and applied to the gel. The gel was run for 40 minutes at 90 volt using PowerPac Basic™ power supply (Bio-Rad) and DNA-bands were visualized by UV light.

2.5.4 Extraction of DNA fragments from agarose gels

Materialer

- NucleoSpin® Extract II kit (Macherey-Nagel)
  - Solubilisation-binding NT buffer
  - Wash- buffer NT3
  - Elution buffer NE
  - NucleoSpin® Extract II columns and 2 ml collection tubes
- Scalpel

Procedure

Extraction of verified PCR products from the agarose gel was done using the NucleoSpin® ExtractII- kit. All centrifugations were carried out at room temperature for 1 minute at 11,000 x g using a Centrifuge 5415R (Eppendorf).

The DNA fragment was excised from the agarose gel by cutting carefully around the band with a clean scalpel. The weight of the gel piece was determined and for every 100 mg of agarose gel, 200 µl NT buffer were added, followed by incubation at 50°C until the gel slice was completely dissolved. The sample was then loaded onto a NucleoSpin® Extract II column placed in a 2 ml collecting tube and centrifuged. The flow-through was discarded and the column washed with 600 µl NT3 buffer by centrifugation. After discarding the flow-through, remaining liquid was removed from silica membrane by re-centrifugation. The NucleoSpin® Extract II column was placed into a clean 1.5 ml micro centrifuge tube and added 25 µl elution buffer NE. After incubation at room temperature for 1 minute, the DNA was eluted by centrifugation. The elution step was repeated once to increase the yield of eluted DNA.

2.6 Plasmid isolation from *E. coli*

2.6.1 Plasmid isolation using NucleoSpin® Plasmid kit

The vector used in cloning of all *Cfcbm33* genes was a pRSET-B-based expression construct used for production of CBP21(Vaaje-Kolstad et al. 2005b), referred to here as pRSET-B/*cbp21*.

Materials

- NucleoSpin® Plasmid/Plasmid (NoLid) kit (Macherey-Nagel)
  - NucleoSpin® Plasmid/Plasmid (NoLid) Column
  - Collection tubes (2 ml)
  - Resuspension buffer A1
  - Lysis buffer A2
  - Neutralization buffer A3
  - Wash buffer A4
  - Elution buffer AE

Procedure

pRSET-B/*cbp21* was isolated from *E.* coli TOP10 cells following the protocol for plasmid DNA preparation with the NucleoSpin® Plasmid/Plasmid (NoLid) kit. All reaction steps were carried out at room temperature and all centrifugations were done at 11,000 x g using a Cetnrifuge 5415R (Eppendorf).

From an *E. coli* overnight culture, 2 ml was harvested by centrifugation for 30 seconds. The medium was completely removed before resuspending the pellet in 250 µl resuspension buffer A1. For cell lysis, 250 µl lysis buffer A2 was mixed into the sample by inverting the tube 6-8 times. The sample was then incubated until the lysate appeared clear. Maximum incubation time was 5 minutes. Lysis was stopped by adding 300 µl neutralization buffer A3 and inverting the tube 6-8 times before the lysate was centrifuged for 5 minutes. The clear supernatant (750 µl) was loaded onto a NucleoSpin® Plasmid/Plasmid (NoLid) column placed in a 2 ml collection tube and spun down for 1 minute. The flow-through was discarded. Washing of the column was performed by adding 600 µl wash buffer A4 to the column and centrifuging for 1 minute. The flow-through was discarded and remaining liquid was removed by an additional 2-minute centrifugation. Prior to elution, the NucleoSpin® Plasmid/Plasmid (NoLid) column was placed in a 1.5 ml micro centrifuge tube followed by addition of 50 µl elution buffer AE. After 1 minute incubation the DNA was eluted by centrifugation for 1 minute.

2.6.2 Plasmid isolation using GeneJET™ Plasmid Miniprep kit

The cloning vector used for transformation in *P. pastoris* was the expression vector pPICZα-A previously transformed into *E. coli* DH5α. pPICZα-A was isolated by the use of GeneJET™ Plasmid Miniprep kit (Fermentas). The same procedure was also used for isolation of pPICZα-A/*Atcbm33A* from transformed *E. coli* XL Gold.

Materials:

- GeneJET™ Plasmid Miniprep Kit (Fermentas)
  - Resuspension solution
  - Lysis solution
  - Neutralization solution
  - Wash solution
- Nuclease-free dH_2O

Procedure:

Plasmids were isolated from 5 ml overnight culture of transformed *E. coli.* All steps were carried out at room temperature.

E.coli cells from a 5 ml overnight culture were harvested by centrifugation at 6,800 x g for 2 minutes using a Centrifuge 5415R (eppendorf). After decanting the supernatant the pellet was completely resuspended in 250 µl resuspension solution and transferred to a micro centrifuge tube. To the resuspended cells, 250 µl lysis solution was added and the solutions were mixed carefully by inverting the tube 6 times. When the solution became viscous and slightly clear, 350 µl neutralization solution was added. The reagents were mixed by inverting the tube 6 times, followed by 5 minutes centrifugation at 10,000 x g. The supernatant was transferred to the supplied GeneJET^{TM} spin column by decanting, followed by centrifugation at 10, 000 x g for 1 minute.

After discarding all flow-through, the column was placed back into the same collection tube and washed by adding 500 µl wash solution followed by centrifugation as above. This was repeated once and, after discarding flow-through the second time, the column was centrifuged again to remove residual wash solution. To elute the plasmid DNA, the GeneJET^{TM} spin column was transferred to a fresh 1.5 ml micro centrifuge tube and 50 µl nuclease-free dH_2O was applied to the centre of the column membrane. Subsequent to 2 minutes incubation, the eluent was separated from the column by a 2 minute centrifugation at 10,000 x g. This elution step was repeated once with 30 µl nuclease-free dH_2O.

The DNA concentration was measured for a 20 x diluted sample in a quartz cuvette with 10.0 mm path using an Ultrospec 2100 pro spectrophotometer (GE Healthcare) at 280 nm.

2.7 Restriction digestion

Restriction endonucleases are sequence-specific enzymes that cleave DNA molecules at particular sites, producing a double-stranded break in the DNA strand. Cleavage may form blunt ends, meaning that both strands are cut at the same position, or overhanging (cohesive) ends, meaning that the two strands are cut at slightly different positions. Different restriction enzymes may require different reaction conditions for optimal digestion.

2.7.1 Double restriction digestion of *Atcbm33A* and pPICZα-A

The expression vector pPICZα-A was prepared for cloning by a double restriction digestion, using *Xho*I and *Xba*I. These restriction enzymes form cohesive ends and to promote integration, *Atcbm33A* was also double-digested with *Xba*I and *Xho*I.

Materials

- Restriction enzymes
  - *Xho*I, 10 U/µl (Fermentas)
  - *Xba*I, 10 U/µl (Fermentas)
- Buffers
  - 10 x Buffer R (Fermentas)
  - 10 x Buffer Tango™ (Fermentas)
- Nuclease-free dH_2O

Procedure

Both pPICZα-A and *Atcbm33A* was first cut by the restriction endonuclease *Xho*I. A 50 µl reaction was set up with 5 µl 10 x buffer R mixed with 5 µl *Xho*I and a total of 1 µg DNA diluted in nuclease-free dH_2O. For optimal digestion, the reaction was incubated at 30°C for 16 hours, followed by 20 minute incubation at 80°C to inactivate the enzyme.

To remove the inactivated *Xho*I enzyme and buffer components, purification was performed according to section 2.7.4. For digestion using *Xba*I, a total of 1 µg DNA diluted in nuclease-free dH_2O was mixed with 5 µl buffer Tango and 5 µl of *Xba*I restriction enzyme in a total volume of 50µl. The reaction was incubated at 37°C for 3 hours before inactivation of the enzyme by incubation at 65°C for 20 minutes, and finally the digested DNA was purified according to section 2.7.4. *Atcbm33A* was then cloned into pPICZα-A (pPICZα-A/*Atcbm33A*) according to section 2.10.2.

2.7.2 Plasmid preparation of pRSET-B

To remove the protein-coding sequence of *cbp21* and enable insertion of *Cfcbm33*-genes, pRSET-B/*cbp21* extracted from *E. coli* TOP10 (section 2.6.1) was digested with *Hin*dIII and

*Bsm*I. By doing so, the signal peptide of CBP21 remains part of the construct and will drive the translocation of the inserted *Cf*CBM33.

Materials

- Restriction enzymes
  - *Hin*dIII, 20 U/µl (New England Biolabs)
  - *Bsm*I, 10 U/µl (New England Biolabs)
- NEbuffer2 (New England Biolabs)
- dH_2O

Procedure

For restriction cutting, 8 µl plasmid extract (see section 2.6.1) was mixed with 6 µl of dH_2O and 4 µl NEBuffer before addition of 2 µl *Hin*dIII restriction endonuclease. The digestion reaction mixture was incubated for 2.5 hours at 37°C in a water-bath. After incubation with *Hin*dIII, 2.2 µl *Bsm*I was added directly to the reaction and the mixture was incubated for 3.5 hours at 65°C in a water-bath. The high temperature inactivates *Hin*dIII. *Bsm*I was finally inactivated by incubating the reaction mixture for 20 minutes at 80°C in a water-bath.

The digested plasmid was visualized on an agarose gel as described in section 2.5.3 and was extracted from the gel as described in section 2.5.4.

2.7.3 Linearization of pPICZ*α*-A/*Atcbm33A*

Prior to transformation into *P. pastoris*, to stimulate recombination, pPICZα-A/*Atcbm33A* (cloned according to 2.10.2) was linearized within the 5'AOX1-region using the *Sac*I endoglucanase.

Materials

- 10 x buffer *Sac*I (Fermentas)
- *Sac*I 10 U/µl (Fermentas)
- Nuclease-free dH_2O

Procedure

Reactions were set up on ice according to Table 2.10 and the linearization was done by 3 hour incubation at 37°C followed by an inactivation of *Sac*I by incubation at 65°C for 20 minutes.

Table 2.10. Reaction set up for linearization of pPICZα-A/*Atcbm33A*.

Reaction component	Volume
Purified plasmid	30 µl (approximately 0.5 - 1 µg DNA)
10 x Buffer *Sac*I	4 µl
Nuclease-free dH2O	2 µl
SacI	4 µl

2.7.4 PCR purification

Restriction digested DNA and PCR-products were purified by the use of GeneJET™ PCR Purification Kit (Fermentas). This removes primers, dNTPs, enzymes and salts from the reaction mixture.

Materials

- GeneJET™ PCR Purification Kit (Fermentas)
  - Binding buffer
  - Eluation buffer, preheated to 45°C
  - Wash buffer
  - GeneJET™ Purification Columns and collection tubes
- 3M sodium acetate
- Ethanol, 100 % (v/v) and 80 % (v/v)
- 2 mM Tris-HCl

Procedure

The GeneJET™ PCR Purification Kit's protocol was followed. All purification steps were carried out at room temperature and all centrifugations at 13,000 x g for 1 minute.

The DNA was mixed thoroughly with 1 volume of binding buffer and transferred to a GeneJET™ purification column. After centrifugation of the column, the flow-through was discarded and the column was washed by adding 700 µl wash buffer followed by centrifugation. All flow-through was discarded and the silica membrane dried by another centrifugation. Then

the column was transferred to a clean 1.5 ml micro centrifuge tube and purified DNA was eluted by centrifugation in two steps with a total of 70 µl pre-heated elution buffer. Eluted DNA was precipitated according to section 2.9.

2.8 DNA sequencing

Plasmids generated according to section 2.10 and isolated from *E. coli* according to section 2.6.1 and 2.6.2, were sequenced to check for mutations before transformation of plasmids into an expression host. For plasmids carrying bacterial genes, the BigDye® Terminator v3.2 Cycle Sequencing Kit (Applied Biosystems) was used for sequencing. Plasmids containing a fungal gene were sent to Eurofins (www.eurofins.com) for sequencing. All data analysis of sequences was performed using GENtle (Manske 2003).

2.8.1 BigDye® Terminator v3.1 Cycle Sequencing Kit, Sequencing PCR

During DNA sequencing by chain-termination, the same principle as for PCR is used, only with the addition of dideoxynucleotides (ddNTPs). ddNTPs lack the 3' hydroxyl group and addition of a ddNTP instead of a dNTP to the growing 3'- end results in termination of the elongation. The four different ddNTPs have different fluorescent tags, absorbing light at different wave-lengths. After amplification, every base from the DNA-strand can be visualized by separating all possible sequence lengths and measuring the absorbance for each amplicon. An ABI Prism 3100- genetic analyzer and BigDye® Terminator Cycle Sequencing Kit v3.1sequencing were used for dye terminator sequencing.

Materials

- BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems)
  - Ready Reaction Premix
  - BigDye Terminator v1.1/3.1 Sequencing Buffer (5x)
- Sequencing primers (see section 2.1.3)
- dH_2O

Four reactions were set up for sequencing, one using the forward pRSET-B-primer, one with the reverse pRSET-B-primer, one with the gene-specific forward and the final with the reverse sequencing primer; the last two annealing in the middle of the gene. Both reactions were set up

on ice in 0.2 ml PCR tubes according to table 2.11. The reaction tubes were placed in a Master cycler gradient 120V (Eppendorf) and the reactions carried out using settings shown in table 2.12.

Table 2.11. Reaction mixture for DNA sequencing.

Reagent	Quantity
Template	250 ng
Primer	3.2 pmol
BigDye Sequencing Buffer	2 µl
dH_2O	To a final reaction volume of 20 µl
Ready Reaction Premix	4 µl

Table 2.12. Thermal cycling conditions for DNA sequencing

Reaction step	Temperature	Time (minutes:seconds)	Number of cycles
Initial Denaturation	96°C	1:00	1
Denaturation	96°C	0:10	25
Annealing	50°C	0:05	
Elongation	60°C	4:00	

After sequencing, the products were precipitated using Pellet Paint®, as described in section 2.9.

2.9 Ethanol Precipitation of DNA using Pellet Paint®

Ethanol precipitation is commonly used for concentrating nucleic acid (DNA or RNA) preparations. DNA has a highly charged phosphate backbone, making the molecule polar and thus soluble in water. A salt is added to the sample to neutralize these charges, in this case sodium acetate. The positively charged sodium ions neutralize the negative charges in DNA but due to the high dielectric constant of water the ionic interactions are weak. To strengthen these interactions ethanol, which has a lower dielectric constant, is added and the DNA is precipitated. The Pellet Paint® Co-Precipitant (Novagen) is a visible dye-labeled carrier formulated specifically for alcohol precipitation of nucleic acids. In this study, precipitation was used for concentrating DNA samples and for precipitating PCR products for sequencing.

Materials

- 3 M sodium acetate pH 5.2
- Ethanol, 96% (v/v) and 70 % (v/v)

- Pellet Paint® Co- Precipitant (Novagen)

Procedure

All reactions were carried out at room temperature and all centrifugations were at 14,000 x g for 5 minutes using Centrifuge 5415R (Eppendorf).

The nucleic acid sample was mixed with 2 µl Pellet Paint® and 0.1 volumes 3 M Na Acetate (both at room temperature) and the solution was mixed briefly. After adding 2 volumes 96 % ethanol the sample was mixed and incubated for 2 minutes at room temperature. Subsequent to centrifugation, a pink pellet was visible in the bottom of the tube and the supernatant was removed. To wash the pellet, 4 volumes 70 % ethanol were added and the sample was vortexed and centrifuged. After removal of the supernatant the pellet was rinsed in 100 % ethanol and centrifuged again. The final supernatant was removed and the pellet was air-dried on the bench for 30 minutes.

2.10 Cloning

2.10.1 In-Fusion™ Cloning of *Cfcbm33*-genes into pRSET-B

In-Fusion™ Advantage PCR Cloning Kits (Clontech) are designed to combine multiple pieces of DNA without the use of restriction enzymes, ligase or phosphatise, and are typically used to clone PCR products into vectors. In-fusion is a PCR based DNA cloning procedure used for cloning a gene directly into a linearized vector. Primers used for the PCR amplification of the gene for insertion extends the gene with 15 base pairs in each end, overlapping each end of the insertion site. The linearized vector and the gene of interest are mixed and incubated with an In-Fusion™ enzyme that creates single-stranded regions at the ends of the vector and the PCR product. These regions are then fused because of the 15 bp of complementary bases. In this work, InFusion™ cloning was used for cloning the *Cfcbm33*-genes, lacking the native signal peptide, into pRSET-B by annealing it to the *cbp21* leader peptide (pRSET-B/*Cfcbm33*).

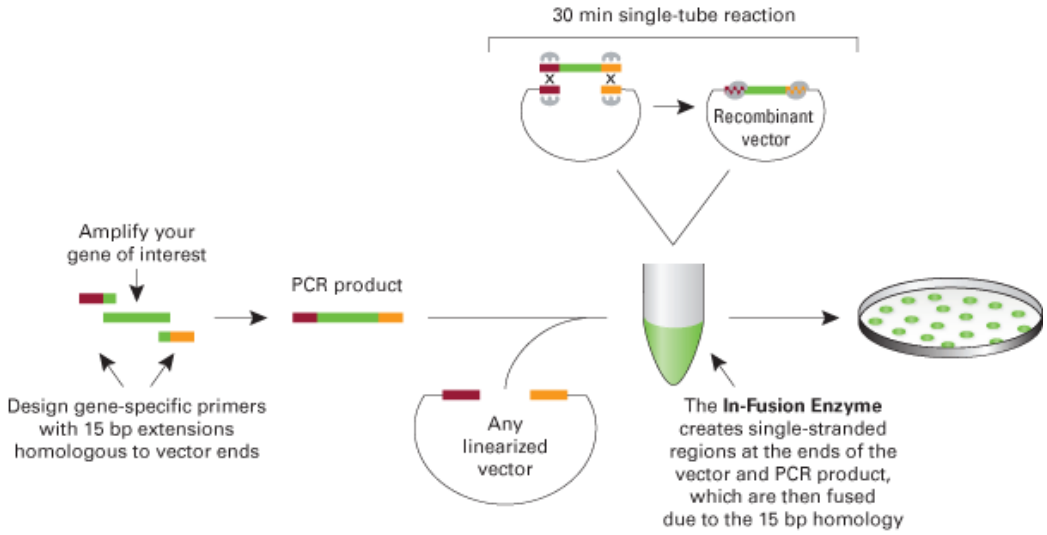


Figure 2.1. The principle of In-fusion™ cloning. The gene of interest is amplified by PCR using primers extending the gene with 15 base pairs in each end, identical with each end of the linearized vector. During 30 minutes of incubation, the In-Fusion™ enzyme generates single-stranded regions at the end of both the insert and the vector. The insert is then ligated with the vector because of the 15 bp complementary regions of the insert and vector. The construct is then transformed into bacterial strains for selective growth of positive transformants (type of selection depends on the vector's specificities). For more information, see www.clontech.com. Figure source: http://openwetware.org/wiki/In-fusion_biobrick_assembly.

Materials

- TE-buffer 8.0:
  157.6 mg Tris-HCl
  29.2 mg EDTA
  Dissolved in 60 ml dH_2O, adjusted to pH 8.0 with 6 M HCl and finally adjusted to 100 ml with dH_2O.
- 5x In-Fusion™ Reaction Buffer (Clontech)
- In-Fusion™ Enzyme (Clontech)
- Nuclease free dH_2O

Procedure

In this case, Protocol II called "In-Fusion™ Cloning Procedure W/Spin-Column Purification" was used. For maximum cloning efficiency, a 2:1 molar ratio of insert vs. vector was used.

Reactions were set up according to table 2.13 and incubated initially for 15 minutes at 37°C followed by 15 minutes at 50°C, after which the reaction mixtures were placed on ice. Before transformation, the reaction volume was brought up to 50 µl with TE buffer (pH 8.0). Only

Cfcbm33A-N and *Cfcbm33B-N* were successfully cloned into pRSET-B, generating pRSET-B/*Cfcbm33A-N* and pRSET-B/*Cfcbm33B-N*.

Table 2.13. Reaction set up for In-Fusion™ cloning of *Cfcbm33* into pRSET-B generating p-PICZα-A/*Cfcbm33*-constructs.

Reaction Component	Cloning Reaction	Negative Control	Positive Control
Purified PCR insert, *Cfcbm33*	2 µl (50 ng)	-	2 µl[2]
Linearized vector, pRSET-B	2 µl (100 ng)	1 µl[1]	1 µl[1]
5x In-Fusion™ Reaction Buffer	2 µl	2 µl	2 µl
In-Fusion™ Enzyme	1 µl	1 µl	1 µl
Deionized water	3 µl	6 µl	4 µl

[1]: linearized pUC19 Control vector (50 ng/µl) included in the kit.
[2]: 2 kb Control insert (40 ng/µl) included in the kit.

2.10.2 Cloning into pPICZα-A

Atcbm33A and pPICZα-A both had cohesive ends as a result of the double restriction digestion performed according to section 2.7.1, and was cloned by using T4 DNA ligase (New England Biolabs) in a ligation reaction with molar vector vs. insert ratio of 3:1.

Materials

- T4 DNA ligase (New England Biolabs)
- 10 x T4 DNA ligase reaction buffer (New England Biolabs)
- Nuclease-free dH_2O

Procedure

Reactions were set up according to Table 2.14 and incubated for 24 hours at 16°C. The DNA ligase was subsequently inactivated by incubating the reactions at 65°C for 10 minutes.

Table 2.14. Reaction set up for cloning of *Atcbm33A* into pPICZα-A (generating pPICZα-A/*Atcbm33A*).

Reaction Component	Cloning Reaction	Negative Control
Restriction digested *Atcbm33A*	5.5 µl (approximately 150 ng)	-
Linearized vector, pPICZα-A	4 µl (100 ng)	4 µl
10 x T4 DNA ligase reaction buffer	2 µl	2 µl
T4 DNA ligase	1,5 µl	1,5 µl
Nuclease-free dH_2O	7 µl	13,5 µl

2.11 Transformation of *E. coli*

Materials

- Super Optimal Broth (SOC) medium:
  2g Bacto tryptone
  0.5g Bacto yeast extract
  0.057g NaCl
  0.019g KCl
  0.247g $MgSO_4$
  All chemicals were dissolved in dH_2O to a total volume of 90 ml before autoclaving. After cooling to room temperature 1 ml 2 M glucose (sterile filtered) was added and the medium was stored at 4°C.
- OneShot® TOP10 chemically competent *E. coli* cells (Invitrogen)
- OneShot® BL21 Star (DE3) chemically competent *E. coli* cells (Invitrogen)
- XL10- Gold® Ultracompetent *E. coli* cells (Stratagene)

Procedure

OneShot® TOP10 chemically competent *E. coli* cells were used for plasmid amplification of p-RSET-B constructs, OneShot® BL21 Star (DE3) chemically competent *E. coli* cells were used for periplasmic expression of *Cf*CBM33A-N and *Cf*CBM33B-N while *E. coli* XL10- Gold® cells were used for plasmid amplification of pPICZα-A/*Atcbm33A*.

For transformation, a 2.5 µl ligation-mixture was added to 25 µl competent cells in sterile cell culturing tubes followed by incubation on ice for 30 minutes. After a heat-shock at 42°C for 30 seconds the cells were cooled on ice for 1 minute. To recover the cells after transformation, 250 µl SOC medium heated to room temperature was added and the transformation mixture was incubated for 1 hour at 37°C with shaking at 220 rpm. Following centrifugation at 6000 rpm and 30°C, the cells were resuspended in 100 µl SOC medium and spread out on preheated LB-agar plates supplied with ampicillin. After overnight incubation at 37°C, pre-cultures of clones were made by inoculation of overnight cultures with transformants as described in 2.2.3. Glycerol stocks of the transformed cells were made according to section 2.3.

To verify that the transformation was successful, a PCR using Red Taq DNA Polymerase Master Mix was set up according to section 2.5.2 using the corresponding sequencing primers listed in section 2.1.3.

2.12 Transformation of *P. pastoris*

Transformants of *P. pastoris* were generated by homologous recombination between pPICZα-A/*Atcbm33A* and regions with homology in the genome.

2.12.1 Preparation of electro-competent *P. pastoris*

Materials

- o dH_2O
- o 1 M sorbitol
- o YPD- medium (see section 2.2.1)

Procedure

From a *P. pastoris* X-33 overnight culture, 1 ml was transferred to a 250 ml baffled cell culture flask containing 20 ml YPD- medium. The culture was grown overnight in an incubator at 30°C while shaking at 200 rpm. The next day, the overnight culture was transferred to a 2 liter baffled cell culture flask containing 500 ml YPD medium. After 18 hours (approximately 7 generation times) with incubation at 30°C and shaking at 160 rpm, the culture had reached an OD_{600}= 1.4 (measured with Spectrophotometer, Ultraspec 2100 (GE Healthcare)). The culture was transferred to a 500 ml centrifugal bottle and harvested by a 5 minute centrifugation at 1,500 x g and 0°C using a Sorvall centrifuge RC-5C, with rotor F10S-6*500Y (Thermo scientific). The pellet was resuspended in 500 ml ice-cold (0°C) sterile water, centrifuged as above and resuspended in 250 ml ice-cold water followed by another centrifugation and resuspension in 20 ml ice-cold 1 M sorbitol. The sorbitol is needed to stabilize the osmotic pressure in the cells.

2.12.2 Electroporation of *P. pastoris*

Transformation of pPICZα-A/*Atcbm33A* into *P. pastoris* was performed by electroporation.

Materials

- 1 M sorbitol
- YPDS- plates supplemented with 100 µg/ml Zeocin™ (see section 2.2.2, preheated to 30°C

Procedure

During transformation, all work was carried out on ice and all solutions and materials used were ice-cold (0°C).

Electro-competent *P. pastoris* X-33 cells, approximately $7*10^7$ cells (prepared according to section 2.12.1), were collected by centrifugation and resuspended in 1 ml ice-cold 1 M sorbitol. From the resuspended cells, 80 µl were mixed with approximately 8 µg linearized pPICZα-A/*Atcbm33A* dissolved in 2 µl of nuclease-free water in an ice-cold 2 mm gap electroporation cuvette (Molecular Bioproducts) and incubated for 30 minutes on ice. Transformation was performed by use of a Micopulser (Bio-Rad) giving 2kV for 5.4 milliseconds. A control without added plasmid was treated in the same way. After the pulse, 1 ml ice-cold 1 M sorbitol was added followed by careful mixing by inverting the cuvettes 4-6 times; the cuvette contents were then transferred to falcon tubes. The transformed cells were incubated at 30°C without shaking. After 1 hour incubation, 1 ml YPD was added to both tubes and the incubation was continued for 2 more hours.

For selective growth of transformed cells, the cell suspension was spread on 3 YPDS plates supplemented with 100 µg/ml of Zeocin™ with varying amount of cells on each plate. 500 µl of the control was spread on one YPDS- plate supplemented with Zeocin™, and all plates were wrapped in aluminum foil and incubated at 30°C for 4 days.

2.12.3 Control of transformation

To check if the gene was incorporated in *P. pastoris'* genome, a PCR was conducted using Zymolyase (Seikagaku Corporation) digested cells. Zymolyase is a β-glucanase that breaks down the cell wall, thus lysing the yeast cells and thereby allowing direct PCR on chromosomal DNA.

Materials

- o dH_2O
- o Zymolyase

Procedure

A small part of a colony with transformed *P. pastoris* X-33 cells was dissolved in 100 µl dH_2O containing 0.5 mg/ml zymolyase and incubated at 35°C for 1.5 hours. Following inactivation of the enzyme at 65°C, the lysate was centrifuged for 5 minutes at 8,000 x g. The viscous supernatant was used as template for Phusion™ High Fidelity PCR according to section2.5.1 using the cloning primers 3320_XhoI_SP and 3320_XbaI_ASP (see section 2.1.3). The PCR product was visualized on an agarose gel according to section 2.5.3.

2.13 Protein expression

2.13.1 Cultivation of transformed *E. coli* BL21 for optimal expression of *Cf*CBM33B-N

For optimal expression of *Cf*CBM33A-N and *Cf*CBM33B-N, *E. coli* Bl21 Star™ (DE3) harboring pRSET-B/*Cfcbm33A-N* and pRSET-B/*Cfcbm33B-N* were cultivated in different media (LB, TB, BHI; section 2.2.1) and different cultivation conditions (temperature and length of cultivation) were tested.

2.13.2 Induction of the *lac*-operon by IPTG

Expression of the gene of interest from pRSET-B is controlled by the strong phage T7 promoter (indicated in the vector map in appendix D, Figure D.1) that drives expression of downstream genes, including the multiple cloning site with the inserted gene. The T7 promoter is transcribed by T7 RNA polymerase and the transcription of this polymerase is controlled by an isopropyl β-D-thiogalactoside (IPTG)-inducible promoter. Therefore, for optimal expression of the inserted gene, T7 RNA polymerase production was induced by addition of 0.5 mM IPTG.

Materials

- 1 M IPTG, sterile filtered

Procedure

5 ml overnight culture was transferred to a culture flask containing 300 ml medium supplemented with 50 µg/ml ampicillin followed by incubation at 37°C with shaking at 220 rpm until OD_{600} reached 0.6. OD_{600} was measured using Biophotometer (Eppendorf). Then, gene expression was induced by adding IPTG to a final concentration of 0.5 mM. After cultivation for 4 more hours, the culture was harvested and a periplasmic extract was made according to section 3.4.

The T7 RNA polymerase promoter in pRSET-B/*cbp21*, and therefore also pRSET-B/*Cfcbm33A-N* and pRSET-B/*Cfcbm33A-N*, is leaking and IPTG induction has no effect (results not shown). IPTG induction was thus only used in expression of ChiC that is cloned into a pRSET-B-vector from another origin.

2.13.3 Protein expression in *P. pastoris*; screening for positive transformants.

P. pastoris is a methylotropic yeast, which implies that it is capable of using methanol as its sole carbon source, in a process that is carried out in the peroxisomes by alcohol oxidases and uses molecular oxygen. Two genes code for these alcohol oxidases, *aox1* and *aox2,* both tightly regulated and induced by methanol. Gene products of *AOX1* accounts for the majority of alcohol oxidase activity in the cell and, since the gene of interest is inserted downstream of the methanol regulated *AOX1* promoter, the gene of interest is also produced in large quantities subsequent to methanol induction (Macauley-Patrick et al. 2005). In *Pichia,* heterologously expressed proteins

can end up either in the intracellular space or be secreted. The major advantage by secretion is that *P. pastoris* secretes very low levels of native proteins, meaning that secreted heterologous proteins are relatively pure

Prior to fermentation, a screening for transformants with the *Atcbm33A*^{His} incorporated (verified by PCR, see section 2.5.1, and agarose gel electrophoresis, see section 2.5.3) was performed. This was done by growing several transformants on methanol as the sole carbon source and then analyzing the medium for the secreted protein of interest.

Materials

- BMGY, see section 2.2.1
- BMMY, see section 2.2.1.72.2.1
- 100 % (v/v) Methanol, sterile- filtered

Procedure

Colonies with transformed *P. pastoris* X-33 (as verified by PCR) were inoculated in 10 ml BMGY in 50 ml tubes and incubated overnight at 30°C, with shaking at 250 rpm, until the OD_{600} was approximately 3 (measured with Ultraspec 2100 pro (GE Healthcare)). *P. pastoris* X-33 harbouring a pPICZα-A vector with the GH61- gene from *Aspergillus nidulans* (*gh61)* integrated was used as a positive control, whereas the negative control was native *P. pastoris* X-33.

The next day, cells were harvested by centrifugation at 3,000 x g for 5 minutes at room temperature using Centrifuge 5430R (Eppendorf). The supernatant was decanted before the pellet was resuspended in 1 ml BMMY and used to inoculate 25 ml BMMY containing 0.5% (v/v) methanol for protein expression. The culture was incubated at 22°C for 96 hours with shaking at 150 rpm. Each 24th hour methanol was added to a final concentration of 0.5 % (v/v) to compensate for consumption and, to some extent, evaporation. Both at the start of the incubation and after 96 hours, 0.5 ml of the expression-cultures was collected and after centrifugation at 5,000 x g the supernatants were concentrated 10 times using Amicon Ultraspin columns containing a cellulose filter with a 10kDa cut-off (Millipore). An SDS-PAGE was then performed according to section 2.18 to identify transformants expressing *At*CBM33A^{His}.

2.13.4 Large scale expression of *At*CBM33A^{His}

For large scale expression, materials and procedures were as in 2.13.3 with some exceptions. The overnight pre-culture was set up in 2 liter baffled cell culture flasks containing 0.5 liter BMGY and inoculation of the production culture was performed by transferring the concentrated cells from the overnight culture (in 10 ml BMMY) to 2 liter baffled cell culture flasks containing 1 liter BMMY. Induction was carried out for 96 hours at 22°C with shaking at 150 rpm. Each 24th hour, starting at 0 hours of induction, the OD_{600} of the culture was measured and the expression culture was added methanol to 5% final concentration and 50 ml BMMY to compensate for consumption and evaporation. Also, each 24th hour, 0.5 ml of the expression culture was collected and, following centrifugation at 5,000 x g, the supernatant was stored at -20°C. After 96 hours of incubation, all supernatants collected were concentrated 10 times using Amicon Ultraspin columns containing cellulose filter with a 10kDa cut-off (Millipore) and analyzed by SDS-PAGE according to section 2.18.

2.13.5 Fermentation of *P. pastoris*

P. pastoris is well-suited for fermentative growth and has the ability to reach high cell density during fermentation. This may improve overall protein yields. Fermentation of *P. pastoris* for protein expression in a bioreactor has three stages: (i) a glycerol growth phase for generation of biomass; (ii) a growth phase with a growth-limiting feed-rate of glycerol, for controlled biomass generation and derepression of the alcohol oxidase promoter, *AOX1*; and (iii) a methanol feed phase, inducing expression of the *AOX1*-controlled genes. The fermentation was carried out in a 5 liter Biostat B plus bioreactor equipped with an additional feed pump, gas mixer, dissolved oxygen tension polarographic electrode and a water cooler. During fermentation, parameters listed in Table 2.15 were monitored and controlled.

Table 2.15. Parameters that should be monitored and controlled during fermentation according to Invitrogen's *Pichia*-fermentation protocol.

Parameter	Reason
Temperature (28.0°C when grown on glycerol and 22.0°C when grown on methanol)	Growth above 32°C is detrimental to protein expression
Dissolved oxygen (>20%)	*Pichia* needs oxygen to metabolize glycerol and methanol
pH (5.0-6.0)	Important when secreting protein into the medium and for optimal growth
Agitation (500 to 1500 rpm)	Maximizes oxygen concentration in the medium
Aeration (0.1 to 1.0 vvm* for glass fermenters)	Maximizes oxygen concentration in the medium
Antifoam (minimum needed to eliminate foam)	Excess foam may cause denaturation of your secreted protein and it also reduces headspace
Carbon source (variable rate)	Must be able to add different carbon sources at different rates during the course of fermentation

* Volume of oxygen (litres) per volume of fermentation culture (litres) per minute

The dissolved oxygen concentration is the relative per cent of oxygen in the medium (O_2-saturated medium has a concentration of dissolved oxygen of 100%). *P. pastoris* consumes oxygen during growth, whether it grows on methanol or glycerol, thereby keeping the dissolved oxygen concentration low. Thus, the dissolved oxygen-values give information about the health and state of the culture. When switching from glycerol to methanol as carbon source, close monitoring of the dissolved oxygen-values is of great importance to determine whether all the glycerol is consumed by the culture prior to induction. Shutting off the carbon source should cause the culture to decrease its metabolic rate, and the dissolved oxygen to rise (spike).

Materials

- PTM, trace salts:

  3.0 g Copper (II) sulphate- $5H_2O$

  0.04g Sodium iodide

  1.50 g Manganese sulphate- H_2O

  0.1 g Sodium molybdate-2 H_2O

  0.01 g Boric acid

  0.25 g Cobalt chloride

  10.0 g Zinc chloride

  37.5 g Ferrous sulphate- $7H_2O$

  0.1 ml Biotin 0.02 % (w/v)

  2.5 ml Sulphuric acid

  All ingredients were thoroughly mixed and dissolved in dH_2O to a final volume of 0.5 liters before the solution was filter- sterilized. PTM is light sensitive and was stored in the dark at room temperature.

- 500 ml Glycerol feed:

  494.2 ml 50 % (w/v) glycerol

  4.8 ml PTM

  1 ml 0.02 % (w/v) biotin

  The listed chemicals were carefully mixed and sterile-filtered through a filter with 0.45µm pores

- 1 liter Methanol feed:

  986 ml 100 % (v/v) methanol

  12 ml PTM

  2 ml 0.02% (w/v) biotin

  100% (v/v) methanol was autoclaved and PTM and biotin were filter-sterilized into the methanol.

- Fermentation basal salts medium:

  18.2 g Potassium sulphate

  3.0 g Magnesium sulphate- 7 H_2O

  4.13 g Potassium hydroxide

  26.7 ml Phosphoric acid, 85% (w/v)

  40.0 g Glycerol

  0.93 g Calcium sulphate

All chemicals were dissolved in listed order in dH_2O to a total volume of 2 liters and autoclaved.

- 0.02 % (w/v) biotin, filter-sterilized through a 0.45 µm pored filter.
- Buffered complex glycerol medium (BMGY, see section 2.2.1)
- Buffered complex methanol medium (BMMY, see section 2.2.1)

2.13.5.1 Procedure

Both the thermometer and the dissolved oxygen tension polarographic electrode were calibrated and disinfected in 96% (v/v) ethanol. Air supply, gas mixer and water cooler were checked before autoclaving. All connection-tubes were clogged and wrapped in foil and, along with the Biostat B plus 5L fermenter (Sartorius Stedim biotech) containing 2 liter fermentation basal salts medium, sterilized by autoclaving. All equipment and solutions used in the fermentation were sterilized either by filter-sterilization or autoclavation. After connecting all the tube-connections between the devices and the steering console, the fermenter was cooled to 28°C. Prior to inoculation, 1 ml 0.02 % (w/v) biotin and 8.7 ml PTM trace salts were added aseptically to the fermenter. The *P. pastoris* transformant showing best expression of *At*CBM33A during screening (2.13.4) was grown overnight at 30°C in 1 liter BMGY. The overnight culture was centrifuged and resuspended in 250 ml BMGY, having an OD_{600}=50, and was used for inoculation of the fermenter. Growth on glycerol was performed at 28°C; the temperature being maintained by water cooling. Continuous addition of 28 % aqueous ammonia served both as a nitrogen source, and to maintain the pH at 5.5 throughout the whole fermentation. After complete consumption of glycerol from the BMGY medium, indicated by an increase in the dissolved oxygen to 100 %, a glycerol feed was initiated. Glycerol feeding was carried out at 28°C with automated control of parameters listed in Table 2.15 for about 4 hours until the culture had reached a cellular yield of 160 g/liter, measured by wet weight of cells from 1 ml cell culture.

Before induction of the *AOX1* promoter, the glycerol in the medium was completely consumed (dissolved oxygen concentration reached ~100%) and the temperature was decreased to 22°C. Induction was initiated with a methanol feed at slow rate to adapt the culture to growth on methanol. The rate was then slowly increased to a final feed rate of 11 g/h/L. Fermentation on methanol was continued for 4 days. After 4 days, the cells were harvested and the supernatant, containing secreted proteins, was separated from the cells by centrifugation at 12,000 x g for 30 minutes at 4°C. During harvesting, cells were kept at 4°C at all times to hinder cell lysis and

proteolysis. The supernatant was sterilized by filtration through a 0.22 µm filter (Millipore) and then concentrated by use of Pellicon ultra-filtration (see section 2.13.6).

2.13.6 Ultrafiltration

The Pellicon ultra-filtration system from Millipore was used for concentrating secreted proteins in the fermentation broth. The system comprises a peristaltic pump and a Pellicon 2 cassette filter holder with a Biomax 10 kDa cut-off membrane- filter. The system uses tangential flow filtration (see Figure 2.1) where the fluid is pumped along the surface of the membrane. The material that does not pass through the membrane (retentate, which contains the protein of interest) is recycled while the permeate, the material that passes through the membrane, is thrown away. This results in an up-concentration of the macromolecules retained in the retentate.

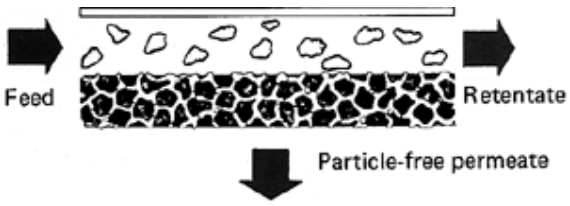


Figure 2.1. Principle of tangential flow filtration. The fluid flows over and parallel to the semipermeable membrane and under pressure, some of the fluid is forced through the filter. The Figure is from "the free dictionary" (http://encyclopedia2.thefreedictionary.com/tangential+flow+filtration).

Materials

- Pellicon cassette acrylic holder and assembly (Millipore) with:
  - Pellicon 2 Cassette filter
  - Biomax® Ultrafiltration Membrane
- Peristaltic pump, 4-13 litre/min
- dH_2O
- 5 L 0.1 M NaOH, room temperature
- 300 ml 10 mM MES buffer pH 6.5
- 1 litre 20 % ethanol

Procedure

All tubes in the pump head were fastened, all three valves opened and the tube ends placed in dH_2O. After washing the filter with 2 liters of dH_2O, 2 liter of 0.1 M NaOH was circulated through the filter to remove any bound proteins. Following this cleaning step, 5 liter dH_2O was used to neutralize the pH in the system (checked using pH indicator paper). Both feed and retentate tubing were placed in the filtrated fermentation supernatant and the permeate outlet placed in a collection flask. The retentate side was slowly clamped so the sample was circulating through the filter system and thereby concentrated by the use of tangential flow. After concentrating the supernatant 10 times (to a total volume of 500 ml), the filter was washed with 100 ml buffer by recirculation, to extract all remaining proteins. To clean the filter, 0.1 M NaOH was run through the system, followed by a neutralization of the pH with dH2O. Finally, the system was washed with 20 % ethanol and stored.

2.14 Periplasmic extracts of *E. coli*

All *Cfcbm33*-genes were cloned into pRSET-B subsequent to a CBP21 signal peptide. This signal peptide drives the translocation of all expressed *Cf*CBM33 proteins to the periplasmic space in *E. coli* and can be extracted by lysing the cells using cold osmotic shock as described below.

Materials

- Spheroplast buffer:
  - 50 µl 0.5 M EDTA, pH 8.0
  - 5 ml 1 M Tris, pH 8.0
  - 8.55 g sucrose
  - 125 µl 50 mM PMSF
  - All ingredients were carefully mixed and dissolved in dH_2O to a final volume of 50 ml, and kept on ice.
- Sterile dH_2O
- 20 mM $MgCl_2$

Procedure

150 ml overnight culture was transferred to 250 ml centrifuge bottles and placed on ice for 20 minutes before centrifuging at 5500 x g for 10 minutes, using Beckman coulter Avanti J-25 centrifuge with a JA-14 rotor, at 4°C. After centrifugation the pellet was resuspended in 15 ml ice-cold spheroplast buffer and kept on ice for 5 minutes. After another centrifugation, this time at 10,000 x g, the pellet was resuspended in 12.5 ml ice-cold dH_2O and left on ice for 45 seconds. 625 µl 20 mM $MgCl_2$ was added to the suspension, followed by another centrifugation as above. The final supernatant, the periplasmic extract, was filtered through a 0.22 µm sterile filter and stored at 4°C.

2.15 Protein purification

For protein characterization it is important to work with purified proteins. In this study, different techniques for protein separation and purification were used, depending on the protein. Chitinases (ChiA, ChiB, ChiC from *S. marcescens*) and the chitin monooxygenase CBP21from *S. marcescens,* were all purified by use of chitin beads (affinity chromatography), while *Cf*CBM33A-N and *Cf*CBM33B-N was purified by a combination of ion exchange chromatography and gel-filtration. His-tagged *At*CBM33A^{His} was purified by the use of a His-trap column.

2.15.1 Ion Exchange Chromatography

Ionic exchange chromatography is frequently used for separation and purification of proteins, polypeptides, nucleic acids and other charged biomolecules. The technique separates proteins based on their surface ionic charge, using beads modified with positively or negatively charged groups. Proteins with low binding-affinity to the column material are washed off the column by using low-salt buffer. To elute the proteins that bind strongly to the column, a high-salt buffer is used. The salt masks the charged groups on the column material allowing the protein to be eluted. Since each protein has a different charge on the surface, proteins are eluted at varying salt concentrations and may thus be separated in different elution fractions.

Materials

- Purification of *Cf*CBM33A-N (pI=5.1) and *Cf*CBM33B-N (pI= 5.2)
  - Binding buffer: 50 mM Tris- HCl pH 7.5
  - Elution buffer: 1 M NaCl in 50 mM Tris- HCl pH 7.5

  All buffers were filtrated through a 0.45 µm sterile- filter
- 10 mM Tris-HCl pH 7.5

Procedure

The ion exchange column, HiTrap™ DEAE FF, 5 ml (GE Healthcare), was connected to an Äkta purifier chromatographic system (GE Healthcare), a fully automated liquid chromatography system, and washed with elution buffer to remove any contamination on the column and equilibrated using the binding buffer. The pH of the sample was adjusted to 7.5 by addition of 50 mM Tris-HCl and the sample (50 ml) was loaded onto the column. To remove unbound proteins, a wash with 2 column volumes of binding buffer was performed. The protein of interest was eluted by a gradient from 0 % to 50 % elution buffer, during 200 minutes at 4 ml/min flow. Eluted proteins were detected by online monitoring absorption at 280 nm and collected using a fraction collector (1 ml per fraction). Fractions were analyzed using SDS-PAGE (see section 3.18). Fractions containing the (partially purified) protein of interest were pooled and concentrated to 1 ml using Amicon Ultra-15 Centrifugal Filter Units with a 3 kDa cut-off cellulose membrane (Millipore), while at the same time changing the buffer to 10 mM Tris-HCl pH 7.5. Further purification was done using size exclusion chromatography (see section 2.15.2).

2.15.2 Size Exclusion Chromatography

Size exclusion chromatography separates proteins on the basis of size and shape. The column matrix is composed of beads with pores in different sizes, where small proteins enter all the pores, thereby using longer time passing through the column and eluting after the larger proteins.

Materials

- Running buffer (50 mM Tris-HCl, pH 7.5; 200 mM NaCl):
  50 ml 1 M Tris-HCl, pH 8.0
  11.6 g NaCl

Dissolved in 900 ml dH_2O, regulated to pH 7.5 with 6 M HCl before the volume was brought up to 1 liter by addition of dH_2O.

Procedure

A HiLoad 16/60 Superdex G-75 column (GE Healthcare) was connected to an Äkta purifier chromatographic system (GE Healthcare). Concentrated protein (1 ml) obtained from the first step of purification (ion exchange chromatography; see section 2.15.1) was applied through a 2 ml loading loop at 0.3 ml/min flow rate, followed by application of 3 column volumes of running buffer. The eluate was collected in fractions of 5 ml. Fractions putatively containing the protein of interest were analyzed bySDS-PAGE according to section 2.18 to visualize the purity of the protein and compared with the periplasmic extract. Since *Cf*CBM33B-N was clearly overexpressed and yielded a highly visible band in the starting material, we considered it sufficient to determine protein identity in the chromatographic fractions by comparison with the periplasmic extract.

2.15.3 Protein purification by immobilized metal ion affinity chromatography

HisTrap™ HP is a ready-to-use column for preparative purification of His-tagged recombinant proteins by immobilized metal ion affinity chromatography (IMAC). The column is pre-packed with charged Ni Sepharose™ High Performance beads. Histidines form complexes with nickel ions and the $(His)_6$- tag on the N- terminus of *At*CBM33A^{His} will bind strongly to the column material.

Materials

- Binding buffer (10 mM Hepes, 15 mM imidazole, 500 mM NaCl)

  2.38 g Hepes

  1.02g Imidazole

  29.2 g NaCl

  117.6 ml 85% (w/v) glycerol

  All chemicals were mixed in dH_2O to a final volume of 1 liter. The pH was adjusted to 7.5 with 6 M HCl and the buffer was sterile-filtered through a 0.22 µm membrane and degassed by sonication for 30 minutes.

- Elution buffer (10 mM Hepes, 15 mM Imidazole, 500 mM NaCl, 10 % (w/v) glycerol):

  2.38 g Hepes

  27.2 g Imidazole

  29.2 g NaCl

  117.6 ml 85% (w/v) glycerol

  All chemicals were mixed and dissolved in dH_2O to a total volume of 1 liter. The pH was adjusted to 7.5 with 6 M HCl and the buffer was sterile-filtered through a 0.22 µm membrane and de-gassed by sonication for 30 minutes.
- Stripping buffer (0,02 mM sodium phosphate, 500 mM NaCl, 50 mM EDTA):

  3.28 g sodium phosphate

  29.2 g NaCl

  100 ml 0.5 M EDTA, pH 8.0

  All ingredients were mixed and dissolved in 500 ml dH_2O and the pH was adjusted to 7.4 with 6 M HCl before dH_2O was added to a final volume of 1 liter. The buffer was sterile-filtered through a 0.22 µm membrane and degassed by sonication for 30 minutes.
- dH_2O
- 1 M NaCl
- 20 % Ethanol

Procedure

The pH of 30 ml of the secreted proteins, concentrated according to section 2.13.6, was adjusted to 7.5 through a buffer exchange with 3 x the sample volume of binding buffer using an Amicon Ultra-15 Centrifugal Filter Units with a 10 kDa cut-off (Millipore). Prior to loading the sample was filtered through a 0.22 µm filter, to remove any particles that could clog the column.

A His-Trap™ HP 1 ml (GE Healthcare) pre-packed column was connected to an Äkta purifier chromatographic system (GE Healthcare). The column was prepared for binding by washing with 5 column volumes of distilled water before equilibration using 5 column volumes of binding buffer. The sample 30 ml was then loaded onto the column and binding buffer was run until the baseline was stable at a low UV signal (i.e. all non-binding proteins had passed through the column). The binding buffer contained imidazole at a low concentration to prevent unspecific binding. The protein of interest was eluted by applying a linear gradient of 20 column volumes from 0% to 100% elution buffer. His-tagged protein was eluted at approximately 10 % of elution buffer and collected in fractions of 1 ml. Elution and fraction collection were

continued until the absorbance at 280 nm had reached the baseline. After elution, the column was washed with 20 column volumes of 1 M NaCl to remove all bound proteins, followed by washing 20 column volumes of dH_2O and application of 20 % ethanol for storage. The purified protein was analyzed by SDS-PAGE according to section 2.18 and stored at 4°C.

If the column was to be used for purification of other proteins, a "stripping" step was performed to remove all bound molecules. The column was washed with 10 column volumes of stripping buffer, followed by 10 column volumes of binding buffer and, finally, 10 column volumes of dH_2O. The column was recharged using 0.5 ml 0.1 M $NiSO_4$ solution followed by a 5 column volume wash with dH_2O and application of 20 % ethanol before storage.

2.15.4 Chitin-affinity chromatography

ChiA, ChiB, ChiC and CBP21 all bind specifically to chitin and were purified by affinity chromatography using chitin beads as a chromatographic material.

Materials

- Binding buffer:
  - In purification of CBP21:
    132 g Ammonium sulphate
    Dissolved in 50 mM Tris-HCl, pH 8.0, yielding a 1 M solution
  - In purification of ChiA, ChiB and ChiC:
    50 mM Tris-HCl, pH 8.0
- Elution buffer:
  20 mM Acetic acid
- 20% (v/v) ethanol
- Tris-HCl pH 7.0, 1 M and 5 mM
- Chitin beads (New England Biolabs)

Procedure

For optimal CBP21 binding to the chitin beads, the periplasmic extract was adjusted to 1 M ammonium sulphate (by carefully adding 4 M ammonium sulphate) and 50 mM Tris-HCl pH 8.0. Periplasmic extracts containing ChiA, -B or C were all adjusted to 20 mM Tris-HCl pH 8.0

prior to purification. Aside from the differences in binding buffer, purifications of CBP21, ChiA, ChiB and ChiC were identical (described below).

To prepare the column, 20 ml of chitin bead suspension was transferred to a 10 cm glass econo-column with an inner diameter of 1.5 cm, yielding a bed volume of approximately 8 ml. The column was inserted into a BioLogic LP (BioRad) system and washed with 20% ethanolat a flow of 2.5 ml/min before equilibrating the column using binding buffer. An UV-reader was used for monitoring the baseline. When the baseline was stable, the UV was reset to zero and adjusted periplasmic extract was loaded onto the column at a flow rate of 1 ml/min. For removal of any unbound proteins, the system was washed with binding buffer until the baseline had stabilized again (approximately 45 minutes).

To elute the bound protein, elution buffer was applied using a flow rate of 1 ml/min and fractions were collected. When all the bound protein was eluted (and collected) the column was washed, first with binding buffer, then with 20% ethanol. The washed column was stored in 20 % ethanol at 4°C. Fractions containing purified protein were pooled and adjusted to pH 7 by adding 1 M Tris-HCl pH 7.0 to a final concentration of 5 mM, followed by a wash with 3 volumes of 5 mM Tris-HCl pH 7.0 and then concentrated using an Amicon Ultra, cellulose filter with 10 kDa cut-off (Millipore). Protein concentration was measured using theBradford method (see section 2.16) and SDS-PAGE (see section 2.18) was performed to verify the purification.

2.16 Protein concentration measurement

Quick Start ™ Bradford Protein Assay is a method for determining the concentration of proteins in a sample, which is based on the binding of Coomassie Brilliant Blue G-250 dye to proteins (Bradford 1976). When the dye binds to protein it is converted from the cationic form (A_{max} = 470 nm) to the stable unprotonated blue form (A_{max} = 595 nm), which can be detected using a spectrophotometer. Protein concentration is calculated by using a standard curve.

Materials

- 5x Dye Reagent, Protein Assay (BioRad)
- BSA Standard (2 mg/ml)
- Polystyrene cuvettes, 1 ml (Brand)
- Sample buffer
- Bovine serum albumin (BSA) for calibration

Procedure

From each sample, 2 µl concentrated protein was diluted in 798µl sample buffer and mixed with 200 µl 5x Dye Reagent. After 5 minutes incubation, the sample absorbance at 595 nm was measured by using the Bradford micro program.

The spectrophotometer, Bio Photometer (Eppendorf), was calibrated with bovine serum albumin (BSA) as standard with the concentrations 2, 1.5, 1, 0.75, 0.5, 0.25 and 0.125 mg/ml ("Bradford micro"). Triplicates of the standards incubated in 5x Dye Reagent for 5 minutes were analyzed and the standard curve used had a standard error less than 10 %.

When measuring samples, a reference containing 800 µl sample buffer was mixed with 200 µl 5x Dye Reagent and incubated for 5 minutes. After the photometer was zeroed using the reference, samples were analyzed in triplicates. The sample concentration was calculated as the mean value of the three replicates.

2.17 Deglycosylation of proteins using EndoH

*At*CBM33A was expressed and secreted by *P. pastoris.* For optimal expression of eukaryotic genes, the host's ability to perform posttranslational modifications is important. For example, some eukaryotic proteins become unstable if they are not glycosylated. *P. pastoris* sometimes hyperglycosylates recombinant proteins, which may have negative effects, including marking them for degradation. Although glycosilation may be necessary to obtain stable proteins, it is not *a priori* known how much glycosilation is needed and hyperglycosilation is not desirable. Therefore, and since improper glycosylation may interfere with enzyme activity and some biochemical assay techniques (see below), it is common to remove most of the glycosylations.

One of the post-translational processing performed in *P. pastoris* is N-glycosylation of the amide nitrogen of asparagine residues in the Asn-Xaa-Thr/Ser consensus sequence. The bacterium also O-glycosylates some proteins, however, the bacterium does not seem to have a preferred amino acid for O glycosylation (no consensus glycosylation sequence is known) and is thus difficult to predict (Cregg et al. 2000). N-glycosylations, on the other hand can easily be predicted by using web-based prediction tools like NetNGlyc (http://www.cbs.dtu.dk/services/NetNGlyc/).

The *At*CBM33A-sequence contains two possible N-glycosylation sites (found by using NetNGLyc) and if glycosylation occurs, the protein may not be visible on the gel as one single

band, and or the molecular weight may be increased in a somewhat unpredictable manner. Endoglucosidase H (EndoH), with a molecular weight of 29 kDa, is a recombinant glycosidase, which cleaves asparagine-linked mannose rich oligosaccharides from N-linked glycoproteins. It cleaves the glycosidic bond in the diacetylchitobiose core of the oligosaccharide, between two N-acetylglucosamine (GlcNAc) subunits, leaving one GlcNAc remaining on the asparagine residue (Figure 3.3).

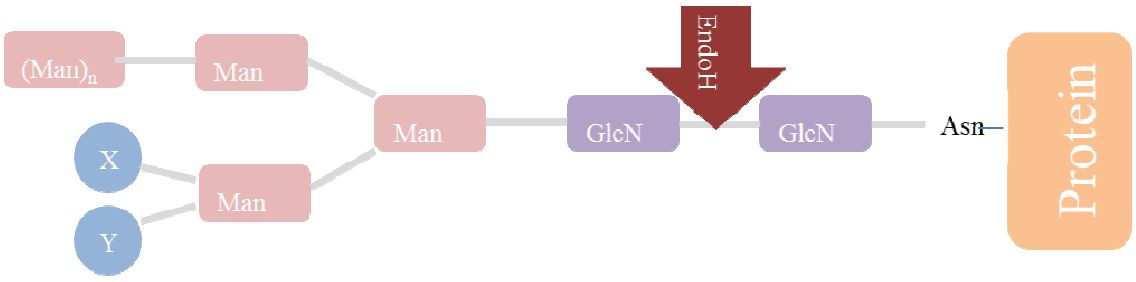


Figure 3.3. Action of EndoH. Endoglucosidase H (EndoH) cleaves the glucan moiety covalently bound to the protein between two core N-acetylglucosamine (GlcNAc) units in high-mannose structures, n= 2-150, X=$(Man)_{1-2}$ and Y=H and in low-mannose structures, n=2, X and/or Y = AcNeu-Gal-GlcNAc.

Materials

- o dH_2O
- o 10 x G5 Reaction Buffer (New England Biolabs)
- o EndoH (New England Biolabs)
- o Amicon Ultra 0.5 ml centrifuge tubes with cellulose filter and a cut-off of 10 kDa (Millipore)

Procedure

After purification, the protein did not appear as a clear band on the SDS-PAGE gel (see Figure 3.11), possibly because of glycosylations of the protein. EndoH treatment was performed on 10 µl of the concentrated purifie protein mixed with 2 µl 10x G5 Reaction buffer, 1 µl EndoH and 7 µl dH_2O. Following 2 hour incubation at 37°C, the EndoH was inactivated by incubation at 65°C for 20 min. The samples were analysed by SDS-PAGE according to section 2.18.

2.18 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)

Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) is a widely used technique for separation of proteins according to their mass. The addition of anionic detergents like SDS or LDS (lithium dodecyl sulphate) and a reducing agent like β-mercaptoethanol (for reducing disulphide bridges) in the sample denatures the protein and provides each protein with a uniform negative charge. The proteins are then separated based on size only by using electrophoresis and can be visualized by protein specific staining, i.e. with Coomassie brilliant blue. A molecular marker containing proteins with known masses is used to estimate the mass of the sample proteins.

Materials

- NuPage® LDS sample buffer 4x (Invitrogen)
- NuPage Sample reducing agent 10x (Invitrogen)
- 20x MES SDS Running buffer pH 7.3:
  97.6 g MES
  60.6 g Tris Base
  10 g SDS
  3.0 g EDTA
  All chemicals were dissolved in dH_2O added to a final volume of 500 ml. Before use, the buffer was diluted 20 times in dH_2O.
- 1x Wash solution:
  23.5 ml 85% (w/v) Phosphoric acid
  Diluted in 976.5 ml dH_2O
- 1x fixer solution:
  300 ml 96% (v/v) Ethanol
  23.5 ml 85% w/v) Phosphoric acid
  Diluted in dH_2O to a final volume of 1 liter
- Coomassie staining solution:
  200 ml 96% (v/v) Ethanol
  117.6 ml 85% (w/v) Phosphoric acid
  100 g Ammonium sulphate
  1.2 g Coomassie Brilliant Blue R 250

All ingredients were dissolved in dH_2O to a final volume of 1 litre.

- Bench Mark™ Protein Ladder (Invitrogen)

Procedure

To 18 µl protein sample, 2 µl NuPage® LDS sample buffer was added and the sample was boiled for 10 minutes. A NuPage® polyacrylamide gel, Bis-Tris 10 % (Invitrogen) was installed in the XCell SureLock™ Mini-Cell Electrophoresis System (Invitrogen) and the chamber was filled with 1x MES SDS running buffer. In the first well, 5 µl protein ladder was applied, followed by application of 15 µl of each sample in the remaining wells. After running the gel using a Power supply (BioRad) for 35 minutes at 200V, the gel was released from the plastic plates and fixed in 1x fixer solution for 1 hour and washed in 1x washing solution for 2 x 10 min. For visualization of proteins, the gel was treated with the Coomassie Brilliant Blue R-250 staining solution for 2 hours. Coomassie blue binds non-specifically to all proteins. Destaining of the gel was done by washing in dH_2O for 3 hours.

2.19 Matrix-Assisted Laser Desorption and Ionization Time of Flight mass spectrometry (MALDI-TOF MS)

MALDI-TOF MS is a non-quantitative analysis method used to determine and identify the molecular weight of small molecules and proteins. The method implies that the analytes are ionized and brought into gas phase in vacuum, using a laser. Ionized molecules are accelerated by a strong magnetic field through an analysing tube with a mass analyser in the end, with ions being separated based on their mass to charge ratios (*m/z)*. The Time-of- flight (TOF) analyser measures the time the ions use to reach the detector, reflecting the molecular mass of the ions. Smaller ions travel at the highest speed (Walsh 2002).

2.19.1 In-gel trypsin digestion of Coomassie-stained protein spots

MALDI-TOF MS is a convenient method for identifying or verifying proteins separated by SDS-PAGE. The most common way of analysis is to excise the spot/band of interest in the SDS-PAGE gel and subject the gel piece to in-gel trypsination in order to release peptides that can be analysed by the mass spectrometer (see section 2.19.2).

Materials

- 40% (v/v) Ethanol
- Acetonitrile, 100% (v/v) and 50% (v/v)
- 100 mM NH_4HCO_3
- 10 mM DL-Dihiothreitol (DTT)-solution:
  - 15.42 mg DTT dissolved in 100 µl 100 mM NH_4HCO_3
- 55 mM iodoacetamide:
  - 1.5 mg iodoacetamide dissolved in 150 µl 100 mM NH_4HCO_3
- Trypsin solution (12.5 ng/µl final concentration):
  - 10 µl sequencing grade modified trypsin (0.1 µg/µlin 10 mM HCl) diluted in 70 µl 25 mM NH_4HCO_3

Procedure

The protein of interest was digested enzymatically (e.g. by trypsin) prior to MALDI- TOF MS analyzes. All chemicals used had a purity grade for HPLC and the procedure was carried out in a laminar air flow (LAF)-bench to keep a dust-free environment.

The desired protein-band was cut out of the SDS-gel using a clean scalpel and transferred to a pure micro centrifuge tube. The gel piece was distained by washing in 100 µl 40 % (v/v) Ethanol. After complete removal of the ethanol by suction, 300 µl 100 % (v/v) acetonitrile was added, and the sample was incubated for 30 minutes with vigorous shaking to dehydrate the gel. After completely removing the acetonitrile, 50 µl DTT-solution was added, and the sample was incubated at 56°C for 45 minutes. DTT reduces disulphide-bridges in proteins. The DTT-solution was discarded and replaced by 50 µl iodoacetamide-solution. The sample was incubated for 30 minutes with vigorous shaking; it was wrapped in foil to protect iodacetamine from exposure to light. Iodoacetamide is an alkyl acting agent used to modify reduced cysteine residues thus hindering reformation of disulphide bonds. Iodacetamine was removed, and the gel washed in 50 % (v/v) acetonitrile for 10 minutes with shaking. The 50 % (v/v) acetonitrile was then replaced by 100 % (v/v) acetonitrile and the gel was incubated for 15 minutes with vigorous shaking for complete dehydration. After removing all acetonitrile, 10 µl trypsin-solution was added and the sample incubated on ice. Trypsin is a serine protease thatcleaves peptide chains at the carboxyl site of the amino acids lysine or arginine. After 45 minutes, 10 µl 25 mM NH_4HCO_3 was added to the sample and the sample incubated overnight at 37°C. The next day, the sample was centrifuged and the supernatant containing tryptic peptides was transferred to a clean tube and stored at -20°C.

2.19.2 Protein identification by MALDI-TOF MS/MS

Materials

- Matrix stock solution:
  100 µg α-cyano-hydroxy-cinnamic acid (CHCA)
  700 µl 100 % (v/v) Acetonitrile
  100 µl 1 % (w/v) Trifluoro acetic acid (TFA)
  200 µl dH_2O
  All ingredients were mixed by thoroughly vortexing.
- Acetonitrile, 100 % (v/v)
- TFA, 0.5 % and 1 % (v/v)
- MALDI-TOF/TOF Mass spectrometer, Ultraflex

Procedure

Prior to use, the MALDI sample plate was washed with dH_2O, dried off with soft tissues and sonicated in 50 % methanol in a water bath for 15 minutes to remove all bound particles. The methanol was washed off with dH_2O and the plate was cleaned with small amounts of 100 % acetone. Finally, the plate was dried with a soft tissue.

For protein identification, 1µl trypsin digested sample (see section 2.19.1) was applied to a MALDI plate and left to dry before 1 µl photoactive matrix, obtained by diluting the stock solution 20 times in freshly made 90 % acetonitrile and 0.1 % TFA, was applied on top of the sample and air-dried. Trypsin digested β-lactoglobulin with known tryptic fragments was used as standard and for this purpose 0.5 µl of a 5 pmol/µl solution of trypsinated β-lactoglobulin was applied to the plate and dried. Then 0.5 µl matrix was added on top of the standard and left to dry. The plate was inserted into the MALDI-TOF and a vacuum was introduced. All instrumental analysis was performed using the program Flexcontrol version 3.3. The programme Biotools was used for interpretation of the data from MS and tandem mass spectrometry (MS/MS) analysis. MS/MS employs multiple MS analysis stages with a fragmentation in between the steps.

First a simple MS-analysis was performed using 30% of the laser intensity to irradiate the samples, resulting in ionized polypeptides that were accelerated through the magnetic field and captured by the TOF analyser. A computationally generated database with predicted peptide-

fragments from trypsin treated proteins was used to search for peptide identification. If no identical sequences were retrieved during MS, MS/MS was used to generate sequence information for the peptides. The largest peak from the MS analysis was captured, and the sample was irradiated again with higher laser intensity to fragment the peptides into a series short oligopeptides from which sequence information could be derived using the Biotools programme.

2.19.3 MALDI-TOF MS analysis for testing enzyme activity

The activity of *Cf*CBM33B-N towards different carbohydrate substrates was analysed using MALDI-TOF MS for product analysis. The different reaction conditions tested are summarized in Table 3.2.

Materials

- Matrix:

  4.5 mg DHB (2,5-Dihydroxybenzoic acid)

  150 µl acetonitrile

  350 µl dH_2O

  DHB and acetonitrile were vortexed until the DHB was completely dissolved before dH_2O was added. The DHB solution was stored at $4^{\circ}C$ for no more than a week.
- MALDI target- plate
- Substrates, all substrates listed in section 2.1.2
- Buffers:
  - 100 mM Sodium acetate, pH 3.0, 4.0 and 5.0
  - 100 mM Bis-Tris, pH 6.0
  - 100 mM Tris HCl, pH 7.0, 7.5, 8.0 and 9.0
- 1 mM copper chloride
- MALDI-TOF/TOF Mass spectrometer, Ultraflex

Procedure

All reactions were set up with 1 mM reducing agent, 20 mM buffer, 1µM enzyme and 2mg/ml substrate (except for β-chitin^{E} which was tested at 1 mg/ml) in cryo-tubes and incubated horizontally in a shaker for 16 hours at 37°C with shaking at 220 rpm. Different reaction

conditions were tested: different substrates, different reaction pH, different buffers, with or without 1 µM copper^{2+} in the reaction, and different reducing agents.

For every reaction set up, a negative control, i.e. a reaction without added enzyme was set-up to account for the presence of carbohydrate products in the substrate solution. A positive control was generated by incubating β-chitin^{A} with CBP21, in a reaction with 1 mM ascorbic acid as reducing agent in Bis-Tris buffer pH 7.0 at room temperature.

Of each sample, 1 µl was applied to a spot on the pre-cleaned MALDI plate (as described in see section 3.17.2) and 2 µl DHB-matrix was applied on top of the sample. The sample/DHB-matrix- spot was completely dried by a warm stream of air before the MALDI plate was placed on a MALDI target-plate and inserted into the MALDI-TOF analyzer. The software FlexControl version 3.3 was used for controlling the system and FlexAnalysis version 3.3 was used for analysis and processing of data. Most samples were analyzed using laser beam intensities between 20% - 30% of maximum capacity. Addition of Cu^{2+} to the reaction decreases the activity (see results section 3.6.1, table 3.2) and, therefore, copper was not used in further enzyme activity assays.

2.20 Binding assays

Binding efficiencies and affinities of *Cf*CBM33A-N and *Cf*CBM33B-N were examined by incubation with various carbohydrate substrates.

Materials

- Substrates (see section 2.1.2)
  - Xylan
  - Cellulose^{C}
  - α-chtin^{A}
  - β-chitin^{A}
- Binding buffers
  - 100 mM Sodium acetate, pH 3.0, 4.0 and 5.0
  - 100 mM Bis-Tris, pH 6.0
  - 100 mM Tris HCl, pH 7.0, 7.5, 8.0 and 9.0

Procedure

All substrates were pre-washed in 3 x 1 ml binding buffer, and the buffer was removed by pipetting. For studies of the binding of proteins in the periplasmic extracts, 100 µl extract was used, whereas purified proteins were diluted to 2 µM in buffer with the appropriate pH prior to addition of substrate. 100 µl protein sample was mixed well with 100 µl followed by a 3 hour incubation at room temperature with shaking at 1,200 rpm. Following the incubation, samples were centrifuged at 10,000 x g for 5 minutes. The supernatant, carrying the unbound protein, was separated from the substrate. For all samples except xylan, the pelleted substrate was washed in 3 x 1 ml buffer and then boiled for 10 minutes in 30 µl SDS-PAGE sample buffer. From xylan-samples, the supernatant, containing unbound protein, was analyzed instead of the pellet. As a control, pure protein samples or periplasmic extracts were used. For SDS-PAGE, 20 µl of sample was mixed with 20 µl 2 x sample buffer and boiled for 10 minutes. Samples and a standard protein ladder were then subjected to SDS-PAGE analysis according to section 2.18.

2.21 High-performance liquid chromatography (HPLC)

HPLC is a chromatographic technique applied for separation of different compounds in a sample, with the aim to identify, quantify or purify the individual compounds. An HPLC apparatus usually consist of a stationary phase (contained in a column), one or more pumps to drive the mobile phase and the samples through the column and a detector (in this case a UV-detector). The column can contain different stationary phases. The retention time is calculated as the time it takes from the sample is injected until the compound is detected. HPLC delivers high performance (high resolution) because the systems are run at high pressure, which again allows the use of stationary phases with very low particle sizes.

*Cf*CBM33B-N activity on various substrates was analysed both by examining oxidized products released by *Cf*CBM33B-N alone (see section 2.21.1) and by determining synergy obtained when combining this enzyme with chitinases and CBP21 derived from *S. marcescens.* After optimizing reaction conditions, synergy experiments were conducted on β-chitin^{A} (highly crystalline) and β-chitin^{E} (less crystalline nano-fiber particles), as described in section 2.21.2.

2.21.1 Analysis of oxidized oligomeric products by UHPLC

Materials

- Reaction buffer
  - 100 mM Bis-Tris pH 6.0
  - Other buffers used during optimization (stock solutions):
    - 100 mM Sodium acetate, pH 3.0, 4.0 and 5.0
    - 100 mM Tris HCl, pH 7.0, 7.5, 8.0 and 9.0
- Reducing agents (stock solutions)
  - 0.1 M ascorbic acid
  - Other reducing agents used during optimization:
    - 0.1 M gallic acid
    - 0.1 M reduced glutathione
    - 0.1 M glucosamine
- Substrate
  - β- chitin[A], squid pen
- Tris-HCl 15 mM, pH 8.0
- 100 % (v/v) Acetonitrile
- dH_2O
- Standard: A mix of oxidized GlcNAcs with a degree of polymerization (DP) between 3 and 8 (produced in-house; Vaaje-Kolstad et al., 2010).
- Column: HELIC, Acquity UPLC® BEH Amide 1.7 µm, 150 x 2.1mm
  - Pre column: BEH Amide VanGuard, 1.7 µm, 5 x 2.1 mm

Procedure

Reactions (100 µl) were set up in triplicates in cryo-tubes with 10 mg/ml substrate, 1 µM enzyme in 20 mM buffer, with or without reducing agents at different concentrations (1 mM, 2.5 mM, 5 mM and 10 mM). All samples were incubated at 37°C with shaking at 990 rpm. At given time points, 40 µl of the reaction was sampled and added acetonitrile to a final concentration of 70 % (v/v). After centrifuging the samples at max speed to sediment all insoluble substrate, 40 µl of each supernatant (containing solubilized products) was transferred to HPLC-tubes and stored at -20°C. Samples were analyzed on an Agilent 1290 Infinity UHPLC system set up with a Waters Acquity UPLC BEH amide column and a BEH Amide pre column. The amount of

sample applied to the column was 5 µl was injected and oligosaccharides were separated using a column temperature 30°C and a flow rate of 0.4 ml/min. The gradient was started by running 72 % acetonitrile and 28 % 15 mM Tris-HCl pH 8.0 for 4 minutes, followed by an 11 minute gradient ending on 62 % acetonitrile, 38 % 15 mM Tris-HCl pH 8.0, followed by reconditioning of the column was obtained by a gradient back to initial conditions which was run for 5 minutes at initial conditions before next injection. An "in-house" produced mix of oxidized GlcNAcs with a degree of polymerization (DP) between 3 and 8 was used as standard.

2.21.2 Analysis of chitin degradation in synergy experiments

For analysis of mono- and disaccharides obtained upon degradation of chitin an HPLC method based on ion exclusion was used. The column used was a Rezex Fast Fruit column (Phenomenex) which was run isocratically with an acidic mobile phase (5 mM H_2SO_4). The Rezex Fast Fruit column separates ionic and non-ionic analytes, repelling the ionic molecules so they elute early, and retaining the non-ionic molecules (i.e. sugars).

Materials

- Buffers
  - 0.1 M Bis-Tris pH 6.0
- Reducing agent
  - 0.1 M ascorbic acid
- Substrates (see section 2.1.2)
  - β-chitin^{A}
  - β-chitin^{E}
- Tris-HCl 15 mM, pH 8.0
- dH_2O
- Standard: *N*-acetyl-D-glucosamine (GlcNAc; Alfa Aesar) and di-acetyl chitobiose ($(GlcNAc)_2$); Megazyme)
- H_2SO_4, 50 mM and 5 mM solutions
- Rezex RFQ_Fast fruit H+, 100 x 7.8 mm column
- Purified ChiA, ChiB, ChiC and CBP21 from *S. marcescens*

Procedure

Reactions (300 µl) were set up in triplicates in 50 mM BisTris pH 6.0 with 2.5 mM ascorbic acid as electron donor and substrate and enzyme concentrations according to table 2.16, in cryo-tubes that were incubated horizontally at 37°C with shaking at 990 rpm. At given time points, 50 µl was sampled and mixed with 50 µl 50 mM H_2SO_4 to stop the enzymatic reaction. Prior to HPLC analysis, the sample was centrifuged at maximum speed and 20 µl supernatant was transferred to a HPLC- tube. At regular intervals during the sample series, GlcNAc and $(GlcNAc)_2$ were run as standards for quantification, using the following concentrations: 0, 62.5, 125, 250, 500, 1000 and 2000 µM. Standard samples were diluteded with 50 mM H_2SO_4 to a final concentration of 25 mM.

To analyse degradation of the substrate to mono- and disaccharides, a Dionex Ultimate 3000 RSLC HPLC system was set up with a Rezex RFQ Fast Fruit column,using Chromeleon version 7.1.1.1127 to record, integrate and analyse chromatograms. Before use, the Rezex column was heated to 85°C while 5 mM H_2SO_4 was run through the system with a flow of 0.3 ml/min. Then the flow rate was set to 1 ml/min, which was the rate used throughout the analyses. Of each sample and standard, 8 µl was injected and separated during the first 4 minutes of a 6 minutes run with 5 mM H_2SO_4 as mobile phase. Eluted oligosaccharides were monitored by measuring absorption at 195 nm.

Table 2.16. Concentrations of enzymes and substrates in reactions set up to analyze the effect of *Cf*CBM33B-N on the degradation of β-chitin by chitinases.

	Enzyme concentration				
	ChiA	ChiB	ChiC	*Cf*CBM33B-N	CBP21
β-chitin^{A} 2 mg/ml	0.2 µM	-	-	-	-
	0.2 µM	-	-	1 µM	-
	0.2 µM	-	-	0.5 µM	0.5 µM
	0.2 µM	-	-	-	1 µM
	-	0.2 µM	-	-	-
	-	0.2 µM	-	1 µM	-
	-	0.2 µM	-	0.5 µM	-
	-	0.2 µM	-	0.5 µM	0.5 µM
	-	0.2 µM	-	-	0.5 µM
	-	0.2 µM	-	-	1 µM
	-	-	0.2 µM	-	-
	-	-	0.2 µM	1 µM	-
	-	-	0.2 µM	0.5 µM	0.5 µM
	-	-	0.2 µM	-	1 µM
β- chitin^{E} 1 mg/ml	0.2 µM	-	-	-	-
	0.2 µM	-	-	1 µM	-
	0.2 µM	-	-	0.5 µM	-
	0.2 µM	-	-	0.5 µM	0.5 µM
	0.2 µM	-	-	-	0.5 µM
	0.2 µM	-	-	-	1 µM
	-	0.2 µM	-	-	-
	-	0.2 µM	-	1 µM	-
	-	0.2 µM	-	0.5 µM	-
	-	0.2 µM	-	0.5 µM	0.5 µM
	-	0.2 µM	-	-	0.5 µM
	-	0.2 µM	-	-	1 µM
	-	0.04 µM	-	-	-
	-	0.04 µM	-	1 µM	-
	-	0.04 µM	-	0.5 µM	-
	-	0.04 µM	-	0.5 µM	0.5 µM
	-	0.04 µM	-	-	0.5 µM
	-	0.04 µM	-	-	1 µM
	-	-	0.2 µM	-	-
	-	-	0.2 µM	1 µM	-
	-	-	0.2 µM	0.5 µM	-
	-	-	0.2 µM	0.5 µM	0.5 µM
	-	-	0.2 µM	-	0.5 µM
	-	-	0.2 µM	-	1 µM

3 RESULTS

3.1 Bioinformatics

For all 5 proteins described in this study, theoretical data were calculated using the ExPASy ProtParam tool (http://web.expasy.org/cgi-bin/protparam/protparam, (Gasteiger E. 2005)). Results are shown in table 3.1.

Table 3.1. Summary of key properties of the five different proteins described in this study. The Table shows IDs in UniProt and GenBank as well as the molecular properties of both the full-length proteins and the isolated N-terminal CBM33-domain. See text for more details on domain structures.

			Full-length protein				N-terminal CBM33 domain			
Protein	UniProt	GenBank	bp	aa	Mw (kDa)	pI	bp	aa	Mw (kDa)	pI
*Cf*CBM33A	D5UGA8	ADG73091	984	327	35.0	5.26	569	190	21.1	5.13
*Cf*CBM33B	D5UGB1	ADG73094	1104	367	39.0	5.22	591	197	21.6	5.22
*Cf*CBM33C	D5UH31	ADG73234	999	332	34.4	4.62	551	187	20.9	4.93
*Cf*CBM33D	D5UHY1	ADG73405	714	271	29.1	6.56	429	146	16.4	5.91
*At*CBM33A	Q0CGA6	EAU32670	1203	400	42.7	4.75				

3.1.1 CBM33s from *C. flavigena*

Cfcbm33A, Cfcbm33B, Cfcbm33C, and *Cfcbm33D* all code for two-domain proteins with similar architectures. A sequence alignment of the four proteins is shown in Figure 3.1 wherethe two histidines are highlighted in pink and the domain arrangement, found in InterPro (http://www.ebi.ac.uk/interpro), is indicated by coloured lines above the sequence. All four *Cf*CBM33s have an N-terminal signal peptide (between 32 and 43 amino acids in length) followed by a CBM33-domain which is connected to a C-terminal CBM2-domain through a low complexity linker as illustrated in Figure 3.1 and 3.2.

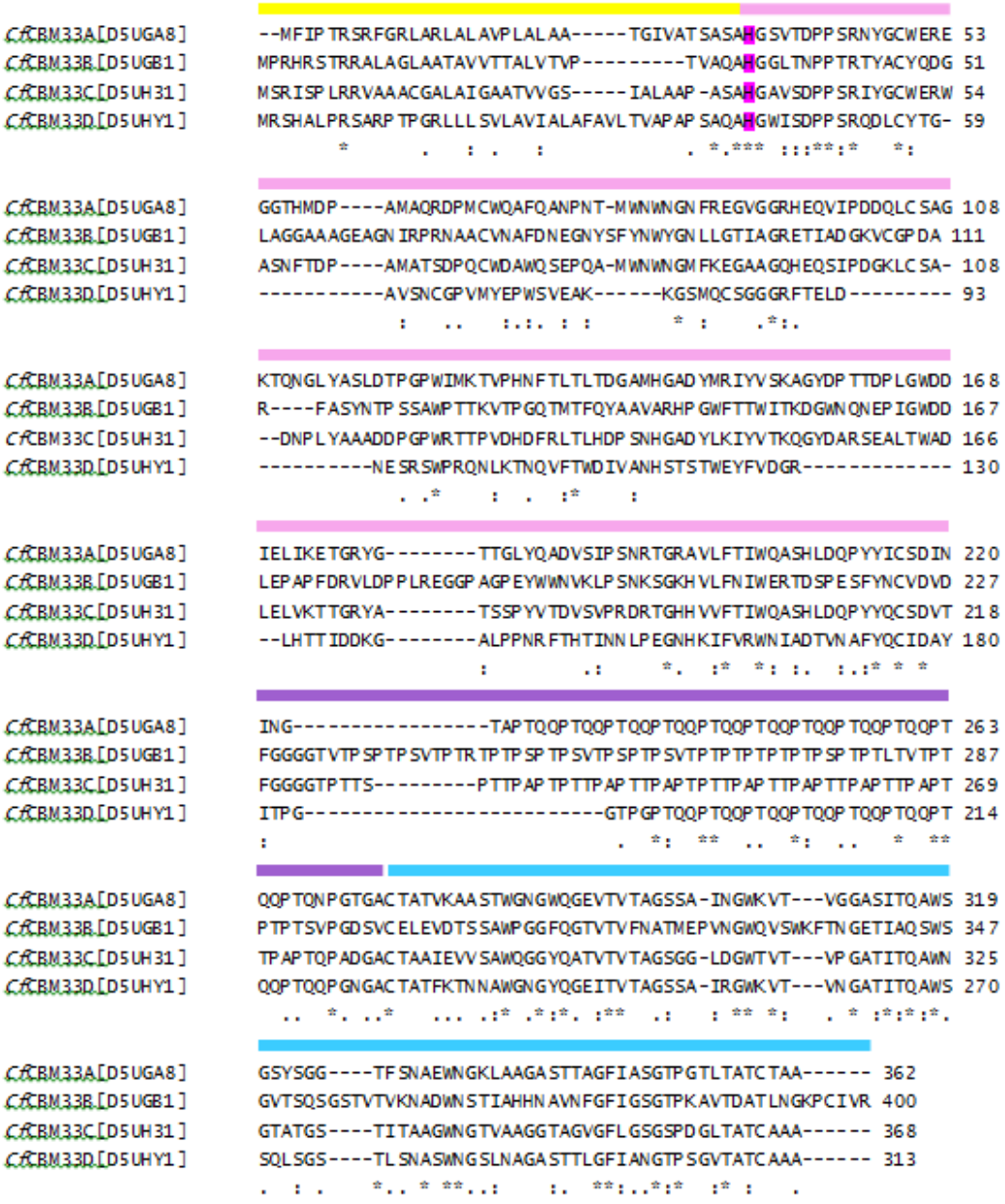


Figure 3.1. Sequence alignment of the full-length CBM33 proteins from *C. flavigena* (see table 3.1 for more details). The four *Cf*CBM33s share a similar domain architecture, presented in Figure3.2 and indicated by coloured lines above the alignment, with an N-terminal signal peptide (yellow) followed by a CBM33-domain (pink) connected though a low complexity linker (purpel) to a C-terminal CBM2 domain. One of the conserved histidines in the active site is highlighted in pink (the corresponding histidines in CBP21 and GH61E are shown in Figure 1.8).

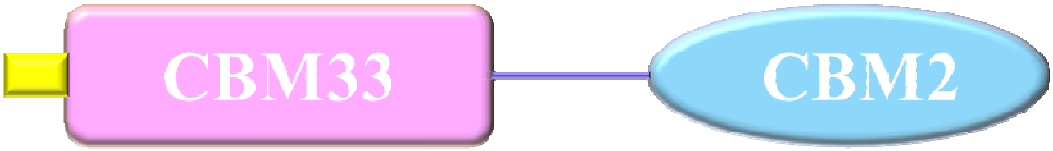


Figure 3.2. Modular scheme of a CfCBM33. All *Cf*CBM33s have the same modules: an N-terminal signal peptide (yellow), a CBM33-domain (pink) and a C-terminal CBM2-domain (blue) connected through a low complexity linker (purple)

The four *Cf*CBM33s all have a C-terminal CBM2-domain. As mentioned in section 1.3.6.1, CBM2s are classified in two groups, CBM2a and CBM2b, referring to a binding affinity towards either cellulose or xylan, respectively. Sequence patterns illustrating the differences between CBM2a and CBM2b may be derived from Prosite (http://prosite.expasy.org/cgi-bin/prosite/prosite-search-ac?PDOC00485) and are shown in Figure 3.3 (top). The Figure also shows a sequence alignment of the CBM2 domains of the four *Cf*CBM33s with the CBM2a-domain of *Cellulomonas fimi* endoglucanase A (CfiCenA; UniProt ID: P07984) and the CBM2b-domain of *C. fimi* xylanase D (CfiXylD; UniProt ID: P54865).

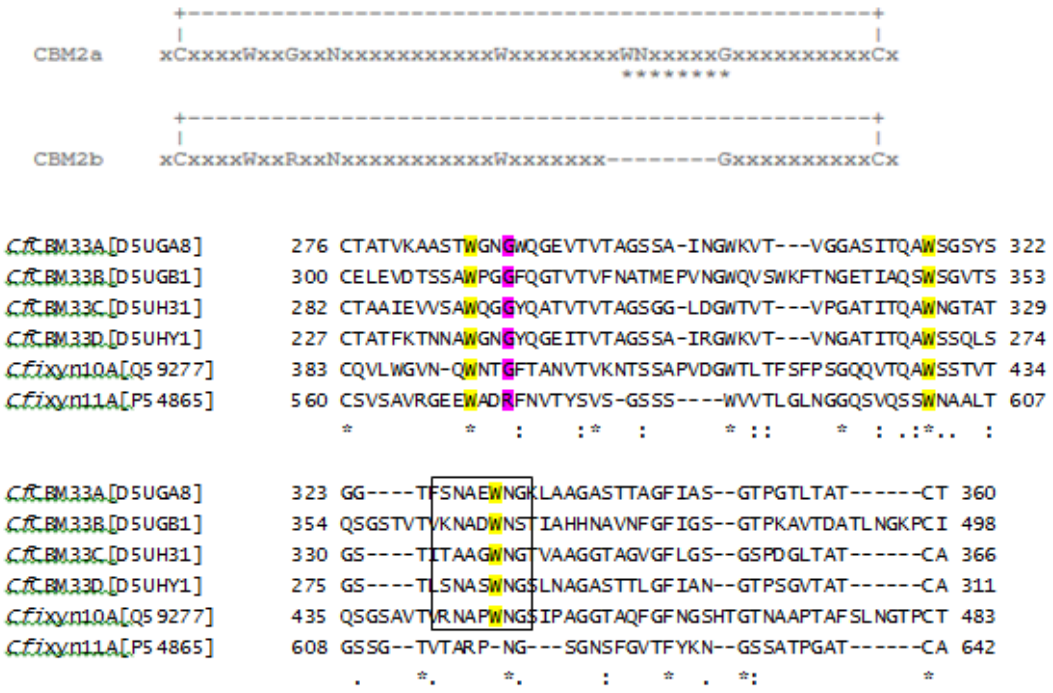


Figure 3.3. Sequence characteristics of *Cf*CBM33s CBM2 domains. Top panel: schematic consensus patterns for CBM2a and CBM2b derived from Prosite. Surface exposed tryptophan residues and conserved residues are indicated by the representative letters, x indicates any amino acid. The sequence pattern that is characteristic for CBM2a and that makes up an extra β-sheet in the structure (see section 1.3.5.1) is indicated by asterisks. Bottom panel: Alignment of *Cf*CBM33-CBM2 sequences with representative sequences of a CBM2a (*Cfi*xyn10A, Figure 1.11, left) and a CBM2b (*Cfi*xyn11A, Figure 1.11 right). All uniprot IDs are indicated in brackets. Conserved surface-located tryptophans are marked in yellow while the amino acids supposed to contribute to substrate specificity by affecting the rotation of the first tryptophan (graphically illustrated in Figure 1.12) is highlighted in pink. The area of the CBM2a-sequence pattern is indicated by a box.

Figure 3.3 shows that all four *Cf*CBM33- CBM2-domains belong to the CBM2a family and most-likely are of the cellulose-binding type, since all have the third conserved tryptophan as well as a glycine in the position that steers the orientation of the first tryptophan. It should be noted that the few CBM2 modules for which binding to chitin has been demonstrated also belong to the CBM2a family (i.e. *Tfu*1665, *Bacillus thuringiensis* chitinase).

3.1.2 CBM33 from *A. terreus*

3.1.2.1 Domain architecture

Atcbm33A (ATEG_07286, Uniprot ID: Q0CGA6) from *A. terreus* encodes a multidomain protein containing 400 amino acids (aa) with an N-terminal signal peptide (aa 1-17) preceding a CBM33-domain (aa 18-117) followed by an unknown domain which is connected to a C-terminal CBM20-domain (aa 298-383) through a low complexity linker (composed of mostly threonine and serine) as illustrated in Figure 3.4. In the present study attempts were made to express the full length protein.

Figure 3.4. Domain organization of *At*CBM33A. The protein consists of an N-terminal signal peptide (yellow), a CBM33-domain (pink) followed by an unknown domain (red) which is connected to a C-terminal CBM20-domain (green) through a serine/threonine linker (blue).
.

3.1.2.2 Codon optimization

To optimize the expression of a protein, efficient DNA transcription alone is not sufficient, since protein production the expression rate also depends on the translation efficiency and the availability of tRNAs with the matching anti-codon. Most amino acids can be coded for by multiple codons and the preferred codon usage varies between organisms (codon usage bias). Codon optimization implies that single nucleotide mutations are introduced into the wild-type DNA sequence to optimise the codon usage without affecting the protein sequence.

Atcbm33A was optimized according to the codon usage bias in *P. pastoris* by GenScript's OptimumGene™ Gene Design software (http://www.genscript.com/codon_opt.html) prior to

gene synthesis. The codon optimized nucleotide sequence (Appendix B) shares 75 % sequence identity with the original sequence.

3.1.2.3 Homology modeling and sequence alignments

In order to explore whether the CBM20 of *At*CBM33A would be structurally compatible with binding starch, a homology model was made using the solution structure of the granular starch binding domain of glucoamylase 1 (G1) from *A. niger* (PDB: 1KUL; (Sorimachi et al. 1996)) as a template. The model was made by use of Modeller (http://salilab.org/modeller/; (Eswar et al. 2006; Sali et al. 1995)), based on sequence alignments generated by ClustalW2, and is presented in Figure 3.5. The sequences share 56 % sequence identity, without insertions and deletions (see below), so the model may be expected to be reasonably accurate.

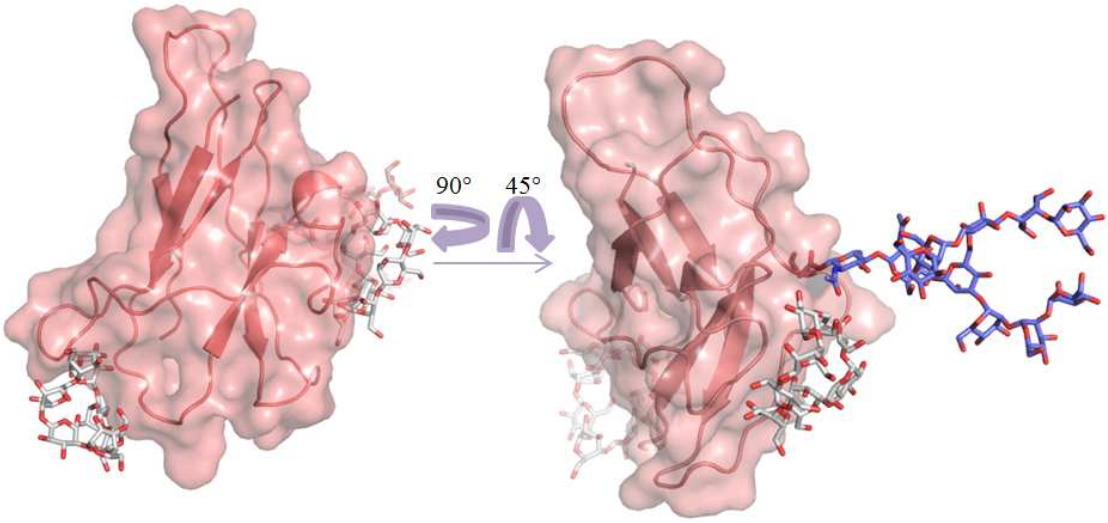


Figure 3.5. Homology model of the CBM20-domain of *At*CBM33A. The model is shown with two bound substrates shown in grey stick representation and oxygen coloured red, one substrate is bound in binding site 1 (in focus in the left model) and one in binding site 2 (in focus in the right model) (see section 1.3.6.3.1 for more details). The right model is rotated 90° to the left and tilted 45° in respect to the left model. An oligomannose is attached to Asn379 (coloured blue) to illustrate a possible N-glycosylation made by *P. pastoris* (see section 3.1.2.4). To position the two bound substrates, the model was superimposed onto the crystal structure of *A. niger* glucoamylase 1 (G1) when bound to β-cyclodextrin (PDB ID: 1AC0). Protein structure graphics were made using PyMol (Schrodinger 2010).

3.1.2.4 Analysis of putative N-glycosylation sites

P. pastoris both O- and N- glycosylates proteins and both modifications could affect the protein expression and function. However, since less is known about the O-glycosylations compared to N-glycosylations, and as no consensus glycosylation sequence is known for O-glycosylations in

P. pastoris, their prediction is not possible. N-glycosylation, on the other hand, is known to occur on asparagine residues which occur in the sequence Asn-Xaa-Ser/Thr, where Xaa is any amino acid except proline with some exceptions depending on the surrounding sequence. For prediction of possible N-glycosylation sites in *At*CBM33 the NetNGlyc (http://www.cbs.dtu.dk/services/NetNGlyc/) server, was used. NetNGlyc is a web-tool for predicting plausible N-glycosylation sites in proteins that based on the surrounding sequence. Analyses show that there is 68% possibility for N-glycosylation to occur at residue 119 and a 51.5 % possibility at residue 379 (see output file in Appendix C) based on the assumption that these amino acids are located on the protein surface.

One of these putative N-glycosylation sites, Asn 379, is located within the CBM20 domain. For illustration purposes, an oligomannose was computationally attached to Asn379 in the homology model shown in Figure 3.5 by the use of the GlyProt (http://www.glycosciences.de/modeling/glyprot/php/main.php) with default settings (see Appendix C). GlyProt is a web-based tool for attaching various N-glycans to potential N-glycosylation sites of a known three-dimensional protein structure.

3.1.2.5 Conserved residues

As stated in section 1.3.4.3, both binding sites of the CBM20-domain have conserved residues proven important for substrate binding. The binding sites of the two proteins were compared by aligning both the sequence and the structures/models of *At*CBM33A_CBM20 and G1_CBM20 (Figure 3.6). The results of these comparisons show that in binding site 1, all residues involved in substrate binding are conserved, while binding site 2 shows more variation.

```
A. niger G1: 509 CTTPTAVAVTFDLTATTTYGENIYLVGSISQLGDWETSDGIALSADKYTSSDPL 563
   AtCBM33A: 298 -----LIPVTFQEFVTTMWGENVFVTGSISQLGSWSTDKAVALSATGYTASNPL 346

A. niger G1: 564 WYVTVTLPAGESFEYKFIRIESDDSVEWESDPNREYTVPQACGTSTATVTDTWR 616
   AtCBM33A: 347 WTTTIDLPAGTTFEYKFIKKETDGSIIWESDPNRSYTVPTGCSGTTATAAASWR 383
```

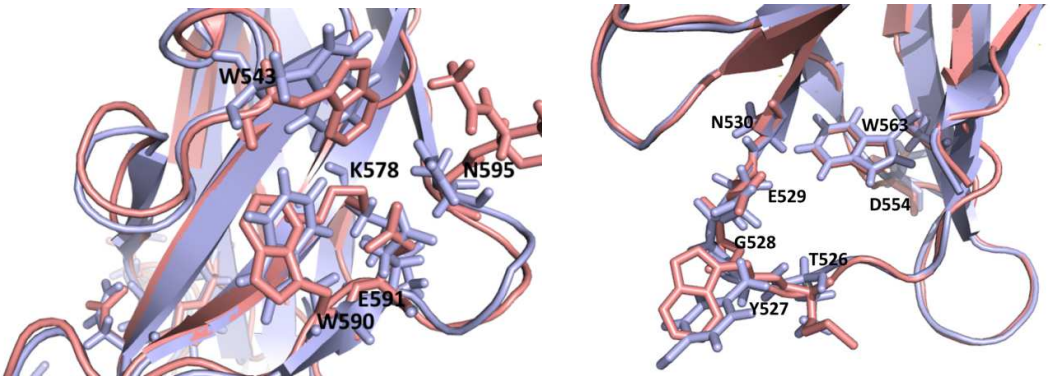


Figure 3.6. Concerved residues in the CBM20 domains of A. niger glucoamylase 1 (PDB ID: 1KUL) and AtCBM33A. (Top) Structure based sequence alignment of the CBM20-domains of A. niger glucoamylase 1 and AtCBM33A. Conserved residues in binding site 1 are highlighted in pink, while conserved residues in binding site 2 are highlighted in green. Residues in binding site 2 that differ between G1 and AtCBM33A are colored in blue and yellow, respectively. (Bottom) Superimposition of the conserved residues in binding site 1 (left) and in binding site 2 (right) of AtCBM33A (pink) and G1 (light blue). In the right picture the flexible loop involved in substrate binding is visible in the upper left corner. The numbering of residues is according to the numbering in 1KUL. Alignments and protein structure graphics were made using PyMol (Schrodinger 2010).

3.2 Cultivation, cloning and transformation

3.2.1 Cultivation of *C. flavigena*

C. flavigena's ability to utilize different carbon sources were examined by cultivation of the bacterium with minimal medium containing distinct different carbon sources (see section 2.2.5). The carbon sources used were cellulose[A], xylan, α-chitin[A], β -chitin[A], all at 1% w/v, and glucose at 0.2 % w/v. We observed no growth on these media, and no secreted proteins were detected in SDS-PAGE of supernatant proteins performed according to the procedure described in section 2.18. No further attempts were made to analyze the susbtrate preferences of the bacterium.

3.2.2 Cloning of *Cf*CBM33s

Chromosomal DNA was extracted from a 2 ml culture of *C. flavigena* grown over night on LB (see section 2.4) and all four *cbm33*-genes from *C. flavigena* (*Cfcbm33A, Cfcbm33B, Cfcbm33C* and *Cfcbm33D*) were amplified by PCR using Phusion™ High-Fidelity DNA polymerase according to section 2.5.1. For each protein, two DNA fragments were generated, one encoding the full length protein and one containing only the N-terminal CBM33 domain without its natural leader sequence, using primer set specific for each gene (Table 2.2; see Figure3.1 for details concerning the truncation points). Full length constructs are indicated by the gene or protein name (e.g. *Cfcbm33A*) whereas constructs comprising only the N-terminal domain have names ending at "-N" (e.g. *Cfcbm33A-N*). PCR fragments of appropriate size were obtained for all eight constructs. Figure 3.7 shows the eight PCR fragments that were successfully cloned (see below).

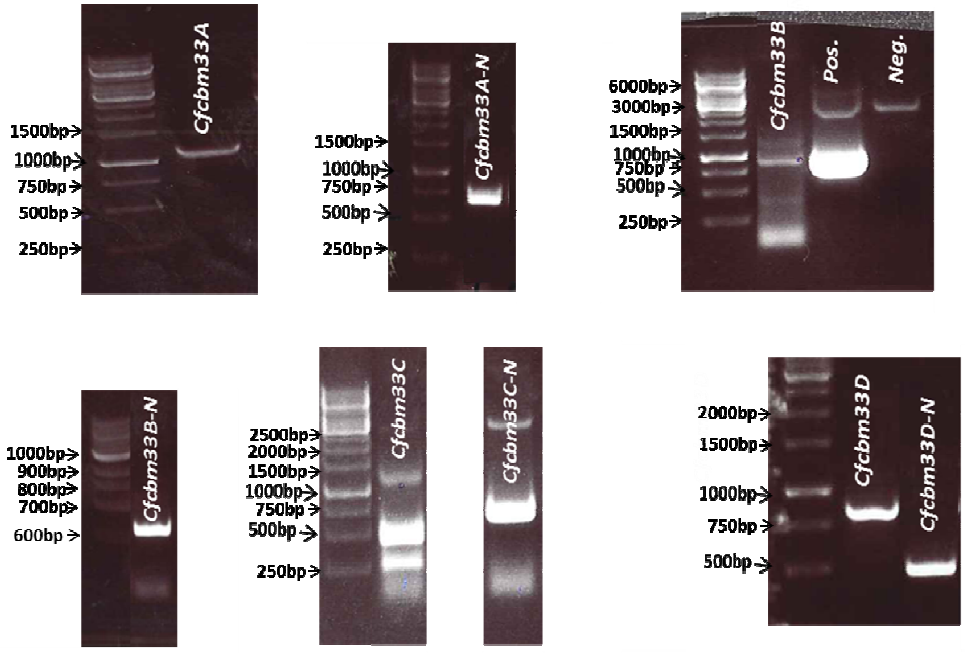


Figure 3.7. PCR amplification of *Cfcbm33* genes. The left lanes show a 1 kb DNA ladder, except for *Cf*cbm33B-N where the left lane show a 100 bp ladder. The size (bp) of all marker-fragments is labeled. The right lanes show PCR products. The predicted sizes of the PCR products are: *Cf*CBM33A 999 bp, *Cf*CBM33A-N 585 bp, *Cf*CBM33B 1117 bp, *Cf*CBM33B-N 607 bp, *Cf*CBM33C 1014 bp, *Cf*CBM33C-N 573 bp, *Cf*CBM33D 931 bp, *Cf*CBM33D-N 445 bp.

The cloning strategy for the *Cfcbm33* genes was based on the use of a variant of pRSET-B (see Appendiks C, figure C.2 for plasmid details) that had the *cbp21* gene from *S. marcescens*

inserted within its multiple cloning site (Vaaje-Kolstad et al. 2005a), referred to as pRSET-B/*cbp21*. The *cbp21* gene includes a signal peptide and it has been shown in previous studies that this pRSET-B construct leads to accumulation of considerable amounts of correctly processed CBP21 in the periplasmic space of the *E. coli* host cells. In the cloning strategy used here the gene fragment encoding mature CBP21 was replaced (by cleaving with *Hin*dIII and *Bsm*I) with *Cfcbm33* gene fragments lacking their natural signal peptide. Thus, vector constructs were generated in which mature *Cf*CBM33s were fused to the leader peptide of CBP21.

All eight amplified gene fragments (shown in Figure 3.7) were attempted cloned into pRSET-B using the In-Fusion™ cloning procedure described in section 2.10.1 and illustrated in Figure 2.1. Ligation mixes were transformed into competent *E. coli* TOP10 according to section 2.11 and transformants were obtained for four of the constructs: *Cfcbm33A, Cfcbm33A-N, Cfcbm33B-N* and *Cfcbm33C*. Each of the four transformants was inoculated overnight in 5 ml LB-medium supplemented with 50 µg/ml ampicillin. The plasmids were isolated from 2 ml of each culture according to section 2.6.1 and the genes sequenced by the use of BigDye version 3.1 according to section 2.8.1 with both forward and reverse sequencing primers (pRSET_B SeqF, RSET-SeqR in addition to the gene-specific forward and revers primers annealing in the middle of the genes, listed in table 2.3). Sequencing results were analysed using GENtle (Manske 2003) and indicated a successful cloning of *Cfcbm33A-N* and *Cfcbm33B-N* while the sequencing results of *Cfcbm33A* and *Cfcbm33C* were incomplete and no further attempts were made to make positive transformants carrying these genes. For protein expression, purified plasmids of *Cfcbm33A-N* and *Cfcbm33B-N* were transformed into BL21 Star (DE3) competent *E. coli* (see section. Transformants containing *Cf*CBM33A-N and *Cf*CBM33B-N were obtained and cultivated for protein expression, as described further below.

3.2.3 Cloning of *Atcbm33A*

The expression vector used for expression of *At*CBM33A in *P. pastoris* was pPICZα-A (see Appendix D for plasmid map of pPICZα-A). pPICZα-A was ligated with the codon optimized synthetic gene encoding *At*CBM33A, both flanked by *Xho*I and *Xba*I, according to section 2.10.2. This strategy would result in the *AtCBM33A* gene replacing the multiple-cloning site preceding the *c-myc* epitope and ending with a stop codon (TGA) prior to the *Xba*I restriction site. The ligation mix was transformed into XL10-Gold competent *E. coli*. Plasmids extracted from positive transformants, selected by growth on Zeocin™, were sequenced by using forward or reverse sequencing primers according to section 2.8. Sequencing results revealed a deletion

of guanine in the stop-codon (TGA) of the synthetic gene. As a result, the stop-codon was lost, but, by coincidence, this mutation led to an in-frame C-terminal fusion between the *AtCBM33A* gene and a strech of nucleotides encoding a C-terminal c-myc epitope followed by a hexa-histidine tag directly followed by a stop codon (Figure 3.8).

Figure 3.8. Alignment of the 3' end of the codon optimized Atcbm33A gene and the Atcbm33A gene present in the pPICZα-A derivative used for AtCBM33A expression. Deletion of one G in the stop-codon of Atcbm33A results in a terminally extended protein with a c-myc epitope and a polyhistidine tag. Additionally, the XbaI restriction site used for cloning (TCTAGA) and the stop-codon subsequent to the polyhistidine tag is marked. Alignments were made using GENtle (Manske 2003).

Because of this deletion, the protein was extended with 23 amino acids and the molecular weight of the protein increased with 2788 Da. This protein is further annotated as *At*CBM33A^{His}. A histidine tag could be beneficial for purification and identification and it was decided to continue using this plasmid throughout the present study. Later, transformants with the correct sequence were obtained (results not shown).

3.3 Protein expression and purification

3.3.1 *Cf*CBM33 expression

For optimal expression of *Cf*CBM33A-N and *Cf*CBM33B-N, different cultivation conditions and media were tested. IPTG, which induce the allegedly inducible T7 promoter, was not used since both previous studies (G. Vaaje-Kolstad, personal communication) and initial experiments carried out in the present study had indicated that induction had no effect on protein expression. Expression of *Cf*CBM33B-N was successful but expression levels depended on the conditions used. Cultivation in TB-medium supplemented with 50 µg/ml ampicillin at 25°C and with shaking at 180 rpm for 3 days gave high expression, as compared to cultivation on e.g. BHI under the same conditions (Figure 3.9). Expression of *Cf*CBM33A-N was also observed, however in lower amounts than *Cf*CBM33B-N (Figure 3.9).

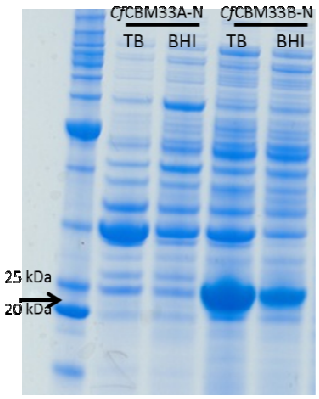


Figure 3.9. Expression of CfCBM33. The picture shows a Coomassie-stained SDS-PAGE gel with a protein marker (left lane) and periplasmic extracts from CfCBM33 producing strains, as indicated. Two of the marker proteins are lebeled by their molecular weight and the approximate position of the expressed proteins (21 kDa for both) is indicated by an arrow. The gel shows that optimal expression of CfCBM33B-N was achieved by cultivation of E. coli BL21 Star™ (DE3) in Terrific broth (TB)-medium for 3 days at 25°C with shaking at 180 rpm. The gel also shows that cultivation in brain heart infusion (BHI) medium gave good expression of CfCBM33B-N, however, to a less extent than cultivation on TB medium (under otherwise same conditions, i.e. 3 days incubation at 25°C with shaking at 180 rpm). Expression of CfCBM33A-N was also achieved after cultivation under the same conditions, however to a less extent than for CfCBM33B-N.

For the larger scale production and subsequent purifcation of *Cf*CBM33 genes, transformed *E. coli* were cultivated in 150 ml TB-medium for three days at 25°C with shaking at 180 rpm. After three days of cultivation, the proteins secreted into the periplasm of the bacterial cells were extracted according to section 2.14.

3.3.2 Expression of *At*CBM33A^{His}

For protein expression, linearized pPICZα-A/ *Atcbm33A*^{His} (approximately 8 µg) was transformed into *P. pastoris* strain X-33 ($7*10^7$ cells) by electroporation according to section 2.12. Transformed cells were selected by cultivation on YPDS-plates supplemented with 100 µg/ml Zeocin™ for four days. This gained approximately 20 transformants, and those carrying the gene of interest were identified by PCR (see section 2.5.1) using the cloning primers (see section 2.1.3). 80 % of the transformants had *Atcbm33A* incorporated in their genomes and a small-scale expression was conducted according to section 2.14.3 to screen for positive transformants able to grow on methanol as sole carbon source and secrete *At*CBM33A^{His}. Approximately $1.5*10^8$ transformed cells were cultivated in 25 ml BMMY containing 0.5 % (v/v) methanol for 96 hours. Secreted proteins were analyzed on SDS-PAGE and transformants secreting *At*CBM33A^{His} (~45.5 kDa) were included in a large-scale expression (500 ml cultures; see section 2.13.4) to determine which transformant to proceed with for fermentation. Growth was monitored by measuring OD_{600} during the four day growth phase with induction by

methanol. The growth curve is shown in Figure 3.10. Out of the 16 transformants carrying the gene, only four could metabolize methanol as sole carbon source and secrete *At*CBM33A^{His} (determined by SDS-PAGE). Out of these four, one was selected for fermentation.

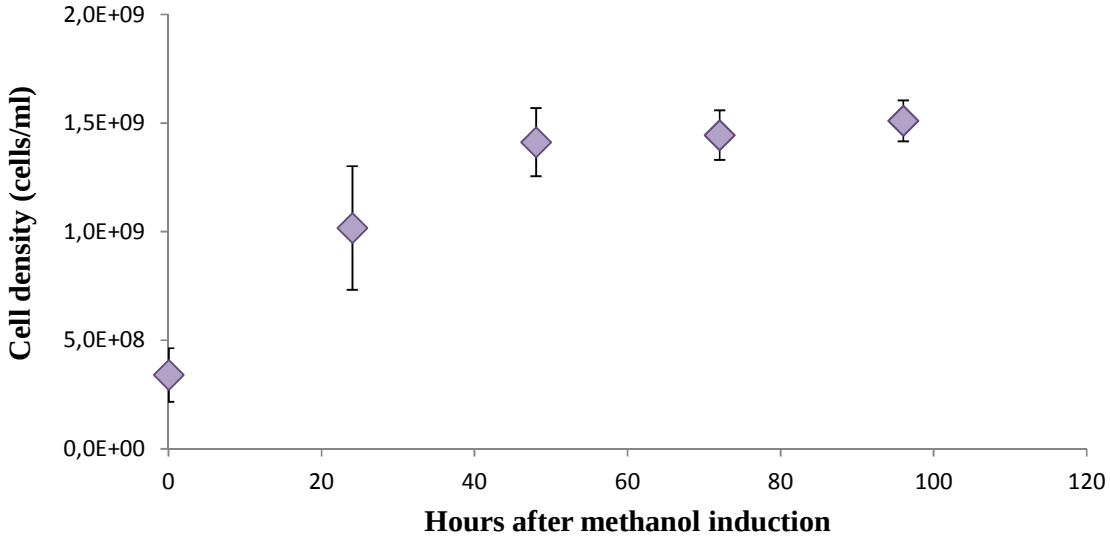


Figure 3.10. Growth of *P. pastoris* X-33 producing *At*cbm33A. Growth of *P. pastoris* X-33 with *Atcbm33A* integrated in the genome during large scale expression (500 ml cultivation medium containing 0.5 % (v/v) methanol as sole carbon source) was monitored by measuring OD_{600} every 24 hours after start of the induction phase (see section 2.13.4 for details on the large-scale expression). The cell density is calculated by multiplying the OD_{600} by $5x10^7$, the approximately cell density/ml when OD_{600}=1 (M. Abou Hachem, personal communication). The graph represents the mean growth of 4 transformants, and the error-bar indicates the standard deviation of the mean.

The selected *P. pastoris* X-33 transformant for expression of secreted *At*CBM33A^{His} was grown in a 5 liter fermentor as described in section 2.13.5. During fermentation, a similar growth-curve as in the large-scale expression was observed, however, with a much higher cell density. The fermentor, containing 2 liter basal salt medium, was inoculated with 250 ml BMGY containing $2.5*10^9$ cells and cultivated with a constant, however minimal, feed of glycerol until the culture reached a cellular yield of 160 g/liter. The glycerol feed was stopped and the culture induced with methanol. Fermentation on methanol was carried out for four days with a more or less constant feed of methanol at 11 g/h/L. After four days the cells were harvested by centrifugation and the supernatant, containing the secreted proteins, were sterile filtered and concentrated 10 times using the Pellicon ultra-filtration system according to section 2.13.6.

3.3.3 Protein purification

Protein purification of *Cf*CBM33A-N and *Cf*CBM33B-N was performed in two steps. First, an ion exchange chromatography was conducted according to section 2.15.1 to exclude molecules with a different isoelectric point than the two *Cf*CBM33s (pI= 5.1/5.2). The periplasmic extract, containing either *Cf*CBM33A-N or *Cf*CBM33B-N retrieved from a 150 ml culture, was pH adjusted to 7.5 by addition of 50 mM Tris-HCl and the sample (50 ml) was loaded onto the column. Elution of *Cf*CBM33s and fraction collection (1 ml/fraction) was performed as described in section 2.15.1 and analyzed on SDS-PAGE (Figure 3.11 Left; chromatograms are shown in Appendix D, Figure D.1). Fractions containing the highest amount of the desired protein (*Cf*CBM33A-N fractions 33-39 and *Cf*CBM33B-N fractions 11-17) were concentrated 7 times and the buffer changed to 10 mM Tris-HCl pH 7.5. The concentrated sample (1 ml) was further purified by size exclusion chromatography according to section 2.15.2 and collected in fractions of 5 ml. Fractions putatively containing *Cf*CBM33A-N (fractions 53-56) or *Cf*CBM33B-N (fractions 53-62) were analyzed by SDS-PAGE (Figure 3.11, Right; chromatograms are shown in Appendix D, Figure D.2-3).

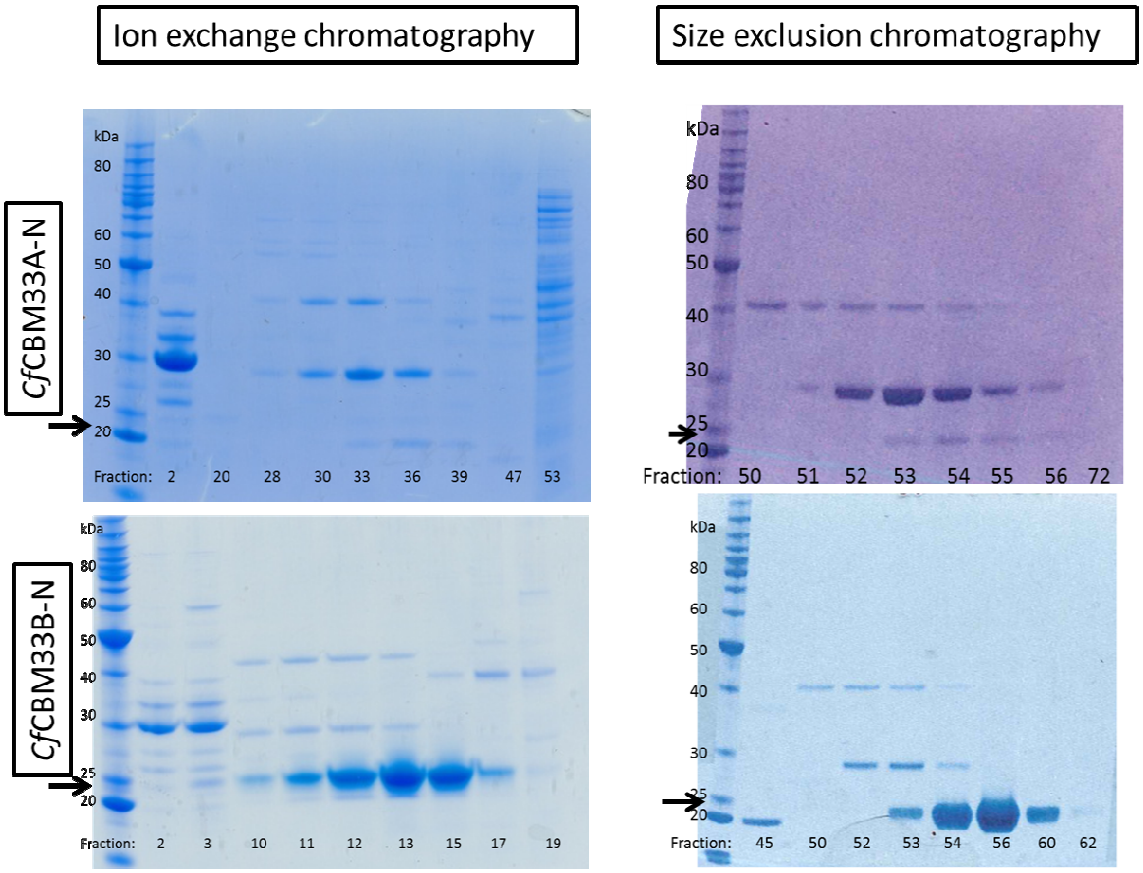


Figure 3.11. Purification of *Cf*CBM33A-N and *Cf*CBM33B-N. The picture shows SDS-PAGE analysis of different fractions from the two-step purification of *Cf*CBM33A-N (Top) and *Cf*CBM33B-N (Bottom), first by ionic exchange chromatography **(Left)** followed by size exclusion chromatography **(Right)**. A protein marker was applied in the left lane in all four gels and the marker proteins are labeled by their molecular weight. The approximate position of the expressed proteins (21 kDa for both) is indicated by arrows. The gels show that both *Cf*CBM33A-N and *Cf*CBM33B-N were not completely purified by ion exchange chromatography. However, after size exclusion chromatography using fractions 33-39 containing *Cf*CBM33A-N and fractions 11-17 containing *Cf*CBM33B-N, pure *Cf*CBM33B-N was obtained from fractions 56 – 62. *Cf*CBM33A-N was not completely purified.

*Cf*CBM33A-N was not completely purified, possibly because of dimerization of proteins (the molecular weight of the largest contaminating protein is approximately 40 kDa) due to intermolecular reformation of cysteine bonds during SDS-PAGE. If so, this could be solved by using reducing agents in the SDS-PAGE running buffer. This is discussed further in the discussion.

*Cf*CBM33B-N, on the other hand, contained purified protein in fractions 58 to 62. These fractions (5 ml/fraction) were pooled and concentrated 25 times before determining the protein concentration according to section 2.16. Total yield of pure *Cf*CBM33B-N was calculated to be approximately 1 mg.

*At*CBM33A^{His} was purified from 30 ml of the secreted proteins (concentrated according to 2.13.6) from the fermentation of transformed *P. Pastoris* (section 2.13.5) using a HisTrap™ HP column according to section 2.15.3. The chromatogram is shown in Appendix E, Figure E.4. Fractions A15 to B3 were pooled, concentrated 10 times and run on an SDS-PAGE for analyzing the purity. The purified protein appeared as a "smear" on the gel (Figure 3.12). Since the protein has known N-glycosylation sites (shown in section0) a treatment of the purified protein with EndoH was performed to remove N-glycosylations (see section 2.18). EndoH-treatment removed some of the glycosylations (visualized by SDS-PAGE; Figure 3.12); however, it did not result in a clear band as should be the case if the protein only were N-glycosylated.

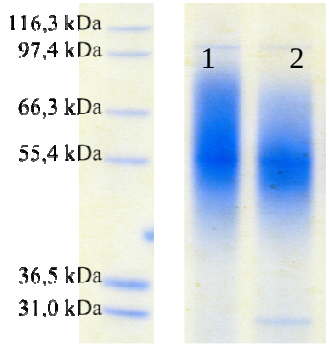


Figure 3.12. Purified AtCBM33AHis. The picture shows a Coomassie-stained SDS-PAGE gel with a protein marker in the left lane and *At*CBM33AHis (Mw: ~45 kDa) first **(1)** purified by immobilized metal ion chromatography followed by **(2)** incubation with EndoH (mw: 29 kDa) at 37°C for two hours. EndoH treatment should remove any N-glycosylations (see section 2.17 and Figure 2.3). Although the gel shows a narrowing of the smear after EndoH treatment, a clear protein band was not achieved. All marker proteins are labelled.

3.3.4 Expression and purification of ChiA, ChiB, ChiC and CBP21

Previously constructed vectors for expression of CBP21 (pRSET-B/*cbp21*) and the three chitinases from *S. marcescens* (*chiA* was cloned in expression vector pMAY20-1, *chiB* in pMAY2-10, *chiC* in pRSET-B) had been transformed into *E. coli* BL21 Star (DE3) for expression (Brurberg et al. 1994; Brurberg et al. 1995; Brurberg et al. 1996; Vaaje-Kolstad et al. 2005b). From these transformants, CBP21, ChiA, ChiB and ChiC were expressed as described in section 2.13.1 in overnight cultures in 300 ml LB-medium supplied with 50 µg/ml ampicillin. ChiC expression was induced by 0.5 mM IPTG when the culture had reached an OD_{600} of 0.6 according to section 2.13.2 while ChiA and ChiB were produced under control of their original promoters. Proteins secreted to the periplasmic space were extracted after lysis of the cells by cold osmotic shock (section 2.14). Following, the three chitinases and CBP21 were purified

from the extract by affinity chromatography with chitin beads according to section 2.15.4. Chromatograms for these one-step purifications are shown in Appendix D, Figure D.5-D.8 and Figure 3.13 shows SDS-PAGE gels illustrating the purity of the purified proteins. The data show that the three chitinases and CBP21 were completely purified. Note that part of the purified ChiC only consist of the GH18-domain and lack the two extra binding domains, CBM12 and fibronectinIII. This truncated version of the protein is referred to as ChiC2, while the full-length protein is known as ChiC1.

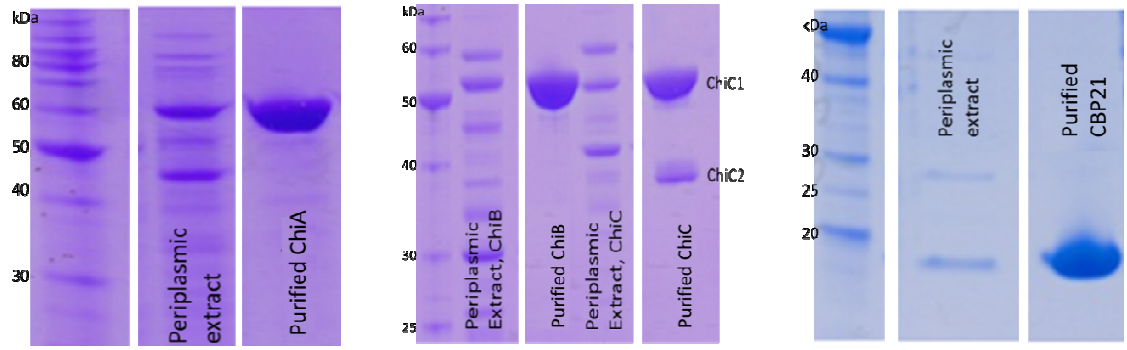


Figure 3.13. Purified chitinases and CBP21 derived from *S. marcescens.* The coomassie-stained SDS-PAGE gels show a protein marker (left lane), the protein contents in the periplasmic extract from witch ChiA (mw: 61 kDa), ChiB (mw: 55.5 kDa), ChiC (ChiC1 mw: 52 kDa, ChiC2: 36 kDa) and CBP21 (mw: 18.8 kDa) was purified and finally the purified proteins. All proteins were completely purified by affinity chromatography using chitin beads (section 2.15.4). Note that purified ChiC results in two protein, one full-length protein (ChiC1) and one truncated variant (ChiC2) lacking the substrate binding modules, CBM12 and fibronectin III-domain.

3.4 Determination of protein identity

Since EndoH treatment of *At*CBM33A^{His} did not result in a clear protein bond on the SDS-PAGE gel (Figure 3.12) protein identification was carried out by MS-analysis. From the SDS-PAGE gel with analyzed purified *At*CBM33A^{His} a piece of the protein-smear (at approximately 55 kDa) was cut out and trypsin digested according to section 2.19.1. The trypsin digested protein was analyzed using MALDI-TOF MS (see section 2.19.2) and gave a positive match in search against the NCBI-database for *At*CBM33A with protein sequence coverage of 30 % and a score of 124 (see Appendix E for output from Mascot; the protein is named ATEG_07286 in this analysis).

3.5 Binding assays

3.5.1 Substrate binding of *Cf*CBM33B-N visualized by SDS-page

Binding properties of the non-purified *Cf*CBM33A-N and *Cf*CBM33B-N, as present in periplasmic extracts of the production strains, were tested with different substrates (α-chitin[A], β-chitin[A], xylan and filter paper; see table 2.1 for details on the substrate) according to section 2.20 at pH 7.5 (Figure 3.).

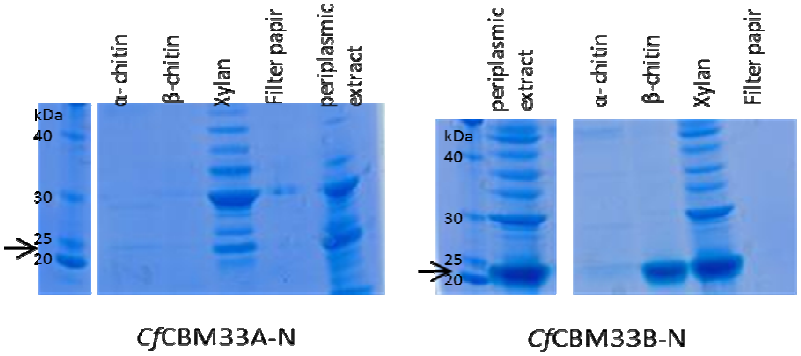


Figure 3.14. Binding of non-purified *Cf*CBM33A-N (left) and *Cf*CBM33B-N (right) as present in periplasmic extracts to α-chitin[A], β- chitin[A], xylan and cellulose[C] at pH 7.5. Samples (100 µl periplasmic extract and an equal amount of substrate) were analysed on SDS-PAGE after incubation for 3 hours at room temperature with shaking at 1200 rpm. Note that the lanes with α-chitin[A], β-chitin[A] and cellulose[C] show proteins bound to the substrates (pelleted by centrifugation), whereas the xylan lanes show the unbound proteins (proteins left in the supernatant after pelleting xylan). A protein marker (left lane in both gels) was used for size-determination of the proteins. All marker proteins in between 20 and 40 kDa are labelled. Periplasmic extract containing the expressed protein is shown as control. The band representing the respective proteins (~21 kDa) is marked with an arrow.

Weak binding to α-chitin[A] was observed for both CfCBM33A-N and CfCBM33B-N. *Cf*CBM33A-N did not show detectable binding to β-chitin[A], while *Cf*CBM33B-N showed strong binding affinity to this substrate. Data for xylan are not conclusive, because, the bound fraction could not be analyzed for technical reasons. Analysis of the unbound material (Fig. 3.13) did indicate that binding to xylan is weak or absent for both proteins (no differences in protein concentrations were observed between the periplasmic extract and the unbound protein sample). Neither of the proteins bound to cellulose[C].

The binding properties of 0.7 µM purified *Cf*CBM33B-N towards 50 mg/ml β-chitin[A] were further examined at varying pH, ranging from 3 to 9, when incubated for 3 hours at room temperature with shaking (1200 rpm). Protein bound to the substrate (+) and protein in the substrate-free sample (÷) were visualized by SDS-page (Figure 3.15**Figure 3.12**).

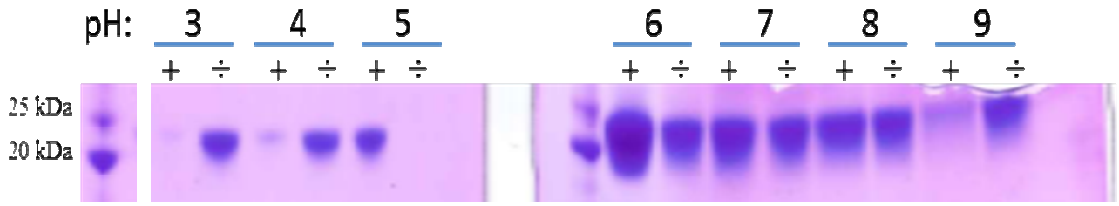


Figure 3.12. Binding of *Cf*CBM33B-N to β-chitin^{A} vary according to pH. 0.7 µM purified *Cf*CBM33B-N was added to samples containing 50 % buffer and 50 % suspension of pre-washed β-chitin^{A}(indicated by +; Final chitin concentration 50 mg/ml) or with 100 % buffer (indicated by ÷) and incubated at 37°C with shaking at 1200 rpm for 3 hours. The gels show protein bound to the substrate (+) and the reference solutions (÷). The reference solutions were tested to check for pH-induced effects on the soluble protein.

The results indicate that binding is clearly pH-dependent. Low substrate binding is observed in the pH range below 5 and above 9, while optimal binding is seen at pH values between 6 and 8. At pH 5, which is close to the proteins calculated isoelctric point, *Cf*CBM33B-N precipitated, as shown by the absence of protein in the reference sample (Figure 3.15) and by the occurrence of a pellet upon centrifugation (not shown).

3.5.2 Binding of *At*CBM33A^{His} to β-cyclodextrin measured by isothermal titration calorimetry

Isothermal titration calorimetry (ITC) experiments with purified *At*CBM33A^{His} were conducted by Maher Abou Hachem and colleagues at the department of systems biology, enzyme and protein chemistry, Denmark Technical University (specifications regarding the experiment are not known). ITC is a technique that measures the released or absorbed heat from a biomolecular binding event. From the measured data, dissociation constants (Kd), reaction stoichiometry (average number of binding sites per mole of protein, N), enthalpy (ΔH) and entropy (ΔS) can be calculated. *At*CBM33A^{His} was titrated with β-cyclodextrin, a cyclic starch analogue consisting of seven glucose units, and the results are shown in Figure 3.. The measured stoichiometry is 0.5 and the calculated dissociation constant 36 µM, indicating a moderate binding affinity towards β-cyclodextrin.

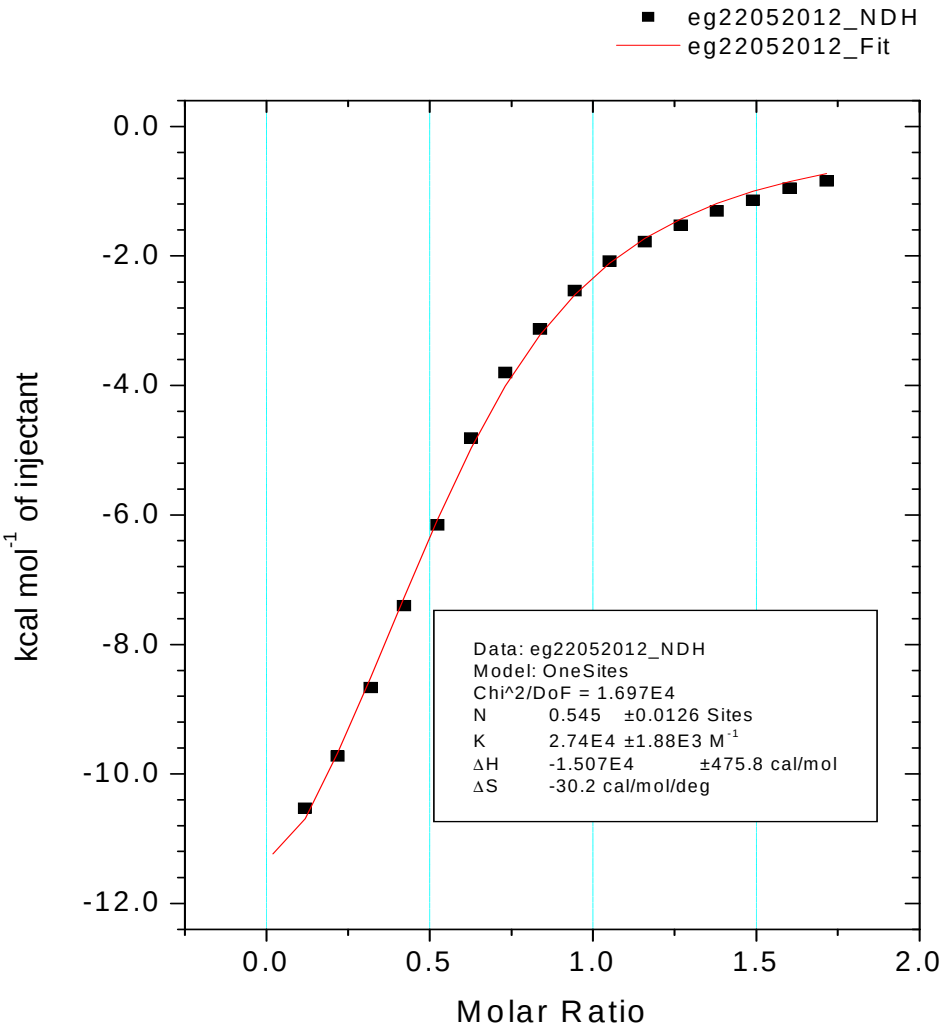


Figure 3.16. ITC experiment of AtCBM33His titrated with β-cyclodextrin. The Figure shows the raw data (black), the fitted line (red) and thermodynamic data derived from the experiment. See Appendix F for raw data.

3.6 Activity assays

3.6.1 Enzyme activity profiling by MALDI-TOF MS

To determine the substrate specificities and the optimal reaction conditions for polysaccharide degradation by *Cf*CBM33B-N, reactions were set up according to Table 3.2 and analyzed by MALDI-TOF MS according to section 2.19.3. The MALDI-TOF analysis was used to detect native and oxidized oligosaccharide products.

Table 3.2. Activity of CfCBM33B-N towards different carbohydrate substrates under different reaction conditions. In addition to the listed variable conditions, all reactions included 1.0 µM CfCBM33B-N, 1.0 mM ascorbic acid and 2.0 mg/ml substrate. See Table 2.1 for detailed information on the various substrates. Reactions were incubated at 37°C with shaking at 900 rpm for 20 hours. Degradation of β-chitinA by CBP21 was used as a positive control, and reactions in absence of enzyme were set up as negative controls. Activity was monitored by MALDI-TOF analysis of oligosaccharide products and detected oligosaccharide products were marked by a + in the table (see section 3.6.2 including Figure 3.6.2 for an example of the interpretation of MALDI-TOF results). The raw data from the MALDI-TOF analyses are provided in Appendix H. Empty boxes mean not determined. The positive controls were done with 1 µM CBP21, 2 mg/ml β-chitinA, 1 mM ascorbic acid in 20 mM Tris-HCl pH 8 and all yielded previous identified oligosaccharides with a degree of polymerization of 4, 6, 8 and 10. No oligomeric products were detected in the negative controls (not shown). The two substrates that were used in further analyses are highlighted in yellow.

Buffer	20 mM Na-acetate	20 mM Bis-Tris	20 mM Bis-Tris	20 mM Bis-Tris	20 mM Tris-HCl
pH	5.0	6.0	7.0	7.0	8.0
Metal	-	-	-	$Cu(II)Cl_2$ 1.0 µM	-
α-chitin^{A}	-	-	-	-	-
α-chitin^{B}	-	-	-	-	-
α-chitin^{C}	-	-	-	-	-
α-chitin^{D}	-	-	-	-	-
α-chitin^{E}	+	+	+	-	+
β-chitin^{A}	+	+	+	+	-
β-chitin^{B}	+	+	+	-	+
β-chtiin^{C}	+	+	+	-	+
β-chitin^{D}	+	+	+	-	+
β-chitin^{E}	+	+	+		
Xylan			-		
Avicel			-		
PASC			-		

The data in table 3.2 and the binding data presented above show that *Cf*CBM33B-N is active on β-chitin in a wide pH range. Table 3.2 also shows some activity towards α-chitin^{E}, which is semi crystalline nanofiber α-chitin. Addition of copper (II) seemed to inhibit the degradation (Table

3.2) and this effect of Cu^{2+} was also seen for the positive control, CBP21 (Appendix G, Figure G.2). β-chitin^{A} and β-chitin^{E} (highlighted in yellow in Table 3.2) were used in further analyses.

3.6.2 A closer view on the activity of *Cf*CBM33B-N towards β-chitin^{A}

A closer view at product profiles obtained with upon incubating β-chitin^{A} with CBP21 or*Cf*CBM33B-N revealed similarities and differences. The two enzymes are similar in that their product profiles show the same periodicity, with a dominance of even numbered products (Figure 3.17). This indicates cleavage at every second glycosidic bond (see Discussion for further comments). Most importantly, the product profiles obtained with *Cf*CBM33B-N showed more peaks than those obtained with CBP21 and the m/z values for these peaks indicated that they represent partially deacetylated oligosaccharides of chitin (Figure 3.1), differing with an atomic mass of 42 from the corresponding completely acetylated products. These partly deacetylated oligosaccharide products seem to be a recurring tendency and were generated from the oxidation of β-chitin with varying crystallinity (see Appendix G and Figure G.9, G.10 and G.11 for zoom-ins of DP4, DP8 and DP10, respectively).

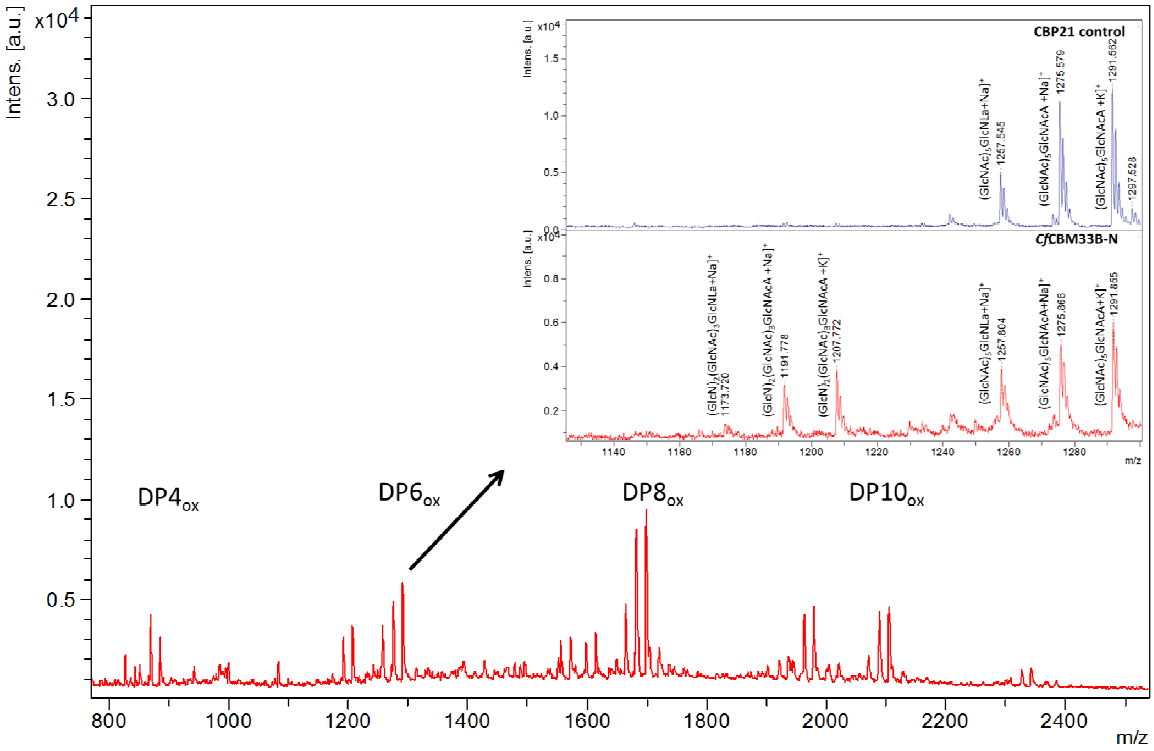


Figure 3.17. MALDI-TOF MS analysis of products detected after treating 2.0 mg/ml β-chitinA with 0.1 µM CfCBM33B-N and 1.0 mM ascorbic acid in Bis-Tris buffer pH 6.0. The upper right corner shows a zoom in of oxidized hexameric products obtained by CfCBM33B-N (red spectrum) and by CBP21 (control; blue spectrum) under the same reaction conditions in reactions carried out at the same time and analyzed in the same analysis

session. In the overview spectrum peaks are labelled by the degree of polymerization (DP) i.e. the number of sugar moieties, including one oxidized sugar. In the zoom in pictures peaks are labelled by the observed atomic mass and the predicted compostion of the oligomer. Abbreviations used: GlcNAc; GlcN, deacetylated GlcNAc; GlcNLa, 1-5 d lactone of GlcNAc; GlcNAcA, aldonic acid of GlcNAc. For similar zoom in pictures (showing similar features) for DP4, DP8 and DP10 see Appendix G, Figure G.9-11.

3.6.3 Optimization of β-chitin oxidation by *Cf*CBM33B-N

Results from MALDI-TOF MS indicated that *Cf*CBM33B-N degrades β-chitin. Therefore, synergy assays with chitinases and CBP21 from *S. marcescens* were performed with β-chitin as substrate. To investigate to what extent substrate crystallinity affects the enzymatic degradation, both the highly crystalline β-chitin^{A} and the less crystalline β-chitin^{E} were used in synergy assays.

Prior to synergy experiments, different posisble conditions for degradation of β-chitin^{A} by *Cf*CBM33B-N and chitinases were screened in a truly quantitative matter, using UHPLC for quantitative product detection according to section 2.21.1. Several parameters were tested.

The buffer capacity of BisTris was tested with 2.5 mM ascorbic acid as reducing agent. Reactions with 10 mg/ml β-chitin^{A} (pre-washed in Bis-Tris with the appropriate pH), 2.5 mM ascorbic acid and 1 µM enzymes (*Cf*CBM33-N, CBP21, ChiA, ChiB or ChiC in separate reactions) were set up in 5, 10, 20 and 50 mM Bis-Tris. The pH of the reactions was measured before and after 24 hours incubation at 37°C with shaking at 900 rpm. A reaction concentration of 20 mM Bis-Tris was necessary to maintain the pH at 6 for all enzymes tested. When analyzed by HPLC using the Rezex column (described in section 2.21.2) a buffer-concentration of 50 mM was used without interfering with the analysis of soluble sugars, as all buffer components eluted prior to mono- and oligosaccharide products.

To determine the optimal pH-conditions for the oxidation of 10 mg/ml β-chitin^{A} by 1 µM *Cf*CBM33B-N, an experiment with 20 mM buffers in the pH-range from 2 to 9 was conducted with specific quantitative detection of oxidized oligomeric products as previously described by Vaaje-Kolstad et al. (2010) and according to section 2.21.1. In this experiment various reductants (ascorbic acid, gallic acid, glucosamine and reduced glutathione) were also tested, all at a final concentration of 2.5 mM. The results are shown in Figure 3.18.

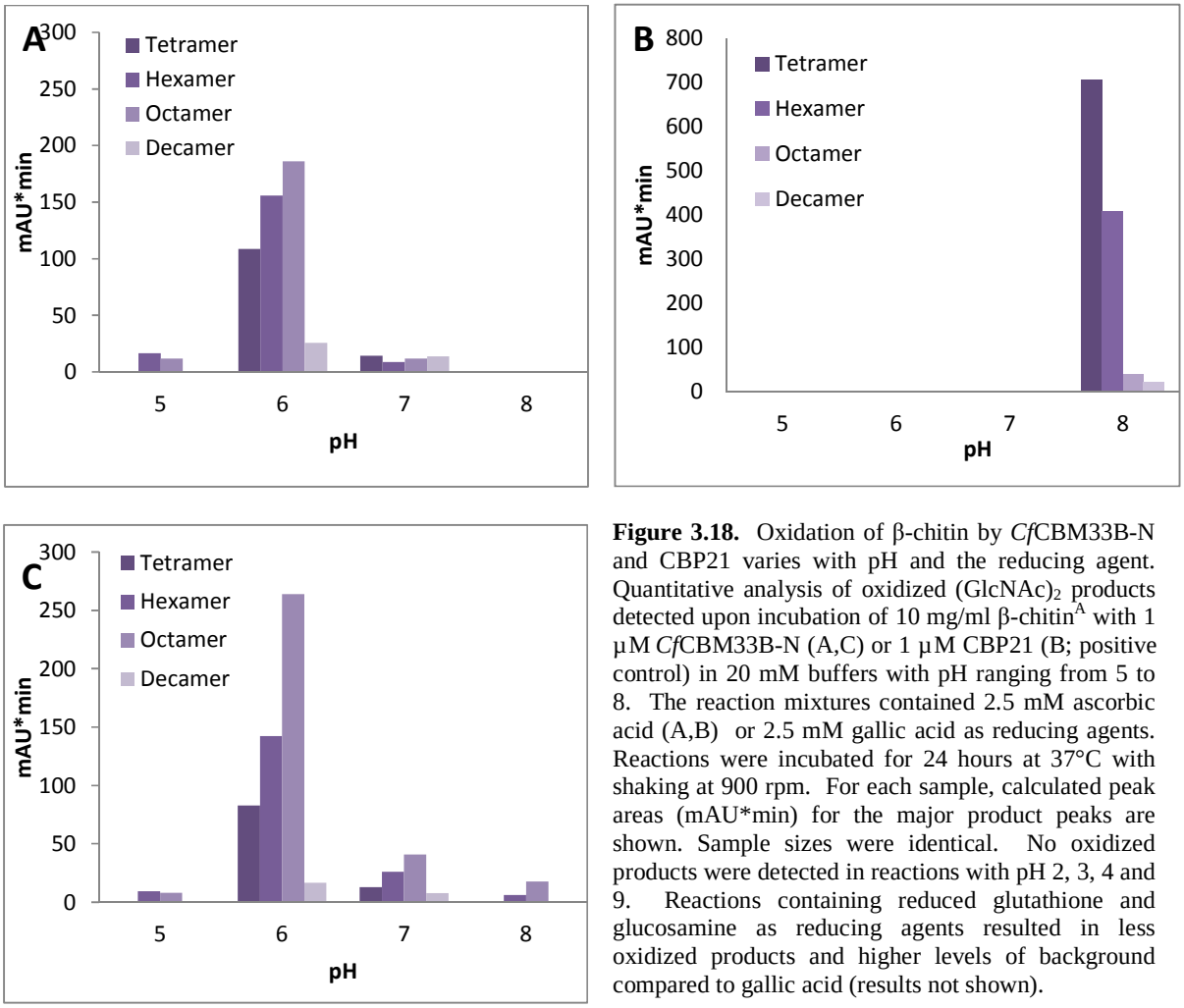


Figure 3.18. Oxidation of β-chitin by *Cf*CBM33B-N and CBP21 varies with pH and the reducing agent. Quantitative analysis of oxidized $(GlcNAc)_2$ products detected upon incubation of 10 mg/ml β-chitin^A with 1 µM *Cf*CBM33B-N (A,C) or 1 µM CBP21 (B; positive control) in 20 mM buffers with pH ranging from 5 to 8. The reaction mixtures contained 2.5 mM ascorbic acid (A,B) or 2.5 mM gallic acid as reducing agents. Reactions were incubated for 24 hours at 37°C with shaking at 900 rpm. For each sample, calculated peak areas (mAU*min) for the major product peaks are shown. Sample sizes were identical. No oxidized products were detected in reactions with pH 2, 3, 4 and 9. Reactions containing reduced glutathione and glucosamine as reducing agents resulted in less oxidized products and higher levels of background compared to gallic acid (results not shown).

The data in Figure 3.18 show that *Cf*CBM33B-N has a pH optimum for activity around pH 6.0 and that this optimum is rather "sharp". No oxidized $(GlcNAc)_2$ products were detected at pH 2, 3, 4 and 9; affirming the observations from binding assays (Figure 3.15). The data also show differences between the reducing agents. Use of ascorbic acid as the reducing agent (Figure 3.8, A) gave a more even distribution of the oxidized products and less noise in the spectra during analysis compared to gallic acid (Figure 3.18 C), reduced glutathione and glucosamine.

Finally, different concentrations of ascorbic acid were tested. The degradation of 10 mg/ml β-chitin^A by 1 µM *Cf*CBM33B-N reactions required a concentration of ascorbic acid between 2.5 and 10 mM (Figure 3., leftpanel). The degradation of 10 mg/ml β-chitin^A by 1 µM CBP21 showed much higher activity, at lower cocnentration ascrobic acid (1 mM) and in this case,

increasing the ascobric acid concentration to 2.5 mM or higher had a clear negative effect on activity (Figure 3.19, right panel). At 2.5 mM ascorbic acid, both enzymes showed an approximately equal activity. Therefore, ascorbic acid was used for all experiments at a concentration of 2.5 mM.

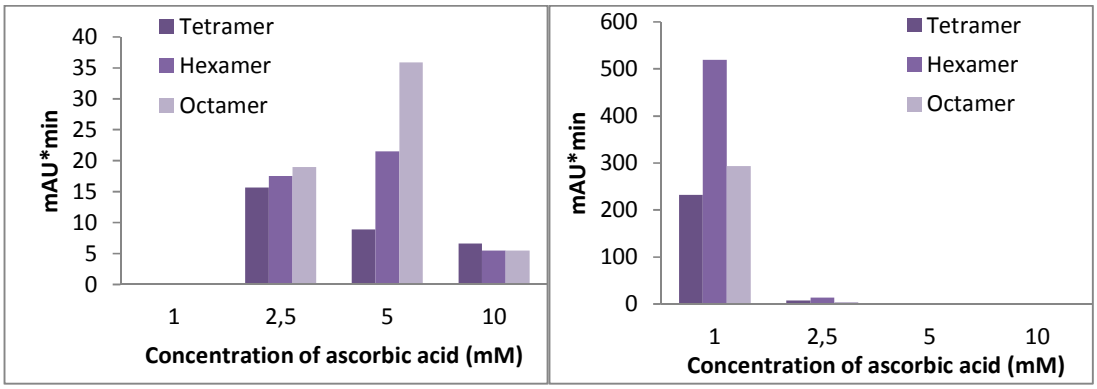


Figure 3.19. Quantitative analysis of oxidized $(GlcNAc)_2$ products detected upon treating 10 mg/ml β-chitin with 1 µM CfCBM33B-N (left) or CBP21 (right) with various concentrations of ascorbic acid (1, 2.5, 5 and 10 mM) as the reducing agent. Reactions were incubated for 24 hours at 37°C with shaking at 900 rpm and oxidized (GlcNAc)2. Products were detected by UHPLC. For raw data see Appendix H, Figure H.1 and H.2.

3.6.4 Testing for synergism: Combining *Cf*CBM33B-N and CBP21 with chitinases to degrade β-chitin

To examine whether or not *Cf*CBM33B-N boosts the degradation of β-chitin^A and β-chitin^E by the *S. marcescens* chitinases ChiA, ChiB or ChiC and to verify whether the CBM33 from *S. marcescens*, CBP21 could work together with *Cf*CBM33B-N synergy experiments were conducted according to section 2.21.2. Reactions (300 µl) set up according to table 2.16 were incubated at 37°C with shaking at 900 rpm for 22-30 hours. Sampling (50 µl) was done at given time points (indicated by points in Figure 3.20) and the amount of $(GlcNAc)_2$ in the sample was measured by UHPLC. The results are presented in Figure 3.3 where the percent conversion equals the fraction of β-chitin in the sample converted to $(GlcNAc)_2$. Data are only based on the levels of generated $(GlcNAc)_2$. Production of the second major product, GlcNAc, was also measured and GlcNAc levels showed the same tendencies. However, since GlcNAc levels were less than 5 % of the $(GlcNAc)_2$, for the sake of simplicity, GlcNAc levels were not taken into account.

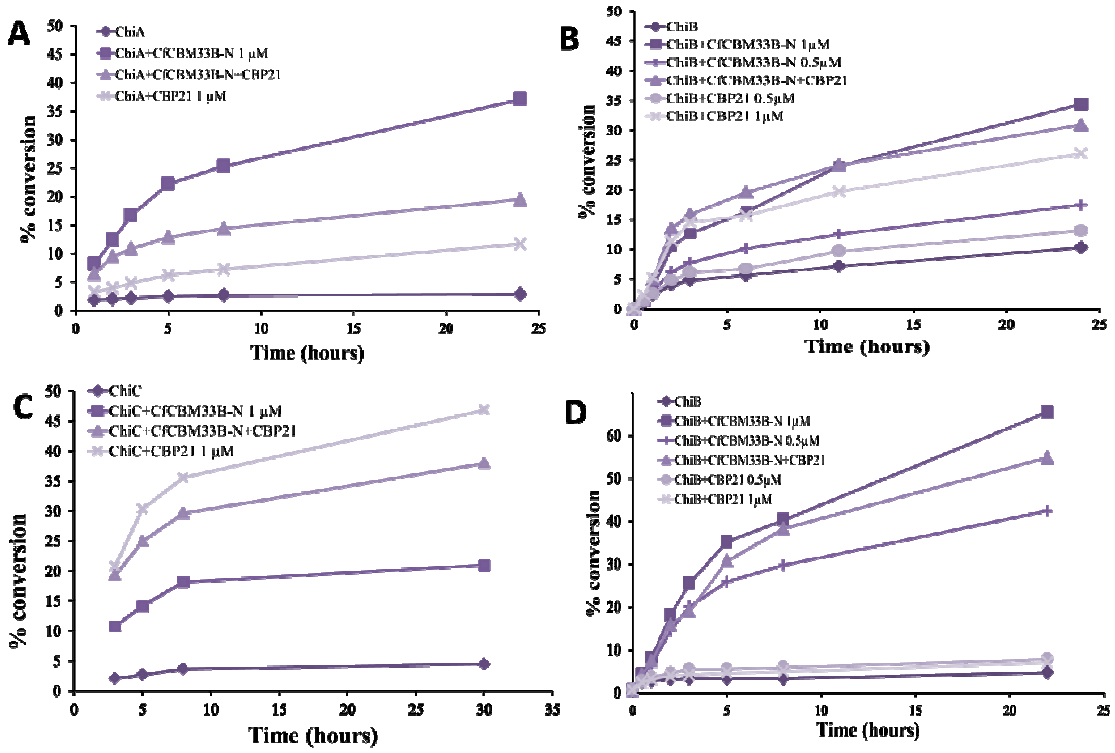


Figure 3.3. Synergy effects of *Cf*CBM33B-N and CBP21 on the degradation of β-chitin by ChiA, ChiB, ChiC. Degradation of 10 mg/ml β-chitin^{A} at pH 6 by 0.2 µM ChiA (A), ChiB (B) and ChiC (C) in the absence or presence of 0.5 µM or 1.0 µM *Cf*CBM33B-N, 0.5 or 1.0 µM CBP21 and equimolar amounts (0.5µM each) of CBP21 and *Cf*CBM33B-N, with 2.5 µM ascorbic acid as external electron donor. (D) Degradation of 2 mg/ml β-chitin^{E} by 0.04 µM ChiB with the same reaction conditions as in B. Reactions were set up in triplicates and incubated at 37°C with shaking at 900 rpm. Degradation is presented as the mean conversion of β-chitin to $(GlcNAc)_2$, as measured by UHPLC. The lines are only drawn between points for illustration. Data for the standard curve used for quantification are shown in Appendix H, Figure H.3.

All Figures indicate an increased degradation rate by addition of *Cf*CBM33B-N to the reaction. CBP21 also appears to increase the degradation rate of the three chitinases towards β-chitin^{A} but does not have any effect on the degradation of β-chitin^{E} by ChiB. Note that the amount of ChiB added to the reactions with β-chitin^{E} is a fifth the amount added to reactions with β-chitin^{A}. Nevertheless, the turnover rate to $(GlcNAc)_2$ remains high. Remarkably, partly in contrast to data presented above, *Cf*CBM33B-N seems more active than CBP21 when combined with ChiA or ChiB, whereas the opposite is the case for ChiC. A statistical model was fitted to the raw data to further be able to test if these alleged synergies are significant (see section 3.6.5).

3.6.5 Estimated regression models for the conversion of β-chitin over time and the effect of synergy

Synergy experiments (section 3.6.2 and Figure 3.20) indicate a boosting activity of *Cf*CBM33B-N on the degradation of β-chitin by ChiA, ChiB and ChiC. However, the graphs in Figure 3.10 is only displaying the experimental results, but is not sufficient for claiming a significant effect of these synergisms. To further analyze the results and test for significance on the observed synergies, a second order regression model was fitted to the observed data presented in section 3.6.4. The model described below, including interactions of enzyme-mix and time of incubation, was adapted to the data using the statistical program R version 2.14.0 (R_Development_Core_Team 2008) with R-commander (Liland 2012) installed.

Model:

$Y_{ijk} = \mu + E_i + (\beta_1 + \beta_{1i})T_j + (\beta_2 + \beta_{2i})T_j^2 + e_{ijk}$ where:

$e_{ijk} \sim N(0,\sigma)$, $\sum_{i=1}^{4} E_i = 0$, $\sum_{i=1}^{4} \beta_{1i} = 0$, $\sum_{i=1}^{4} \beta_{2i} = 0$

Y_{ijk} is the per cent conversion of the substrate by enzyme-mix i, at time j and in replicate k. i=1,2,3,4, j=1,…,6 and k=1,2,3
μ is a constant.
E_i is the effect of enzyme-mix i on the conversion of the substrate.
β_1 and β_2 is the general effect of time (hours) on the substrate conversion
β_{1i} and β_{2i} is the synergistic effect on the conversion of the substrate of time (hours)
T_j is the time j, measured in hours, for sampling
e_{ijk} represents each sample's discrepancy from the mean. All e_{ijk} are independent.
For all models, an analysis of variance (ANOVA) was conducted, and the results showed that all interaction terms included in this model are significant for the estimation of β-chitin conversion over time.

Using this model, the time-dependent conversion of the substrates was calculated (estimated) for each enzyme-mix and plotted in Figure 3.4. The output from the estimation of all model parameters using R is shown in Appendix J along with the model-diagnostic plots. All p-values (p) lower than 0.05 ($p<0.05$) are considered significant. A p-value of 0.05 entails that it is a 5 % possibility to observe this high interaction in the data, assuming there is no real interaction.

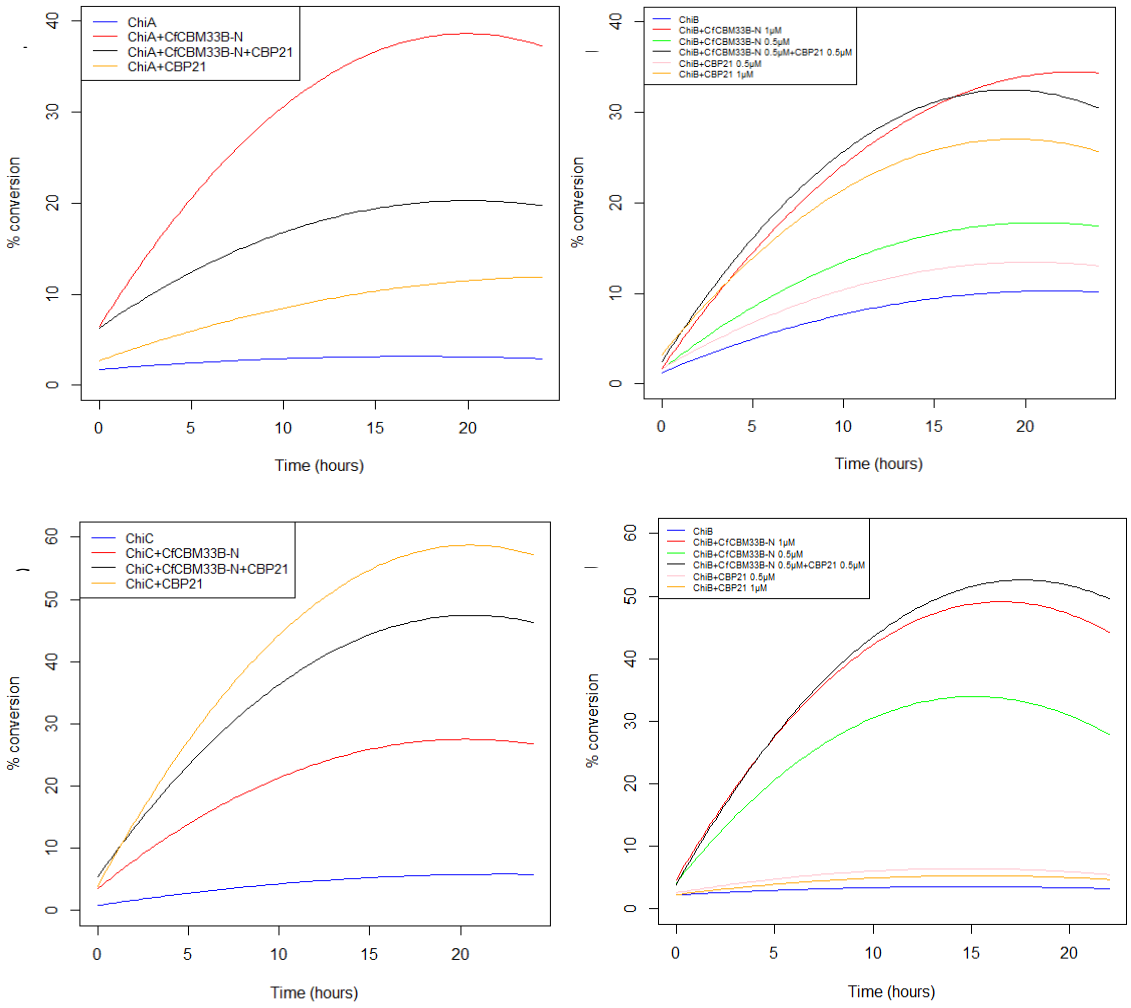


Figure 3.4. Effect plots showing the estimated models of per cent conversion over time for each synergy experiment including degradation of β-chitin^{A} by ChiA (A), ChiB (B) and ChiC (C), and degradation of β-chitin^{E} by ChiB (D). A second order model was fitted to the data presented in Figure 3.19 for the time period from 1 to 24 hours. Detailed results from the statistical analysis of all models and model-diagnostic plots are shown in Appendix I. Effect plots were made using the statistical software R.

To further analyze differences between the different enzyme treatments a post hoc pairwise test, the Tukey test, was conducted (results are shown in Appendix I). From the Tukey test it is clear that the degradation rate over time of most synergies differ from each other. The degradation of β-chitin^{A} by ChiA and ChiC is significantly ($p<0.01$) increased by both *Cf*CBM33B-N and CBP21. Interestingly, while *Cf*CBM33B-N increases the degradation of β-chitin^{A} by ChiA more than CBP21, the opposite is observed in the degradation by ChiC, where addition of CBP21 gains a higher degradation rate compared to addition of *Cf*CBM33B-N. In both cases, equimolar amounts of the two LPMs in combination give a degradation rate that is in between the rates obtained with each of the single LPMs. On the degradation of β-chitin^{A} by ChiB there is a significant effect ($p<0.01$) of addition of *Cf*CBM33B-N and *At*CBM33A, but in this case they

increase the degradation rate equally. Conversely, on the degradation of β-chitin^{E} by ChiB there is a major difference. While CBP21 has no effect on the degradation, *Cf*CBM33B-N most significantly ($p<0.01$) boosts the degradation rate. In total, *Cf*CBM33B-N and *At*CBM33A show different effect on the degradation of β-chitin by ChiA, ChiB and ChiC. This is very interesting and will be further discussed in the following section.

4 DISCUSSION

The objective of this study was to clone, express, purify and characterize CBM33s from two different species, *C. flavigena* and *A. terreus*, primarily to determine whether or not the enzymes possess a lytic polysaccharide monooxygenase activity and secondly to understand *C. flavigena'*s need for as many as four CBM33s.

Despite the fact that the genome of *S. marcescens* encodes only one CBM33, the LPM CBP21, this bacterium is considered one of the most efficient chitin degraders (Fuchs et al. 1986) indicating that CBP21 alone is adequate to complement its chitinases. Many organisms, including *C. flavigena*, encode multiple CBM33s, all with similar domain architecture. The reasons for having several CBM33s are unknown and remain a matter of speculation. Do they complement each other in the degradation of complex biomasses such as lignocellulosic materials comprising both cellulose and hemicellulose? Do they target different substrates or different types of surfaces within the same substrate? Answers to these questions could hopefully give more insight into carbohydrate degradation performed by microorganisms in the soil and could possibly be used for optimization of industrial biomass conversion.

Cloning and expression of all four CBM33s from *C. flavigena*, both in the full length version and in the truncated version lacking the CBM2-domain (Figure 3.1), was attempted. However, due to a desire to at least characterize a few of these enzymes, attempts to clone and express them were not exhaustive. Four of the eight cloning steps failed for reasons that were not further investigated (cloning was attempted only once). Two of the four protein variants that were successfully cloned were not expressed. As is commonly observed for CBM33 proteins (MSC thesis Torfinn Nome; G Vaaje-Kolstad, personal communication) expression is difficult and often fails, especially for multi-domain proteins.

Two *Cf*CBM33s, *Cf*CBM33A-N and *Cf*CBM33B-N, i.e. both in truncated forms, were successfully cloned and expressed. *Cf*CBM33B-N was produced at high levels; it was completely purified and studied in detail further. *Cf*CBM33A-N was produced at lower levels and appeared not to be completely purified after size exclusion chromatography (Figure 1.10). This impurity was unexpected, since the proteins in the starting material differ highly in molecular mass. It is thus plausible that some of the other proteins expressed have an affinity for the column material that is not blocked by the salt concentration in the buffer. A possible solution would be to increase the salt concentration, lower the flow and increase the length of the column. Another possible explanation is that the protein has formed dimers after purification, as the size of one of the contaminating bands is about 40 kDa. If this is the case, this might be because of intermolecular reformation of cysteine bonds during SDS-PAGE, resulting in dimerization (CfCBM33A-N contains four cysteins). A possible solution would then be to use reducing agents also in the running buffer.

C. flavigena is known to be a cellulose and xylan degrading bacterium (Abt et al. 2010). The presence of a CBM2a-domain at the C-terminus of all *Cf*CBM33s indicates binding affinity towards cellulose (Simpson et al. 2000) but binding to chitin is also possible, as seen for the CBM2 domain of chitinase C from *Streptomyces lividans* (Fujii & Miyashita 1993). From both binding experiments and activity assays, *Cf*CBM33B-N shows no specificity towards xylan or cellulose, only for chitin. Obviously, since these domains lack their CBM2s it cannot be excluded that the complete proteins bind to and cleave e.g. cellulose. The genome of *C. flavigena* DSM 20109 contains no identified chitinases (GH18 or GH19) or known chitosanases but it encodes four proteins belonging to family four of the carbohydrate esterases (CE4). The family of CE4 includes acetyl xylan esterases, chitooligosaccharide deacetylases and peptidoglycan GlcNAc deacetylases, all with a deacetylase activity. While three of the four CE4s are linked to CBM- and GH-domains, indicating xylan-specific activity, the fourth CE4 found in *C. flavigena* (Uniprot ID: D5UC80) has no auxiliary domains. A search through the databases using the basic local alignment search tool (BLAST) revealed sequence similarities with polysaccharide deacetylases, among them chitin deacetylases.

Although speculative, it is possible that one role of the CfCBM33s is to make chitin more accessible to deacetylation. Such an effect, which could be part of an otherwise not yet known alternative chitin degradation route, has indeed been observed for CBP21 (Z. Liu et al., manuscript in preparation). Moreover, intriguingly, in addition to generating oxidized acetylated chitin oligosaccharides as observed in the degradation of β-chitin[A] by CBP21 (Vaaje-Kolstad et

al. 2010), *Cf*CBM33B-N also generates partly deacetylated products from the oxidation of the same substrate. These partly deacetylated products are also observed from the degradation of other, less crystalline, β-chitin substrates by *Cf*CBM33B-N (Appendix H), displaying a recurring tendency, and possibly a new function or binding specificity compared to CBP21. Nevertheless, both CBP21 and *Cf*CBM33B-N tend to cleave every second glycosidic bond in the chitin chain, thereby generating even-numbered oligosaccharides reflecting the periodicity seen in the structure (Figure 1.6).

A possible difference in function between CBP21 and *Cf*CBM33B-N was also observed in the synergy experiments (Figure 3.3). Previous studies have demonstrated that CBP21 from *S. marcescens* promotes hydrolysis of crystalline β-chitin^{A} by ChiA and ChiC, and is essential for full chitin degradation by ChiB (Vaaje-Kolstad et al. 2010). This is consistent with the observations found in this study and presented in Figure 3.3 and Figure 3.21. Interestingly, on the less crystalline nanofiber chitin (β-chitin^{E}) CBP21 did not affect the degradation by ChiB. *Cf*CBM33B-N, on the other hand, demonstrates the same increase of ChiB's degradation of all β-chitin substrates. This further indicates a difference in substrate binding between CBP21 and *Cf*CBM33B-N.

The production of partially deacetylated products suggests that *Cf*CBM33B-N has a preference for deacetylated regions of the substrate. Combined to the opposing activities observed in the synergy experiments, this leads to the hypothesis that *Cf*CBM33B-N has a new type of specificity or a possible loss of specificity; i. e. *Cf*CBM33B and CBP21 may target different substrates or have activity on different surfaces or regions on the same substrate.

The activity of *Cf*CBM33B compared to that of CBP21 is clearly different as shown by the difference in synergy obtained when combined with ChiA and ChiC. This difference may result from inequalities in the surface targeting. Such dissimilarities in surface targeting have previously been described for cellulose binding proteins (Carrard et al. 2000; Lehtiö et al. 2003). Furthermore, as elaborated on in section 1.3.3, the three chitinases from *S. marcescens* show different, but complementary activities. ChiB and ChiC show little synergy in degradation of chitin, while ChiA shows strong synergy with both ChiB and ChiC on the same substrate (Suzuki et al. 2002; Vaaje-Kolstad et al. 2005a). Taken into consideration that these three chitinases have different auxiliary domains (ChiA contains a fibronectinIII–like domain, ChiB contains a CBM5 and ChiC contains a CBM12 and an fibronectin III-like domain), Horn et al. (2006b) postulated that auxiliary carbohydrate binding domains potentiate the catalytic activity

or directs it to different parts of the substrate. One possible explanation is the competitiveness regarding substrate binding as CBM5 and CBM12 are classified in the same family in the protein family database, Pfam (http://pfam.sanger.ac.uk/) (Finn et al. 2008), as “distantly related”. All in all, it might seem that both chitinases and CBM33s vary in their affinity for specific types of chitinous substrates. Notably, considering the lack of chitinases in the genome of *C. flavigena*, it remains to be seen what the natural substrate of *Cf*CBM33B actually is.

Considering the possibility of LPMs to be surface specific, it is likely that while CBP21 only acts on surfaces with protruding N-acetyl groups, *Cf*CBM33B-N can oxidize chitin chains on multiple surfaces (illustrated in Figure 4.1). If also the chitinases display differences in surface targeting and the target surface of CBP21 is the same as the target surface of ChiB, it does not matter whether LPM is added to the reaction; they both make the substrate available for ChiB. This is consistent with the observations presented in Figure 3.3 (B) and the tukey tests (Appendix J) showing that CBP21 and *Cf*CBM33B-N increase the β-chitin degradation by ChiB equally. If ChiA then targets another surface than ChiB and CBP21, its β-chitin degradation will be greatly affected by which LPM is added to the reaction. Addition of *Cf*CBM33B-N would increase the substrate degradation rate by ChiA, while addition of CBP21 would have a lower impact, as observed in the presented data (Figure 3.4(A)) and proven significant by tukey testing (Appendix I).

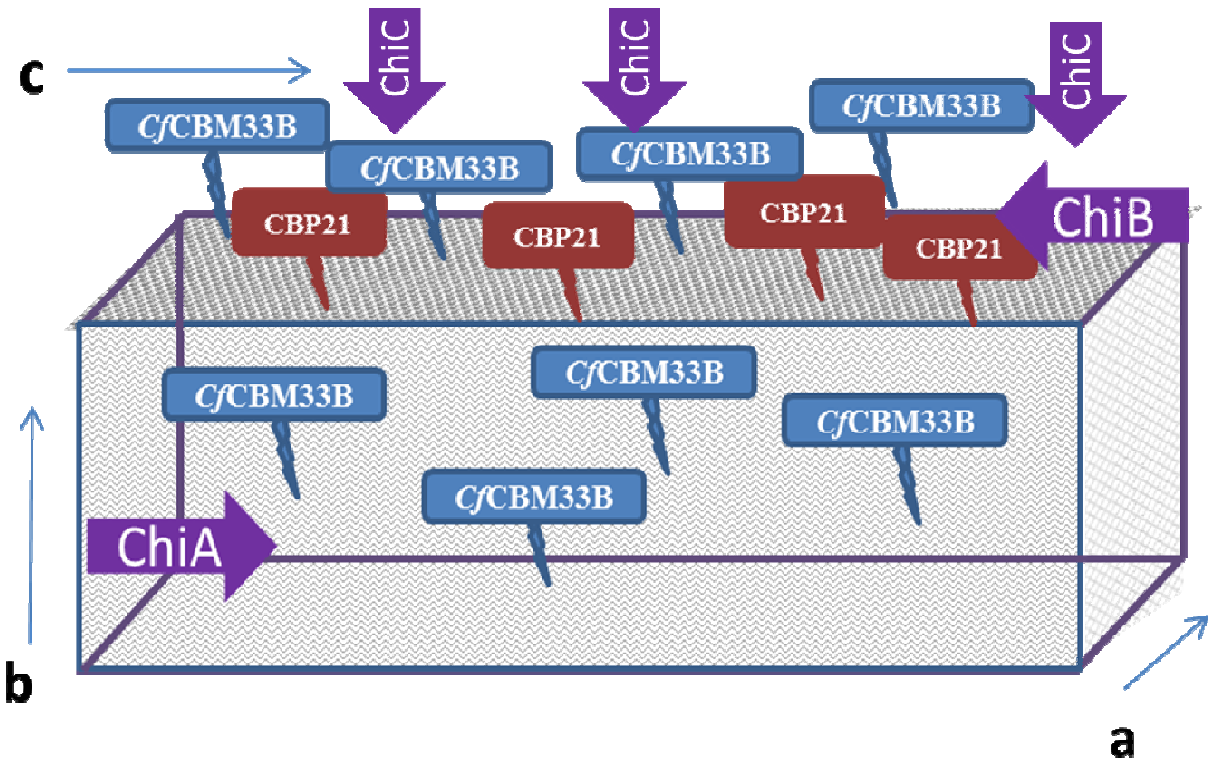


Figure 4.1. Illustration of one possible explanation for the different functionalities of CBP21 and CfCBM33B-N in the synergistic degradation of β-chitin by ChiA, ChiB and ChiC. The side ab corresponds to the reducing end, bc to the planar side with only hydroxyl groups and hydrogen atoms sticking out, while ac corresponds to the side where every second N-acetyl group is sticking out. CBP21 and CfCBM33B may bind at the same surface as ChiB explaining equal increase of the degradation of β-chitin by ChiB by addition of either of the two LPM. If then, ChiA binds to another surface than ChiB, and if CfCBM33B binds multiple surfaces; addition of CfCBM33B-N would increase the degradation of β-chitin by ChiA to a greater extent than by addition of CBP21.

Finally, assuming CBM5 and CBM12 targets the same surface; ChiC would target the same surface as ChiB, and should thus be affected by both *Cf*CBM33B and CBP21. This is also observed in Figure 3.3, however CBP21 increases the degradation rate significantly more than *Cf*CBM33B-N does. A possible explanation for this could be that, in addition to possessing different surfaces specificities, CBP21 and *Cf*CBM33B target different regions of the substrate crystal. Assuming that LPMs are adapted to different regions of the substrate crystal, the *Cf*CBM33B-N could have a preference for oxidizing amorphous areas of the crystal, whereas CBP21 oxidize in the more crystalline parts. The latter activity would probably have the largest effect because it may be safely assumed that chitinase more easily gain access to amorphous regions than to the more crystalline regions that are being made accessible by CBP21.

LPMs acting on chitin and cellulose cleave β(1-4) bonds while leaving the substrate in an oxidized form according to the mechanisms shown in Figure 1.9. What has mainly been observed up to now is the generation of an "oxidized reducing end", i.e oxidation of C1 Vaaje-

Kolstad et al., 2010; Westereng et al., 2011; Forsberg et al., 2011; Quinlan et al. 2011). For some fungal LPMs it has been shown that they oxidize C4 (Phillips et al., 2011; Beeson et al. 2012), and perhaps even C6 in what may be a side reaction (Horn et al. 2012; Quinlan et al. 2011). Although alternative oxidations could take place, this has not been emphasized in our study. The MS data presented do not permit easy discrimination between the various types of oxidation. Since C4- or even C6-oxidation seems much less common than C1 oxidation, and since only C1 oxidation has been found for by LPMs working on chitin, through the rest of the discussion we suppose that all chain breaks are made by C1-oxidation.

It would be of great interest to further study the differences in surface targeting and substrate oxidation among the different CBM33s and LPMs. Additionally, as CBM33s previously have been shown to degrade crystalline substrates like chitin (Vaaje-Kolstad et al. 2010) and cellulose (Forsberg et al. 2011) it would also be interesting to explore to which extent CBMs are active on other crystalline substrates, like starch. Finding an enzyme streamlining the starch degradation would target a whole new area in the use of LPMs, namely the food-industry, as well as, possibly, production of first generation bio-ethanol from corn starch. Therefore, the CBM33 of the starch degrading yeast *A. terreus*, anticipated to have starch binding properties as it is linked to a CBM20, was included in the present study. *At*CBM33A was expressed in *P. pastoris* (work performed at the group for enzyme and protein chemistry lead by Birte Svensson at the Department of systems biology, Technical university of Denmark) with the intention to express the protein with all its natural features, foremost glycosylations.

*At*CBM33A has two possible sites for N-glycosylations (Appendix C, Figure C.1.). Whether such glycosylations are essential for functionality is not easy to predict. It is interesting to note that in a recent study, Taylor et al (2012) claim that N-glycosylations improve the binding of CBM1 towards cellulose. The study implies that glycosylations function to lead the substrate into the binding site. This may also be the case for other CBMs and as the glycosylation site in the homology model of the CBM20 domain (Figure 3.5.) is near binding site 2 it is likely to have an impact on the substrate binding of *At*CBM33A. In addition to N-glycosylations of the amide nitrogen of asparagine residues in the Asn-Xaa-Thr/Ser consensus sequence, *P. pastoris* also adds O-linked oligosaccharides made up of mannose residues (Daly & Hearn 2005). *P. pastoris* does not seem to have a preferred amino acid for O-glycosylation, in other words, no consensus glycosilation sequence is known (Cregg et al. 2000). From studies on other fungal carbohydrate degrading enzymes, it is well known that Serine/Threonine rich interdomain linkers tend to be O-glycosylated (Letourneur et al. 2001), and this could also be the case for the CBM20 domain of

*At*CBM33A. In a review on glycosylation of *P. pastoris*-derived proteins, Bretthauer and Castellino (Bretthauer & Castellino 1999)point out that N-glycosilation may be demonstrated by shong that treatment with enzymes such as EndoH leads to clear band shifts in SDS-PAGE gels. EndoH treatment removed some of the glycosylations on *At*CBM33A (Figure 3.1). Nevertheless, if these simply were N-glycosylations, treatment with EndoH would have resulted in a clearer and sharper band. This might indicate that the protein is also O-glycosylated.

In comparison to CBM20 from glucoamylase 1 from *A. niger*, the predicted model of the *At*CBM33A CBM20 shows only small differences in conserved residues at binding site 2, while binding site 1 is completely conserved (Figure 3.5). As the CBM20 of glucoamylase 1 has two substrate-binding sites (Figure 3.5) and the similarities with the predicted binding sites of *At*CBM33A are high, *At*CBM33A was expected to have a binding stoichiometry of 2 (i.e. binding two β-.cyclodextrins per enzyme). Contradictory, results from the ITC experiments showed a binding stoichiometry of 0.5 for β-cyclodextrin (Figure 3.). As discussed previously, it is possible, however doubtful, that the glycosylation near the second binding site could be of steric hindrance for substrate binding. Also unlikely is the possibility that the CBM33-domain conflicts with the binding as the domains are separated by a long linker. Nonetheless, it is a possibility. It is also possible that the histidine-tag at the C-terminus interferes with the binding. Importantly, the ITC experiments did yield a binding constant lower than expected, but still large enough to confirm the functionality of the CBM20 domain.

Time-limitations prohibited further functional analysis of *At*CBM33A^{His}, such as analyses of the proteins possible contribution to starch degradation by amylases. Obviously, such experiments need to be done and will be of great interest.

4.1.1 Concluding remarks and Future work

The present study has revealed functional differences between *Cf*CBM33B-N and CBP21. However, this study does not explain why *C. flavigena* encodes four CBM33s, neither deos it explain the structural basis and the in vivo functional implications of the differences between *Cf*CBM33B-N and CBP21. To better understand *C. flavigena*'s substrate specificities and substrate degrading properties, cloning, expression and characterization of the remaining *Cf*CBM33s, both in the truncated variant (lacking the C-terminal CBM2-domain) and in full-length would be of considerable interest. This was attempted; however, due to time-limitations, no further efforts were made to clone and express the remaining *Cf*CBM33s. A deeper

understanding of these proteins will also require deeper functional and structural analysis of each individual protein. Studies like this could be complemented by transcriptional data showing which CBM33s are expressed when *C. flavigena* is cultivated with different carbon sources. This was attempted during this study, but with no success.

The generation of partly deacetylated products from the oxidation of β-chitin by *Cf*CBM33B-N was not further investigated. Future work should endeavour to characterize these products in detail and further understand the function of *Cf*CBM33B. This may reveal a new function in the group of LPMs. The most likely explanation for the present observation is that *Cf*CBM33B oxidizes the substrate in deacetylated areas and would then increase the degradation rate in synergy with deacetylases.

An interesting approach would also be to investigate the difference in degradation rates observed when adding CBP21 and *Cf*CBM33B-N, alone or in synergy, to reactions with various chitinases. As the crystalline β-chitin displays different surfaces available for the enzymes, the possibility of chitinases and or various CBM33s to target these different surfaces for optimal degradation should be considered in future studies.

Making the statistical model (section 3.6.5), one assumption was that the variance was constant. From the residual plots in Appendix J it is clear that this assumption was wrong. The variance is increasing by the time of incubation and a model with weighted regression, weighting the variances differently according to the time of the reaction. However, this would not affect the results remarkably, as all interactions are highly significant.

Regarding *At*CBM33A, the present work was only the start of a big study on CBM33s with putative starch affinity. ITC experiments were only conducted on glycosylated proteins and to exclude the possibility of interference by N-glycosylations, it would be preferable to also perform the analysis on deglycosylated proteins. Although glycosylations have proven to be of great importance of the binding affinity of other CBMs, a glycosylation near the starch binding domain could possibly be of steric hindrance for substrate binding. Also unknown is the effect of the extended C-terminal His-Tag on the binding affinity. Native *At*CBM33A, lacking this His-tag, was also expressed and further work should be conducted on this variant of the protein. Clearly, functional studies, with the aim of detecting possible oxidative cleavage in starch need to be conducted.

While much further work needs to be done, the present work on *At*CBM33A does lead to two important conclusions. Firstly, the protein does indeed show affinity for starch that is lower than expected but still in the same range as e.g. the affinity of CBP21 for chitin (Vaaje_Kolstad et al, første 2005 paper i JBC). Secondly, the data show that it is in fact possible to produce a complete recombinant version of this potentially complicated multi-domain protein.

5 REFERENCES

Aachmann, F., Eijsink, V. & Vaaje-Kolstad, G. (2011). ^{1}H, ^{13}C, ^{15}N resonance assignment of the chitin-binding protein CBP21 from *Serratia marcescens*. *Biomolecular NMR Assignments*, 5 (1): 117-119.

Aam, B. B., Heggset, E. B., Norberg, A. L., Sorlie, M., Varum, K. M. & Eijsink, V. G. (2010). Production of chitooligosaccharides and their potential applications in medicine. *Mar Drugs*, 8 (5): 1482-517.

Abt, B., Foster, B., Lapidus, A., Clum, A., Sun, H., Pukall, R., Lucas, S., Glavina Del Rio, T., Nolan, M., Tice, H., et al. (2010). Complete genome sequence of Cellulomonas flavigena type strain (134). *Standards in genomic sciences*, 3 (1): 15-25.

Asare, E. K., Jaiswal, S., Maley, J., Baga, M., Sammynaiken, R., Rossnagel, B. G. & Chibbar, R. N. (2011). Barley grain constituents, starch composition, and structure affect starch in vitro enzymatic hydrolysis. *Journal of agricultural and food chemistry*, 59 (9): 4743-54.

Baban, J., Fjeld, S., Sakuda, S., Eijsink, V. G. & Sorlie, M. (2010). The roles of three Serratia marcescens chitinases in chitin conversion are reflected in different thermodynamic signatures of allosamidin binding. *J Phys Chem B*, 114 (18): 6144-9.

Bennett, J. W. (2010). An overview of the Genus *Aspergillus*. I: Machida, M. & Gomi, K. (red.) *Aspergillus - Molecular biology and genomics*, s. 238: Caister Academic Press.

Bergey, D. H., Breed, R. S., Hammer, R. W., Harrison, F. C. & M., H. F. (1923). *Bergey's manual of determinative bacteriology*. 1st ed. utg. Baltimore: The Williams and Wilkins.

Blackwell, J. (1988). Physical methods for the determination of chitin structure and conformation. I: Willis A. Wood, S. T. K. (red.) b. Volume 161 *Methods in Enzymology*, s. 435-442: Academic Press.

Bostock, J. M., Miller, K., O'Neill, A. J. & Chopra, I. (2003). Zeocin resistance suppresses mutation in hypermutable Escherichia coli. *Microbiology*, 149 (4): 815-816.

Bretthauer, R. K. & Castellino, F. J. (1999). Glycosylation of Pichia pastoris-derived proteins. 30 (Pt 3): 193-200.

Brurberg, M. B., Eijsink, V. G. H. & Nes, I. F. (1994). Characterization of a chitinase gene (chiA) from Serratia marcescens BJL200 and one-step purification of the gene product. *FEMS Microbiology Letters*, 124 (3): 399-404.

Brurberg, M. B., Eijsink, V. G., Haandrikman, A. J., Venema, G. & Nes, I. F. (1995). Chitinase B from Serratia marcescens BJL200 is exported to the periplasm without processing. *Microbiology*, 141 (Pt 1): 123-31.

Brurberg, M. B., Nes, I. F. & Eijsink, V. G. (1996). Comparative studies of chitinases A and B from Serratia marcescens. *Microbiology*, 142 (Pt 7): 1581-9.

Buléon, A., Colonna, P., Planchot, V. & Ball, S. (1998). Starch granules: structure and biosynthesis. *International Journal of Biological Macromolecules*, 23 (2): 85-112.

Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V. & Henrissat, B. (2009). The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic acids research*, 37 (Database issue): D233-8.

Carlström, D. (1957). The crystall structure of alpha chitin. *J. Biophysic and Biochem. Cytol*, 3 (5): 669-683.

Carrard, G., Koivula, A., Söderlund, H. & Béguin, P. (2000). Cellulose-binding domains promote hydrolysis of different sites on crystalline cellulose. *Proceedings of the National Academy of Sciences*, 97 (19): 10342-10347.

Christiansen, C., Abou Hachem, M., Glaring, M. A., Viksø-Nielsen, A., Sigurskjold, B. W., Svensson, B. & Blennow, A. (2009a). A CBM20 low-affinity starch-binding domain from glucan, water dikinase. *FEBS Letters*, 583 (7): 1159-1163.

Christiansen, C., Abou Hachem, M., Janecek, S., Vikso-Nielsen, A., Blennow, A. & Svensson, B. (2009b). The carbohydrate-binding module family 20--diversity, structure, and function. *The FEBS journal*, 276 (18): 5006-29.

Cregg, J., Cereghino, J., Shi, J. & Higgins, D. (2000). Recombinant protein expression in *Pichia pastoris*. *Molecular Biotechnology*, 16 (1): 23-52.

Daly, R. & Hearn, M. T. W. (2005). Expression of heterologous proteins in Pichia pastoris: a useful experimental tool in protein engineering and production. *Journal of Molecular Recognition*, 18 (2): 119-138.

Davies, G. & Henrissat, B. (1995). Structures and mechanisms of glycosyl hydrolases. *Structure*, 3 (9): 853-859.

DeLano, W. L. (2002). The PyMOL molecular graphics system.

Durand, A., Hughes, R., Roussel, A., Flatman, R., Henrissat, B. & Juge, N. (2005). Emergence of a subfamily of xylanase inhibitors within glycoside hydrolase family 18. *FEBS Journal*, 272 (7): 1745-1755.

Eijsink, V. G., Vaaje-Kolstad, G., Varum, K. M. & Horn, S. J. (2008). Towards new enzymes for biofuels: lessons from chitinase research. *Trends Biotechnol*, 26 (5): 228-35.

Finn, R. D., Tate, J., Mistry, J., Coggill, P. C., Sammut, S. J., Hotz, H.-R., Ceric, G., Forslund, K., Eddy, S. R., Sonnhammer, E. L. L., et al. (2008). The Pfam protein families database. *Nucleic acids research*, 36 (suppl 1): D281-D288.

Forsberg, Z., Vaaje-Kolstad, G., Westereng, B., Bunaes, A. C., Stenstrom, Y., MacKenzie, A., Sorlie, M., Horn, S. J. & Eijsink, V. G. (2011). Cleavage of cellulose by a CBM33 protein. *Protein Sci*, 20 (9): 1479-83.

Fuchs, R. L., McPherson, S. A. & Drahos, D. J. (1986). Cloning of a *Serratia marcescens* gene encoding chitinase. *Applied and environmental microbiology*, 51: 504-509.

Fujii, T. & Miyashita, K. (1993). Multiple domain structure in a chitinase gene (chiC) of *Streptomyces lividans*. *journal of General Microbiology*, 139: 677-686.

Gardner, K. H. & Blackwell, J. (1975). Refinement of the structure of β-chitin. *Biopolymers*, 14 (8): 1581-1595.

Gasteiger E., H. C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A. (2005). Protein Identification and Analysis Tools on the ExPASy Server. I: Walker, E. M. (red.) *The Proteomics Protocols Handbook*, s. 571-607: Humana Press.

Gooday, G. W. (1990). *The ecology of chitin degradation*, b. 11. New York, NY, ETATS-UNIS: Plenum. 44 s.

Harris, P. V., Welner, D., McFarland, K. C., Re, E., Navarro Poulsen, J. C., Brown, K., Salbo, R., Ding, H., Vlasenko, E., Merino, S., et al. (2010). Stimulation of lignocellulosic biomass hydrolysis by proteins of glycoside hydrolase family 61: structure and function of a large, enigmatic family. *Biochemistry*, 49 (15): 3305-16.

Hashimoto, M., Ikegami, T., Seino, S., Ohuchi, N., Fukada, H., Sugiyama, J., Shirakawa, M. & Watanabe, T. (2000). Expression and Characterization of the Chitin-Binding Domain of Chitinase A1 from Bacillus circulans WL-12. *Journal of Bacteriology*, 182 (11): 3045-3054.

Henrissat, B. (1991). A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.*, 280: 309-316.

Henrissat, B. & Davies, G. (1997). Structural and sequence-based classification of glycoside hydrolases. *Current Opinion in Structural Biology*, 7 (5): 637-644.

Hoell, I., Vaaje-Kolstad, G. & Eijsink, V. G. H. (2010). Structure and function of enzymes acting on chitin and chitosan. *Biotechnol Genet Eng Rev*, 27: 331-66.

Hori, C., Igarashi, K., Katayama, A. & Samejima, M. (2011). Effects of xylan and starch on secretome of the basidiomycete Phanerochaete chrysosporium grown on cellulose. *FEMS Microbiol Lett*, 321 (1): 14-23.

Horn, S. J., Sikorski, P., Cederkvist, J. B., Vaaje-Kolstad, G., Sorlie, M., Synstad, B., Vriend, G., Varum, K. M. & Eijsink, V. G. (2006a). Costs and benefits of processivity in enzymatic degradation of recalcitrant polysaccharides. *Proc Natl Acad Sci U S A*, 103 (48): 18089-94.

Horn, S. J., Sorbotten, A., Synstad, B., Sikorski, P., Sorlie, M., Varum, K. M. & Eijsink, V. G. (2006b). Endo/exo mechanism and processivity of family 18 chitinases produced by Serratia marcescens. *FEBS J*, 273 (3): 491-503.

Horn, S. J., vaaje-Kolstad, G., Westereng, B. & Eijsink, V. (2012). *Novel enzymes for the degradation of cellulose*. Biotechnology for Biofuels. 5.

Hult, E., Katouno, F., Uchiyama, T., Watanabe, T. & Sugiyama, J. (2005). Molecular directionality in crystalline β-chitin: hydrolysis by chitinases A and B from Serratia marcescens 2170. *Biochem. J.*, 388: 851-856.

Jeuniaux, C. (1972). Chitinous structures. *Comprehensive biochemistry*, 26 (C): 595-632.

Krulwich, T. A. & Pate, J. L. (1971). Ultrastructural Explanation for Snapping

Postfission Movements in *Arthrobacter. crystallopoietesJ. Bacteriol.*, 105 (1): 408-412.

Kurašin, M. & Väljamäe, P. (2011). Processivity of Cellobiohydrolases Is Limited by the Substrate. *Journal of Biological Chemistry*, 286 (1): 169-177.

Langston, J. A., Shaghasi, T., Abbate, E., Xu, F., Vlasenko, E. & Sweeney, M. D. (2011). Oxidoreductive Cellulose Depolymerization by the Enzymes Cellobiose Dehydrogenase and Glycoside Hydrolase 61. *Applied and environmental microbiology*, 77 (19): 7007-7015.

Lehtiö, J., Sugiyama, J., Gustavsson, M., Fransson, L., Linder, M. & Teeri, T. T. (2003). The binding specificity and affinity determinants of family 1 and family 3 cellulose binding modules. *Proceedings of the National Academy of Sciences of the United States of America*, 100 (2): 484.

Letourneur, O., Gervasi, G., Gaïa, S., Pagès, J., Watelet, B. & Jolivet, M. (2001). Characterization of Toxoplasma gondii surface antigen 1 (SAG1) secreted from Pichia pastoris: evidence of hyper O-glycosylation. 33 (Pt 1): 35-45.

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

Merino, S. & Cherry, J. (2007). Progress and Challenges in Enzyme Development for Biomass Utilization

Biofuels. I: Olsson, L. (red.) Advances in Biochemical Engineering/Biotechnology, b. 108, s. 95-120: Springer Berlin / Heidelberg.

Merzendorfer, H. & Zimoch, L. (2003). Chitin metabolism in insects: structure, function and regulation of chitin synthases and chitinases. *Journal of Experimental Biology*, 206 (24): 4393-4412.

Minke, R. & Blackwell, J. (1978). The structure of α-chitin. *Journal of Molecular Biology*, 120 (2): 167-181.

Monreal, J. & Reese, E. T. (1969). The chitinase of Serratia marcescens. *Canadian journal of microbiology*, 15 (7): 689-96.

Morris, V. J., Gunning, A. P., Faulds, C. B., Williamson, G. & Svensson, B. (2005). AFM images of complexes between amylose and Aspergillus niger glucoamylase mutants, native and mutant starch binding domains: a model for the action of glucoamylase. *Starch-Stärke*, 57 (1): 1-7.

Papanikolau, Y., Tavlas, G., Vorgias, C. E. & Petratos, K. (2003). *De novo* purification scheme and crystallization conditions yield high-resolution structures of chitinase A and its complex with the inhibitor allosamidin. *Acta Crystallographica, Section D*, 59: 400-403.

Penninga, D., van der Veen, B. A., Knegtel, R., van Hijum, S. A. F. T., Rozeboom, H. J., Kalk, K. H., Dijkstra, B. W. & Dijkhuizen, L. (1996). The raw starch binding domain of cyclodextrin glycosyltransferase from Bacillus circulans strain 251. *Journal of Biological Chemistry*, 271 (51): 32777.

Perrakis, A., Tews, I., Dauter, Z., Oppenheim, A. B., Chet, I., Wilson, K. S. & Vorgias, C. E. (1994). Crystal structure of a bacterial chitinase at 2.3 Å resolution. *Structure*, 2 (12): 1169-1180.

Perrakis, A., Ouzounis, C. & Wilson, K. S. (1997). Evolution of immunoglobulin-like modules in chitinases: their structural flexibility and functional implications. *Folding and Design*, 2 (5): 291-294.

Peters, W. (1972). Occurrence of chitin in Mollusca. *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry*, 41 (3): 541-544, IN15-IN20, 545-550.

Phillips, C. M., Beeson, W. T., Cate, J. H. & Marletta, M. A. (2011). Cellobiose dehydrogenase and a copper-dependent polysaccharide monooxygenase potentiate cellulose degradation by Neurospora crassa. *ACS chemical biology*, 6 (12): 1399-406.

Quinlan, R. J., Sweeney, M. D., Lo Leggio, L., Otten, H., Poulsen, J. C., Johansen, K. S., Krogh, K. B., Jorgensen, C. I., Tovborg, M., Anthonsen, A., et al. (2011). Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. *Proceedings of the National Academy of Sciences of the United States of America*, 108 (37): 15079-84.

Ravi Kumar, M. N. V. (2000). A review of chitin and chitosan applications. *Reactive and Functional Polymers*, 46 (1): 1-27.

Rinaudo, M. (2006). Chitin and chitosan: Properties and applications. *Progress in Polymer Science*, 31 (7): 603-632.

Sanchez-Herrera, L. M., Ramos-Valdivia, A. C., de la Torre, M., Salgado, L. M. & Ponce-Noyola, T. (2007). Differential expression of cellulases and xylanases by Cellulomonas flavigena grown on different carbon sources. *Applied microbiology and biotechnology*, 77 (3): 589-95.

Simpson, P. J., Xie, H., Bolam, D. N., Gilbert, H. J. & Williamson, M. P. (2000). The structural basis for the ligand specificity of family 2 carbohydrate-binding modules. *J Biol Chem*, 275 (52): 41137-42.

Smith, A. M. (2001). Starch and Starch Granules. I: *eLS*: John Wiley & Sons, Ltd.

Sorimachi, K., Jacks, A. J., Le Gal-Coëffet, M. F., Williamson, G., Archer, D. B. & Williamson, M. P. (1996). Solution structure of the granular starch binding domain of glucoamylase from Aspergillus niger by nuclear magnetic resonance spectroscopy. *Journal of Molecular Biology*, 259 (5): 970-987.

Sorimachi, K., Gal-Coëffet, M.-F. L., Williamson, G., Archer, D. B. & Williamson, M. P. (1997). Solution structure of the granular starch binding domain of Aspergillus niger glucoamylase bound to β-cyclodextrin. *Structure*, 5 (5): 647-661.

Southall, S. M., Simpson, P. J., Gilbert, H. J., Williamson, G. & Williamson, M. P. (1999). The starch-binding domain from glucoamylase disrupts the structure of starch. *FEBS Letters*, 447 (1): 58-60.

Suzuki, K., Taiyoji, M., Sugawara, N., Nikaidou, N., Henrissat, B. & Watanabe, T. (1999). The third chitinase gene (*chiC*) of *Serratia marcescens* 2170 and the relationship of its product to other bacterial chitinases. *Biochem. J.*, 343: 587-596.

Suzuki, K., Sugawara, N., Suzuki, M., Uchiyama, T., Katouno, F., Nikaidou, N. & Watanabe, T. (2002). Chitinases A, B and C1 of *Serratia marcescens* 2170 produced by recombinant *Escherichia coli*: enzymatic properties and synergism n chitin degradation. *Biosci Biotechnol Biochem*, 66 (5): 1075-1083.

Sørlie, M., Zakariassen, H., Norberg, A. L. & Eijsink, V. G. H. (2012). Processivity and substrate-binding in family 18 chitinases. *Biocatalysis and Biotransformation*: 1-13.

Taylor, C. B., Talib, M. F., McCabe, C., Bu, L., Adney, W. S., Himmel, M. E., Crowley, M. F. & Beckham, G. T. (2012). Computational investigation of glycosylation effects on a family 1 carbohydrate-binding module. *The Journal of biological chemistry*, 287 (5): 3147-55.

Vaaje-Kolstad, G. (2005). *The chitinolytic machinery of Serratia marcescens : the catalytic mechanism of chitinase B and the function of the chitin-binding protein, CBP21*. Ås: Protein Engineering and Proteomics Group, Department of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences. 1 b. (flere pag.), ill. s.

Vaaje-Kolstad, G., Horn, S. J., van Aalten, D. M., Synstad, B. & Eijsink, V. G. (2005a). The non-catalytic chitin-binding protein CBP21 from Serratia marcescens is essential for chitin degradation. *J Biol Chem*, 280 (31): 28492-7.

Vaaje-Kolstad, G., Houston, D. R., Riemen, A. H., Eijsink, V. G. & van Aalten, D. M. (2005b). Crystal structure and binding properties of the Serratia marcescens chitin-binding protein CBP21. *J Biol Chem*, 280 (12): 11313-9.

Vaaje-Kolstad, G., Westereng, B., Horn, S. J., Liu, Z., Zhai, H., Sorlie, M. & Eijsink, V. G. (2010). An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. *Science*, 330 (6001): 219-22.

Vaaje-Kolstad, G., Bøhle, L. A., Gåseidnes, S., Dalhus, B., Bjørås, M., Mathiesen, G. & Eijsink, V. G. H. (2012). Characterization of the Chitinolytic Machinery of Enterococcus faecalis V583 and High-Resolution Structure of Its Oxidative CBM33 Enzyme. *Journal of Molecular Biology*, 416 (2): 239-254.

van Aalten, D. M. F., Synstad, B., Brurberg, M. B., Hough, E., Riise, B. W., Eijsink, V. G. H. & Wierenga, R. K. (2000). Structure of a two-domain chitotriosidase from Serratia marcescens at 1.9-Å resolution. *Proceedings of the National Academy of Sciences*, 97 (11): 5842-5847.

Watanabe, T., Ito, Y., Yamada, T., Hashimoto, M., Sekine, S. & Tanaka, H. (1994). The roles of the C-terminal domain and type III domains of chitinase A1 from Bacillus circulans WL-12 in chitin degradation. *Journal of Bacteriology*, 176 (15): 4465-4472.

Westereng, B., Ishida, T., Vaaje-Kolstad, G., Wu, M., Eijsink, V. G. H., Igarashi, K., Samejima, M., Ståhlberg, J., Horn, S. J. & Sandgren, M. (2011). The Putative Endoglucanase PcGH61D from <italic>Phanerochaete chrysosporium</italic> Is a Metal-Dependent Oxidative Enzyme that Cleaves Cellulose. *PLoS ONE*, 6 (11): e27807.

Xu, G. Y., Ong, E., Gilkes, N. R., Kilburn, D. G., Muhandiram, D., Harris-Brandts, M., Carver, J. P., Kay, L. E. & Harvey, T. S. (1995). Solution structure of a cellulose-binding domain from Cellulomonas fimi by nuclear magnetic resonance spectroscopy. *Biochemistry*, 34 (21): 6993-7009.

Zakariassen, H., Aam, B. B., Horn, S. J., Vårum, K. M., Sørlie, M. & Eijsink, V. G. H. (2009). Aromatic Residues in the Catalytic Center of Chitinase A from Serratia marcescens Affect Processivity, Enzyme Activity, and Biomass Converting Efficiency. *Journal of Biological Chemistry*, 284 (16): 10610-10617.

Zakariassen, H., Eijsink, V. G. H. & Sørlie, M. (2010). Signatures of activation parameters reveal substrate-dependent rate determining steps in polysaccharide turnover by a family 18 chitinase. *Carbohydrate Polymers*, 81 (1): 14-20.

Appendix A

Codon optimized *Atcbm33A*. *Atcbm33A* was codon optimized according to the codon usage bias in *Pichia pastoris* and synthesized by GenScript's OptimumGene™ Gene Design software (http://www.genscript.com/codon_opt.html). Following is the codon optimized sequence.

ATGTTGCTGACTGTATTGGCGGTCGTCGGTTGTTTCACGGCTGTAAATGGTCACGGCTACTTGACTATTCCTGCGTCCAGAACTCGCCTGGGTTTTGAAACCGGCATTGACACTTGTCCGGAATGC
AGCATCCTGGAACCTGTTACCGCTTGGCCAGATCTGGAAGCTGCACAAGTGGGCCGCTCAGGTCCTTGTGGCTATAATGCACGTGTTTCCGTGGATTATAACCAGCCGTCGGAATACTGGGGTAA
TGAACCTGTTGTGACATACACGTCTGGCGAAGTCGTAGAAGTTCAATGGTGTGTGGATGCCAACGGTGACCATGGTGGCATGTTTACCTATGGCATTTGCCAAAACCAGACTTTAGTGGATAAGT
TCTTGACACCAGGTTACCTGCCGACGAACGAAGAAAAGCAAGCCGCGGAAGATTGCTTCTTGGACGGCGAACTGAAGTGTAAGGATGTCTCTGGTCAGACATGTGGCTATAACCCAGACTGCAC
GGAAGGTGCTGCATGTTGGAGAAATGATTGGTTTACTTGCAATGCTTTCCAAGCAAACACAGCACGTGCATGTCAGGGTGTCGATGGTGCTAGTTTAAACTCCTGCAAAACCACTATCGCAGGTG
GCTATACCGTAACTAAACGTATTAAGATCCCGGATTACTCTTCAGACCATACCCTGTTACGTTTTAGATGGAATAGTTTCCAAACTGCCCAGGTTTACCTGCACTGTGCGGATATTGCTATCGCAG
GTTCCGGTGGCGGTACAACGTCCAAATCGACCACTTCGACAACGTCTACCACTTCTACCTCAAGAAGCACAAGTACGTCCGCCCCTACAACGACCAGCAGTGCCTCAACCGCGACTCCTATTTGC
ACTACACAAGCGAGCCTGATCCCAGTCACTTTTCAGGAATTCGTAACGACCATGTGGGGTGAAAACGTCTTTGTAACAGGTTCGATTTCTCAGTTGGGCTCATGGAGCACTGACAAAGCCGTGGC
GCTGTCCGCTACAGGTTATACGGCATCTAATCCATTATGGACTACAACGATTGATTTGCCGGCTGGTACCACTTTCGAATACAAGTTCATCAAAAAGGAAACCGATGGCTCAATTATCTGGGAAA
GCGACCCGAATCGCAGTTACACCGTCCCGACTGGTTGTTCTGGCACTACCGCTACCGCAGCCGCTTCTTGGCGTTGA

Appendix B

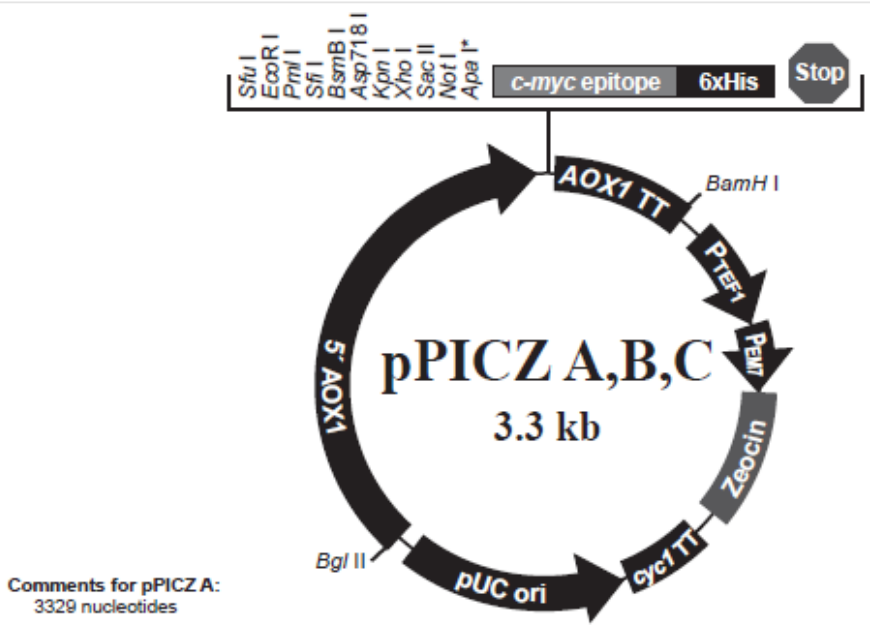
Name: cfcbm33A Length: 400
MLLTVLAVVGCFTAVNGHGYLTIPASRTRLGFETGIDTCPECSILEPVTAWPDLEAA
QVGRSGPCGYNARVSVDYNQPSE 80
YWGNEPVVTYTSGEVVEVQWCVDANGDHGGMFTYGICQNQTLVDKFLTPGYLPTN
EEKQAAEDCFLDGELKCKDVSGQTC 160
GYNPDCTEGAACWRNDWFTCNAFQANTARACQGVDGASLNSCKTTIAGGYTVTKR
IKIPDYSSDHTLLRFRWNSFQTAQV 240
YLHCADIAIAGSGGGTTSKSTTSTTSTTSTSRSTSTSAPTTTSSASTATPICTTQASLIPV
TFQEFVTTMWGENVFVTGS 320
ISQLGSWSTDKAVALSATGYTASNPLWTTTIDLPAGTTFEYKFIKKETDGSIIWESDP
NRSYTVPTGCSGTTATAAASWR 400
..... 80
.....N..... 160
..... 240
..... 320
.....N..... 400

(Threshold=0.5)

SeqName	Position	Potential	Jury agreement	N-Glyc result
cfcbm33A	119	NQTL	0.6798 (9/9)	++
cfcbm33A	379	NRSY	0.5152 (5/9)	+

Figure B.1. Output from prediction of N-glycosylations in *At*CBM33A from NetNGlyc (http://www.cbs.dtu.dk/services/NetNGlyc/; left) and the selections made when using GlyProt (www.glycoscience.de/modeling/glyprot/php/main.php) to predict the connection of an oligomannose to Asn379 in the homologymodel of the CBM20 domain in *At*CBM33A (right).

Appendix C



Comments for pPICZ A:
3329 nucleotides

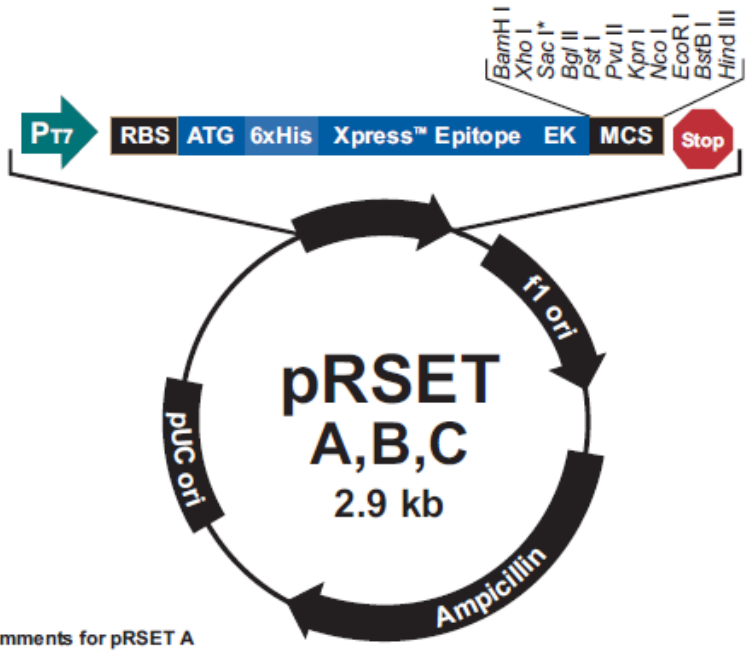
5′ AOX1 promoter region: bases 1-941
5′ end of *AOX1* mRNA: base 824
5′ AOX1 priming site: bases 855-875
Multiple cloning site: bases 932-1011
c-myc epitope tag: bases 1012-1044
Polyhistidine tag: bases 1057-1077
3′ AOX priming site: bases 1159-1179
3′ end of mRNA: base 1250
AOX1 transcription termination region: bases 1078-1418
Fragment containing TEF1 promoter: bases 1419-1830
EM7 promoter: bases 1831-1898
Sh ble ORF: bases 1899-2273
CYC1 transcription termination region: bases 2274-2591
pUC origin: bases 2602-3275 (complementary strand)

* The restriction site between *Not* I and the *myc* epitope is different in each version of pPICZ:
Apa I in pPICZ A
Xba I in pPICZ B
*Sna*B I in pPICZ C

pPICZα A MCS

```
                                5' end of AOX1 mRNA                                          5' AOX1 priming site
811  AACCTTTTTT TTTATCATCA TTATTAGCTT ACTTTCATAA TTGCGACTGG TTCCAATTGA

871  CAAGCTTTTG ATTTTAACGA CTTTTAACGA CAACTTGAGA AGATCAAAAA ACAACTAATT

931  ATTCGAAACG ATG AGA TTT CCT TCA ATT TTT ACT GCT GTT TTA TTC GCA GCA
                Met Arg Phe Pro Ser Ile Phe Thr Ala Val Leu Phe Ala Ala

                                                                                             α-factor signal sequence
983  TCC TCC GCA TTA GCT GCT CCA GTC AAC ACT ACA ACA GAA GAT GAA ACG GCA
     Ser Ser Ala Leu Ala Ala Pro Val Asn Thr Thr Thr Glu Asp Glu Thr Ala

1034 CAA ATT CCG GCT GAA GCT GTC ATC GGT TAC TCA GAT TTA GAA GGG GAT TTC
     Gln Ile Pro Ala Glu Ala Val Ile Gly Tyr Ser Asp Leu Glu Gly Asp Phe

1085 GAT GTT GCT GTT TTG CCA TTT TCC AAC AGC ACA AAT AAC GGG TTA TTG TTT
     Asp Val Ala Val Leu Pro Phe Ser Asn Ser Thr Asn Asn Gly Leu Leu Phe

                                                                                                  Xho I*
1136 ATA AAT ACT ACT ATT GCC AGC ATT GCT GCT AAA GAA GAA GGG GTA TCT CTC
     Ile Asn Thr Thr Ile Ala Ser Ile Ala Ala Lys Glu Glu Gly Val Ser Leu

                Kex2 signal cleavage                    EcoR I   Pml I          Sfi I                BsmB I  Asp718 I
1187 GAG AAA AGA GAG GCT GAA GCT CGAATTCAC GTGGCCCAG CCGGCCGTC TCGGATCGGT
     Glu Lys Arg Glu Ala Glu Ala
                          Ste13 signal cleavage
     Kpn I  Xho I      Sac II  Not I              Xba I                    c-myc epitope
1244 ACCTCGAGCC GCGGCGGCCG GCCAGCTTTC TA GAA CAA AAA CTC ATC TCA GAA GAG
                                                Glu Gln Lys Leu Ile Ser Glu Glu
                                                          polyhistidine tag
1299 GAT CTG AAT AGC GCC GTC GAC CAT CAT CAT CAT CAT CAT TGA GTTTGTAGCC
     Asp Leu Asn Ser Ala Val Asp His His His His His His ***

1351 TTAGACATGA CTGTTCCTCA GTTCAAGTTG GGCACTTACG AGAAGACCGG TCTTGCTAGA

                                3' AOX1 priming site
1411 TTCTAATCAA GAGGATGTCA GAATGCCATT TGCCTGAGAG ATGCAGGCTT CATTTTTGAT

                                                                    3' polyadenylation site
1471 ACTTTTTTAT TTGTAACCTA TATAGTATAG GATTTTTTTT GTCATTTTGT TTCTTCTCGT
```

Figure C.1. Summary of the pPICZα-A, B and C vectors (left) and the specific sequence of the multiple cloning site in pPICZα-A with labeled restriction sites and functions in or near the cloning site (right). Both figures are retrieved from Invitrogen (www.invitrogen.com).



Comments for pRSET A
2897 nucleotides

*Version C does not contain *Sac* I

T7 promoter: bases 20-39
6xHis tag: bases 112-129
T7 gene 10 leader: bases 133-162
Xpress™ epitope: bases 169-192
Multiple cloning site: bases 202-248
T7 reverse priming site: bases 295-314
T7 transcription terminator: bases 256-385
f1 origin: bases 456-911
bla promoter: bases 943-1047
Ampicillin (*bla*) resistance gene (ORF): bases 1042-1902
pUC origin: bases 2047-2720 (C)

pRSET B Multiple Cloning Site

T7 promoter — RBS

21 AATACGACTC ACTATAGGGA GACCACAACG GTTTCCCTCT AGAAATAATT TTGTTTAACT TTAAGAAGGA

Polyhistidine (6xHis) region

91 GATATACAT **ATG** CGG GGT TCT CAT CAT CAT CAT CAT CAT GGT ATG GCT AGC ATG ACT
Met Arg Gly Ser His His His His His His Gly Met Ala Ser Met Thr

T7 gene 10 leader — Xpress™ Epitope — *Bam*H I — *Xho* I *Sac* I

148 GGT GGA CAG CAA ATG GGT CGG GAT CTG TAC GAC GAT GAC GAT AAG GAT CCG AGC TCG
Gly Gly Gln Gln Met Gly Arg Asp Leu Tyr Asp Asp Asp Asp Lys Asp Pro Ser Ser

EK recognition site — EK cleavage site

Bgl II *Pst* I *Pvu* II *Kpn* I *Nco* I *Eco*R I *Bst*B I *Hind* III

205 AGA TCT GCA GCT GGT ACC ATG GAA TTC GAA GCT TGA TCCGGCTGCT AACAAAGCCC
Arg Ser Ala Ala Gly Thr Met Glu Phe Glu Ala ***

T7 reverse priming site

261 GAAAGGAAGC TGAGTTGGCT GCTGCCACCG CTGAGCAATA ACTAGCATAA

Figure C.2. Summary of the pRSET A, B and C vectors (left) and the specific sequence of the multiple cloning site in pPICZα-A with labeled restriction sites and functions in or near the cloning site (right). Both figures are retrieved from Invitrogen (www.invitrogen.com).

Appendix D

Chromatograms from protein purification using ion exchange chromatography, size exclusion chromatography, immobilized metal affinity chromatography or affinity chromatography with chitin beads.

E

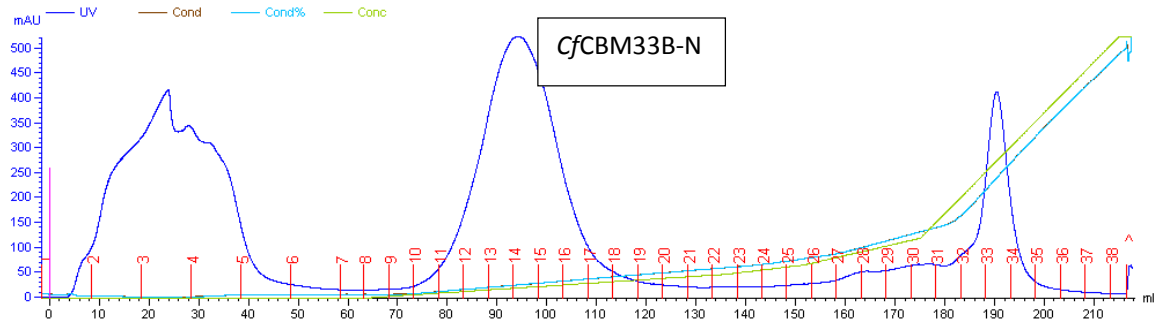


Figure D.1. Ion exchange chromatography used for purification of *Cf*CBM33B-N. Fraction numbers are shown in red. The chromatogram for the purification of *Cf*CBM33A-N by ion exchange chromatography showed a similar elution pattern. Both *Cf*CBM33s were purified using a HiTrap™ 5 ml column. Sterile-filtered periplasmic extract *Cf*CBM33s (50 sample volume) was applied to the column at pH 7.5 with a flow of 1 ml/min, followed by 2 column volumes of 50 mM Tris-HCl pH 7.5 and the protein was eluted during a gradient from 0 % to 50 % elution buffer during 200 minutes at 0.4 ml/min flow rate. Eluted proteins were detected by online monitoring of the absorption at 280 nm and collected in 1 ml fractions.

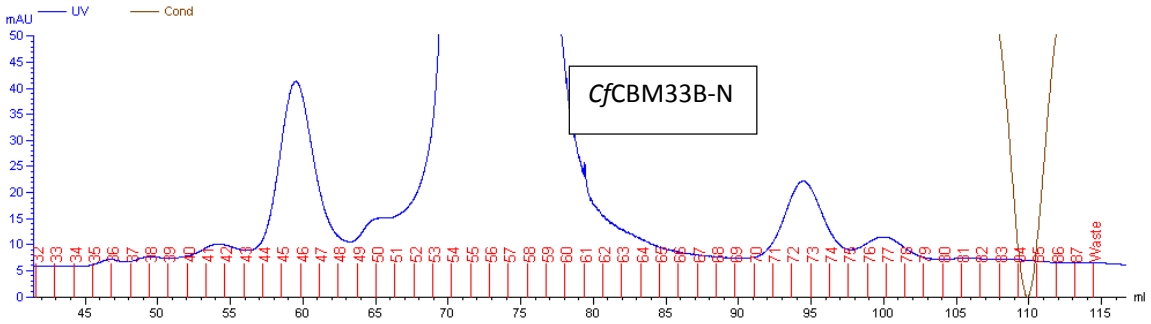


Figure D.2. Size exclusion chromatography used for further purification of *Cf*CBM33B-N, after ion exchange chromatography. Fraction numbers are shown in red. Fractions containing the protein of interest from ion exchange chromatography was pooled and concentrated 7 times. The sample (1 ml) was applied through a 2 ml loading loop to a HiLoad 16/60 Superdex G-75 column at a 0.3 ml/min flow rate, followed by application of 3 column volumes of running buffer to elute the protein (collected in fractions of 5 ml).

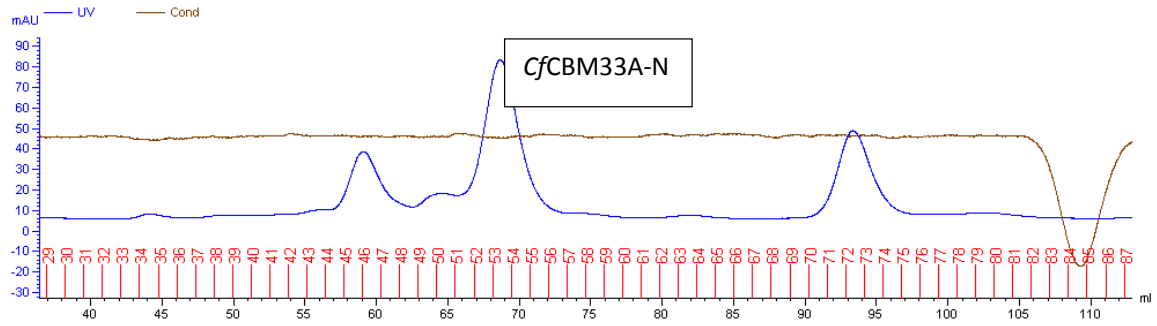


Figure D.3. Size exclusion chromatography used for purification of *Cf*CBM33A-N, semi-purified by ion exchange. Fraction numbers are shown in red. Fractions containing the protein of interest from ion exchange chromatography was pooled and concentrated 7 times. The sample (1 ml) was applied through a 2 ml loading loop to a HiLoad 16/60 Superdex G-75 column at a 0.3 ml/min flow rate, followed by application of 3 column volumes of running buffer to elute the protein (collected in fractions of 5 ml).

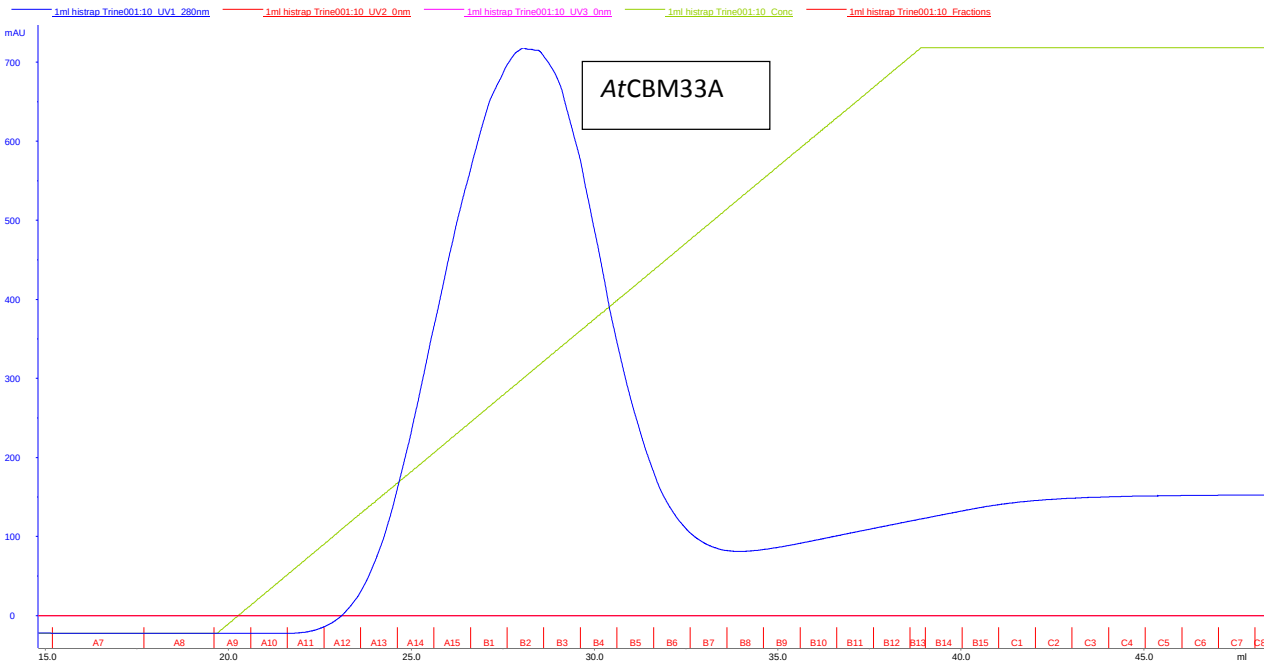


Figure D.4. Chromatogram from the purification of *At*CBM33A^{His} using a His-Trap™ HP column. *At*CBM33A^{His} was purified from 30 ml of the secreted proteins, concentrated 10 times using a HisTrap™ HP (1 ml) column. The sample was loaded onto the column and binding buffer was run until the baseline was stable at a low UV signal. The protein was eluted through a linear gradient from 0 % to 100 % by 20 column volumes elution buffer. The His-tagged protein was eluted at approximately 10% elution buffer and was collected in fractions of 1 ml.

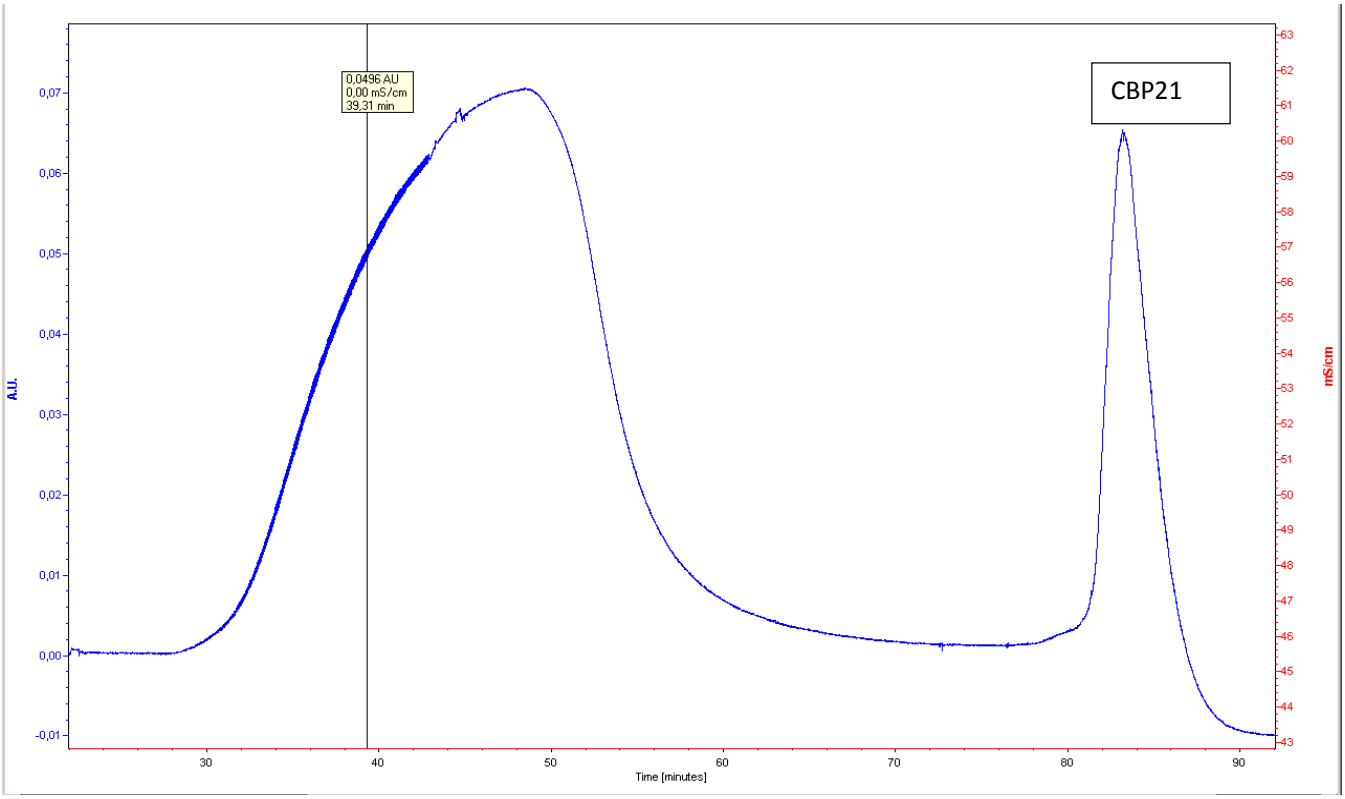


Figure D.5. Chromatogram from the purification of CBP21 by affinity chromatography using chitin beads. CBP21 was purified from periplasmic extract adjusted to 1 M ammonium sulphate and 50 mM Tris-HCl (both final concentration) at pH 8.0 (according to section 2.15.4). The sample was loaded onto the column at a flow rate of 1 ml/min, followed by 3 column volumes of binding buffer (50 mM Tris-HCl and 1 M ammonium sulphate pH 8.0). CBP21 was then eluted (and collected) by during application of one column volume of 20 mM acetic acid.

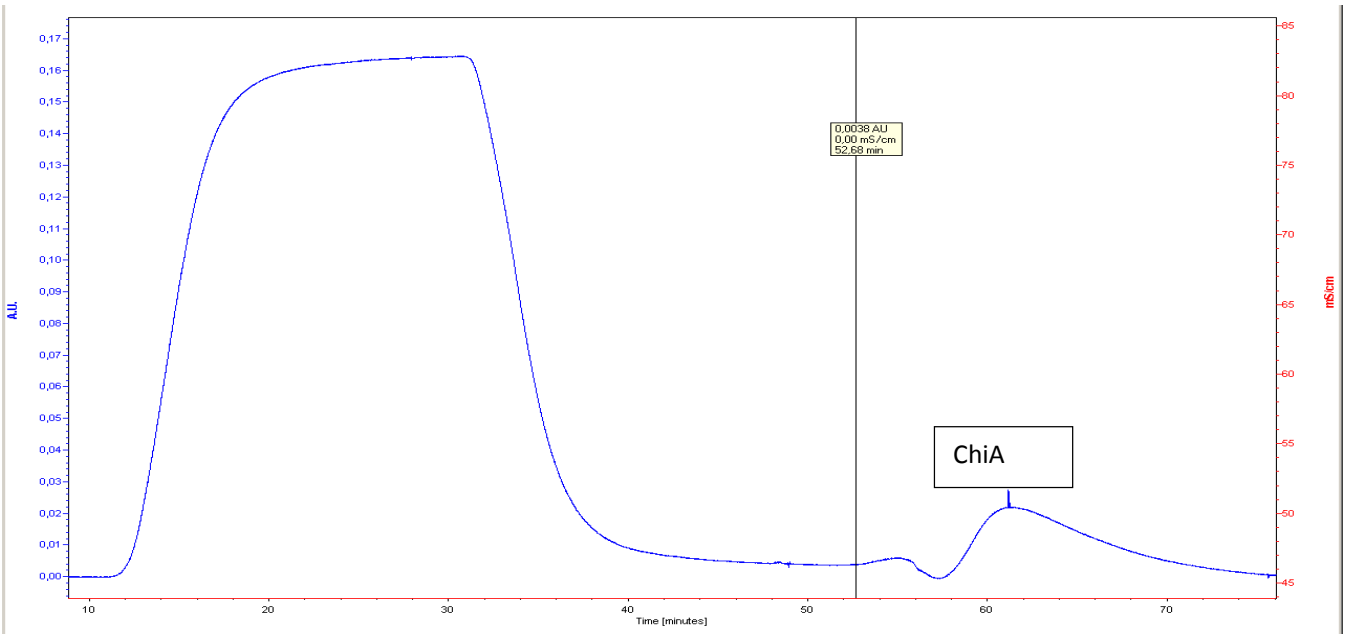


Figure D.6. Chromatogram from the purification of ChiAby affinity chromatography using chitin beads. ChiA was purified from periplasmic extract adjusted to 20 mM Tris-HCl (both final concentration) at pH 8.0 (according to section 2.15.4). The sample was loaded onto the column at a flow rate of 1 ml/min, followed by 3 column volumes of 50 mM Tris-HCl pH 8.0. CBP21 was then eluted (and collected) by during application of one column volume of 20 mM acetic acid.

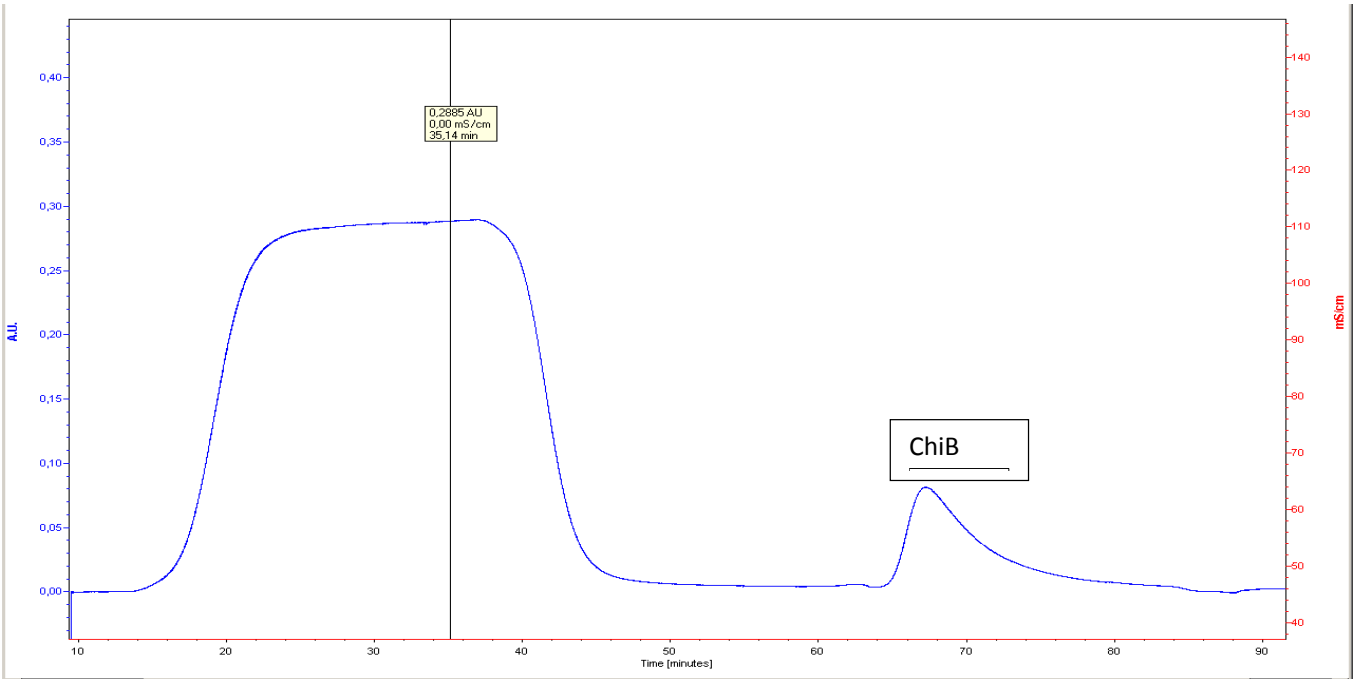


Figure D.7. Chromatogram from the purification of ChiB by affinity chromatography using chitin beads. ChiB was purified from periplasmic extract adjusted to 20 mM Tris-HCl (both final concentration) at pH 8.0 (according to section 2.15.4). The sample was loaded onto the column at a flow rate of 1 ml/min, followed by 3 column volumes of 50 mM Tris-HCl pH 8.0. CBP21 was then eluted (and collected) by during application of one column volume of 20 mM acetic acid.

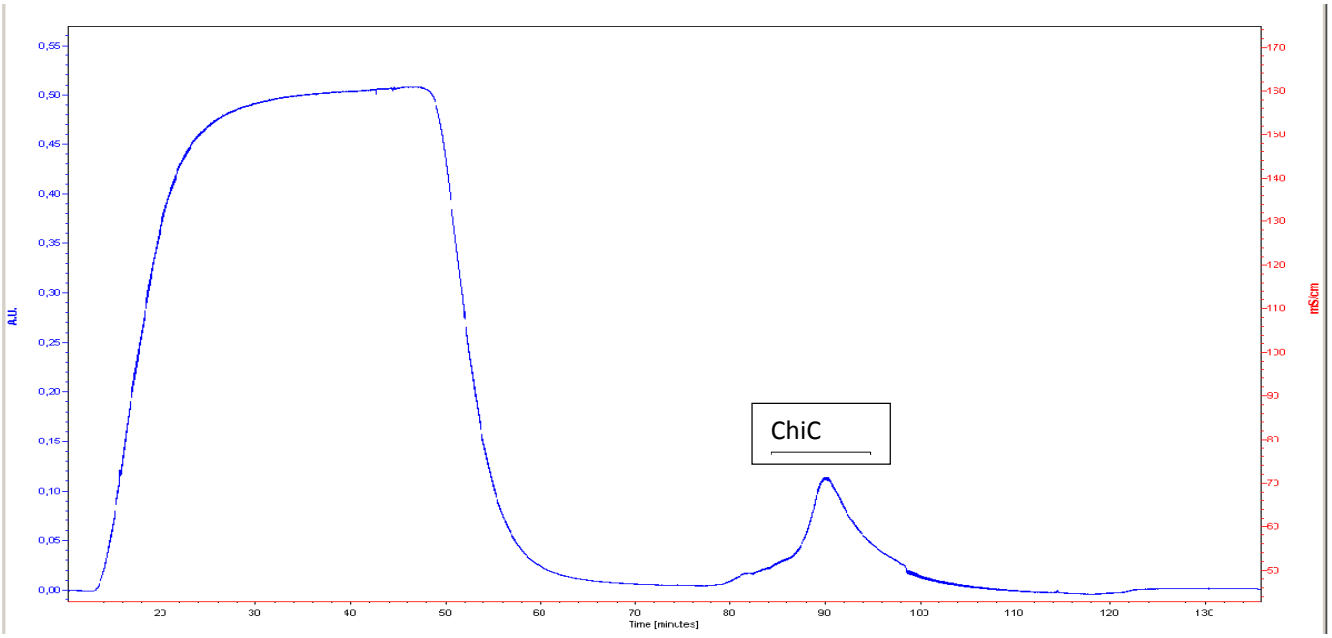
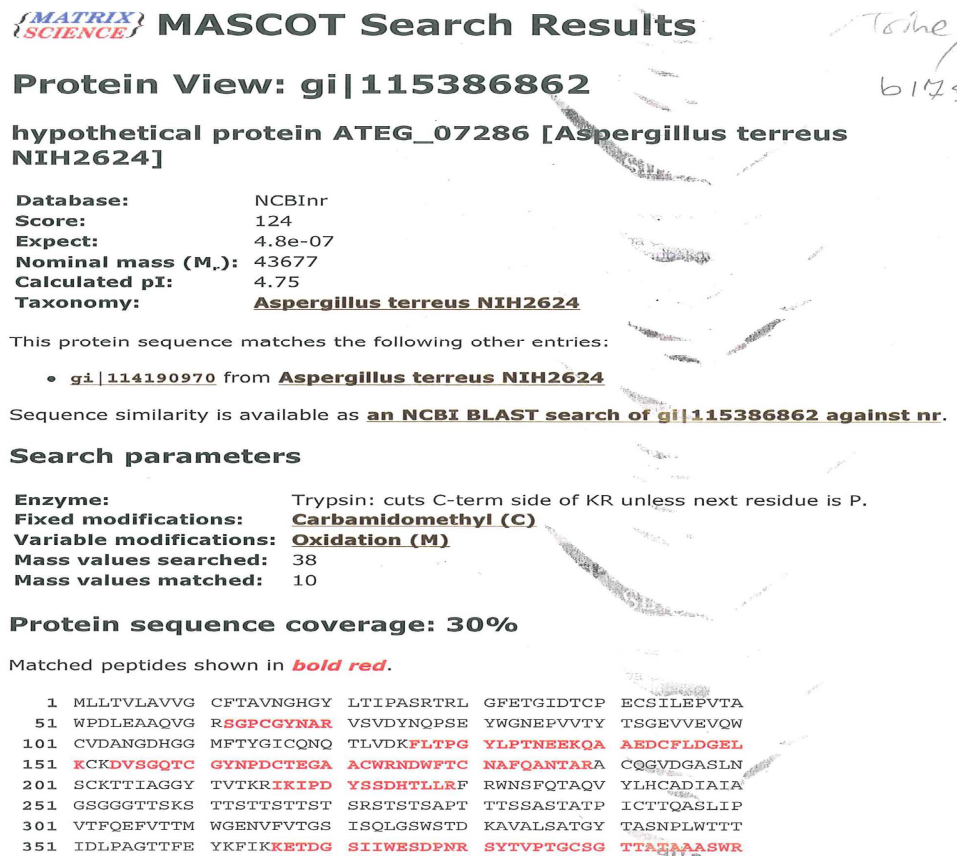


Figure D.8. Chromatogram from the purification of ChiC by affinity chromatography using chitin beads. ChiC was purified from periplasmic extract adjusted to 20 mM Tris-HCl (both final concentration) at pH 8.0 (according to section 2.15.4). The sample was loaded onto the column at a flow rate of 1 ml/min, followed by 3 column volumes of 50 mM Tris-HCl pH 8.0. CBP21 was then eluted (and collected) by during application of one column volume of 20 mM acetic acid.

Appendix E: Protein identification by fragment analysis



MASCOT Search Results

Protein View: gi|115386862

hypothetical protein ATEG_07286 [Aspergillus terreus NIH2624]

Database:	NCBInr
Score:	124
Expect:	4.8e-07
Nominal mass (M_r):	43677
Calculated pI:	4.75
Taxonomy:	Aspergillus terreus NIH2624

This protein sequence matches the following other entries:

- gi|114190970 from Aspergillus terreus NIH2624

Sequence similarity is available as an NCBI BLAST search of gi|115386862 against nr.

Search parameters

Enzyme:	Trypsin: cuts C-term side of KR unless next residue is P.
Fixed modifications:	Carbamidomethyl (C)
Variable modifications:	Oxidation (M)
Mass values searched:	38
Mass values matched:	10

Protein sequence coverage: 30%

Matched peptides shown in ***bold red***.

```
  1  MLLTVLAVVG  CFTAVNGHGY  LTIPASRTRL  GFETGIDTCP  ECSILEPVTA
 51  WPDLEAAQVG  RSGPCGYNAR  VSVDYNQPSE  YWGNEPVVTY  TSGEVVEVQW
101  CVDANGDHGG  MFTYGICQNQ  TLVDKFLTPG  YLPTNEEKQA  AEDCFLDGEL
151  KCKDVSGQTC  GYNPDCTEGA  ACWRNDWFTC  NAFQANTARA  CQGVDGASLN
201  SCKTTIAGGY  TVTKRIKIPD  YSSDHTLLRF  RWNSFQTAQV  YLHCADIAIA
251  GSGGGTTSKS  TTSTTSTTST  SRSTSTSAPT  TTSSASTATP  ICTTQASLIP
301  VTFQEFVTTM  WGENVFVTGS  ISQLGSWSTD  KAVALSATGY  TASNPLWTTT
351  IDLPAGTTFE  YKFIKKETDG  SIIWESDPNR  SYTVPTGCSG  TTATAAAASWR
```

Figure E.1. Output from the Mascot-search towards the NCBI database after MALDI-.TOF MS analysis of trypsinated *At*CBM33A. The results show a positively match against *At*CBM33A (ATEG_07286)

Appendix F: Results from ITC.

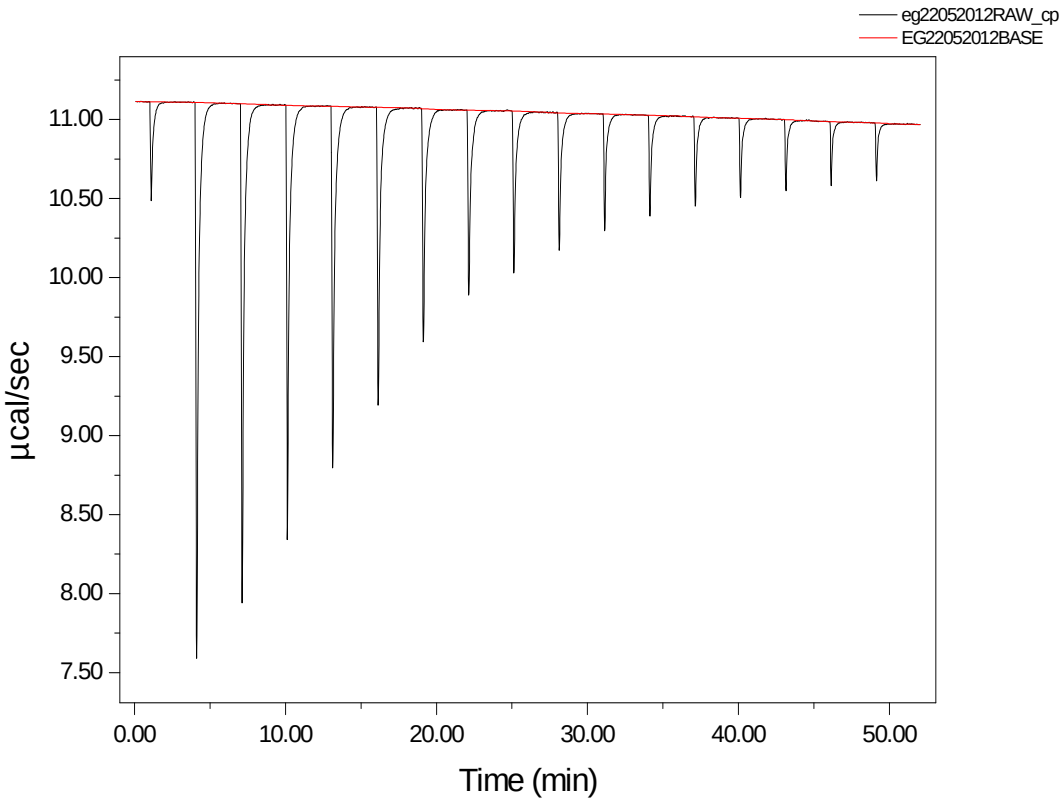


Figure F.1. Raw data from ITC experiment showing the released heat upon binding of *At*CBM33A^{His} to β-cyclodextrin. This work was performed by M. Abou Hachem, Technical University of Denmark.

Appendix G: Results from MALDI-TOF MS.

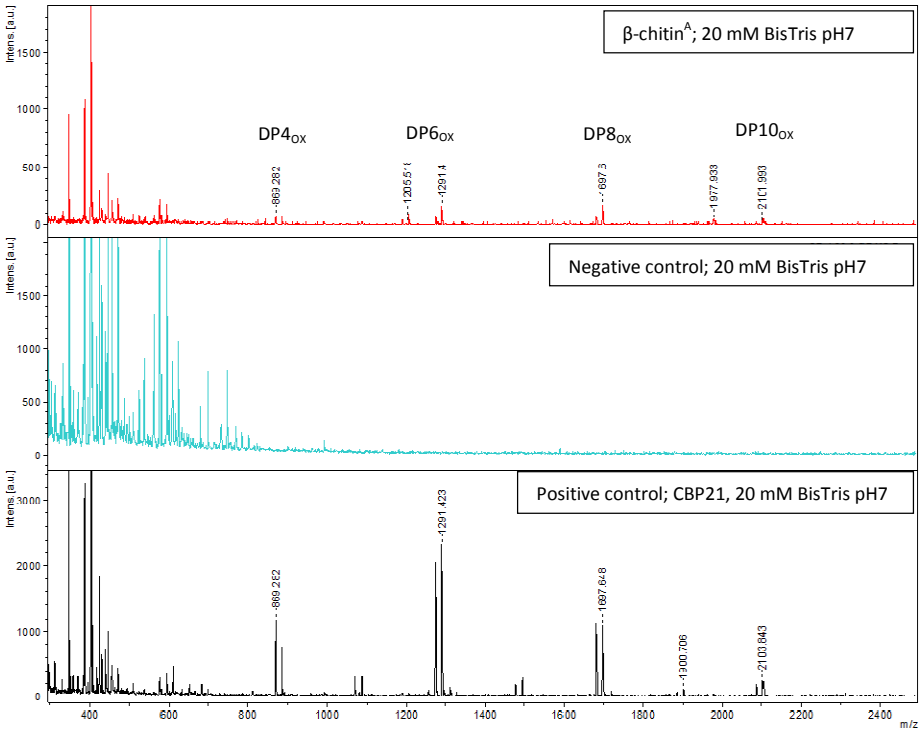


Figure G.1. MALDI-TOF analysis of oligomeric products released upon incubation of 2 mg/ml β-chitin^{A} with *Cf*CBM33B-N(upper panel) or CBP21 (lower panel; positive control)in 20 mM Bist-tris pH 7.0. The middle spectrum shows a negative control, with no enzyme added. In all reactions 2.5 mM ascorbic acid was added as reducing agent. CBP21 degradation of β-chitin^{A} as in this experiment was used as positive control in all subsequent experiments. The peaks are labeled by their m/z values and/or the degree of polymerization (DP). The subscript "ox" means that the oligomer is oxidized. Predicted masses were taken Table S1in Vaaje-Kolstad et al., 2010.

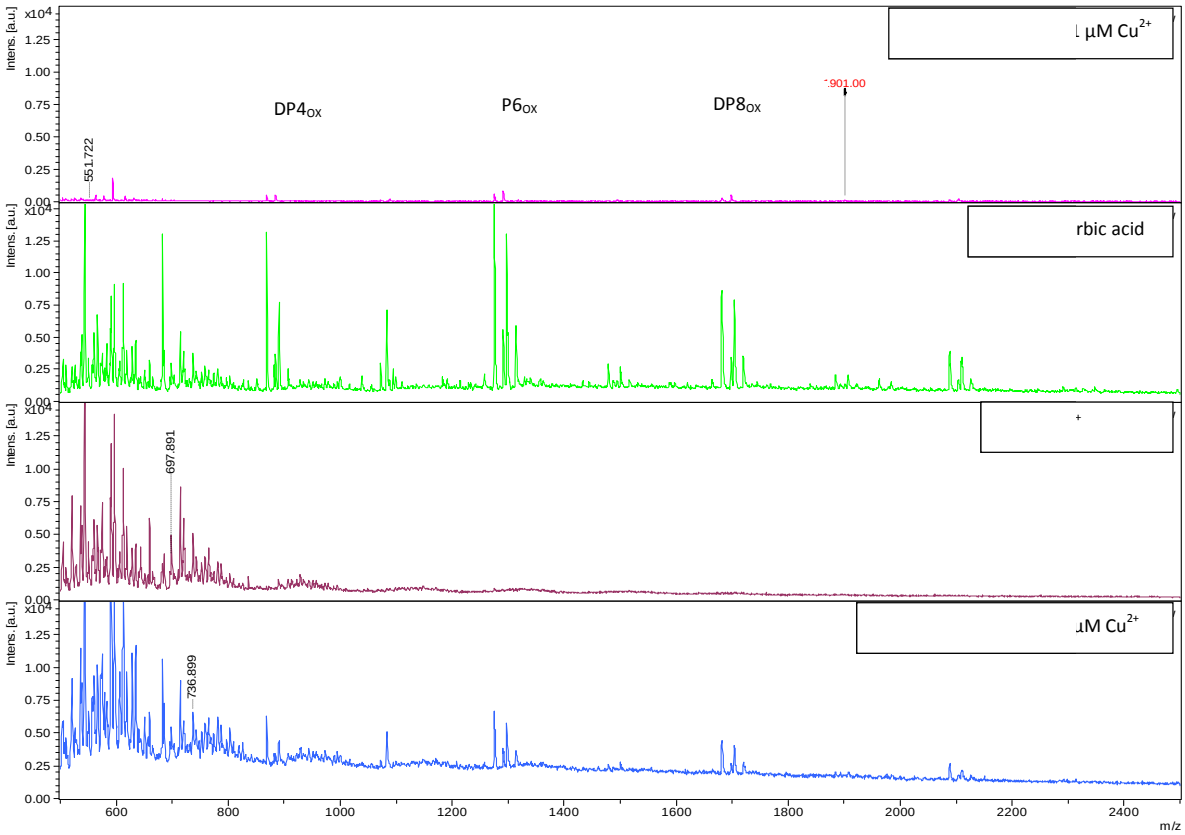


Figure G.2. MALDI-TOF analysis of oligomeric products released upon incubation of 2 mg/ml β-chitin^{A} with 1 µM CBP21 in 20 mM Bist-tris pH 7.0 with 1 mM ascorbic acid as reducing agent (the upper and the two lower panels) and without reducing agent (second lowest panel). 1 µM $^{Cu2+}$ was added to the reactions wich spectra is shown in the upper and the second lowest panel. The peaks are labeled by the degree of polymerization (DP, shown in upper panel) and for corresponding m/z-values, see Figure G.1.. The subscript "ox" means that the oligomer is oxidized.

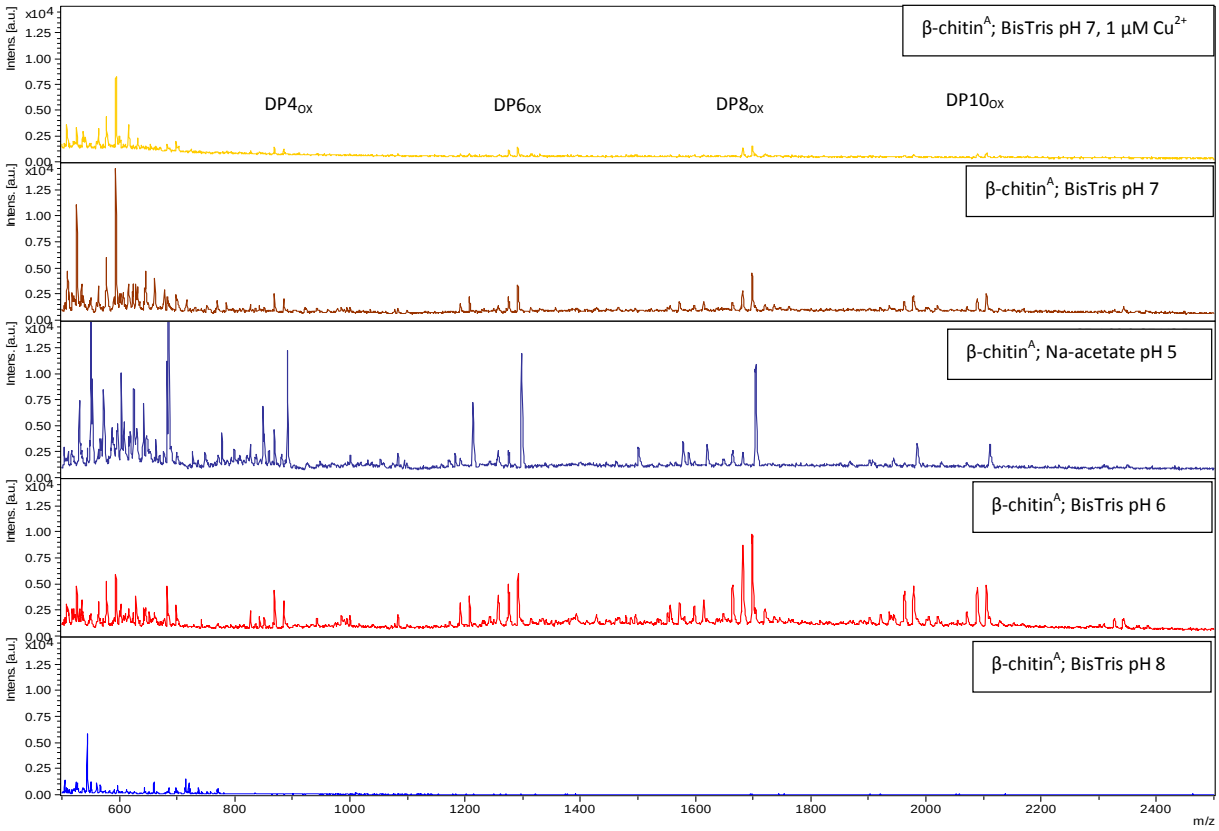


Figure G.3. The degradation of β-chitin^{A} by 1 µM *Cf*CBM33B-N varies with pH and is affected by the presence of Cu^{2+}. MALDI-TOF analysis of oligomeric products released upon incubation of 2 mg/ml β-chitin^{A} with 1 µM *Cf*CBM33B-N in 20 mM Na-acetate pH 5 (middle panel) and Bis-Tris pH 7.0 (to upper panels) pH 6 (second lowest panel) and pH 8 (lower panel) with 1 mM ascorbic acid as reducing agent. The upper panel show a reaction incubated in the presence of 1 µM Cu^{2+}. The peaks are labeled by the degree of polymerization (DP, shown in upper panel), and for corresponding m/z-values, see Figure G.1.

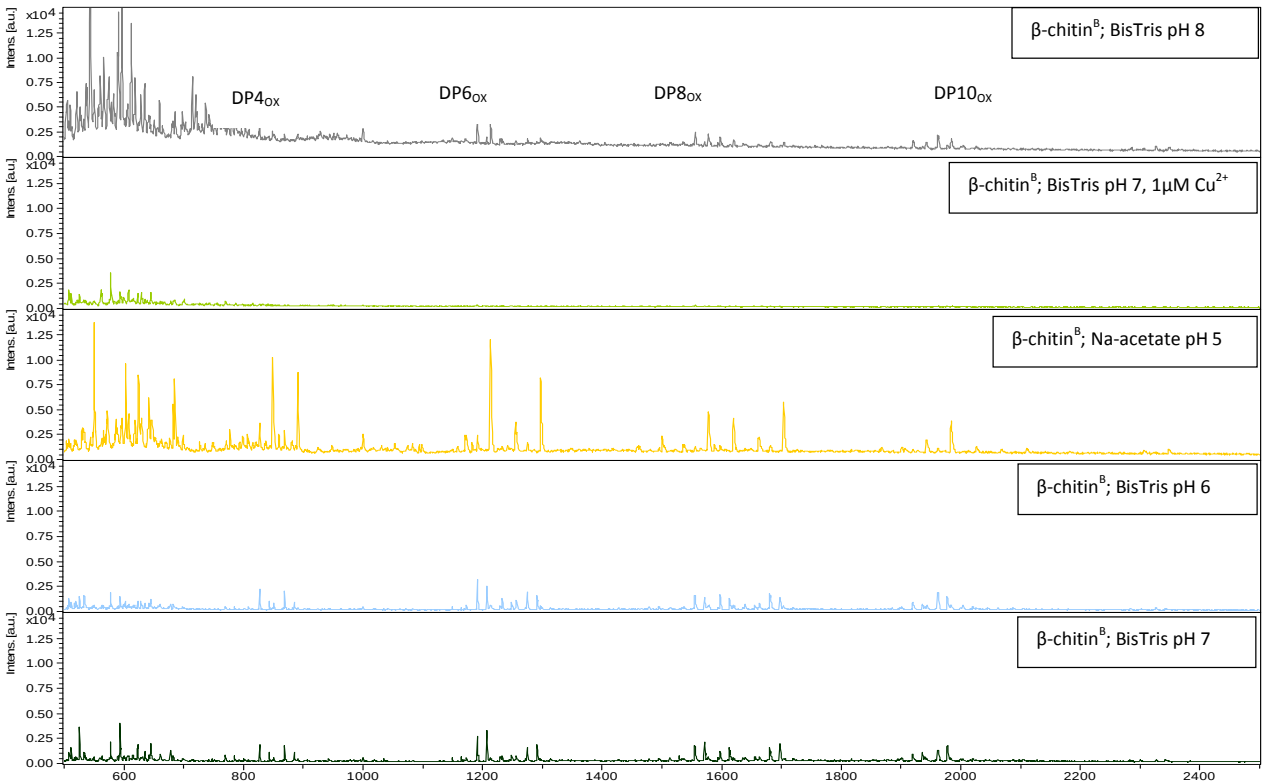


Figure G.4. The degradation of β-chitin^{B} by 1 µM *Cf*CBM33B-N varies with pH, and is affected by the presence of Cu^{2+}. MALDI-TOF analysis of oligomeric products released upon incubation of 2 mg/ml β-chitin^{B} with 1 µM *Cf*CBM33B-N in 20 mM Na-acetate pH 5 (middle panel) and Bis-Tris pH 6.0 (second lowest panel) pH 7.0 (second upper and lower panels) and pH 8 (upper panel) with 1 mM ascorbic acid as reducing agent. The second upper panel show a reaction incubated in the presence of 1 µM Cu^{2+}. The peaks are labeled by the degree of polymerization (DP, shown in upper panel), and for corresponding m/z-values, see Figure G.1.

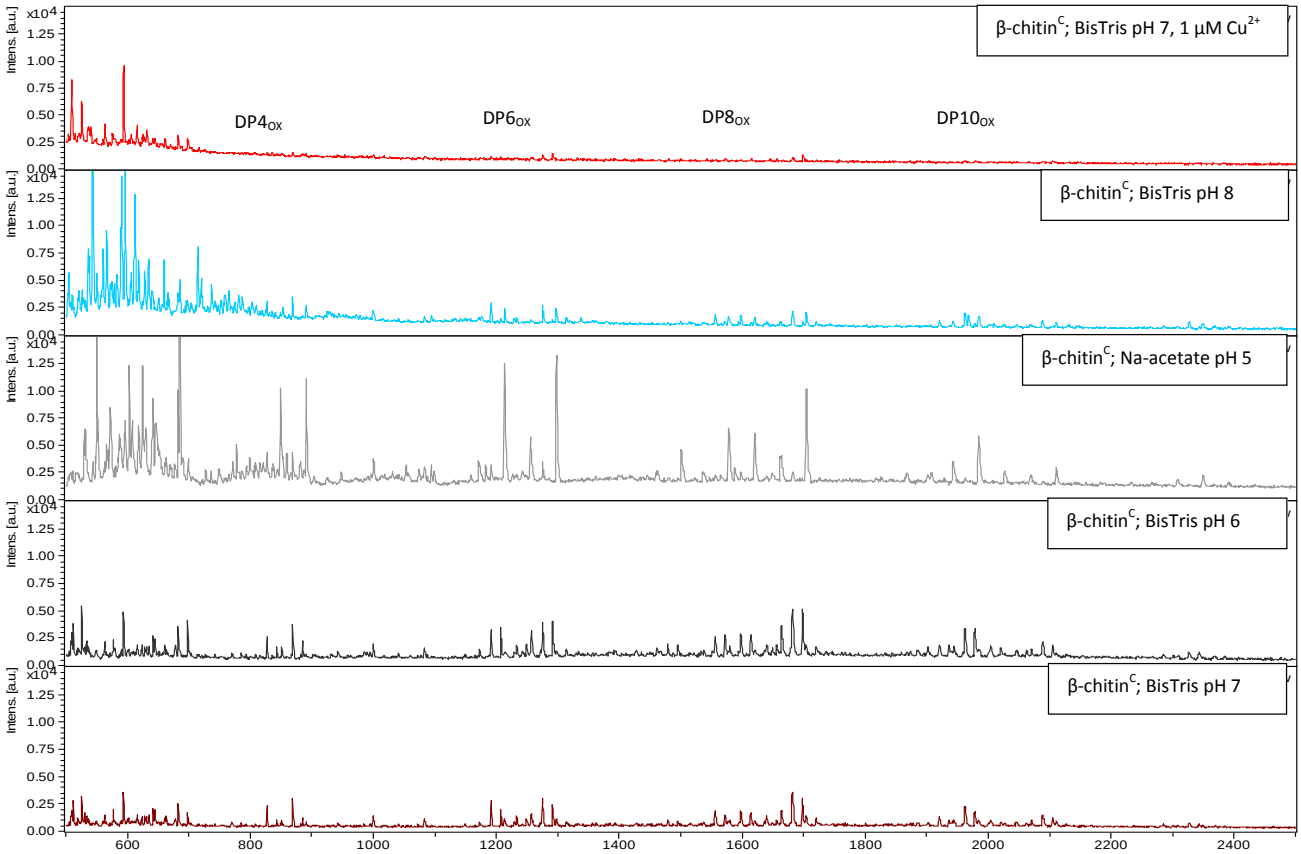


Figure G. 5. The degradation of β-chitin^{C} by 1 µM *Cf*CBM33B-N varies with pH, and is affected by the presence of Cu^{2+}. MALDI-TOF analysis of oligomeric products released upon incubation of 2 mg/ml β-chitin^{C} with 1 µM *Cf*CBM33B-N in 20 mM Na-acetate pH 5 (middle panel) and Bis-Tris pH 6.0 (second lower panel) pH 7 (upper and lower panels) and pH 8 (second upper panel) with 1 mM ascorbic acid as reducing agent. The upper panel show a reaction incubated in the presence of 1 µM Cu^{2+}. The peaks are labeled by the degree of polymerization (DP, shown in upper panel), and for corresponding m/z-values, see Figure G.1.

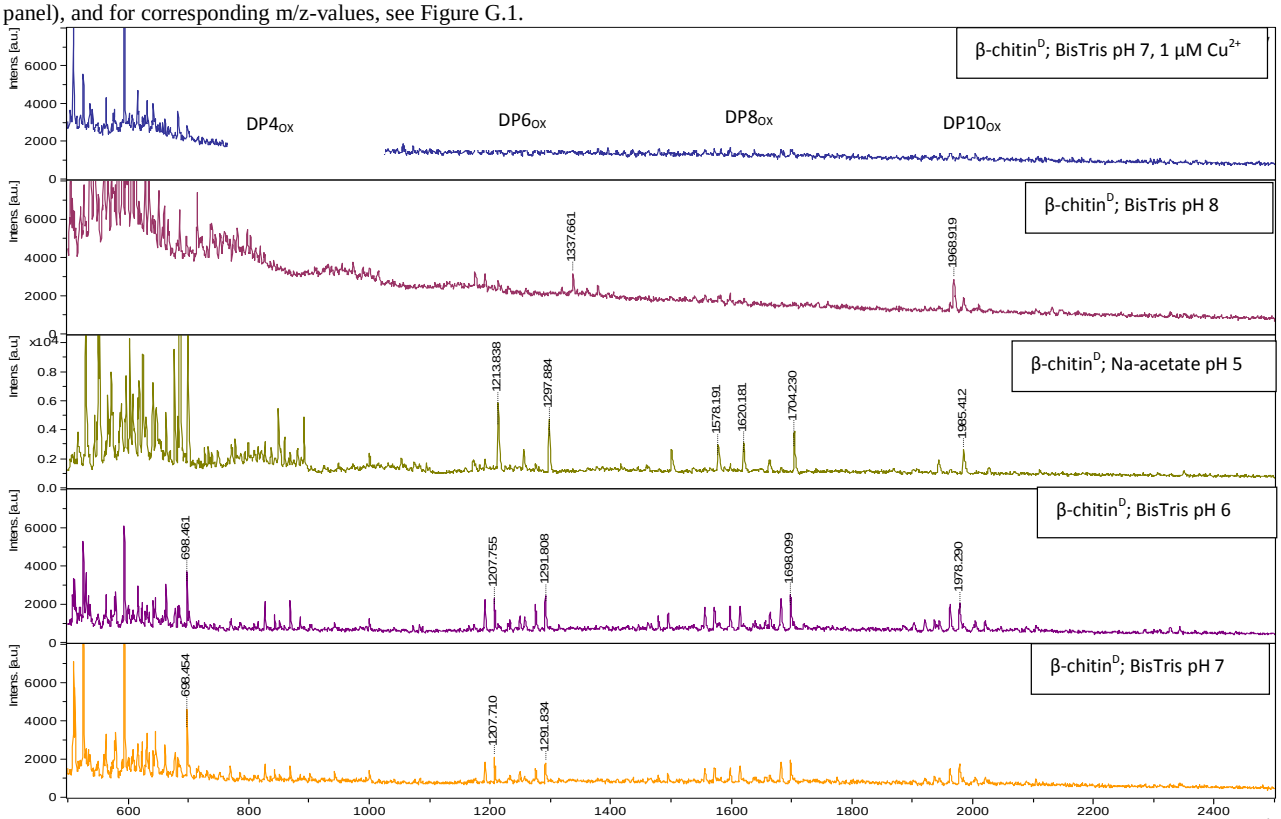


Figure G.6. The degradation of β-chitin^{D} by 1 µM *Cf*CBM33B-N varies with pH, and affected by the presence of Cu^{2+} . MALDI-TOF analysis of oligomeric products released upon incubation of 2 mg/ml β-chitin^{D} with 1 µM *Cf*CBM33B-N in 20 mM Na-acetate pH 5 (middle panel) and Bis-Tris pH 6.0 (second lower panel) pH 7 (upper and lower panels) and pH 8 (second upper panel) with 1 mM ascorbic acid as reducing agent. The upper panel show a reaction incubated in the presence of 1 µM Cu^{2+}. The peaks are labeled by the degree of polymerization (DP, shown in upper panel), and for corresponding m/z-values, see Figure G.1.

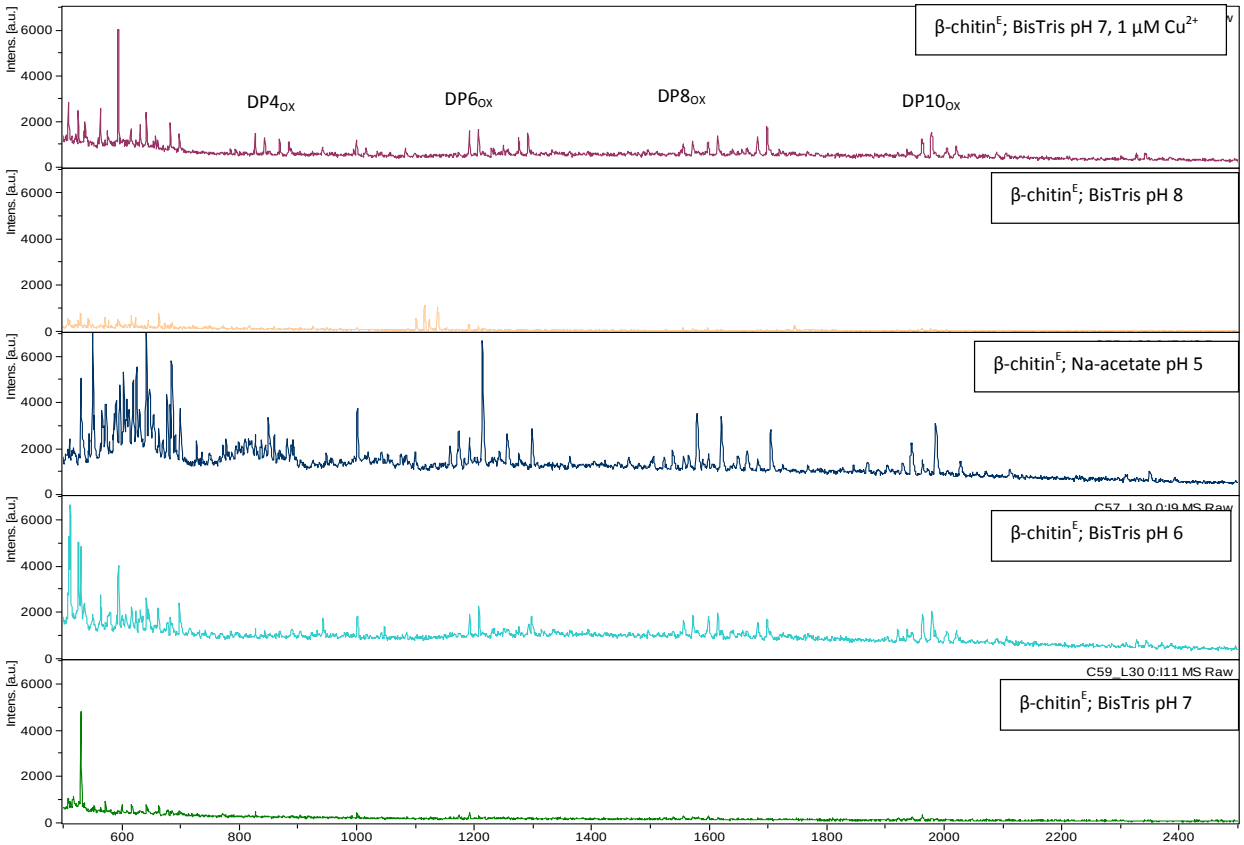


Figure G.7. The degradation of β-chitin^{E} by 1 µM *Cf*CBM33B-N varies with pH, and is affected by the presence of Cu^{2+}. MALDI-TOF analysis of oligomeric products released upon incubation of 2 mg/ml β-chitin^{E} with 1 µM *Cf*CBM33B-N in 20 mM Na-acetate pH 5 (middle panel) and Bis-Tris pH 6.0 (second lower panel) pH 7 (upper and lower panels) and pH 8 (second upper panel) with 1 mM ascorbic acid as reducing agent. The upper panel show a reaction incubated in the presence of 1 µM Cu^{2+}. The peaks are labeled by the degree of polymerization (DP, shown in upper panel), and for corresponding m/z-values, see Figure G.1.

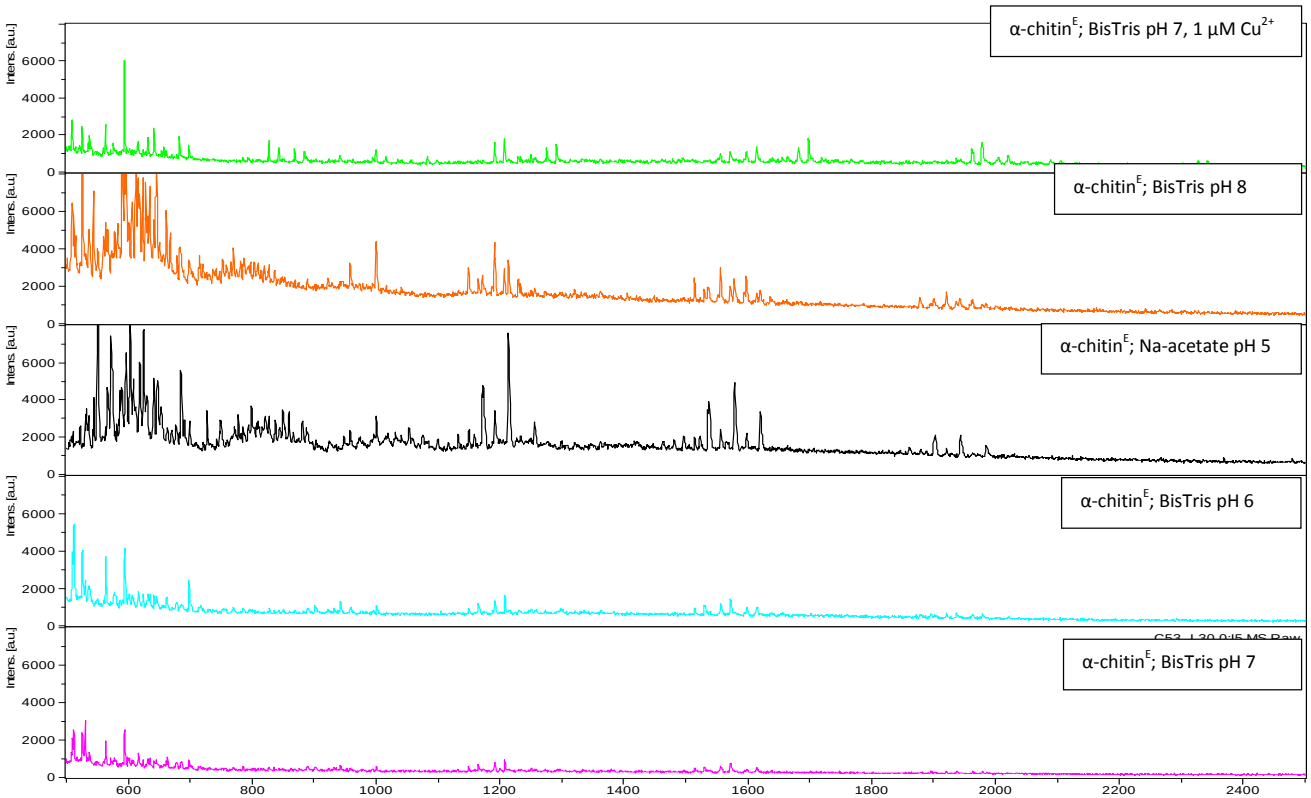


Figure G.8. The degradation of α-chitin^{E} by 1 µM *Cf*CBM33B-N varies with pH, and is affected by the presence of Cu^{2+}. MALDI-TOF analysis of oligomeric products released upon incubation of 2 mg/ml α-chitin^{E} with 1 µM *Cf*CBM33B-N in 20 mM Na-acetate pH 5 (middle panel) and Bis-Tris pH 6.0 (second lower panel) pH 7 (upper and lower panels) and pH 8 (second upper panel) with 1 mM ascorbic acid as reducing agent. The upper panel show a reaction incubated in the presence of 1 µM Cu^{2+}. For labeleling corresponding m/z-values, see Figure G.1.

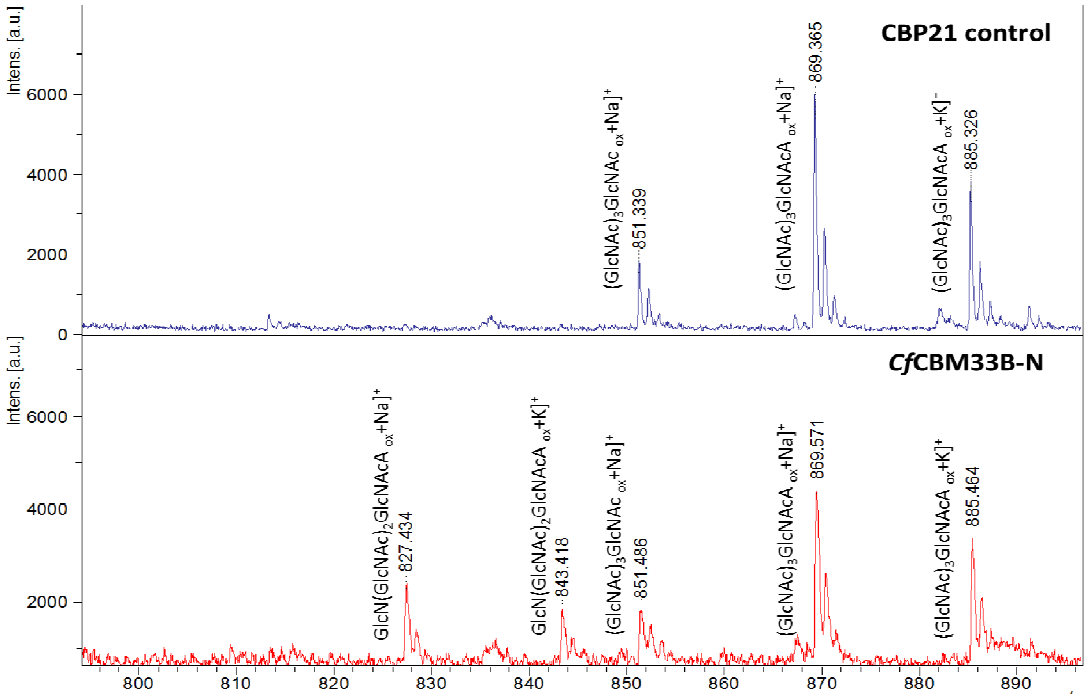


Figure G.9. Zoom in of DP4. MALDI-TOF MS analysis of products detected after treating 2.0 mg/ml β-chitinA with 0.1 µM CfCBM33B-N and 1.0 mM ascorbic acid in Bis-Tris buffer pH 6.0. The figure shows a zoom in of oxidized hexameric products obtained by CfCBM33B-N (red spectrum) and by CBP21 (control; blue spectrum) under the same reaction conditions in reactions carried out at the same time and analyzed in the same analysis session. In the overview spectrum peaks are labeled by the degree of polymerization (DP) i.e. the number of sugar moieties, including one oxidized sugar. In the zoom in pictures peaks are labeled by the observed atomic mass and the predicted composition of the oligomer. Abbreviations used: GlcNAc; GlcN, deacetylated GlcNAc; GlcNLa, 1-5 d lactone of GlcNAc; GlcNAcA, aldonic acid of GlcNAc.

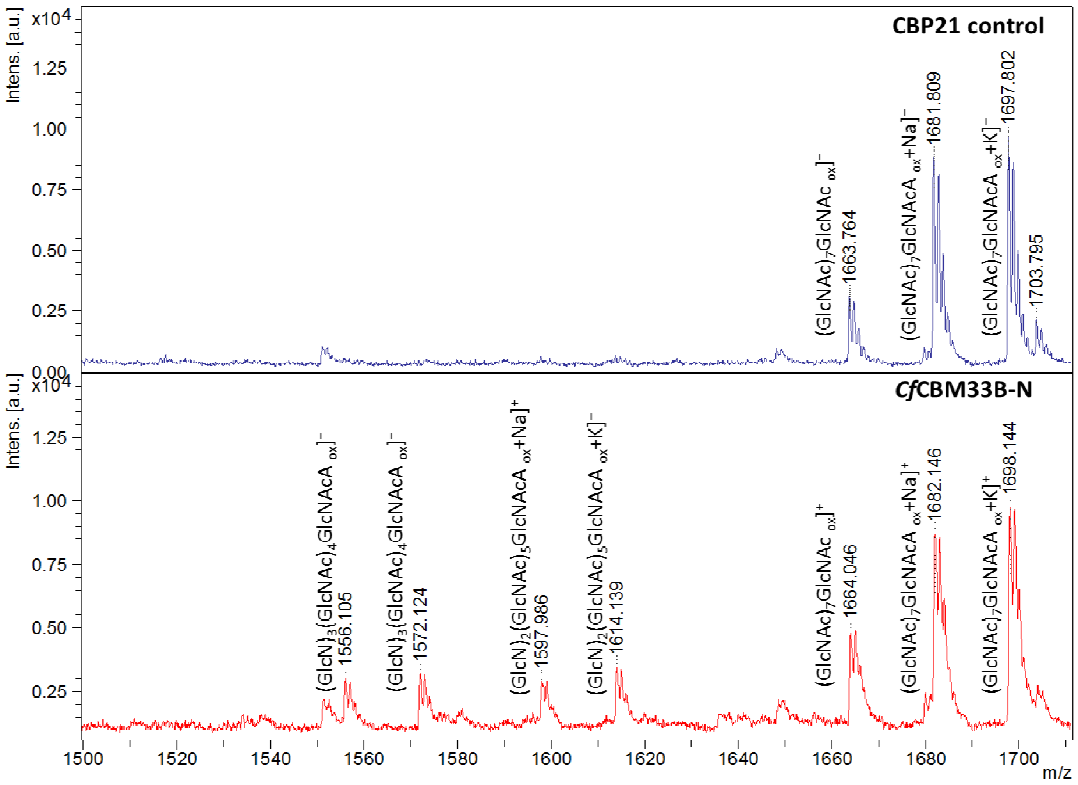


Figure G.10. Zoom in of DP8. MALDI-TOF MS analysis of products detected after treating 2.0 mg/ml β-chitinA with 0.1 µM CfCBM33B-N and 1.0 mM ascorbic acid in Bis-Tris buffer pH 6.0. The figure shows a zoom in of oxidized hexameric products obtained by CfCBM33B-N (red spectrum) and by CBP21 (control; blue spectrum) under the same reaction conditions in reactions carried out at the same time and analyzed in the same analysis session. In the overview spectrum peaks are labeled by the degree of polymerization (DP) i.e. the number of sugar moieties, including one oxidized sugar. In the zoom in pictures peaks are labeled by the observed atomic mass and the predicted composition of the oligomer. Abbreviations used: GlcNAc; GlcN, deacetylated GlcNAc; GlcNLa, 1-5 d lactone of GlcNAc; GlcNAcA, aldonic acid of GlcNAc.

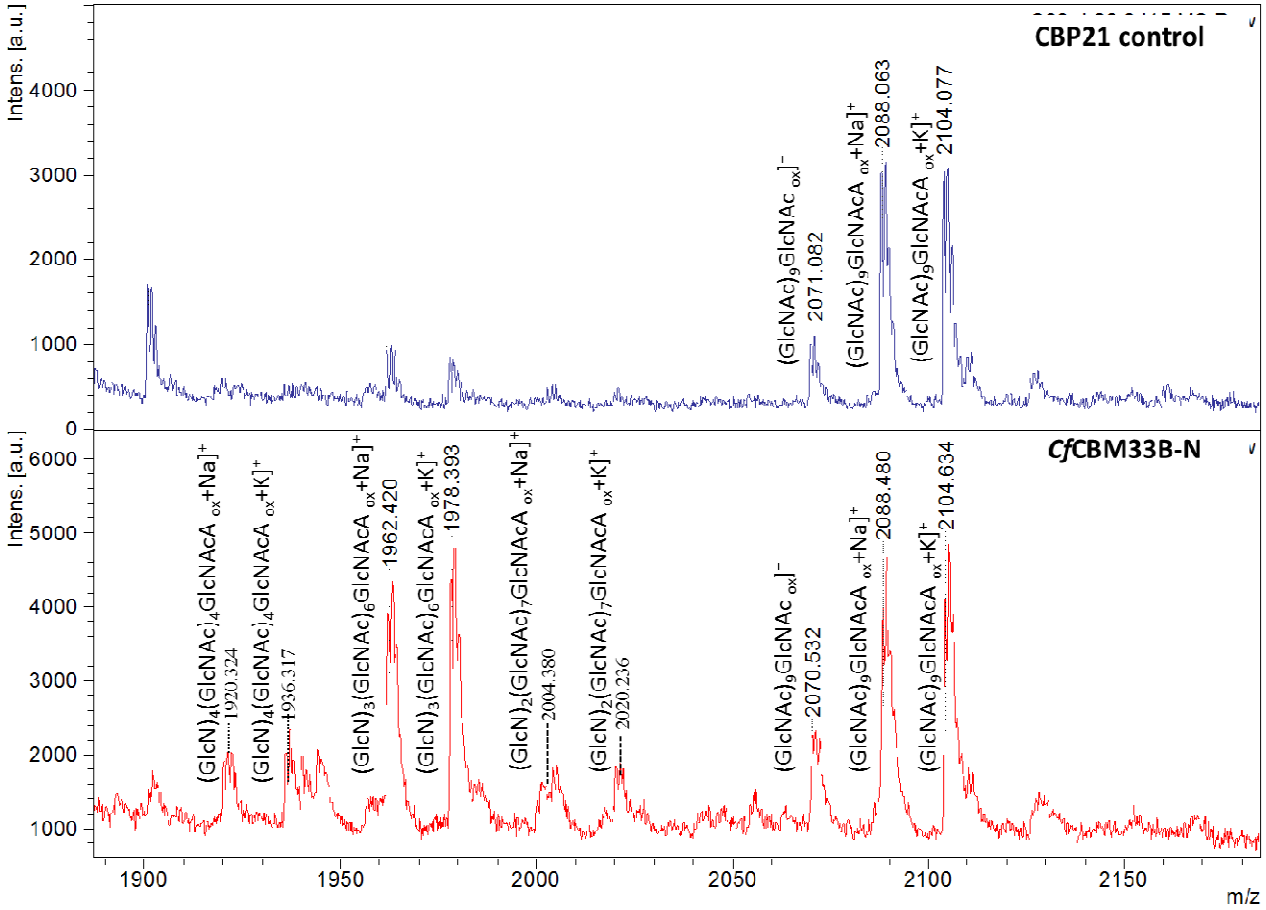


Figure G.11. Zoom in of DP10. MALDI-TOF MS analysis of products detected after treating 2.0 mg/ml β-chitinA with 0.1 µM CfCBM33B-N and 1.0 mM ascorbic acid in Bis-Tris buffer pH 6.0. The figure shows a zoom in of oxidized hexameric products obtained by CfCBM33B-N (red spectrum) and by CBP21 (control; blue spectrum) under the same reaction conditions in reactions carried out at the same time and analyzed in the same analysis session. In the overview spectrum peaks are labeled by the degree of polymerization (DP) i.e. the number of sugar moieties, including one oxidized sugar. In the zoom in pictures peaks are labeled by the observed atomic mass and the predicted composition of the oligomer. Abbreviations used: GlcNAc; GlcN, deacetylated GlcNAc; GlcNLa, 1-5 d lactone of GlcNAc; GlcNAcA, aldonic acid of GlcNAc.

Appendix H: Results from HPLC-measurements

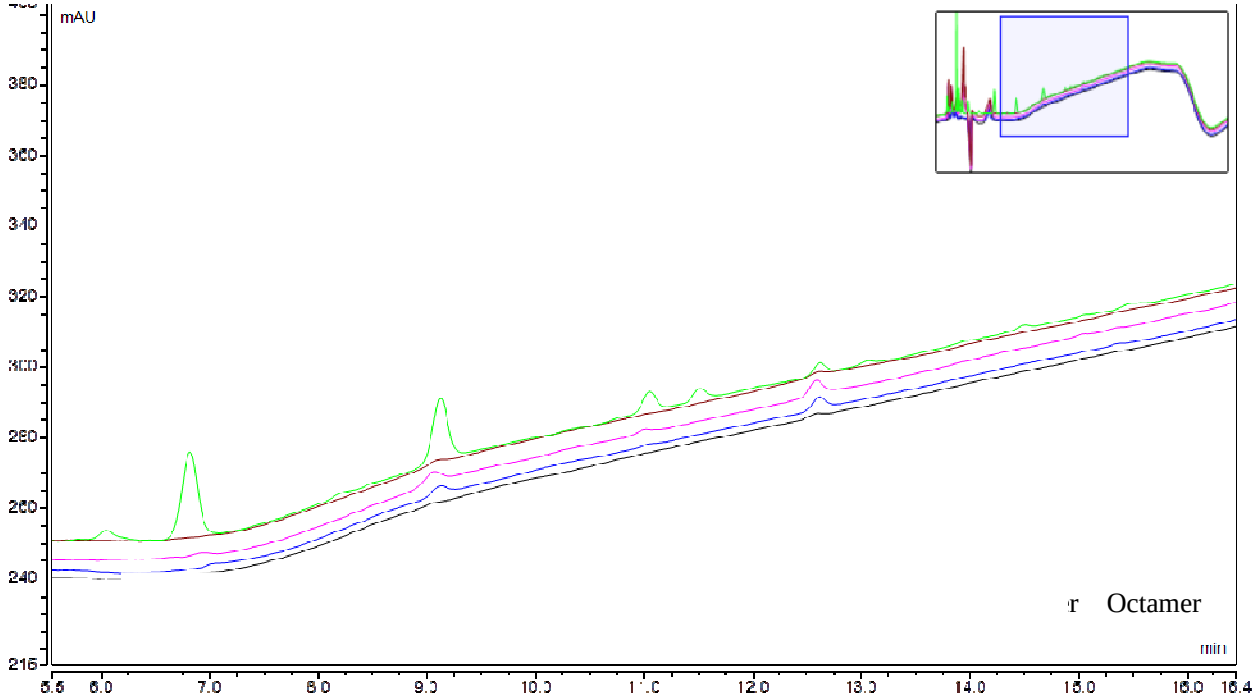


Figure H.2. HPLC- analysis of oxidized products from degradation of 10 mg/ml β-chitin[A] by 1 µM *Cf*CBM33B-N in 20 mM Na acetate pH 5. Reactions were set up with varying concentrations of ascorbic acid as reducing agent: 1 mM (black line), 2.5 mM (blue line), 5 mM (pink line), 10 mM (brown line). An in-house standard mixture of oxidized chitio-oligosaccharides with a degree of polymerization between 3 and 8 was used for comparison (green line).

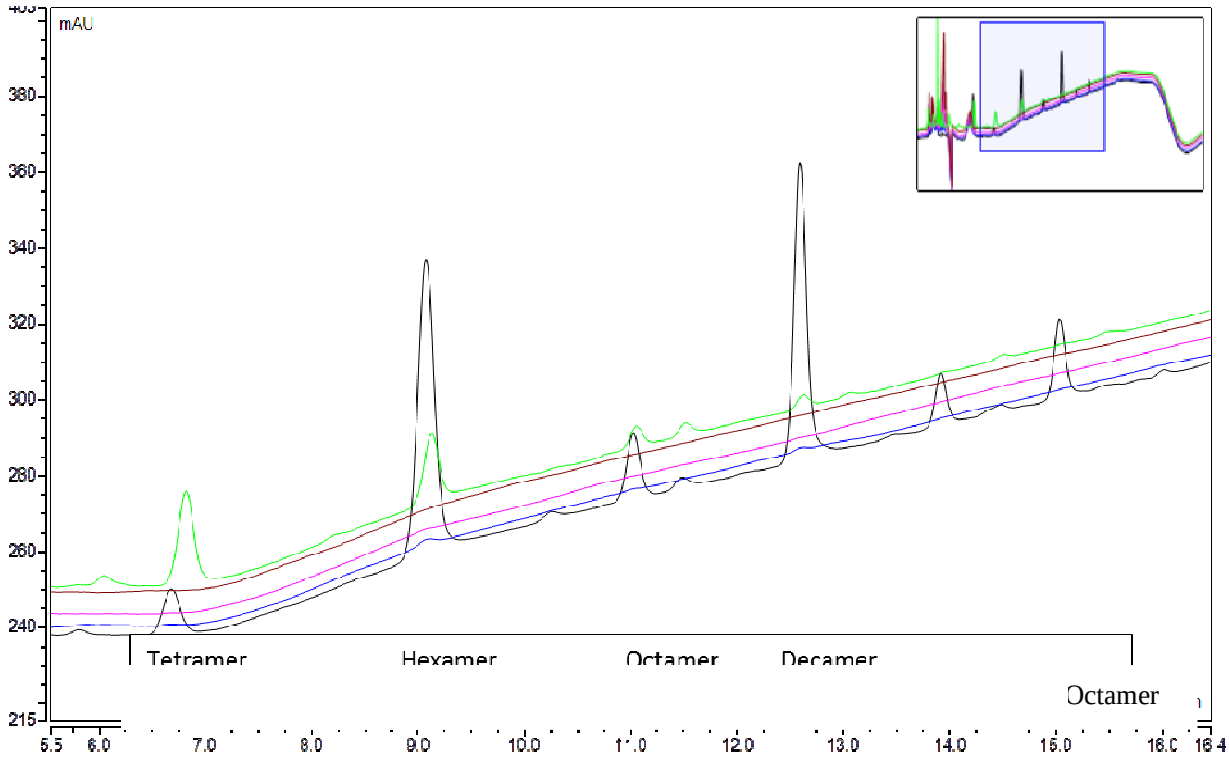


Figure H.3. HPLC- measurements of the oxidized products from the degradation of 10 mg/ml β-chitin[A] by CBP21 with 20 mM Na acetate pH 5 as buffer. The reactions were set up with varying concentration of ascorbic acid as reducing agent; 1 mM (black line), 2.5 mM (blue line), 5 mM (pink line), 10 mM (brown line). An in-house standard mixture of chito-oligosaccharides with a degree of polymerization between 3 and 8 was used for comparison (green line).

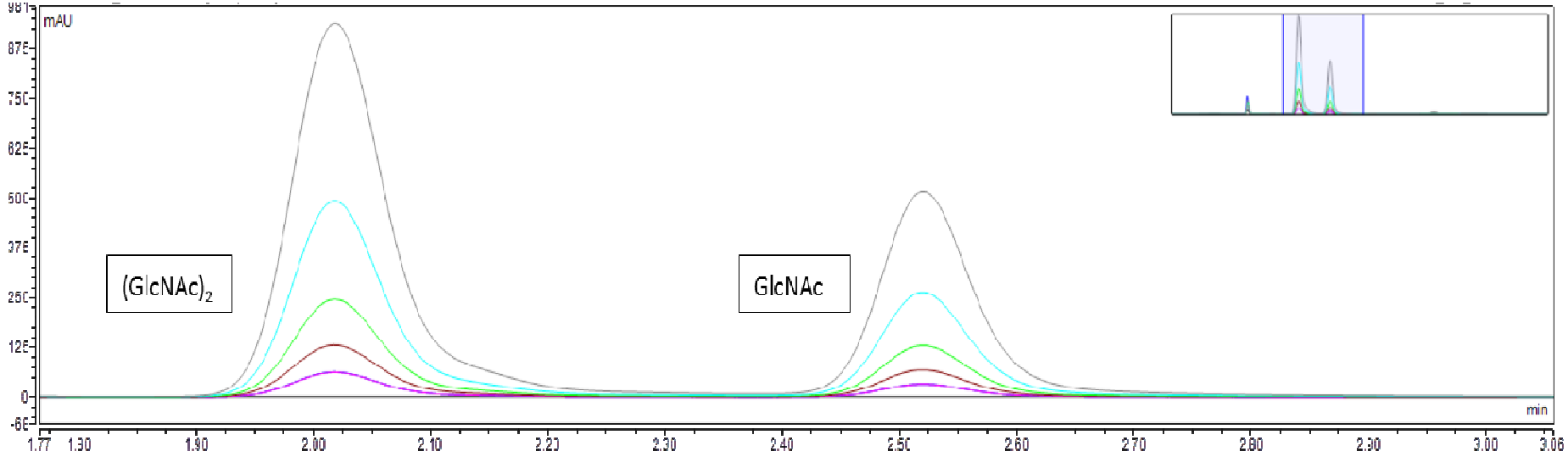


Figure H.4. Overlay of the chromatograms of the standard curve used for determination of the concentration of GlcNAc and $(GlcNAc)_2$. All standards were set up with concentrations of (black), 62.5 (blue), 125 (pink), 250 (brown), 500 (green), 1000 (cyan) and 2000 (grey) µM of each GlcNAc and $(GlcNAc)_2$. All standards were run in triplicates evenly distributed througl sample-run.

Appendix I:

Statistical analysis on the synergy effect of *Cf*CBM33B-N and CBP21 on the degradation of β-chitin by chitinases from *S. marcescens*.

A=ChiA, B=ChiA+*Cf*CBM33B-N 1µM, C=ChiA+*Cf*CBM33B-N 0.5µM+CBP21 0.5µM, D=ChiA+CBP21 0.5µM

ChiA, β-chitin[A]

Estimation of the coefficients and anova-analysis of the effect of degradation over time

Call: lm(formula = Conversion ~ Enzyme + Time + Time2 + Enzyme * Time + Enzyme * Time2, data = Dataset)

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	1.725883	1.729299	0.998	0.32228	
Enzyme[T.B]	4.705500	2.445598	1.924	0.05909	.
Enzyme[T.C]	4.550369	2.445598	1.861	0.06770	.
Enzyme[T.D]	0.967406	2.445598	0.396	0.69383	
Time	0.170357	0.503496	0.338	0.73628	
Time2	-0.005021	0.019024	-0.264	0.79273	
Enzyme[T.B]:Time	3.056005	0.712050	4.292	6.56e-05	***
Enzyme[T.C]:Time	1.228733	0.712050	1.726	0.08956	.
Enzyme[T.D]:Time	0.548036	0.712050	0.770	0.44452	
Enzyme[T.B]:Time2	-0.076346	0.026904	-2.838	0.00619	**
Enzyme[T.C]:Time2	-0.030396	0.026904	-1.130	0.26306	
Enzyme[T.D]:Time2	-0.009246	0.026904	-0.344	0.73230	

Signif. codes:	0 '***'	0.001 '**'	0.01 '*'	0.05 '.'	0.1 ' '

Anova Table (Type II tests)

Response: Conversion

	Sum Sq	Df	F value	Pr(>F)	
Enzyme	3316.2	3	105.6711	< 2.2e-16	***
Time	313.7	1	29.9857	9.036e-07	***
Time2	133.8	1	12.79	0.0006961	***
Enzyme:Time	219.1	3	6.98	0.0004170	***
Enzyme:Time2	100.4	3	3.2	0.0295718	*
Residuals	627.6	60			

> options(contrasts=c('contr.Treatment', 'contr.poly'))

Tukey; multiple comparisons of the means

95% family-wise confidence level

$Enzyme

	diff	lwr	upr	p adj
B-A	17.967072	15.1181789	20.815964	0.0000000
C-A	9.916474	7.0675809	12.765366	0.0000000
D-A	3.848663	0.9997699	6.697555	0.0038648
C-B	-8.050598	-10.8994907	-5.201705	0.0000000
D-B	-14.118409	-16.9673017	-11.269516	0.0000000
D-C	-6.067811	-8.9167037	-3.218918	0.0000030

ChiB, β-chitin[A]

A=ChiB, B=ChiB+*Cf*CBM33B-N 1 µM, C= ChiB+*Cf*CBM33B-N 0.5µM, D=ChiB+*Cf*CBM33B-N 0.5µM+CBP21 0.5µM, E=ChiB+CBP21 0.5µM, F=ChiB+CBP211µM

Estimation of the coefficients and anova-analysis of the effect of degradation over time

Call:lm(formula = conversion ~ Enzyme * Time + Enzyme * Time2, data = Dataset)
Coefficients:

	Estimate	Std, Error	t value	Pr(>|t|)	
(Intercept)	1,2533	1,0785	1,162	0,24737	
Enzyme[T,B]	0,4493	1,5252	0,295	0,76878	
Enzyme[T,C]	0,3713	1,5252	0,243	0,80808	
Enzyme[T,D]	1,1941	1,5252	0,783	0,43514	
Enzyme[T,E]	0,4152	1,5252	0,272	0,78591	
Enzyme[T,F]	1,9622	1,5252	1,287	0,20062	
Time	0,8387	0,3389	2,475	0,01467	
Time2	-0,0195	0,0139	-1,409	0,16123	
Enzyme[T,B]:Time	2,0466	0,4793	4,270	3,81E-05	
Enzyme[T,C]:Time	0,7215	0,4793	1,505	0,13473	
Enzyme[T,D]:Time	2,3055	0,4793	4,810	4,22E-06	*
Enzyme[T,E]:Time	0,3160	0,4793	0,659	5,11E-01	
Enzyme[T,F]:Time	1,622315	0,479281	3,385	9,50E-04	***
Enzyme[T,B]:Time2	-0,044391	0,019606	-2,264	0,02527	
Enzyme[T,C]:Time2	-0,018448	0,019606	-0,941	0,34854	***
Enzyme[T,D]:Time2	-0,062842	0,019606	-3,205	0,00171	
Enzyme[T,E]:Time2	-0,009108	0,019606	-0,465	0,64306	***
Enzyme[T,F]:Time2	-0,044127	0,019606	-2,251	0,02614	*

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Anova Table (Type II tests)

Response: conversion

	Sum Sq	Df	F value	Pr(>F)	
Enzyme	1824,55	5,00	32,75	< 2,2e-16	
Time	2345,01	1,00	210,49	< 2,2e-16	
Time2	847,22	1,00	76,05	<1,38e-14	
Enzyme:Time	442,48	5,00	7,94	0,0000015580	
Enzyme:Time2	171,29	5,00	3,08	0,0118400000	***
Residuals	1403,73	126,00			***

Tukey; multiple comparisons of the means

95% family-wise confidence level
$Enzyme

	diff	lwr	upr	p adj
B-A	8,45	5,67	11,24	0,000000
C-A	2,93	0,14	5,72	0,033197
D-A	9,01	6,22	11,80	0,000000
E-A	1,44	-1,35	4,23	0,667888
F-A	7,47	4,68	10,26	0,000000
C-B	-5,52	-8,31	-2,73	0,000001
D-B	0,56	-2,23	3,35	0,992225
E-B	-7,01	-9,80	-4,23	0,000000
F-B	-0,98	-3,77	1,81	0,910775
D-C	6,08	3,29	8,87	0,000000
E-C	-1,49	-4,28	1,30	0,634223
F-C	4,54	1,75	7,33	0,000092
E-D	-7,57	-10,36	-4,78	0,000000
F-D	-1,54	-4,33	1,25	0,601103
F-E	6,0321608	3,2436458	8,820676	0,0000001

ChiC, β-chitin[A]

A=ChiC, B=ChiC+*Cf*CBM33B-N 1µM, C= ChiB+*Cf*CBM33B-N 0.5 µM + CBP21 0.5µM, D= ChiB+CBP21 1µM

Estimation of the coefficients and anova-analysis of the effect of degradation over time

Call:lm(formula = Conversion ~ Enzyme * Time + Enzyme * Time2, data = Dataset3)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	0.78669	1.74353	0.451	0.65347	
Enzyme[T.B]	2.74384	2.46572	1.113	0.27024	
Enzyme[T.C]	4.63736	2.46572	1.881	0.06486	.
Enzyme[T.D]	3.09766	2.46572	1.256	0.21388	
Time	0.44966	0.47889	0.939	0.35151	
Time2	-0.01087	0.01462	-0.743	0.46010	
Enzyme[T.B]:Time	1.91285	0.67725	2.824	0.00642	**
Enzyme[T.C]:Time	3.62825	0.67725	5.357	1.41e-06	***
Enzyme[T.D]:Time	4.88718	0.67725	7.216	1.07e-09	***
Enzyme[T.B]:Time2	-0.04854	0.02067	-2.348	0.02217	*
Enzyme[T.C]:Time2	-0.08899	0.02067	-4.305	6.26e-05	***
Enzyme[T.D]:Time2	-0.11940	0.02067	-5.776	2.90e-07	***
Signif. codes:	0 '***'	0.001 '**'	0.01 '*'	0.05 '.'	0.1 ' '

Anova Table (Type II tests)

Response: Conversion

	Sum Sq	Df	F value	Pr(>F)	
Enzyme	5729.5	3	169.756	< 2.2e-16	***
Time	1833.5	1	162.968	< 2.2e-16	***
Time2	1188.0	1	105.599	7.546e-15	***
Enzyme:Time	663.3	3	19.652	5.420e-09	***
Enzyme:Time2	422.8	3	12.526	1.836e-06	***
Residuals	675.0	60			

Tukey; multiple comparisons of the means

95% family-wise confidence level

Fit: aov.default(formula = LinearModel.10)

$Enzyme

	diff	lwr	upr	p adj
B-A	10.251254	7.2967771	13.205731	0.000000
C-A	19.391592	16.4371149	22.346069	0.000000
D-A	23.050638	20.0961612	26.005116	0.000000
C-B	9.140338	6.1858607	12.094815	0.000000
D-B	12.799384	9.8449069	15.753861	0.000000
D-C	3.659046	0.7045691	6.613523	0.009336

ChiB, β-chitin[E]

A=ChiB, B=ChiB+*Cf*CBM33B-N 1 µM, C= ChiB+*Cf*CBM33B-N 0.5µM, D=ChiB+*Cf*CBM33B-N 0.5µM+CBP21 0.5µM, E=ChiB+CBP21 0.5µM, F= ChiB+CBP21 1µM

Estimation of the coefficients and anova-analysis of the effect of degradation over time

Call:lm(formula = conversion ~ Enzyme * Time + Enzyme * Time2, data = J145_162)

Coefficients:

	Estimate	Std, Error	t value	Pr(>|t|)	
(Intercept)	2,2385	1,3465	1,6620	0,09890	,
Enzyme[T,B]	2,4167	1,9043	1,2690	0,20670	
Enzyme[T,C]	1,8987	1,9043	0,9970	0,32060	
Enzyme[T,D]	1,6540	1,9043	0,8690	0,38670	
Enzyme[T,E]	0,3537	1,9043	0,1860	0,85300	
Enzyme[T,F]	0,0268	1,9043	0,0140	0,98880	
Time	0,1784	0,4821	0,3700	0,71200	
Time2	-0,0064	0,0210	-0,3030	0,76250	
Enzyme[T,B]:Time	5,1560	0,6818	7,5620	<1e-05	***
Enzyme[T,C]:Time	3,7637	0,6818	5,5200	<1e-05	***
Enzyme[T,D]:Time	4,7534	0,6818	6,9720	<1e-05	***
Enzyme[T,E]:Time	0,3457	0,6818	0,5070	0,61310	
Enzyme[T,F]:Time	0,2240	0,6818	0,3290	0,74310	
Enzyme[T,B]:Time2	-0,1570	0,0297	-5,2870	<1e-05	***
Enzyme[T,C]:Time2	-0,1241	0,0297	-4,1780	<1e-05	***
Enzyme[T,D]:Time2	-0,1514	0,0297	-5,0970	<1e-05	***
Enzyme[T,E]:Time2	-0,0123	0,0297	-0,4140	0,67980	
Enzyme[T,F]:Time2	-0,0071	0,0297	-0,2390	0,81160	

Signif, codes: 0 '***' 0,001 '**' 0,01 '*' 0,05 ',' 0,1 ' ' 1

Anova Table (Type II tests)

Response: conversion

	Sum Sq	Df	F value	Pr(>F)	
Enzyme	7836,2	5	99,604	< 2,2e-16	***
Time	2645,6	1	168,138	< 2,2e-16	***
Time2	1427,7	1	90,738	< 2,2e-16	***
Enzyme:Time	2010,9	5	25,559	< 2,2e-16	***
Enzyme:Time2	1039,5	5	13,213	2,39E-10	***
Residuals	1982,6	126			

Signif, codes: 0 '***' 0,001 '**' 0,01 '*' 0,05 ',' 0,1 ' ' 1

Tukey; multiple comparisons of the means

95% family-wise confidence level

$Enzyme

	diff	lwr	upr	p adj	
B-A	17,63536	14,3214	20,94932	<1e-05	
C-A	12,31338	8,999424	15,62734	<1e-05	
D-A	15,19845	11,88449	18,51241	<1e-05	
E-A	1,244626	-2,06933	4,558585	0,886	
F-A	0,668039	-2,64592	3,981999	0,992	
C-B	-5,32197	-8,63593	-2,00801	<1e-05	
D-B	-2,4369	-5,75086	0,877057	0,280	
E-B	-16,3907	-19,7047	-13,0768	<1e-05	
F-B	-16,9673	-20,2813	-13,6534	<1e-05	
D-C	2,88507	-0,42889	6,199029	0,126	
E-C	-11,0688	-14,3827	-7,7548	<1e-05	
F-C	-11,6453	-14,9593	-8,33139	<1e-05	
E-D	-13,9538	-17,2678	-10,6399	<1e-05	
F-D	-14,5304	-17,8444	-11,2165	<1e-05	
F-E	-0,57659	-3,89055	2,737373	0,996	

Basic diagnostic plots; residual plots and normal distribution plots for the four estimated models.