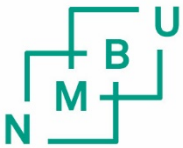


Norges miljø- og biovitenskapelige universitet

Master's Thesis 2019 60 ECTS

Faculty of Chemistry, Biotechnology and Food Science (KBM)

Gene cloning, purification and characterization of novel hemicellulases for production of tailor made prebiotics and biochemicals

Lars Jordhøy Lindstad

Biotechnology

Acknowledgment

This master thesis was completed at the Faculty of Chemistry, Biotechnology, and Food Science at the Norwegian University of Life Sciences (NMBU), with Dr. Bjørge Westereng and Dr. Sabina Leanti la Rosa as my supervisors.

First, I would like to thank my main supervisor Bjørge for his guidance and availability throughout the work on this master thesis. I have really appreciated our discussions, both professionals and others. Your explanatory answers and positive way of being has made this year much easier. For that I am thankful.

I would also like to express my great appreciation to Sabina, for always having an open door and taking the time to answer all my questions. All your guidance in the lab and help during the writing process has been most valuable, and I deeply appreciate it.

Dr. Shaun Allan Leivers – you deserve a whole page of appreciation for your endless guidance in the lab, and for always being available and taking the time to help me.

I am also thankful to Dr. Leszek Michalak for all your guidance, support and teaching during the lab work.

Thanks to everyone in the BioRef Group for making it a great place to work.

Finally, I want to thank my family, friends and girlfriend for all your encouragements and support throughout this process.

Ås, May 15th

Lars Jordhøy Lindstad

Abstract

The human gut contains billions of microbes which play a key role in human health. They provide us with important short-chain fatty acids, like butyrate, through the fermentation of dietary fibers. A disruption in the gut microbiota have been linked to several diseases, like colorectal cancer, Crohn's disease, obesity and irritable bowel syndrome. Two of the most important phyla in the gut, *Bacteroidetes* and *Firmicutes*, are well equipped with enzymes capable of breaking down complex carbohydrates. Several butyrate producing *Firmicutes* are generally acknowledged to be beneficial to our health. The growth of these bacteria can be selectively promoted through the use tailored of prebiotics. Complex β-mannans, acetylated galactoglucomannan, from Norway spruce (*Picea abies*) are a new possible prebiotic substrate. This resembles mannans found in our diet, but the complex structure requires a sophisticated degradation system to be utilised.

In this thesis, we have looked into two esterases, *Fp*CE2 and *Fp*CEXX, found in a mannan degrading gen cluster in *Faecalibacterium prausnitzii*. *F. prausnitzii* has been shown to utilise the acetylated galactoglucomannan when supplemented in a feeding trial on pigs. The esterases were successfully cloned, expressed and purified. Both enzymes were active on mannan-based substrates, and not on other substrates tested. They show a complementary deacetylation of the substrates. Testing of specificity have revealed that the *Fp*CE2 removes 3-*O*-, 4-*O*- and 6-*O*-acetylations, while the *Fp*CEXX is 2-*O* specific. The complementary reactivity of These two esterases enables production of tailor made acetylated mannan.

Another part of this thesis was to make mutants of a structurally characterised esterase from *Roseburia intestinalis*, to look into the possibility of making more specific substrates. The *Ri*CEX esterase removes 2-*O*-acetylation with high specificity. Mutations made aimed at; A) blocking the galactose pockets found on each side of the active site, or B) open up the catalytic site pocket to possibly fit larger groups, like propionate and butyrate, to the substrate in a transesterification reaction. All mutants were active on the same substrate as the native *Ri*CEX, but we were not able to identify any differences from the native *Ri*CEX on the substrates tested so far. This is subject to further research in our group, and this study has provided several targeted approaches to reveal the potential effects of the different mutants.

Sammendrag

Menneskets tarm inneholder millioner av mikrober som spiller en nøkkelrolle i menneskers helse. De gir oss viktige kortkjedede fettsyrer, som butyrat, gjennom gjæring av diettfibre. En forstyrrelse i tarmmikrobiotaens sammensetning har vært knyttet til flere sykdommer, som tarmkreft, Crohns sykdom, fedme og irritabel tarmsyndrom. To av de viktigste phylum i tarmen, *Bacteroidetes* og *Firmicutes*, er godt utstyrt med enzymer som kan bryte ned komplekse karbohydrater. Flere butyrat produserende *Firmicutes* er generelt anerkjent for å være gunstige for vår helse. Vekst av disse bakteriene kan selektivt fremmes gjennom bruk skreddersydde prebiotika. Komplekse β-mannaner, som acetylert galactoglukomannan fra Norsk gran (*Picea abies*) er et nytt mulig prebiotisk substrat. Dette ligner mannan funnet i kostholdet vårt, men den komplekse strukturen krever et sofistikert nedbrytningssystem for å kunne utnyttes.

I denne oppgaven har vi sett på to esteraser, *Fp*CE2 og *Fp*CEXX, funnet i en mannan nedbrytende gen klynge i *Faecalibacterium prausnitzii*. *F. prausnitzii* har vist seg å kunne benytte acetylert galaktoglukomannan når det suppleres i et fôringsforsøk på griser. Esterasene ble klonet, uttrykt og renset. Begge enzymer var aktive på mannan baserte substrater, og ikke på andre substrater testet. De viser en komplementær deacetylering av substratene. Test av spesifisitet har vist at *Fp*CE2 fjerner 3-*0*-, 4-*0*- og 6-*0*-acetyleringer, mens *Fp*CEXX er 2-*0*-spesifikk. Den komplementære aktiviteten til disse to esterasene muliggjør produksjon av skreddersydd acetylert mannan.

En annen del av denne oppgaven var å lage mutanter av en strukturelt karakterisert esterase fra *Roseburia intestinalis*, for å se på muligheten for å lage mer spesifikke substrater. *Ri*CEX-esterasen fjerner 2-*0*-acetylering med høy spesifisitet. Mutasjoner ble laget rettet mot; A) blokkering av galaktoselommer funnet på hver side av det aktive setet, eller B) åpne det katalytiske setet for å muligens passe større kjemiske grupper, som propionat og butyrat, til substratet i en transesterifikasjons reaksjon. Alle mutanter var aktive på samme substrat som den native *Ri*CEX, men vi kunne ikke identifisere noen forskjeller fra den opprinnelige *Ri*CEX på de undersøkte substratene hittil. Dette er tema for videre forskning i vår gruppe, og denne studien har gitt flere målrettede tilnærminger for å avdekke de potensielle effektene av de forskjellige mutantene.

Abbreviations

AcGGM	Acetylated galactoglucomannan
ACN	Acetonitrile
CBM	Carbohydrate-Binding Module
CE	Carbohydrate Esterase
dGTP	Deoxyguanosine Triphosphate
dH_2O	Milli-Q® Sterile Water
DNA	Deoxyribonucleic Acid
dNTP	Deoxynucleoside Triphosphates
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
HILIC-LC-MS	Hydrophilic Interaction Liquid Chromatography Mass Spectrometry
HPLC	High-Performance Liquid Chromatography
IMAC	Ion Metal Affinity Chromatography
IPTG	Isopropyl b-D-1-thiogalactopyranoside
kb	Kilobases
LB	Lysogeny broth
LDS	Lithium dodecyl sulfate
LIC	Ligation-Independent Cloning
MALDI-ToF MS	Matrix-Assisted Laser Desorption/Ionizaton Time of Flight Mass Spectrometry
MS	Mass spectrometry
MWCO	Molecular Weight Cut-off
nm	Nanometers
PCR	Polymerase Chain Reaction
pNP	4-nitrophenyl
rpm	Revolutions per minute
SDS-PAGE	Sodium Dodecyl Sulfate - Polyacrylamide Gel Electrophoresis
SEC	Size Exclusion Chromatography
SOC	Super Optimal Broth with Catabolite repression
TAE	Tris-Acetate-EDTA

TB	Terrific broth
UV	Ultraviolet

Table of contents

Acknowledgment ..... **I**

Abstract ..... **II**

Sammendrag ..... **III**

Abbreviations ..... **IV**

1 Introduction **1**

1.1 Gut microbiota 1

1.2 Prebiotics 2

1.3 Carbohydrates 3

1.3.1 Carbohydrates 3

1.3.2 Hemicelluloses 3

1.3.3 Acetylation 5

1.4 Bacterial degradation of plant polysaccharides in the gut 6

1.5 Carbohydrate-active enzymes 7

1.5.1 Carbohydrate-active enzymes (CAZymes) 7

1.5.2 Carbohydrate Esterases 8

1.6 *Faecalibacterium prausnitzii* and *Roseburia intestinalis* 9

1.7 Wood prebiotics project 10

1.8 Aim of this study 10

2 Materials **12**

2.1 Chemicals 12

2.2 Laboratory equipment and materials 13

2.3 Instruments 15

2.4 Growth media and agar 16

2.5 Kits 16

2.6 Substrates 16

2.7 Bacterial strains 17

2.8 DNA and plasmids 17

2.9 Proteins and enzymes 18

2.10 Primers 18

3 Methods 19
3.1 Methods working with microbiology 19
3.2 Preparation of culture media and agar 19
3.3 Buffers 21
3.4 Cloning 22
3.4.1 Primers: 22
3.4.2 Polymerase Chain Reaction: 23
3.4.3 Agarose Gel Electrophoresis: 24
3.4.4 PCR clean-up: 25
3.4.5 DNA concentration: 25
3.4.6 T4 DNA digestion for cloning into pNIC-CH vector: 26
3.4.7 Transformation of One Shot® TOP10 Chemically Competent *E. coli*: 27
3.4.8 Colony PCR: 27
3.4.9 Plasmid purification: 28
3.4.10 Sequencing: 29
3.4.11 Transformation of One Shot® BL21 Star™ (DE3) Chemically Competent *E. coli*: 29
3.4.12 Glycerol stocks: 29
3.5 *Ri*CEX Mutants 30
3.6 Protein expression: 30
3.6.1 Harbinger: 31
3.6.2 Cell harvesting: 31
3.7 Protein purification: 32
3.7.1 Protein extraction: 32
3.7.2 Immobilized Metal Affinity Chromatography: 32
3.7.3 Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis: 33
3.7.4 Filtration and buffer exchange: 34
3.7.5 Size Exclusion Chromatography: 34
3.7.6 Protein concentration: 35
3.7.7 Storage of enzymes: 35
3.8 Enzyme characterization 35
3.8.1 MALDI-Tof MS 35
3.8.2 Activity on 4-nitrophenyl acetate 37
3.8.3 High-Performance Liquid Chromatography (HPLC) 38

3.8.4 Time course HPLC 38
3.8.5 Hydrophilic Interaction Liquid Chromatography Mass Spectrometry 39
3.8.6 Transesterification 40
3.8.7 Preferred substrate 41
3.8.8 Protein thermal shift assay 42
3.9 Protein Crystallization 43
4 Results 46
4.1 Bioinformatics 46
4.1.1 F. prausnitzii 46
4.1.2 R. intestinalis mutants 49
4.2 Cloning, Expression and Purification 52
4.2.1 *F. prausnitzii* 52
4.2.2 *Ri*CEX Mutants 54
4.3 Characterization of *F. prausnitzii* enzymes 58
4.3.1 Activity test on mannose and non-mannose containing substrates with *Fp*CE2 and *Fp*CEXX 58
4.3.2 Catalytic activity 60
4.3.3 Transesterification 62
4.3.4 Crystallization 66
4.4 Characterization of *R. intestinalis* mutants 67
4.4.1 *Ri*CEX_CatD and *Ri*CEX_CBM 67
4.4.2 *Ri*CEX mutants # 1-16 67
5 Discussion 70
5.1 Cloning, expression and purification 70
5.2 Characterization 72
5.3 Concluding remarks and future perspectives 74
6 References 76
7 Appendix 81
Appendix A 81
Appendix B 81
Appendix C 83

1 Introduction

1.1 Gut microbiota

The human gastrointestinal tract contains billions of microbes of which many are helpful and necessary for our health. Important features of the microbes are; energy source, synthesis of vitamins, bile salt transformation, development of the innate and adaptive immune system and barrier to colonization of pathogens (Dave et al., 2012). A disruption in the gut microbial community balance can result in different illnesses. Several diseases have been linked to dysbiosis of the gut microbiota, such as obesity, type 2 diabetes, irritable bowel syndrome, Chron's disease and colorectal cancer (Sankar et al., 2015).

More than 30 bacterial phyla are detected in the human gut, where most of the species belongs to the phyla *Firmicutes, Bacteroidetes, Actinobacteria, Proteobacteria, Fusobacteria, Cyanobacteria,* and *Verrucomicrobia*. *Firmicutes* and *Bacteroidetes* are the most abundant phyla (Sankar et al., 2015). Many of these bacteria, especially *Firmicutes* and *Bacteroidetes*, are equipped with carbohydrate-active enzymes able to utilize polysaccharides from our diet which are indigestible for humans. Through the fermentation of these carbohydrates, some of the gut bacteria can produce short chain fatty acids (SCFAs), like acetate, propionate, butyrate and succinate, which displays many important roles for our health (Flint et al., 2008). SCFAs are the main metabolic products from the fermentation by gut microbes (Louis & Flint, 2017). All these SCFAs can be absorbed by our body and metabolised systemically (R. Gibson et al., 2010). Acetate is used to generate ATP in muscles, and it may be involved in appetite regulation (Frost et al., 2014). Propionate is mainly metabolised in the liver where it regulates cholesterol synthesis (Al-Lahham et al., 2010). Butyrate is primarily an energy source for the gut mucosa (Louis & Flint, 2017). Propionate and butyrate also have an effect in regulating the immune system and on anti-inflammatory response (Morrison & Preston, 2016).

The gut microbiotas' ability to break down and utilize complex carbohydrates, non-digestible for humans, gives the possibility to selectively change the composition and abundance of some bacterial groups through the use of substrates that promote the growth of the health beneficial bacteria; these substrates are called prebiotics (Flint et al., 2008).

1.2 Prebiotics

There has been no common definition of what a prebiotic is, and the meaning and definition have varied among scientists throughout the years. The first use of the term prebiotic dates back to 1995 (Gibson & Roberfroid, 1995), and has been changed several times. There is currently consensus that a prebiotic should selectively promote growth of bacteria in the gastrointestinal tract giving a health benefit for the host (Gibson et al., 2017).

The International Scientific Association for Probiotics and Prebiotics (ISAPP) is a non-profit association that work for scientific progress in the fields of probiotics and prebiotics. In 2016, the ISAPP collected a panel of experts to review the definition of prebiotics and make a consensus statement. Academic and industrial scientists in this panel covered research fields in microbiology, biochemistry, nutrition and clinical research. The goal was to come up with a new definition of prebiotic, making the concept clearer in use, and adopted to all research fields. This was based on former definitions, while taking new knowledge and research into account, and dealing with various problems with earlier definitions, which often were narrower than expedient (Gibson et al., 2017).

This led to the new definition of a prebiotic: "*a substrate that is selectively utilized by host microorganisms conferring a health benefit*" (Gibson et al., 2017). This new definition gives the term prebiotic a wider meaning and could address bacteria living elsewhere in or on the body, other than just the gastrointestinal tract, and including other substrates than carbohydrates. In addition, this also applies for animals, like pets and farm animals (Gibson et al., 2017).

Prebiotics on the market have mainly been fructans (fructooligosaccharides (FOS) and inulin) and galactans (galactooligosaccharides (GOS)). These have been recognized as prebiotics for a long time, and are targeting, either or both, *Lactobacillus* and *Bifidobacterium* species. More recently, there is a growing number of studies and methods showing the diversity of species in our gut, implying that also other species may be affected by the prebiotics (Gibson et al., 2017). Relating to this, there is a big potential for new prebiotics based on oligosaccharides and polysaccharides (Rastall & Gibson, 2015).

1.3 Carbohydrates

1.3.1 Carbohydrates

Carbohydrates makes up a big portion of our diet and is the main source of energy for many living organisms. Most of the carbohydrates in diets come from plants, but also from other sources, for example milk. This study focuses on plant-based carbohydrates. Plants contains a large number of carbohydrates, supplying them long polymers for structure and strength. These carbohydrates can be grouped into four major classes; cellulose, hemicellulose, lignin and pectin. Other energy rich carbohydrates are used as energy storage, like starch, inulin, fructose, glucose and sucrose.

Carbohydrates are grouped after molecular size as degree of polymerization (DP), meaning how many monosaccharide units they are made of. They can be separated into; monosaccharides (DP1); disaccharides (DP2); oligosaccharides (DP3-9); and polysaccharides (DP≥10) (Cummings & Stephen, 2007). Saccharide chains can vary from linear polymers to highly branched with various side groups, of either carbohydrate units or different chemical groups. In carbohydrate chains, the monosaccharide units are covalently linked together with glycosidic bonds that are formed between the hemiacetal group of one unit and one hydroxyl group of the other unit. In D-sugars, which are the most common monosaccharides, the glycosidic bond can either be α-conformation (axial orientation) or β-conformation (equatorial orientation), and vary between the carbon atoms of two monosaccharides, for example (1→1), (1→2), (1→3), (1→4) and (1→6).

1.3.2 Hemicelluloses

Hemicelluloses are found in the cell walls of plants and is the second most abundant plant polysaccharide after cellulose. Their main biological role is to strengthen the cell walls by interaction with cellulose and/or lignin. Hemicelluloses have a high variation in their structure, with the most prevalent being xylans and mannans (Scheller & Ulvskov, 2010). Other examples are xyloglucans, glucomannans and β-(1→3,1→4)-glucans. These structures vary a lot through the different branching of the polysaccharide backbone. Their backbone units are linked with β-(1→4) bonds in an equatorial configuration at C1 and C4 (Scheller & Ulvskov, 2010).

Xylans are found in edible consumables like fruits, vegetables and cereal grains, where they may contribute to more than 30 % of the dry weight. Mannans can be found in certain beans, nuts and legume seeds, and are used extensively as food thickeners (Scheller & Ulvskov, 2010). In addition to be found in food products, mannan is also used in, for example, cosmetics and nutraceuticals. Some types of mannans also show health promoting effects, such as immunostimulating, anti-inflammatory and anti-tumour properties (Singh et al., 2018). Xylans and mannans are also abundant in woody biomass, like birch and spruce, and share many structural similarities with those found in edible plants. Figure 1.1 shows a variety of mannan based substrates, with some common structures and some used in this study.

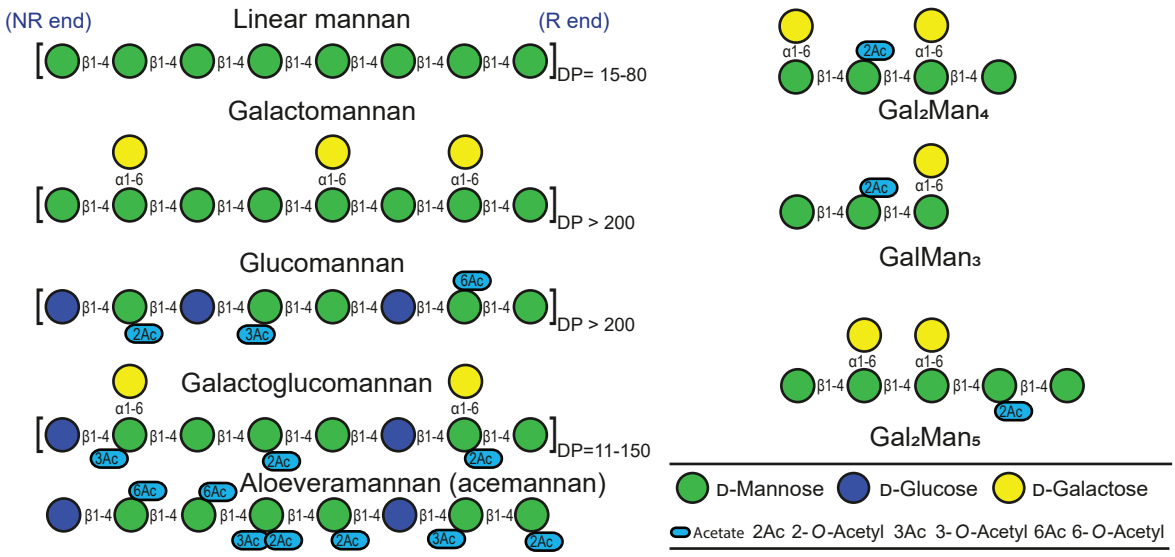


Figure 1.1. A variety of mannans found in plants, and three acetylated manno-oligosaccharides, Gal_2Man_4, $GalMan_3$ and Gal_2Man_5, used in this study.

In this study acetylated galactoglucomannan (AcGGM) extracted from Norway spruce (*Picea abies*) was used as the carbohydrate source of interest (figure 1.2). In softwood like spruce, AcGGM is the most abundant hemicellulose in the secondary cell wall and constitutes about 20 % of the dry wood mass (Lundqvist et al., 2002). The AcGGM backbone consists of randomly distributed *β*-(1→4)-D-mannopyranosyl and *β*-(1→4)-D-glucopyranosyl residues, with α-(1→6)-D-galactopyranosyl units and 2-*O*, 3-*O* and/or 6-*O*-acetylations attached to the mannosyl units. The galactose:glucose:mannose ratio varies and two types have been described (Lundquist et al); one galactose rich (1:1:3) and one galactose poor (0.1:1:3), with dry weight

fractions of 5-8 % and 10-15 %, respectively. This can vary depending on the extraction and purification method (Lundqvist et al., 2002; Timell, 1967).

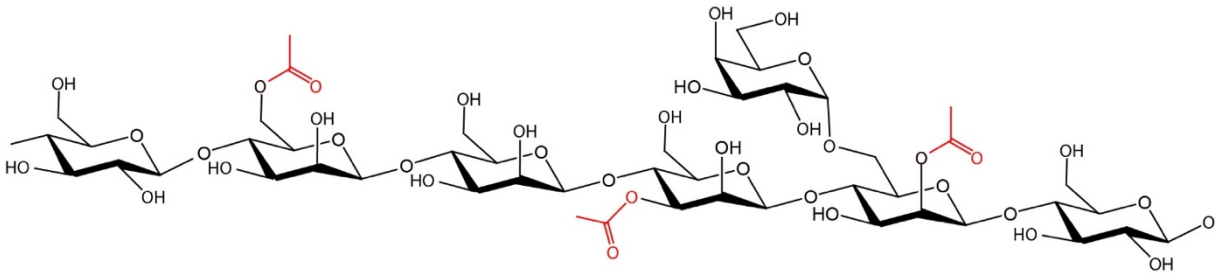


Figure 1.2. Acetylated galactoglucomannan from Norway spruce. The backbone consists of randomly distributed *β*-(1→4)-D-mannopyranosyl and *β*-(1→4)-D-glucopyranosyl. The mannosyl units are branched with α-(1→6)-D-galactopyranosyl and 2-O, 3-O and 6-O acetylations.

1.3.3 Acetylation

Hemicelluloses in plant cell walls are often, to some degree, branched with acetylations (Biely et al., 2014). Esterification, with acetic acid and other chemical groups, of plant polysaccharides are important for physical properties and functionalities. For example, esterification with phenolic acids is important for the cross-linking between polymers in the cell wall (Williamson et al., 1998). The acetylation of polysaccharides is important for the plants ability to resist degradation by microbial enzymes during infections (Biely, 2012). Acetylated polymers will be bulkier, thus hindering degrading enzymes from interacting with the substrate.

The degree of acetylation affects the substrate's properties. The water solubility will increase with a partially acetylated polysaccharide, while a full acetylation makes it water-insoluble (Biely et al., 2014). This has a large effect on viscosity, which is important for the use of mannans as additives as thickeners and stabilizers in, for example, cosmetics and food (Singh et al., 2018). The methods used to extract mannans from plants influence their degree of acetylation. For example, the pH used in the treatment during extraction from woody biomass can alter the number of acetyl groups present, and a high pH has been shown to deacetylate the substrate (Michalak et al., 2018). Acetyl group migration may occur during treatment of acetylated substrates, affected by pH changes in the solution. Acetyl groups can migrate to a neighbouring hydroxyl group, and thus change the compositional structure of acetylated substrates. This is important for, and needs consideration through, the production and testing

of acetylated oligo- and polysaccharides (Roslund et al., 2008). It has also been shown that acetyl migration can occur over the glycosidic bond between two monosaccharides (Lassfolk et al., 2018).

1.4 Bacterial degradation of plant polysaccharides in the gut

Humans, with our own enzymes, can only break down and utilize simple sugars, like sucrose, lactose and starch (Cockburn & Koropatkin, 2016). More complex polysaccharides from plants are going through the upper gastrointestinal tract to the large intestine. Here, the non-digestible dietary carbohydrates will become available for the microbiota (Louis et al., 2007). The organisms living in the gut are competing for many of the oligo- and polysaccharides, but cross-feeding also occur. In fact, some bacteria can utilize metabolic products released by other species during their catabolism of a substrate, or from products released from enzymatic break down extracellularly. In this way bacteria that don't have enzymes to break down larger carbohydrates can still live in the gut (Cockburn & Koropatkin, 2016; Louis et al., 2007).

Our diet has a significant role in the composition of the bacterial species in our gut. A change in diet can supply the gut microbiota with other substrates, which can lead to an altered bacterial community. A dietary change could thus provide some species an advantage over others and outcompete them over time. Introducing new substrates can have either a positive or negative effect, depending on the species that benefit of the altered diet. Other co-products from fermentation or extracellular break down of substrates may also affect some species. The different metabolism between bacterial species, whose abundance depends on the substrates available, leads to the possible selectivity for or against certain families or species. In addition, competition for the available glycans may cause changes in environmental conditions in the gut, like pH, oxygen or hydrogen availability as well as production of antimicrobial agents. This may further affect the composition of species (Louis et al., 2007).

While the *Bacteroidetes* and *Firmicutes* are the dominating phyla in the gut microbiota, their composition and abundance vary extensively between individuals and geographically because of diet, lifestyle and disease (El Kaoutari et al., 2013). The *Bacteroidetes* are considered to be generalists as they can utilize a large number of glycans, while the *Firmicutes* tend to be

specialists and target a more limited source of substrates (Cockburn & Koropatkin, 2016). Figure 1.3 shows an example of the break down of mannan substrates in *Roseburia intestinalis*.

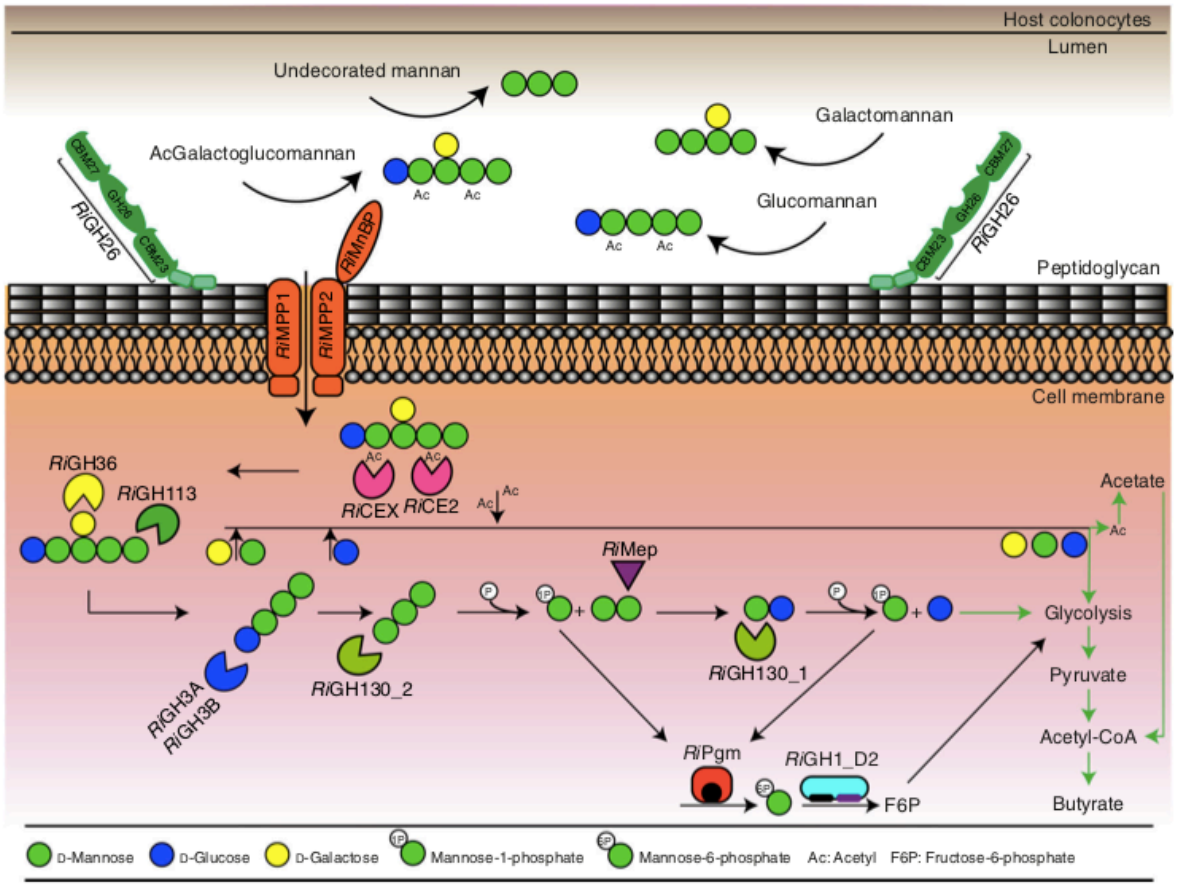


Figure 1.3. The break down of mannan substrates in the commensal gut bacteria *Roseburia intestinalis*, showing a variety of enzymes needed to utilise these complex substrates. Figure from La Rosa et al. (2019b).

The gut microbiota encodes a large number of enzymes, that enables them to utilize the glycans that can't be used directly by humans. Remarkably, for the gut microbiota encodes hundreds of carbohydrate active enzymes, while humans only produces 17 enzymes for digestion of glycans in food (El Kaoutari et al., 2013).

1.5 Carbohydrate-active enzymes

1.5.1 Carbohydrate-active enzymes (CAZymes)

Carbohydrate-active enzymes (CAZymes) are collected in the CAZy database (http://www.cazy.org), and grouped into families based on amino acid sequence similarity

(Lombard et al., 2013). The enzymes (catalytic modules) are categorised in five classes and one associated modules class; Glycoside Hydrolases (GHs), Glycosyl Transferases (GTs), Polysaccharide Lyases (PLs), Carbohydrate Esterases (CEs), Auxiliary Activities (AAs) and Carbohydrate-Binding Modules (CBMs). All of these classes, except for the glycosyl transferases, are relevant in the research of carbohydrate degrading enzymes.

The glycoside hydrolases class is the most abundant, with 162 GH families in the CAZy database (May 2019). Glycoside hydrolases are hydrolysing the glycosidic bond between carbohydrates or a carbohydrate and a non-carbohydrate. El Kaoutari et al. (2013) found that in the human gut microbiota, more than half of the genes coding for carbohydrate-active enzymes were glycoside hydrolases. The polysaccharide lyases cleaves uronic acid-containing polysaccharide chains, and the auxiliary activities enzymes are lytic polysaccharide mono-oxygenases (LPMOs) or ligninolytic enzymes using redox reactions. The carbohydrate-binding modules are non-catalytic domains with carbohydrate-binding activity found attached to the catalytic domain of some carbohydrate-active enzymes. CBMs bind carbohydrate substrates to the enzyme within the range of the catalytic domain, thus promoting catalytic activity (Boraston et al., 2004).

1.5.2 Carbohydrate Esterases

Carbohydrate esterases are a group of esterases which remove ester-based groups on substituted carbohydrates, by catalysing the de-*O* or de-*N*-acylations (Cantarel et al., 2008; Nakamura et al., 2017). The carbohydrate esterases are classified in 15 (earlier 16) families in the CAZy database, as the CE 10 family doesn't appear to have esterases active on carbohydrates (Nakamura et al., 2017). The CAZy database also have a group of non-classified sequences, which are not assigned to the present families. Carbohydrate esterases in family 1 – 7 and 16 are assumed to degrade acetylated plant hemicelluloses (Biely, 2012). Esterases are important for the removal of acyl side-groups, so that other enzymes, like GHs, will get access to the substrate (Nakamura et al., 2017).

The carbohydrate esterase family 2 (CE2) is classified as acetyl xylan esterases, although some enzymes in this family has been showed to have a preference for acetylated glucomannan over acetylated xylan (Montanier et al., 2009). The CE2 family consists of 455 sequences in total

(May 2019) in the CAZy database, and the majorities comes from bacteria (440), while 12 comes from Eukaryotes and three are unclassified. Some characterized members of this family deacetylate the 3-*O*-, 4-*O*- and/or the 6-*O*-position of their substrates (Adesioye et al., 2016; Topakas et al., 2010). Structurally, members of the CE2 family usually have two domains: a C-terminal SGNH domain and N-terminal jellyroll domain. The SGNH domain contains the catalytic site, consisting of a conserved catalytic dyad of histidine and serine, while the jellyroll domain contains two β-sheets associated to carbohydrate binding (Nakamura et al., 2017; Till et al., 2013).

In the non-classified group of carbohydrate esterases, some enzymes are able to deacetylate the acetylated galactoglucomannan in hemicellulose from softwoods are found (Biely, 2012).

1.6 *Faecalibacterium prausnitzii* and *Roseburia intestinalis*

In this thesis, the primary focus was on two esterases from a mannan degrading gene cluster found in *F. prausnitzii* and on mutants of an esterase from *R. intestinalis*.

F. prausnitzii and *R. intestinalis* belong to the *Firmicutes* phyla. *F. prausnitzii* is found in the large intestine, and is one of the most abundant commensal bacteria in healthy humans (Lopez-Siles et al., 2012; Miquel et al., 2013). It is an important butyrate producer (Pryde et al., 2002), and exerts other beneficial health effects, like anti-inflammatory effects (Lopez-Siles et al., 2017). The abundance of *F. prausnitzii* have also been linked to some diseases, where a low abundance is related to Crohn's disease (Miquel et al., 2013). The role of *F. prausnitzii* in the gut have thus gained interest in using this species as a potential probiotic and/or using prebiotics to promote its growth (Lopez-Siles et al., 2017; Miquel et al., 2013). As for the use of prebiotics to promote growth of *F. prausnitzii*, many strains have been showed to grow on pectin (Lopez-Siles et al., 2012), and on inulin, despite only a few strains grew well on that substrate (Lopez-Siles et al., 2017; Ramirez-Farias et al., 2008). La Rosa et al. (2019a) have shown that *F. prausnitzii* from human faeces utilised and increased in growth when AcGGM and acetylated arabinoglucuronoxylan (AcAGX) were used as substrates. *R. intestinalis* is a butyrate producer found in the colon (Duncan et al., 2002). *Roseburia* species are linked to health benefits, like anti-inflammatory effects and immunity maintenance, and a low abundance is associated with diseases like irritable bowel syndrome (Tamanai-Shacoori et al., 2017). *R. intestinalis* have the

genes to utilise β-mannans from plant cell walls, which can be used to selectively promote growth of this species (La Rosa et al., 2019b).

1.7 Wood prebiotics project

This study was conducted as a part of the project called "Wood Prebiotics: Exploring tree-derived hemicelluloses as a source for prebiotics". The project explored the possibility of using hemicellulose from wood as prebiotics and to increases our understanding of prebiotics. As the structure of hemicellulose in wood resembles those found in plants used as food, the potential for an abundantly available (and low cost) prebiotic source is substantial. As a part of this project, Michalak et al. extracted AcGGM from Norway spruce, which were used as substrate for testing enzymes of interest. AcGGM was tested on live animals in a feeding trial on piglets with AcGGM supplemented to the feed. Piglets with a 4 % AcGGM supplement in their feed, showed an increased population of *Faecalibacterium* in their gut microbiota.

1.8 Aim of this study

The pig feeding trial (Leszek Michalak, 2019), which showed that *Faecalibacterium* population increased when AcGGM was added to the feed, lead to the hypothesis that *Faecalibacterium* have enzymes able to utilize mannan-based substrates. Two potential mannan acetyl esterases from *F. prausnitzii* were of interest, hereafter referred to as *Fp*CE2 and *Fp*CEXX. Previous work on two homologous esterases from *R. intestinalis* (*Ri*CE2 and *Ri*CEX) showed that *Ri*CE2 was a broad specific esterase, working on 3-*O*, 4-*O* and 6-*O* acetylations, while the *Ri*CEX was 2-*O* specific (Michalak et al., 2019)

The aim of this study was to clone, express and purify the *F. prausnitzii* esterases, as well as to test their functionality. This will increase our understanding of how this health beneficial bacteria can utilize AcGGM, and thus shed light on the possibility to use this substrate to be used as a prebiotic.

In addition, a truncated version of the *Ri*CEX, separating the catalytic domain and carbohydrate binding module, was generated. In addition to provide deeper insight on how this esterase

interacts with its substrates, changes in the catalytic domain could potentially lead to a broader specific activity. A goal was to look into the possibility of making more specific oligosaccharide substrates, by the use of rationally designed mutants of the *Ri*CEX. The structure of this 2-*O* specific esterase has been solved previously as a part of the Wood Prebiotic project (Michalak et al., 2019) and it was used as template for designing mutants.

2 Materials

2.1 Chemicals

Table 2.1. List over chemicals used in this study.

Chemicals	Supplier
2,5-dihydroxybenzoic acid ($C_7H_6O_4$)	Sigma-Aldrich, Germany
Acetonitrile (CH_3CN)	VWR, USA
Antifoam 204	Sigma-Aldrich, Germany
CHES ($C_8H_{17}NO_3S$)	Sigma-Aldrich, Germany
Dithiothreitol (DTT) ($C_4H_{10}O_2S_2$)	Invitrogen, USA
DNA Gel Loading Dye 6X	New England Biolabs, USA
Ethanol 96 % vol	VWR, USA
Ethylenediaminetetraacetic acid (EDTA) ($C_{10}H_{16}N_2O_8$)	Merck, Germany
Glycerol 85% ($C_3H_8O_3$)	Merck, Germany
HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)	Sigma-Aldrich, Germany
Hydrochloric acid 32 % (HCl)	Merck, Germany
Imidazole ($C_3H_4N_2$)	Sigma-Aldrich, Germany
Isopropanol (C_3H_8O)	Sigma-Aldrich, Germany
Isopropyl β-D-1-thiogalactopyranoside (IPTG) ($C_9H_{18}O_5S$)	Sigma-Aldrich, Germany
Kanamycin ($C_{18}H_{36}N_4O_{11}$)	Sigma-Aldrich, Germany
Methanol	VWR, USA
NEBuffer ™ 2.1	New England Biolabs, USA
NuPAGE® LDS Sample Buffer (4X)	Invitrogen, USA
NuPAGE® Sample Reducing Agent (10X)	Invitrogen, USA
peqGreen DNA/RNA Dye	Peqlab, Germany
Potassium phosphate dibasic (K_2HPO_4)	Sigma-Aldrich, Germany
Potassium phosphate monobasic (KH_2PO_4)	Sigma-Aldrich, Germany
Protein Assay Dye Reagent Concentrate	Bio-Rad, USA
Sodium acetate (NaOAc)	Sigma-Aldrich, Germany
Sodium Chloride (NaCl)	VWR, USA

Sodium hydroxide (NaOH)	VWR, USA
Sodium phosphate dibasic (Na_2HPO_4)	Sigma-Aldrich, Germany
Sodium phosphate monobasic (NaH_2PO_4)	Sigma-Aldrich, Germany
Sucrose ($C_{12}H_{22}O_{11}$)	VWR, USA
Sulfuric acid (H_2SO_4)	Merck, Germany
Tris tris(hydroxymethyl)aminomethane	Sigma-Aldrich, Germany
Tris-Acetate-EDTA (TAE) Buffer (50X)	Thermo Scientific, USA
Tris-Glycine-SDS (TGS) Buffer (10X)	Bio-Rad, USA

2.2 Laboratory equipment and materials

Table 2.2. List over laboratory equipment and materials used in this study.

Laboratory equipment and materials	Supplier
8-Well Comb	Bio-Rad, USA
ARE Aluminum Hot Plate Stirrer	VELP Scientifica, Italy
Cellstar® Centrifuge Tubes (15 and 50 ml)	Greiner Bio-One, Germany
Nalgene®, Centrifuge bottles	Sigma-Aldrich, Germany
Centrifuge Avanti™ J-25	Beckman Coulter, USA
Mega Star 1.6R	VWR, USA
Heraeus™ Pico™ 21 Centrifuge	Thermo Scientific, USA
Centrifuge 5418R	Eppendorf, Germany
CertoClav Sterilizer GmbH A-4050	CertoClav, Austria
Concentrator plus	Eppendorf, Germany
Cuvettes (disposable)	Eppendorf, Germany
Eppendorf tubes, 1.5 mL	Axygen, USA
Gel Casting Tray	Bio-Rad, USA
Gel Doc™ EZ Imager	Bio-Rad, USA
Glassware	Duran Group, Germany
Innova® 44 Incubator Shaker Serie	New Brunswick Scientific, USA
Inoculation loop, 1 uL, white	Sarstedt, Germany
LEICA MZ7.5	Leica, Germany

Magnet, Teflon Stirring Bar	SP Scienceware
Microtiter® 96-Well UV Microplates	Thermo Scientific, USA
Milli-Q® Direct Water Purification System, Direct 16	Merck Millipore, Germany
Mini-PROTEAN® Tetra System	Bio-Rad, USA
Mini-PROTEAN® TGX Stain-Free□	Bio-Rad, USA
Mini-Sub® Cell GT	Bio-Rad, USA
PCR Tubes, 0.2 mL	Axygen, USA
Petri dishes, 90 mm	Heger, Norway
pH meter large	WTW inoLab, Germany
pH meter small	Sentron, Netherlands
Pipette tips	VWR, USA
PowerPac™ Basic	Bio-Rad, USA
Scale Entris 3202I-1S	Sartorius, Germany
Scale Quintix 124-1S	Sartorius, Germany
Screw Cap Micro Tubes, 2 mL	Sarstedt, Germany
Siliconized Glass Circle Cover Slides, 12 and 18 mm	Hampton Research, USA
Snap Ring Micro-Vials, 0.3 mL, Snap Ring Caps	VWR, USA
Sonics Vibra-Cell™ Ultrasonic Processor, VC750	Sonics & Materials, Inc., USA
Syringe Filtration Unit, Filtropur S 0.22 □m PES Membrane	Sarstedt, Germany
Syringes, BD Plastipak™	BD, Germany
ThermoMixer C	Eppendorf, Germany
Transferpette	Brand, Germany
Ultrafiltration Unit Vivaspin® 20, 10,000 MWCO PES	Sartorius, Germany
Vacuum Filtration Systems, 0.22 □m PES Membrane	VWR, USA
Vacuum pump	Millipore, USA
VDXm Plate-With Sealant, 12 and 18 mm	Hampton Research, USA
Vortex	IKA, Germany
Water bath large 100 °C (SSB	Grant, Great Britain
Water bath small 42 °C	Julabo, Germany

2.3 Instruments

Table 2.3. Instruments used in this study.

Instruments	Supplier
NGC™ Chromatography system	Bio-Rad, USA
BioPhotometer® D30	Eppendorf, Germany
bioZen 2.6 µm Glycan 100 x 2.1 mm column	Phenomenex
Corona Ultra CAD (charged aerosol detector)	ESA
HiLoad™ 16/600 Superdex™ 75 pg	GE Healthcare Life Sciences, USA
IMAC Column, HisTrap™ HP, 5 mL	GE Healthcare Life Sciences, USA
LEX-48 Bioreactor	Epiphyte3, Canada
Mastercycler gradient	Eppendorf, Germany
MTP 384 ground steel target plate	Bruker Daltonics GmbH, Germany
Qubit™ fluorometer	Invitrogen, USA
REZEX ROA-Organic Acid H+ 300x7.8mm ion exclusion column	Phenomenex
HPLC UltiMate 3000	Thermo Scientific, USA
StepOnePlus™ Real-Time PCR System	Thermo Fisher Scientific, USA
Stormy Pro S	MARK, Italy
Synergy H4 Hybrid Reader	BioTek®
UltrafleXtreme™ MALDI-TOF	Bruker Daltonics GmbH, Germany
Velos Pro ion trap mass spectrometer	Thermo Scientific, USA
Xbridge Amide 3.5 µm particle size 4.6x250 mm column	Waters, USA
Äkta Pure	GE Healthcare Life Sciences, USA
Äkta purifier	GE Healthcare Life Sciences, USA

2.4 Growth media and agar

Table 2.4. Growth media and agar used in this study.

Growth media and agar	Supplier
Agar powder	VWR, USA
Bacto™ Tryptone	Becton, Dickinson and Co, USA
Bacto™ Yeast extract	Becton, Dickinson and Co, USA
SeaKem® LE Agarose	Lonza, Switzerland
SOC ready-to-use	Invitrogen, USA

2.5 Kits

Table 2.5. Kits used in this study.

Kits	Supplier
Clear Strategy™ Screen 1 MD1-14	Molecular Dimension, England
JCSG-*plus*™ Screen MD1-37	Molecular Dimension, England
NucleoSpin Gel and PCR Clean-up Kit	MACHEREY-NAGEL, Germany
NucleoSpin Plasmid Kit	MACHEREY-NAGEL, Germany
PEG/Ion Screen HR2-126	Hampton Research, USA
Qubit™ dsDNA BR Assay Kit	Thermo Fisher Scientific, USA
Wizard™ Classic 1 & 2	Rigaku, Japan

2.6 Substrates

Table 2.6. Substrates used in this study.

Substrates	Supplier
4-nitrophenyl	Sigma-Aldrich, Germany

4-nitrophenyl acetate	Sigma-Aldrich, Germany
6^1-α-D-Galactosyl-mannotriose	Megazyme, Ireland
6^3, 6^4-α-D-Galactosyl-mannopentaose	Megasyme, Ireland
Aloe vera mannan (Acemannan)	Elicityl, France
Birch xylan	Made in house
Mannohexaose	Megazyme, Ireland
Mannotetraose	Megazyme, Ireland
Mannotriose	Megazyme, Ireland
Spruce galactoglucomannan (R5K)	Made in house
Spruce galactoglucomannan fluorescent (R5K 2AB)	Made in house

2.7 Bacterial strains

Table 2.7. Bacterial strains used in this study.

Bacterial strains	Supplier
One Shot® BL21 Star™ (DE3) Chemically Competent *E. coli*	Thermo Scientific, USA
One Shot® TOP10 Chemically Competent *E. coli*	Thermo Scientific, USA

2.8 DNA and plasmids

Table 2.8. DNA and plasmids used in this study.

DNA and plasmids	Supplier
dGTP Mix, 10 mM	New England Biolabs, USA
F. prausnitzii genomic DNA	Provided by Sylvia Duncan
pET-28a(+)	GenScript®
pNIC-CH	Addgene
Quick-Load® 1 kb DNA Ladder	New England Biolabs, USA
R. intestinalis genomic DNA	DSMZ, Germany

2.9 Proteins and enzymes

Table 2.9. Proteins and enzymes used in this study.

Protein/Enzyme	Supplier
BenchMark™ Protein Ladder	Invitrogen, USA
Q5® Hot Start High-Fidelity Master Mix 2X	New England Biolabs, USA
T4 DNA Polymerase	New England Biolabs, USA

2.10 Primers

Table 2.10. Primers used in this study.

Primers	Supplier
CEx_down (*R. intestinalis*)	Eurofins Genomics, Germany
CEx_up (*R. intestinalis*)	Eurofins Genomics, Germany
CEXcatD_rev1 (*R. intestinalis*)	Eurofins Genomics, Germany
CEXcbm_up2 (*R. intestinalis*)	Eurofins Genomics, Germany
*Fp*CE2_Forw	Eurofins Genomics, Germany
*Fp*CE2_Rev	Eurofins Genomics, Germany
*Fp*CEXX_Forw	Eurofins Genomics, Germany
*Fp*CEXX_Rev	Eurofins Genomics, Germany

3 Methods

3.1 Methods working with microbiology

When working with bacteria, and when there's a risk of contamination, sterile equipment in sterile cabinets, gloves and lab coat were used. The equipment and tools were either used directly from sterile wrapping, autoclaved or sterilized with 70 % ethanol. For non-disposable equipment and tools, sterilizing was performed before and after use.

3.2 Preparation of culture media and agar

For cultivation of the *E. coli* strains different media were used. Lysogeny broth (LB) was used for overnight inoculation of cells and in agar plates, while Terrific broth (TB) was used for expression of proteins.

LB medium:

- Bacto™ Tryptone 2.5 g
- NaCl 2.5 g
- Bacto™ Yeast Extract 1.25 g

The components were mixed with 250 mL dH_2O, and then autoclaved. Prior to use, kanamycin was added to a final concentration of 50 μg/mL.

Agar plates:

- Agar 7.5 g
- Bacto™ Tryptone 5 g
- NaCl 5g
- Bacto™ Yeast extract 2.5 g
- Sucrose 25 g

Agar, tryptone and NaCl were mixed with 400 mL dH_2O, and the sucrose was mixed with 100 mL dH_2O before both bottles were autoclaved. After cooling down to around 50 °C, the solution with sucrose was added to the rest. Kanamycin was added when the temperature reached approximately 30 °C to a final concentration of 50 μg/mL. The medium was then poured into petri dishes in a sterile cabinet to cool down and solidify.

For agar plates without sucrose, the same amount of agar, tryptone and NaCl were used and mixed with 500 mL dH_2O and autoclaved, before following the same procedure as outlined above. All agar plates were stored at 4 °C prior to use.

TB broth:

- Bacto™ Tryptone 6 g
- Bacto™ Yeast extract 12 g
- Glycerol 85 % v/w 2 ml

Tryptone, yeast extract and glycerol 85 % were mixed with dH_2O to a finale volume of 450 mL and autoclaved. Before use of TB, 50 mL of phosphate solution was added to medium.

Super Optimal Broth with Catabolite repression (SOC) medium:

Pre-made by Invitrogen.

Antibiotics:

All vectors used in this study contain genes for Kanamycin resistance. This enables selectivity for cells containing the cloned vector. Kanamycin was made by dissolving 500 mg in 10 mL dH_2O to a final concentration of 50 mg/mL. The solution was sterile filtered using a 0.22 μM Filtropur filter and stored in 1 mL aliquots at -20 ºC.

Isopropyl β-D-1-thiogalactopyranoside (IPTG):

1 M solution of IPTG was made by dissolving 2.383 g in 10 mL dH_2O. The solution was sterile filtered using a 0.22 µM Filtropur filter and stored in 1 mL aliquots at -20 ºC.

3.3 Buffers

Potassium phosphate solution for TB Broth Medium:

- 0.17 M KH_2PO_4 23.14 g
- 0.72 M K_2HPO_4 125.41 g

The components were dissolved in dH_2O to a final volume of 1 L and autoclaved.

Sodium phosphate buffer:

- 0.2 M NaH_2PO_4 24 g
- 0.2 M Na_2HPO_4 28.4 g

NaH_2PO_4 and Na_2HPO_4 were made up in separate bottles and dH_2O added to a final volume of 1 L. The solutions were sterile filtered using a 0.22 µM Filtropur filter and stored at room temperature. These solutions were used as stocks to make sodium phosphate solutions with different concentrations and pH's using table at Sigma-Aldrich.com.

Tris buffer:

- 1 M Tris tris(hydroxymethyl)aminomethane 121.14 g

121.14 g Tris base was dissolved in dH_2O to a final volume of 1 L. The 1 M Tris solution was used as a basis for making different buffers containing Tris. Buffers were made by diluting with dH_2O from the 1 M Tris stock and pH were adjusted with HCl. Tris buffers were sterile filtered using a 0.22 µM Filtropur filter and stored at room temperature.

Ion Metal Affinity Chromatography (IMAC) buffers:

Binding buffer (Buffer A)

- 50 mM Tris-HCl pH 8.0
- 5 mM Imidazole 0.34 g
- 250 mM NaCl 14.61 g

Elution buffer (Buffer B)

- 50 mM Tris-HCl pH 8.0
- 500 mM Imidazole 34 g
- 250 mM NaCl 14.61 g

The components of Buffer A and Buffer B were dissolved in dH_2O in separate bottles to a final volume of 1 L, sterile filtered using a 0.22 µM Filtropur filter and stored at room temperature.

Size Exclusion Chromatography (SEC) Buffer

- 50 mM Sodium Phosphate pH 5.9
- 250 mM NaCl 14.61 g

224.625 mL 0.2 M NaH_2PO_4 and 25.375 mL 0.2 M Na_2HPO_4 were diluted with dH_2O to a volume of 1 L, and 14.61 g NaCl was dissolved into the sodium phosphate solution. The solution was then sterile filtered using a 0.22 µM Filtropur filter and stored at room temperature.

3.4 Cloning

3.4.1 Primers:

Primers (Table 3.1) were designed so the genes could be cloned into the pNIC-CH vector (Appendix A). Each of the genes to be cloned needs two primers, working from each end of the gene sequence and with an added sequence before and after the gene sequence that is complementary with the pNIC-CH vector.

The primers were ordered from Eurofins Genomics. They were provided in solid form and were dissolved and diluted to 10 pM in dH_2O prior to use.

Table 3.1. Primers used during cloning.

Gene	Forward primer	Reverse primer
*Fp*CE2	*Fp*CE2_Forw	*Fp*CE2_Rev
*Fp*CEXX	*Fp*CEXX_Forw	*Fp*CEXX_Rev
*Ri*CEX_CATD	CEx-up	CEXcatD_rev1
*Ri*CEX_CBM	CEXcbm_up2	CEx_down

3.4.2 Polymerase Chain Reaction:

The genes were amplified using Polymerase Chain Reaction (PCR). This technique makes many copies of the DNA sequence of interest using specific primers. The PCR machine use heat to denature the DNA strand, and a thermostable polymerase and primer annealing use dNTP to synthesize new double stranded DNA from the DNA template. In this study a Q5® High-Fidelity 2X Master Mix was used, which contains the thermostable DNA polymerase, deoxynucleoside triphosphates, Mg^{++} and reaction buffer.

Materials:

- Q5® Hot Start High-Fidelity Master Mix 2X
- Primers
- DNA template (genomic DNA)

Each of the genes have two primers that were used for amplifying the specific gene from the genomic DNA (Appendix B). For each gene, a mix with 25 µL Q5 X2 MasterMix, 19 µL dH_2O, 2.5 µL 10 pM Forward primer, 2.5 µL pM Reverse primer and 1 µL genomic DNA (DNA template) were made in 0.2 ml PCR tubes on ice. This was run in the Mastercycler gradient PCR with the program in table 3.2.

Table 3.2. PCR program for amplification of the genes from genomic DNA.

Step	Temperature	Time	Cycles
Initial denaturation	98 °C	30 sec	1
Denaturation	98 °C	10 sec	35
Primer annealing	57 °C	15 sec	
Primer extension/ Elongation	72 °C	45 sec	
Final extension	72 °C	10 min	1
Storage?	4 °C	Hold	-

3.4.3 Agarose Gel Electrophoresis:

To verify the PCR product agarose gel electrophoresis was conducted. This is a common technique used to separate DNA molecules and determine their size. Agarose gel allows DNA fragments to travel through the gel when an electric current is applied, and the negatively charged DNA will move against the positive pole in the electric field. Agarose gels have pores that allows smaller molecules to travel faster through the gel, thus separating them by size. By marking the DNA samples with dye and using a ladder with known band sizes, the size of DNA can be determined by using UV light.

Materials:

- Agarose powder
- Quick-Load® 1 kb DNA Ladder
- DNA Gel Loading Dye 6X
- peqGreen DNA/RNA Dye

A 1.6 % agarose gel was used for these genes. 0.8 g agarose was mixed with 50 mL 1X TAE buffer and heated in a microwave until the agarose had dissolved. After cooling to 55 °C, 2.5 µL peqgreen DNA dye was added, and the solution was poured onto a gel tray, with a comb attached to make the wells, for further cooling to solidify. 5 µL Quick-Load® 1 kb DNA Ladder was added to the first well and 5 µL samples, premixed with 1 µL 6X Gel Loading Dye, were added to the other wells (for colony or plasmid, 15 µL sample and 3 µL 6X Gel Loading Dye

was used). The gel was run on 90 V for approximate 45 min. The gel was visualized using a Gel Doc EZ Imager, to show the DNA bands.

3.4.4 PCR clean-up:

After confirming the results from the PCR, the DNA in the samples were isolated using a PCR clean-up kit with premade buffers (NT1 and NT3).

Materials:

- NucleoSpin Gel and PCR Clean-up Kit
- DNA samples from PCR reaction

45 µL from the PCR reaction and 90 µL NT1 was mixed in a new Eppendorf tube, before transferring it to the NucleoSpin and PCR clean-up tube. This was centrifuged at 11,000 g for 1 minute, and the flow-through was discarded. 700 µL of the washing buffer NT3 was added and centrifuged at 11,000 g for 1 minute, and the flow-through was discarded. This step was repeated once, before drying the membrane by centrifuging it at 11,000 g for 2 minutes. The column was then placed in a new Eppendorf tube and 25 µL dH_2O was added. This was incubated for 1 minute at room temperature before centrifugation at 11,000 g for 1 minute. The eluted DNA in the flow-through was put in -20 °C prior to use. A Heraeus™ Pico™ 21 Centrifuge was used for all the centrifugations.

3.4.5 DNA concentration:

After PCR clean-up the DNA concentration in the samples were measured using the Qubit™ BR Kit.

Materials:

- Qubit™ dsDNA BR Assay Kit
- DNA sample

First a Qubit™ working solution was made with Qubit™ dsDNA BR Reagent and Qubit™ dsDNA BR Buffer in 1:200 solution. A standard curve for the Qubit™ instrument was made by using 10 µL of standard 1 and 2 in 190 µL of the working solution, for further calculations of the DNA concentration. 199 µL of working solution were mixed with 1 µL DNA for measurement. All samples were prepared in 0.5 mL Qubit™ tubes, vortexed and incubated for 2 minutes at room temperature before they were measured with the Qubit™ fluorometer.

3.4.6 T4 DNA digestion for cloning into pNIC-CH vector:

The pNIC-CH plasmid and gene requires treatment with the T4 DNA polymerase to be able to insert the gene into the vector. The T4 DNA polymerase make single-stranded overhangs on the vector and the gene, with complementary bases (G on the gene, and C on the vector) to anneal them. For gene digestion, with dGTPs added to the reaction, the T4 polymerase 3'->5' exonuclease activity is favoured over the polymerase activity until the first guanin is reached.

Materials:

- T4 DNA Polymerase
- Dithiothreitol (DTT)
- dGTPs
- NEBuffer™ 2.1
- Genes
- pNIC-CH plasmids (pre-digested)
- 25 mM Ethylenediaminetetraacetic acid (EDTA)

To prepare the genes to LIC cloning they were digested with T4 DNA Polymerase. The genes were used in concentrations of 0.2 pmol. A mix with 1 µL of the gene, 2 µL dGTP, 2 µL NEB Buffer 2.1, 1 µL DTT, 1 µL T4 DNA Polymerase and 13 µL dH_2O were mixed in 0.2 mL PCR tubes on ice. This mix was run in the PCR machine for 60 minutes at 22 °C, before additional 21 minutes at 75 °C to deactivate the T4 DNA Polymerase. 4.5 µL of the genes digested with the T4 DNA Polymerase were mixed with 0.5 µL T4 pre-digested plasmids and incubated for 60 minutes at room temperature. Then 2 µL 25 mM EDTA was added and incubated for another 10 minutes at room temperature before transformation into One Shot® Top10 Chemically Competent *E. coli*.

3.4.7 Transformation of One Shot® TOP10 Chemically Competent *E. coli*:

One Shot® TOP10 Chemically Competent *E. coli* cells were used for the cloning, as they have replication of high-copy number plasmids.

Materials:

- One Shot® TOP10 Chemically Competent *E. coli*
- Plasmid/gene solution
- SOC medium
- Agar plates (containing kanamycin and sucrose)

The One Shot® TOP10 Chemically Competent *E. coli* cells were thawed on ice for 10 min, before transferring 50 µL cells and adding the DNA solution to empty, prechilled Eppendorf tubes. This was put on ice for 5 minutes, then heat shocked in water bath at 42 °C for 30 seconds, before additional 2 minutes on ice. 200 µL SOC medium was mixed with the cells and incubated at 37 °C with 220 rpm shaking for 60 minutes. The cells were then plated on agar plates containing a concentration of 50 µg/mL kanamycin and 5 % sucrose. For each gene two agar plates were prepared, one with 75 µL of cells and another with the rest. The agar plates were incubated at 37 °C overnight. The SOC medium and the agar plates were pre warmed at 37 °C before use.

3.4.8 Colony PCR:

Colony PCR was used to verify that the genes had been inserted into the One Shot® TOP10 Chemically Competent *E. coli* colonies.

Materials:

- Transformed One Shot® TOP10 Chemically Competent *E. coli*
- Q5® Hot Start High-Fidelity Master Mix 2X
- Forward and reverse primers (for the specific gene)

A small portion of the colony was put in PCR tubes using a toothpick, and heat shocked in a microwave for 1 minute to rupture the cell wall. Then it was put on ice and mixed with 7.5 µL Q5 X2 MasterMix, 5 µL dH_2O and 0.75 µL Forward Primer and Reverse Primer (10 pM), and the PCR program was used as earlier in section 3.4.2. The PCR products were run on a gel to verify the gene insert (procedure as in section 3.4.3).

The genes could eventually be verified after plasmid purification, using 1 µL plasmid instead of cells from a colony and run the same PCR program and a gel.

Colonies with the gene inserted were inoculated in 10 mL LB medium with kanamycin (50 µg/mL concentration) in 15 mL Falcon tubes at 37 °C with 220 rpm shaking overnight.

3.4.9 Plasmid purification:

Plasmid were purified from One Shot® TOP10 Chemically Competent *E. coli* cultures for sequencing and further transformation into One Shot® BL21 Star™ (DE3) Chemically Competent *E. coli*.

Materials:

- NucleoSpin Plasmid Kit
- Overnight cell culture

From the 10 mL overnight cultures 1 mL cell culture were put aside in 1.5 mL Eppendorf tubes for glycerol stocks, and the rest were spun down at 4500 rpm for 10 minutes at 4 °C in the Mega Star 1.6R centrifuge. The supernatants were discarded, and the cell pellets dissolved in 250 µL buffer A1 from the NucleoSpin Plasmid Kit and transferred to Eppendorf tubes. Continuing the purification, the protocol was followed as described in the manual. The plasmid concentrations were then examined using Qubit™ (section 3.4.5).

A Heraeus™ Pico™ 21 Centrifuge was used for all the centrifugations.

3.4.10 Sequencing:

To verify that the cloning was successful the genes needed to be sequenced. Samples with 5 µL plasmid and 5 µL 10pM pLIC forward or reverse primer were mixed in Eppendorf tubes, labelled with barcodes and sent for sequencing at Eurofins Genomics. Plasmids with correct gene sequence were used further in the study.

3.4.11 Transformation of One Shot® BL21 Star™ (DE3) Chemically Competent *E. coli*:

The plasmids were transformed into One Shot® BL21 Star™ (DE3) Chemically Competent *E. coli* cells for expression of proteins.

Materials:

- One Shot® BL21 Star™ (DE3) Chemically Competent *E. coli*
- pNIC-CH vector containing the gene of interest
- SOC medium
- Agar plates (containing kanamycin and sucrose)

The transformation of One Shot® BL21 Star™ (DE3) Chemically Competent *E. coli* cells were performed with the same procedure as for One Shot® TOP10 Chemically Competent *E. coli* (section 3.4.7) with some changes in the protocol. 25 µL of cells and 1 µL plasmid were used prior to heat shock, and 75 µL SOC medium was added afterwards. After the incubation 25 µL of cells were used for agar plates.

3.4.12 Glycerol stocks:

For long-term storage of the *E. coli* strains, glycerol stocks were made. Glycerol protect the bacterial cell walls from damage during the freezing temperatures. 765 µL of bacterial culture and 235 µL sterile glycerol 85 % were mixed and frozen at -80 °C. All *E. coli* clones made in this study were frozen in glycerol stocks for storage. For later use, strains can be inoculated by scraping cells off the frozen stock and transferring it to growth medium.

3.5 *Ri*CEX Mutants

*Ri*CEX mutants were designed using the structure of the original *Ri*CEX in Pymol to visualize how specific amino acids can change the enzyme. The rational design of the mutants was achieved with help from Dr. Åsmund Kjendseth Røhr. A selection of mutations that can interfere and possibly block the galactose pockets or change the cavity in the active site were chosen.

The amino acid sequences with mutations were made in the software program Jalview, and the sequences were sent to, and ordered from, GenScript®, with the mutant gene pre inserted into the pET-28a(+) vector (Native amino acid sequence in appendix C). The pET-28a(+) plasmid contains a gene for kanamycin resistance and have a T7 promoter. The *Ri*CEX mutant genes were cloned in the cloning site Nco I/Xho I with a 6X His tag on the C-terminal.

Materials:

- pET-28a(+) vector containing gene of interest
- One Shot® BL21 Star™ (DE3) Chemically Competent *E. coli*
- Agar plates (containing kanamycin)

The plasmids were spun down at 6000 g for 1 minute at 4 °C in the Centrifuge 5418R, and 20 µL dH_2O were added and vortexed for 1 minute. The plasmids were transformed into One Shot® BL21 Star™ (DE3) Chemically Competent *E. coli* cells following the same protocol as in section 3.4.11.

3.6 Protein expression:

For protein expression, One Shot® BL21 Star™ (DE3) Chemically Competent *E. coli* containing the gene of interest were inoculated from either colony on agar plate or glycerol stock. This was inoculated in 10 mL LB medium containing kanamycin (50 µg/mL concentration) over night at 37 °C with shaking.

3.6.1 Harbinger:

The Harbinger system was used to express the proteins. This gave good growth of *E. coli* cells prior to inducing protein expression.

Materials:

- Overnight culture of cloned One Shot® BL21 Star™ (DE3) Chemically Competent *E. coli*
- TB medium, containing phosphate solution, antifoam and kanamycin
- 1 M IPTG

500 µL kanamycin, to a finale concentration of 50 µg/mL, and 150 µL antifoam were added to the TB broth medium containing phosphate solution, before inoculation with 1.5 mL of One Shot® BL21 Star™ (DE3) Chemically Competent *E. coli* culture. Aeration caps were used on the bottles for the Harbinger, and incubated overnight at 23 °C.

Then 100 µL 1 M IPTG was added to each bottle, to a finale concentration of 0.2 mM. The bottles were incubated overnight at 23 °C.

3.6.2 Cell harvesting:

The cell culture from each bottle was divided into two centrifuge bottles and centrifuged at 6000 rpm for 10 minutes at 4 °C in an Avanti™ J-25 Centrifuge with a JA-10 rotor. The supernatant was collected, and 30 mL was put back to resuspend the pellet, before transferring the suspended cells to 50 mL Falcon tubes and centrifuged at 4500 rpm for 30 minutes at 4 °C in the Mega Star 1.6R centrifuge. The supernatant was then discarded, and the pellet frozen at -80 °C for a few hours before it was moved to -20 °C freezer until use.

3.7 Protein purification:

3.7.1 Protein extraction:

Frozen cell pellets were thawed, filled with 30 mL IMAC Buffer A and vortexed to resuspend the pellet. The cells were then lysed with sonication in the Sonics Vibra-Cell™ Ultrasonic Processor at 30 % amplitude for 6 minutes with 5 seconds pulses, with the cells kept on ice. To separate the proteins from larger cell components, the solution was centrifuged in an Avanti™ J-25 Centrifuge with a JA-25 rotor at 15,000 g for 15 minutes at 4 °C. The supernatant was then filtered into 50 mL Falcon tubes with Filtropur S 0.22 μm PES Membrane syringe filters.

3.7.2 Immobilized Metal Affinity Chromatography:

The proteins of interest containing the 6xHis-tag was purified with the Immobilized Metal Affinity Chromatography (IMAC) method. In general, the IMAC column contain a resin with a metal ion which can bind proteins. In this study, a HisTrap™ High-Performance (HP) column with Ni Sepharose matrix was used and the target proteins had 6 histidines cloned into the C-terminal. The histidine binds to the Ni^{2+} in the column with high affinity. The proteins were loaded to the column with a binding buffer which contains a low concentration of imidazole that prevent other proteins with low affinity for the column to bind. The elution buffer contains a high concentration of imidazole which will elute the 6xHis-tagged proteins by binding to Ni^{2+}.

Materials:

- HisTrap™ HP 5 mL column
- IMAC Binding Buffer (A) and Elution Buffer (B)
- Protein solution

The HisTrap™ HP 5 mL column was used with the ÄKTA Pure chromatography system. The system, with the column connected, was flushed through with 5 column volumes (CV) dH_2O at 2.5 mL/minute to remove the 20 % EtOH storage buffer, and then 5 CV with Buffer A to prepare the system for protein binding. The sample was added to the column at 1 mL/minute to let the target protein bind to the column, and additional Buffer A was added to remove other

proteins until a stable baseline was reached. Buffer B was applied to the column in an increasing concentration gradient from 0 % to 75 % to elute the target protein, and then increased to 100 % Buffer B for 5 CV for removal of all proteins. In the elution step, 1 mL fractions were collected containing the protein of interest, until the signal reached baseline. New runs were started with 5 CV Buffer A to reequilibrate the system, removing Buffer B and prepare the system again. After purification, the system and column were first flushed with 5 CV dH_2O and then 5 CV 20 % EtOH storage buffer. The fractions containing proteins were run on SDS-PAGE to find the protein of interest and collect these fractions. The fractions were combined, and the buffer exchanged to either SEC buffer or storage buffer.

3.7.3 Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis:

After purification with IMAC, Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis (SDS-PAGE) gels were run on fractions from section 3.7.2. to determine if proteins were of correct size were purified. This method separate proteins by size (molecular mass). Proteins are made negatively charged by mixing with sodium dodecyl sulfate (SDS) or lithium dodecyl sulfate (LDS). Together with dithiothreitol (DTT), SDS or LDS and boiling will denature the proteins, allowing the proteins to travel through the gel using an electric current. The negatively charged proteins will go towards the positively charged anode, separating them relative to their size.

Materials:

- Mini-PROTEAN® TGX Stain-Free Gel
- 4X NuPAGE® LDS Sample Buffer
- 10X NuPAGE® Sample Reducing Agent
- Protein sample
- 1X TGS buffer
- Benchmark™ Protein Ladder

A stock with 65 µL 4X NuPAGE® LDS Sample Buffer and 25 µL 10X NuPAGE® Sample Reducing Agent was made (if many samples were to be tested, a larger stock were made after calculation of the amount needed). 7 µL of this stock were mixed with 13 µL of the sample, or a diluted sample, in 1.5 mL Eppendorf tube and boiled for 10 minutes. The Mini-PROTEAN®

TGX Stain-Free Gel was placed in a Mini-PROTEAN® Tetra Electrode Assembly and filled with 1X TGS buffer and put in a buffer tank that was then filled with 1X TGS buffer. 4 µL of Benchmark™ Protein Ladder and 10 µL of the samples were applied to the gel walls (for 10 walls gels 8 µL of ladder and 20 µL of sample were used). The gels were run for 17 minutes at 270 V, using a PowerPac™ 300 power supply. After the run, gels were visualized by placing them on a Stain-Free Sample Tray and using a Gel Doc EZ Imager and Image Lab Software. This confirmed that the proteins of the correct size had been produced, and fractions containing them were collected.

3.7.4 Filtration and buffer exchange:

Buffer from the purification step was replaced with a new buffer for the next purification step or a storage buffer by using Vivaspin 20, 10,000 MWCO PES filters. The enzyme solutions were spun down at 4500 rpm and 4 °C in the Mega Star 1.6R centrifuge. It was topped up with new buffer three times to exchange the old buffer, and to concentrate the enzyme solution.

3.7.5 Size Exclusion Chromatography:

After IMAC, a second purification step was conducted. Size Exclusion Chromatography (SEC) separates the proteins based on their size. The SEC column consist of beads that contains pores, where small proteins will enter, thus using longer time to move through the column than bigger proteins.

Materials:

- HiLoad™ 16/600 Superdex™ 75 prep grade column
- SEC Buffer
- Protein sample

A HiLoad™ 16/600 Superdex™ 75 prep grade column was used on the Äkta purifier. The 20 % EtOH storage buffer was washed out with 120 mL dH_2O, before an additional 120 mL buffer to condition the system. IMAC samples were buffer exchanged with the SEC eluent followed by upconcentration and added to the 2 mL sample loop and loaded onto the column. SEC

fractionation was conducted by running 120 mL elution buffer to elute the proteins at a flow rate of 0.8 mL/minute. Eluted proteins were monitored by absorbance at 280 nm and collected in 1 mL fractions. After purification, the system and column were washed with dH_2O and then put in 20 % EtOH for storage. The fractions containing proteins were examined with SDS-PAGE, and proteins with the correct size and purity were pooled for buffer exchange and up concentration.

3.7.6 Protein concentration:

The protein concentration was measured with the Bradford method, after purification and buffer exchange. This was done by mixing 799 µL dH_2O, 200 µL Protein Assay Dye Reagent Concentrate and 1 µL of the protein in triplicates, and a blank sample containing dH_2O instead of protein. Absorbance was measured at 595 nm in the BioPhotometer® D30 after 5 minutes incubation.

3.7.7 Storage of enzymes:

After purification and measurement of protein concentration, the enzymes were either stored at 4 °C for immediate use or at -20 °C in 100 – 500 µL aliquots for later use.

3.8 Enzyme characterization

Galactoglucomannan from spruce was made by Dr. Leszek Michalak and Dr. Bjørge Westereng, (La Rosa et al., 2019a; Michalak et al., 2018). Fluorescent labelled galactoglucomannan was made by Dr. Shaun Allan Leivers (Ruhaak et al., 2010).

3.8.1 MALDI-Tof MS

Matrix-Assisted Laser Desorption/Ionizaton Time of Flight Mass Spectrometry (MALDI-ToF MS) was used for analysis of enzyme activity on various substrates. MALDI-ToF MS can determine the weight of large molecules, which can then be identified. The sample is mixed

with a matrix as a liquid and fixed to a metal plate by drying. By using a laser in pulses and in vacuum, the molecules in the sample matrix are ionized with a soft ionization method that prevents fragmentation of the molecules. The laser heats up a part of the sample matrix solution, which is then vaporized, and brought into a gas phase. The ionized molecules are accelerated into the mass spectrometer by a magnetic field, separating the ions based on their mass to charge ratio (*m/z*). The Time of Flight measures the time the ion uses to travel a certain distance to the detector, were small molecules travels faster than bigger ones, giving the value of the mass to charge ratio.

Materials:

- UltrafleXtreme™ MALDI-ToF with a nitrogen 337 nm laser beam
- Matrix solution (2,5-dihydroxybenzoic acid and acetonitrile)
- MTP 384 ground steel target plate
- Reaction components (enzyme, substrate and buffer)

In general, for MALDI-ToF MS reactions, 1 µL reaction sample was mixed with 2 µL matrix solution, by pipetting a few times, on the MTP 384 ground steel target plate in a clean spot. The matrix solution was prepared by mixing 4.5 mg 2,5-dihydroxybenzoic acid, 150 µL acetonitrile and 350 µL dH_2O. The samples were dried with a warm stream of air (using a hairdryer). For use of the UltrafleXtreme™ and analysis of the data, the software FlexControl and FlexAnalysis were used, respectively. Samples often needed 10 to 100-fold dilutions with dH_2O to achieve optimum data quality.

The activity of *Fp*CE2 and *Fp*CEXX were tested on spruce GGM (GH26 treated), *Aloe vera* mannan, konjac glucomannan, chemically acetylated konjac glucomannan, birch xylan, acetylated cellulose and chitopentaose. The substrates, 10 µL of 10 mg/mL stocks, were mixed with 89 µL 20 mM sodium phosphate pH 5.9 buffer and 1 µL 1 mg/mL (24 µM) *Fp*CE2 and *Fp*CEXX were added, separately and combined for each substrate. Reactions ran in a thermomixer with 600 rpm shaking at 30 °C overnight.

*Ri*CEX mutants, including the truncated version *Ri*CEX_CATD and *Ri*CEX_CBM, were screened for activity on R5K and R5K 2AB to see if all were active on those substrates or if some mutations had given a knock out effect. Reactions were set up mixing 89 µL 50 mM

sodium acetate pH 6 buffer, 10 µL 10 mg/mL R5K or 1 mg/mL R5K 2AB substrates and 1 µL 1 mg/mL enzyme in Eppendorf tubes in a thermomixer with 600 rpm shaking at 37 °C for 1 hour. Blanks with buffer and substrate were treated in the same manner as the reactions containing enzyme.

3.8.2 Activity on 4-nitrophenyl acetate

To determine pH optimum of acetylesterases 4-nitrophenyl (pNP) acetate was used. This synthetic substrate is used to examine esterase activity by measurement of acetate release during reactions.

Materials:

- 4-nitrophenyl
- 4-nitrophenyl acetate
- 50 mM Sodium phosphate buffers (pH 5.0-8.0)
- *Fp*CE2 and *Fp*CEXX
- Microtiter® 96-Well UV Microplates

A standard plot for 4-nitrophenyl (pNP) was set up using 50 mM sodium phosphate buffer in the pH range 5.0 – 5.5 – 6.0 – 6.25 – 6.5 – 6.75 – 7.0 – 7.5 – 8.0. A stock of 2 mM pNP were made with the buffers in the different pH ranges and used to make dilutions ranging from 0.001 to 2.0 mM pNP. For each pH, eight concentrations of pNP were used to get linear plots. The samples were tested in triplicates on a Synergy H4 Hybrid plate reader, using a Microtiter® 96-well plate at ambient temperature, and measuring the absorbance at 405 nm.

The pH optimum was determined using reactions with 0.5 mM pNP acetate with the different pH buffers. A 50 mM stock of pNP acetate was made by mixing it with methanol. A stock with 980 µL buffer and 10 µL 50 mM pNP acetate was mixed for each pH and 99 µL of the solution was added into eight wells of the Microtiter® 96 well plate. 1 µL of each enzyme was added to three wells and absorbance was measured at 1-minute intervals for 10 minutes, with continuous shaking. Stocks of 10 µM *Fp*CEXX and 100 nM *Fp*CE2 were used, giving a final concentration of 0.1 µM and 1 nM, respectively, in the reactions. Blanks, containing the buffer for each protein, were run under the same conditions as for enzyme reactions.

3.8.3 High-Performance Liquid Chromatography (HPLC)

High-Performance Liquid Chromatography (HPLC) was used to quantify and purify samples. HPLC is a chromatographic method used for separation of different compounds in samples, which can then be identified, purified or quantified. The high pressure gives good separation of the components in the sample because of the small particle sizes on the packing materials used in the column. Samples are going through the column were the components are separated. The components then reach a detector, for example a UV-detector, after passing the column. The time a component in the sample uses from it is injected to it reaches the detector is called the retention time, is used together with various detection principles and external standards for identification of the analytes. This gives separate peaks showing the different compounds. The size of the peak, or the changes in size over several samples, can be used to quantify the amount of the component, usually with the use of standards with known concentrations.

3.8.4 Time course HPLC

Time course was performed with *Fp*CE2 and *Fp*CEXX, individual or both enzymes combined, to quantify the acetate release over time from R5K.

Materials:

- RSLC Ultimate 3000 UHPLC
- REZEX ROA-Organic Acid H+ 300x7.8mm ion exclusion column
- *Fp*CE2 and *Fp*CEXX, 10 uM stocks
- Galactoglucomannan degraded by an endo-mannanase GH26
- 50 mM Sodium phosphate pH 6.75 buffer
- 4 mM H_2SO_4
- Acetonitrile

11 samples with 200 µL of 100 mg/mL R5K GH26 digest with 50 mM sodium phosphate pH 6.75 buffer were prepared in Eppendorf tubes. The substrate samples were put in the thermomixer at 35 °C with 600 rpm shaking before 1 µL of each of 10 µM *Fp*CE2 and 10 µM *Fp*CEXX and 1 µL from a 10 µM mix of both combined were added to three samples each.

One sample was used as a blank control, and the last one was treated with NaOH to estimate the total acetylation content in the sample. When the enzymes were added, and mixed by pipetting a few times, the first sample (20 μL) was immediately taken out and mixed with 20 μL acetonitrile (Acn) and put on a 100 °C heating plate for 2 minutes (until visible boiling was seen) to quench the reaction and denature the enzyme. New samples were taken out and treated the same way in 5 minutes intervals for 30 minutes. The control samples were taken in the same time intervals and treated in the same manner as enzyme reactions.

The acetate content in the samples were analysed on a RSLC Ultimate 3000 HPLC, using a REZEX ROA-Organic Acid H+ 300x7.8mm ion exclusion column. The injection volume was 5 μL, and separation was conducted at 65 °C with 0.6 mL/minute isocratic elution of 4 mM H_2SO_4 as mobile phase. The UV detector was set to 210 nm. Chromeleon software was used.

3.8.5 Hydrophilic Interaction Liquid Chromatography Mass Spectrometry

Hydrophilic Interaction Liquid Chromatography Mass Spectrometry (HILIC-LC-MS) was conducted on native R5K and fluorescent labelled R5K-2AB treated with the *Ri*CEX mutants, number #5, 6, 7, 8, 9 and 12, (Table 4.2) and then GH26 treated afterwards.

Materials:

- Dionex UltiMate 3000 RSLC UHPLC system
- Xbridge Amide 3.5 μm particle size 4.6x250 mm column (R5K)
- Phenomenex bioZen 2.6 μm Glycan column (R5K-2AB)
- Acetonitrile
- 50 mM Ammonium formate pH 4.5
- Reaction components (enzyme, substrate, buffer)
- Vivaspin 3K 500 filters

Reactions were set up with *Ri*CEX mutants, native *Ri*CEX and blanks without enzyme. 178 μL 50 mM sodium acetate pH 6.0 buffer was mixed with 20 μL 10 mg/mL R5K 2AB substrate and 2 μL 1 mg/mL enzymes in Eppendorf tubes. The reactions were left in a thermomixer for 1 hour at 37 °C with 600 rpm shaking. The samples were then filtered with pre-washed Vivaspin 3K 500 filters to remove the enzymes. 99 μL of the samples was added to new Eppendorf tubes

and 1 μL GH26 was added, before a new run in the thermomixer for 1 hour at 37 °C with 600 rpm shaking. Two samples of blanks were prepared, one without any enzyme treatment and one with the GH26 treatment, both were run two times in the thermomixer as the other reactions.

The R5K samples were analysed on a Dionex UltiMate 3000 RSLC UHPLC system with an Xbridge Amide 3.5 μm particle size 4.6x250 mm column running at 0.6 mL / minute. Samples (5 μL) were eluted from the column using the following gradient: 0 – 5 minutes 75 % Acn, 5 – 30 minutes 75 – 50 % Acn, 30 – 35 minutes 75 – 50 % Acn. Post-column, samples were divided using a T-split (50:50), between a Velos Pro ion trap mass spectrometer and a Corona Ultra CAD detector. The samples were analysed by Chromeleon 7.2 Software for the CAD signal and the Excalibur software was used to interpret the mass spectrometry data.

The R5K-2AB samples were analysed on the Phenomenex bioZen 2.6 μm Glycan column, with 0.3 mL/minute flowrate and 50 °C column temperature. The eluents were 50 mM ammonium formate pH 4.5 and acetonitrile. Samples (2 μL) were eluted with acetonitrile using the gradient: 90 – 72 %, 0 – 16 minutes; 72 – 40 %, 16 – 20 minutes; 40 % 20 – 25 minutes; 40 – 76 %, 25 – 27 minutes; 76 %, 27 – 40 minutes. Samples were analysed with fluorescent detection, with excitation and emission wavelengths at 320 nm and 420 nm, respectively.

3.8.6 Transesterification

Transacetylation was performed on mannotriose and mannotetraose with vinyl acetate as an acetate donor. 1 mg/mL substrate stocks were prepared in 10 mM sodium phosphate pH 5.9 buffer. 1 μL of 1 mg/mL *Fp*CE2 and *Fp*CEXX were mixed with 99 μL of each substrate in an Eppendorf tube, and 50 μL vinyl acetate was then added. The reactions were left overnight in a thermomixer at 600 rpm shaking at ambient temperature. After 24 hours the samples were vortexed, to get the vinyl on the top and the aqueous layer in the bottom, and then frozen. The vinyl on the top was then removed using a pipette, and 100 μL 96 % ethanol was added to precipitate the carbohydrates and to deactivate the enzymes. The samples were then dried using an Eppendorf Concentrator plus rotovac at 30 °C. 100 μL dH_2O was added to the dried samples, giving transacetylated oligos. MALDI-ToF MS was used to analyse the results.

The same procedure was performed on *Fp*CE2 and *Fp*CEXX using vinyl propionate. For *Ri*CEX mutants, with mutations in the active site pocket, vinyl acetate, vinyl propionate and vinyl butyrate were also tested in the same manner on mannotriose, mannohexaose, galactosylmannotriose and digalactosylmannopentaose.

3.8.7 Preferred substrate

A transesterification reaction with different types of oligosaccharides were set up to test what substrates the *F. prausnitzii* esterases prefer. In addition to a mannose source, glucose and xylan (all containing β-1,4-linkages) were added, and vinyl acetate, vinyl propionate and vinyl butyrate were used as an ester source.

Materials:

- *Fp*CE2 and *Fp*CEXX
- Mannotriose, cellotetraose and xylopentaose
- Vinyl acetate, vinyl propionate and vinyl butyrate
- 10 mM Sodium phosphate pH 5.9
- Thermomixer
- MALDI-ToF
- MALDI target plate
- Isopropanol (IPA)
- Acetonitrile (Acn)
- Eppendorf Concentrator plus rotovac

The substrates were mixed equally to 1 mg/mL final concentration in 100 µL stocks with 10 mM sodium phosphate pH 5.9 buffer. 1 µL of 100 µM *Fp*CE2 and 100 µM *Fp*CEXX were added separately to 100 µL of each of the substrate stocks. The last 100 µL of sample was used as a blank. All of the vinyl esters, vinyl acetate, vinyl propionate and vinyl butyrate, were added to each sample, 30 µL of each. The samples were left overnight in a thermomixer with 600 rpm shaking at ambient temperature. After 16 hours the samples were frozen and the top layer with the vinyl esters were removed. 500 µL IPA:Acn (50:50 mix) were added and the samples were thawed by vortexing, and then dried on the Eppendorf Concentrator plus rotovac at ambient

temperature. The samples were resolubilized with 100 µL dH_2O and analysed with MALDI-ToF MS.

Acetyl position specificity

*Fp*CE2 and *Fp*CEXX were tested on substrates with acetyl groups in either 2-*O* or 3-*O*, 4-*O* and/or 6-*O* position. Transacetylated mannohexaose were made by using *Ri*CE2 and *Ri*CEX in a transesterification reaction with vinyl acetate as acetate donor, as described in section X.

Materials:

- *Fp*CE2 and *Fp*CEXX
- Mannohexaose with acetyl groups in the 2-*O* position
- Mannohexaose with acetyl groups in the 3-*O*, 4-*O* and/or 6-*O* position
- 10 mM Sodium phosphate pH 5.9

*Fp*CE2 and *Fp*CEXX were tested on both substrates. 1 mg/mL substrate stocks were prepared in 10 mM sodium phosphate pH 5.9 buffer. 1 µL 1 mg/mL of each esterase were mixed with substrate solution and left overnight in a thermomixer at ambient temperature with 600 rpm continuous shaking. MALDI-ToF MS was used to analyse the result.

3.8.8 Protein thermal shift assay

The *Fp*CE2 and *Fp*CEXX melting temperatures in different pH's were tested with a protein thermal shift assay. 50 mM Sodium phosphate with the pH range 5.0 – 8.0 were tested as buffers. 1 mg/mL esterase solutions with the different buffers were mixed with the Protein Thermal Shift™ buffer and dye, and further steps was performed according to the manual (Protein Thermal Shift™ Kit). The reactions were run on the real-time PCR and analysed with the Protein Thermal Shift™ software.

3.9 Protein Crystallization

Attempts to get protein crystals to solve the structures of *Fp*CE2 and *Fp*CEXX. Protein structures can be solved exposing high purity protein crystals with high energy X-ray beams, a technology named X-ray crystallography. The beams from X-rays radiates the protein crystal, which have a fixed three-dimensional array of the protein molecules and produces scattered beams which hits a detector. Rotating the crystal during X-ray exposure, changes the angle and intensity of the beams, a two-dimensional diffraction pattern of spots is generated. A computer program is then used to interpret the diffraction pattern and convert it into a 3D-structure of the protein by advanced calculations. There are several methods to produce protein crystals, while vapor diffusion is the most common and was also used in this study. This can be either by the hanging-drop or sitting-drop method. The protein and crystallization solution, containing buffer, salt and/or precipitant, are mixed in a droplet and placed in a sealed well containing more of the crystallization solution. A pure protein solution with a high concentration, up to 10-30 mg/mL, should be used. Different conditions of buffer, pH, salt and precipitant are often needed before a successful protein crystal is obtained. Commercially available crystallization kits contain many different pre-mixed conditions to screen for crystals. Over time, the protein and precipitant concentration will increase in the drop, as the drop and crystallization solution in the well will equilibrate, and crystal growth may occur. Finding the appropriate crystallization solution for a specific protein, will give crystal growth. A big challenge with this method is to obtain protein crystals, and crystals with of sufficient quality to yield diffraction spectra of high quality.

List of crystallization sets:

- Hampton PEG/Ion Screen HR2-126
- Rigaku Wizard™ Classic 1 & 2
- Molecular Dimensions Clear Strategy™ Screen 1 MD1-14
- Molecular Dimensions JCSG-*plus*™ Screen MD1-37

Materials:

- Screening set
- Plate with sealant, 24 and 48 well
- Cover slides, 12 and 18 mm

- Enzymes, 10 and 20 mg/mL solutions
- Solutions for optimization conditions
- Substrates; Man_3, Man_6, $GalMan_3$ and Gal_2Man_5, in 5 mg/mL Solution
- Stereomicroscope
- Vials (storage of crystals in liquid nitrogen)

Screening conditions were set up for *Fp*CE2 and *Fp*CEXX, using four commercial kits. 1 µL enzyme, both *Fp*CE2 and *Fp*CEXX in 10 and 20 mg/mL solutions, were mixed with 1 µL crystallization solution on the cover slide. On the 24 well plates, all four drops were placed on the same cover slide and 1 mL crystallization solution was added to the well before placing the cover slide on top, with the drops hanging under, to seal it. Two drops and 300 µL crystallization solution was used on the 48 well plates (these were used because of long delivery time for new cover slides). The plates were stored at room temperature, and regularly inspected with a stereomicroscope.

For the conditions with the most promising crystals, all with *Fp*CEXX, initial crystal optimization was executed. The conditions were changed in small steps, changing the buffer pH and/or the concentration of precipitant or salt. In addition, a 5 mg/mL solution of the enzymes were used, and four substrates were tried out for some conditions, attempting to get substrate-enzyme interaction. Conditions from Rigaku Wizard™ 1 nr. 29, containing 0.1 M CHES/Sodium hydroxide pH 9.5, 10 % PEG 8000 and 0.2 M NaCl, and nr. 38, containing 0.1 M CHES/Sodium hydroxide pH 9.5, 1.0 M Potassium Sodium tartrate and 0.2 M Lithium sulfate were selected for optimization. Number 29 was set up with 0.1 M CHES pH 9.0, 9.5 and 10.0, 5 %, 10 %, 15 % and 20 % PEG 8000 and 0.2 M NaCl. Number 38 was set up with 0.1 M CHES pH 7.0, 7.5, 8.0, 8.5, 9.0, 9.5 and 10.0, 1.0 M Potassium Sodium tartrate and 0.2 M Lithium sulfate. Rigaku Wizard™ 2 number 15, containing 0.1 M HEPES pH 7.5 and 1.26 M Ammonium sulfate, was set up with HEPES pH 7.0, 7.5 and 8.0 with 1.0 M or 1.26 M Ammonium sulfate. The conditions for Wizard™ 2 number 15 was also set up with Man_3, Man_6, $GalMan_3$ and Gal_2Man_5, mixed with 1 µL substrate, 1 µL crystallization solution and 1 µL 10 mg/mL *Fp*CEXX.

Selected crystals were flash frozen in liquid nitrogen in vials for storage under transport. The crystals were analysed at the European Synchrotron Radiation Facility in Grenoble, France, by Dr. Bjørge Westereng and Dr. Åsmund Kjendseth Røhr.

4 Results

4.1 Bioinformatics

4.1.1 F. prausnitzii

We identified a cluster of genes encoding carbohydrate-active enzymes potentially involved in mannan metabolism in the genome of *F. prausnitzii* SL3/3 (figure 4.1). The cluster lacks a GH26 endomannanase, as would be required to hydrolyze the mannan backbone into smaller oligosaccharides, but it equips *F. prausnitzii* with the ability to utilize manno-oligosaccharides released by other colonic microorganisms. Indeed, the cluster contains genes encoding the full enzymatic arsenal to utilize acetylated galactoglucomanno-oligosaccharides, including two esterases, *Fp*CE2 and *Fp*CEXX.

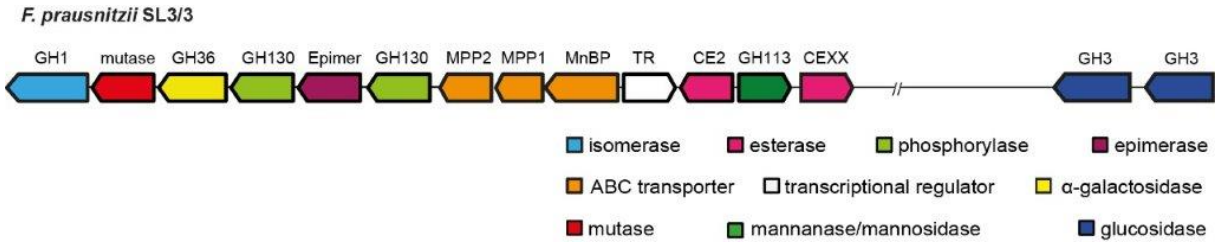


Figure 4.1. Gene cluster of mannan degrading enzymes in F. prausnitzii. The cluster contains all enzymes needed to break down complex mannans, like acetylated galactoglucomannan.

The amino acid sequence of *Fp*CE2 and *Fp*CEXX were compared with *Ri*CE2 and *Ri*CEX in a sequence alignment (Figure 4.2 and 4.3). The *R. intestinalis* esterases have previously been characterised on the same substrate as used in this study.

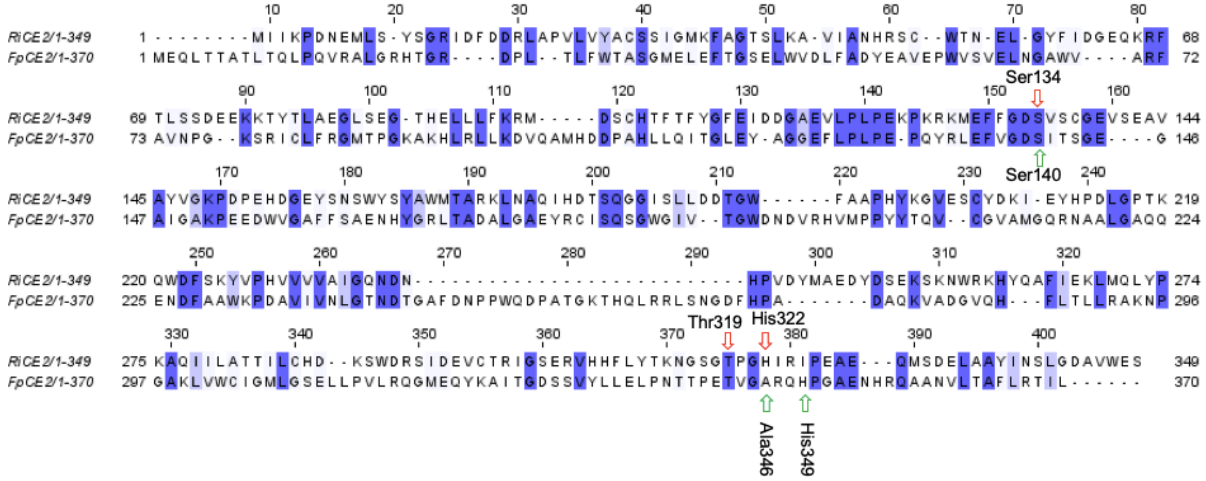


Figure 4.2. Amino acid sequence alignment of *Fp*CE2 and *Ri*CE2. InterPro analysis identified the catalytic triad residues of *Fp*CE2 as Ser140, Ala346 and His349 (marked with green arrows). The active residues of *Ri*CE2 are Ser134, Thr319 and His322 (marked with red arrows).

The *Fp*CE2 is a CAZy family 2 esterase (as defined by the CAZy database, Lombard V (2014)). Analysis with InterPro identified the residues from 18-128 in the N-terminal as a carbohydrate esterase 2, and residue 133-367 as an SGNH hydrolase-type esterase domain. It has a conserved catalytic triad, characteristic for the SGNH domain, residue annotated to Ser140, Ala346 and His349 (marked in figure 4.2), with the corresponding residues in *Ri*CE2 being Ser134, Thr319 and His322, respectively. The *Fp*CE2 share a 40 % coverage, 47 % similarity and 32 % identity with *Ri*CE2.

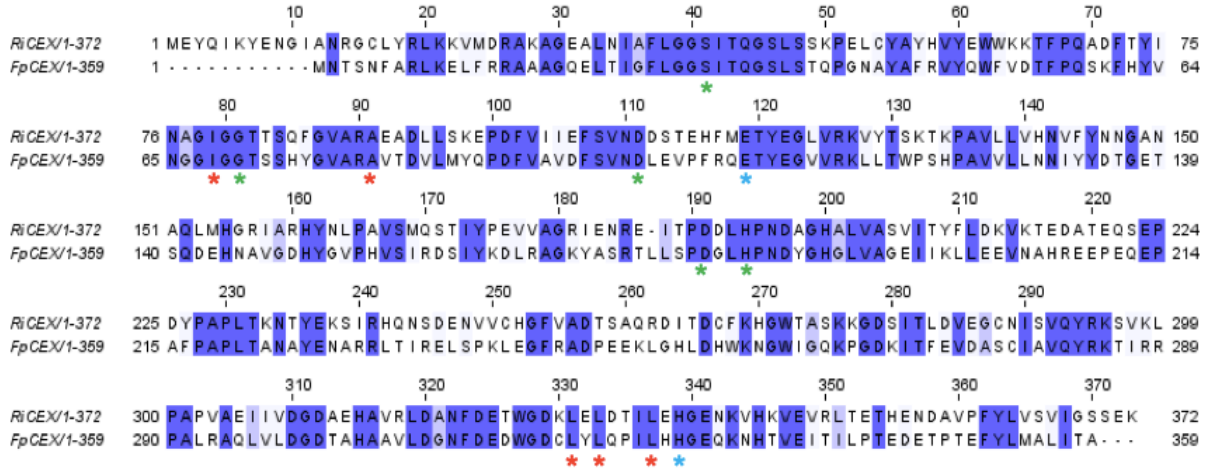


Figure 4.3. Amino acid sequence alignment of *Fp*CEXX and *Ri*CEX. Important amino acids in *Ri*CEX that share identity with *Fp*CEXX are marked. Active residues (green asterices), and amino acids that are possibly involved in stabilisation of the inter-domain interaction are labelled with red asterices (being hydrophobic amino acids) and blue asterices (charged or polar amino acids).

The residues 26-189 of *Fp*CEXX are identified as an SGNH hydrolase-type esterase domain when doing a similar analysis as above. However, this analysis did not link this fold to any carbohydrate esterase family. In figure 4.3, active site residues in *Ri*CEX are marked (green) were they share identity with *Fp*CEXX. Amino acids which are contributing to stabilize the inter-domain interaction are shown with red (hydrophobic amino acids) and blue (charged or polar amino acids). *Fp*CEXX has a 99 % coverage, 63 % similarity and 46 % identity with *Ri*CEX. Table 4.1 shows the theoretical physical and chemical properties of the two *F. prausnitzii* esterases.

Table 4.1. Theoretical physicochemical properties of *Fp*CE2 and *Fp*CEXX. The number of amino acids (aa), molecular weights (MW), isoelectric point (pI) and extinction coefficient (ε), were calculated with ExPASy ProtParam. The 6xHis-tag were included in the analysis.

Protein	aa	MW (kDa)	pI	ε
*Fp*CE2	377	41.48	5.76	68,660
*Fp*CEXX	366	41.06	5.61	51,465

4.1.2 R. intestinalis mutants

*Ri*CEX is a two domain mannan esterase that has a catalytic domain and a carbohydrate-binding module (CBM) working specific on 2-*O* acetylation's. It has a distinct clamp with the active site and the CBM contributing to ligand binding (figure 4.4). Truncated versions of either the catalytic domain or the CBM of *Ri*CEX were generated using specific primers. The amino acid sequence of *Ri*CEX, with the catalytic domain and the CBM highlighted in blue and red, respectively, is:

MEYQIKYENGIANRGCLYRLKKVMDRAKAGEALNIAFLGGSITQGSLSSKPELCYAY
HVYEWWKKTFPQADFTYINAGIGGTTSQFGVARAEADLLSKEPDFVIIEFSVNDDSTE
HFMETYEGLVRKVYTSKTKPAVLLVHNVFYNNGANAQLMHGRIARHYNLPAVSMQ
STIYPEVVAGRIENREITPDDLHPNDAGHALVASVITYFLDKVKTEDATEQSEPDYPAP
LTKNTYEKSIRHQNSDENVVCHGFVADTSAQRDITDCFKHGWTASKKGDSITLDVEG
CNISVQYRKSVKLPAPVAEIIVDGDAEHAVRLDANFDETWGDKLELDTILEHGENKV
HKVEVRLTETHENDAVPFYLVSVIGSSEK.

The native *Ri*CEX structure was used as template for making the mutants in this study (Figure 4.4 A). How each mutation may affect the enzyme and the substrate interaction is visualized *in computo* using Pymol (Example in figure 4.4 C). *Ri*CEX have cavities on each side of the active site, which could accommodate galactose side chains that are connected to the 6-*O* position on mannose units. Different amino acids in these cavities were selected for mutations, with the aim of blocking the cavities and potentially change substrate specificity. Mutations were also made in the active site, with the purpose of making a deeper pocket; this may open up the possibility for transesterification of substrates with larger molecules than acetate, like propionate and butyrate.

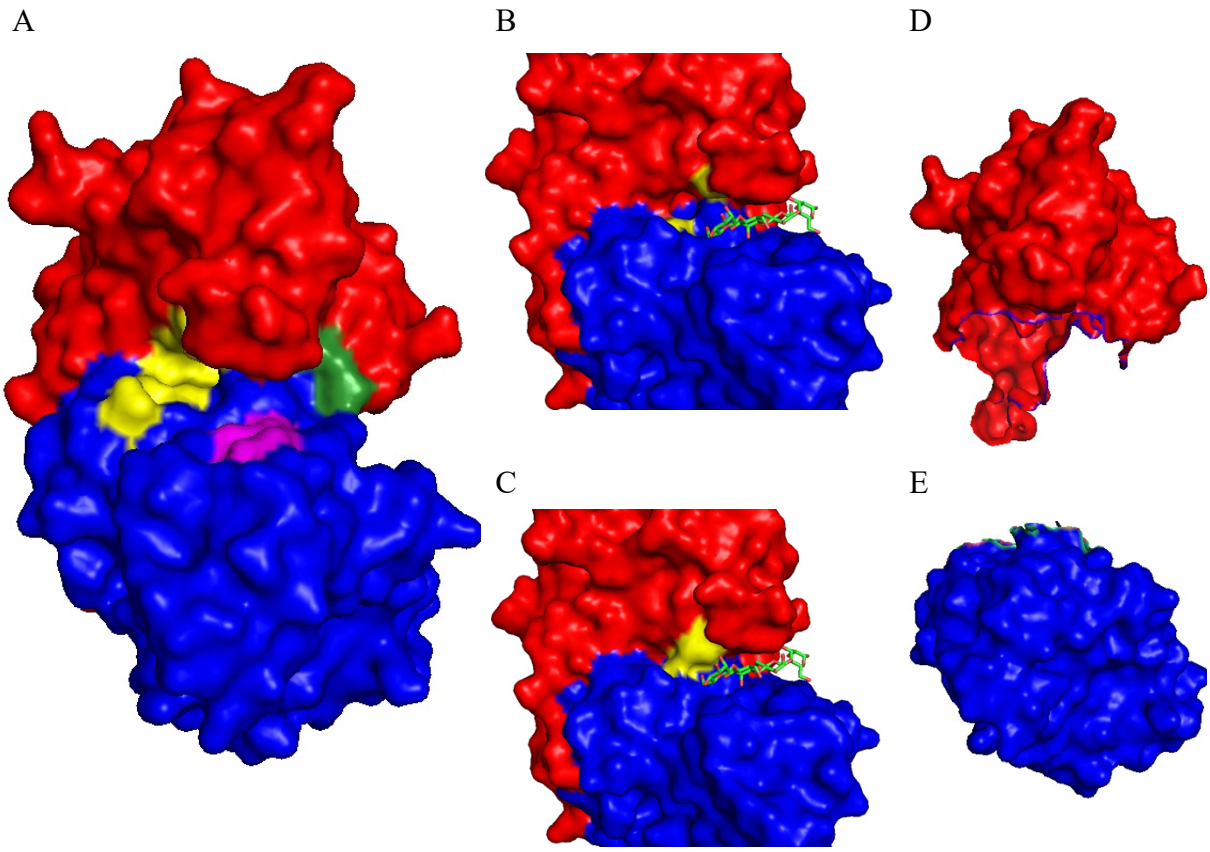


Figure 4.4. A) The structure of *Ri*CEX. The catalytic domain is showed in blue, with the active site (magenta), and the CBM in red, and the left and right galactose pocket in yellow and green, respectively. B) The native structure with two amino acids, Thr83 and Lys329 (both marked yellow), in the left galactose pocket, which were changed to Glu83 and Arg329 (mutant #7) seen in C). The aim was to block this cavity for the galactose side chains connected to the 6-*O* position on mannose. B) and C) are shown with mannotetraose. D) and E) shows the CBM- and catalytic domain, respectively, of the truncated version of *Ri*CEX.

Table 4.2 shows the mutants, including the amino acids that were changed. For the left (B) and right (C) galactose pockets seven and five mutants were made, respectively. Furthermore, four mutants in the active site were made (D). All were single mutations, except for one double. The theoretical chemical and physical properties, molecular weight, isoelectric point and extinction coefficient, were calculated using the ExPASy ProtParam software.

Table 4.2. Theoretical physicochemical properties of *Ri*CEX, including truncated version, and *Ri*CEX mutants. The name and amino acids changed for each *Ri*CEX mutants are shown in addition to the native *Ri*CEX and the truncated version, *Ri*CEX_CatD and *Ri*CEX_CBM. The number of amino acids (aa), molecular weight (MW), isoelectric point (pI) and extinction coefficient (ε), including 6xHis-tag, were calculated with ExPASy's ProtParam. In the purpose of mutation column is the truncated domains marked with A, the left side galactose pockets with B, right galactose pockets with C and the active site with D. These are also marked in figure 4.4 A.

Protein	aa	MW (kDa)	pI	ε	Aa changed in mutant (from→to)	Purpose of mutation
*Ri*CEX	379	42.48	5.58	47,580		
*Ri*CEX_CatD	228	25.68	6.21	30,495		A
*Ri*CEX_CBM	156	17.54	5.53	17,085		A
*Ri*CEX_Ser112 (#1)	379	42.45	5.65	47,580	Asp112→Ser112	B
*Ri*CEX_Thr112 (#2)	379	42.47	5.65	47,580	Asp112→Thr112	B
*Ri*CEX_Ala112 (#3)	379	42.44	5.65	47,580	Asp112→Ala112	B
*Ri*CEX_Lys83 (#4)	379	42.51	5.65	47,580	Thr83→Lys83	B
*Ri*CEX_Arg83 (#5)	379	42.54	5.65	47,580	Thr83→Arg83	B
*Ri*CEX_Arg329 (#6)	379	42.51	5.58	47,580	Lys329→Arg329	B
*Ri*CEX_Glu83_Arg329 (#7)	379	42.54	5.52	47,580	Thr83+Lys329→ Glu83+Arg329	B
*Ri*CEX_Gln266 (#8)	379	42.51	5.58	47,580	Cys266→Gln266	C
*Ri*CEX_Arg266 (#9)	379	42.53	5.65	47,580	Cys266→Arg266	C
*Ri*CEX_His266 (#10)	379	42.52	5.63	47,580	Cys266→His266	C
*Ri*CEX_Ala266 (#11)	379	42.45	5.58	47,580	Cys266→Ala266	C
*Ri*CEX_Asn264 (#12)	379	42.45	5.58	47,580	Thr264→Asn264	C
*Ri*CEX_Thr41 (#13)	379	42.45	5.58	47,580	Ser41→Thr41	D
*Ri*CEX_Val42 (#14)	379	42.47	5.58	47,580	Ile42→Val42	D
*Ri*CEX_Ala42 (#15)	379	42.44	5.58	47,580	Ile42→Ala42	D
*Ri*CEX_Val192 (#16)	379	42.47	5.58	47,580	Leu192→Val192	D

4.2 Cloning, Expression and Purification

4.2.1 *F. prausnitzii*

The *Fp*CE2 and *Fp*CEXX genes were isolated from the genomic DNA from *F. prausnitzii* using the specific primers described in section 3.4.1. The genes of *Fp*CE2 (1113 bp) and *Fp*CEXX (1080 bp) were successfully amplified using PCR (Figure 4.5 A), and then cloned and transformed into competent *E. coli* cells containing the plasmid with the corresponding gene of the correct size (Figure 4.5 B).

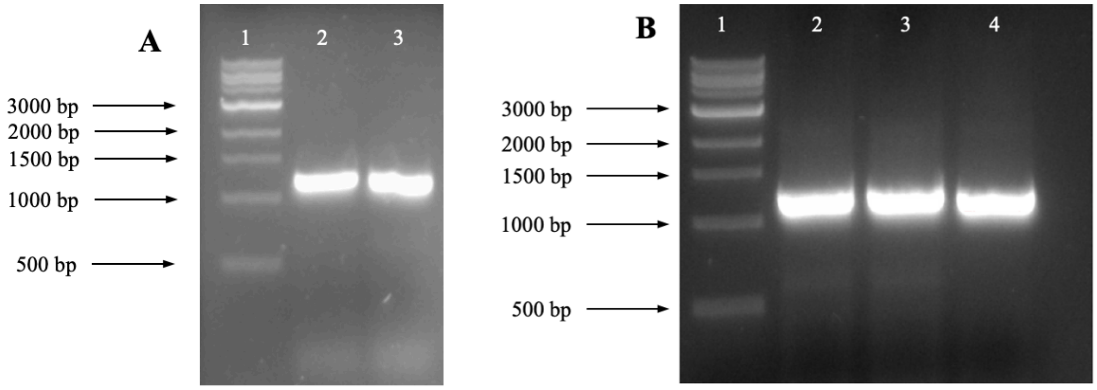


Figure 4.5. A shows the amplified genes after PCR of *Fp*CE2 (1113 bp) and *Fp*CEXX (1080 bp) in lane 2 and 3, respectively, obtained from the F. prausnitzii genomic DNA. B shows that the genes were cloned into the plasmid and transformed and produced in the One Shot® TOP10 cells successfully. Lane 2 and 3 contain the *Fp*CE2 and lane 4 the *Fp*CEXX.

Transformants were verified by sequencing and plasmids were then transferred into One Shot® BL21 Star™ (DE3) Chemically Competent *E. coli* cells. They were grown using the Harbinger system as described in section 3.6.1. Protein expression was induced using IPTG. The first purification step was conducted with IMAC to separate the target proteins from other proteins in the sample. The *Fp*CE2 and *Fp*CEXX containing a C-terminal 6xHis-tag bound successfully to the column and were collected during the eluting step (peak 2 in figure 4.6). Fractions containing the target proteins were combined for the next purification step.

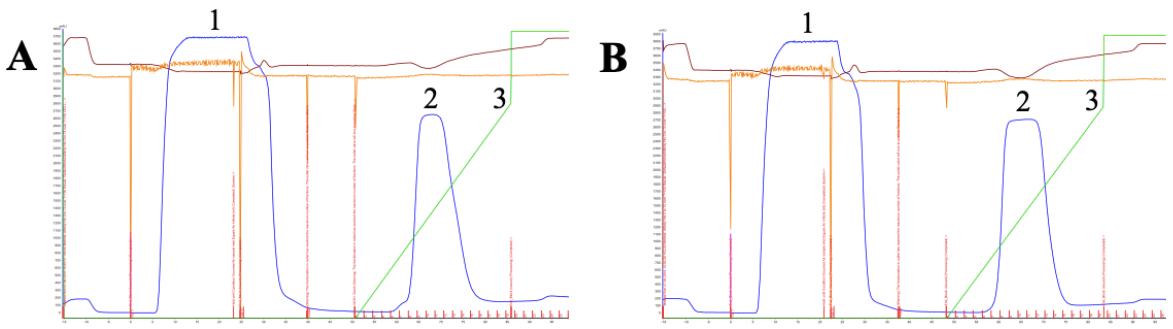


Figure 4.6. Chromatograms of the purification of *Fp*CE2 (A) and *Fp*CEXX (B) using IMAC. Both proteins bound to the column while the other proteins were washed out (peak 1); *Fp*CE2 and *Fp*CEXX were collected (peak 2) during the eluting step (3).

A 'polishing' purification step with size exclusion chromatography (SEC) followed the IMAC. During elution, *Fp*CE2 showed two peaks and *Fp*CEXX one peak (figure 4.7), and the results from SDS-PAGE confirmed that proteins with the seemingly correct size were expressed, approximately 41 kDa. Both peaks in *Fp*CE2 contained protein of the correct size. This purification step resulted in samples with the target proteins, purified with relatively small amounts of other contaminating proteins. Fractions containing the highest amounts of proteins were combined for buffer exchange and up concentration. The two peaks for *Fp*CE2 were collected separately for further examination.

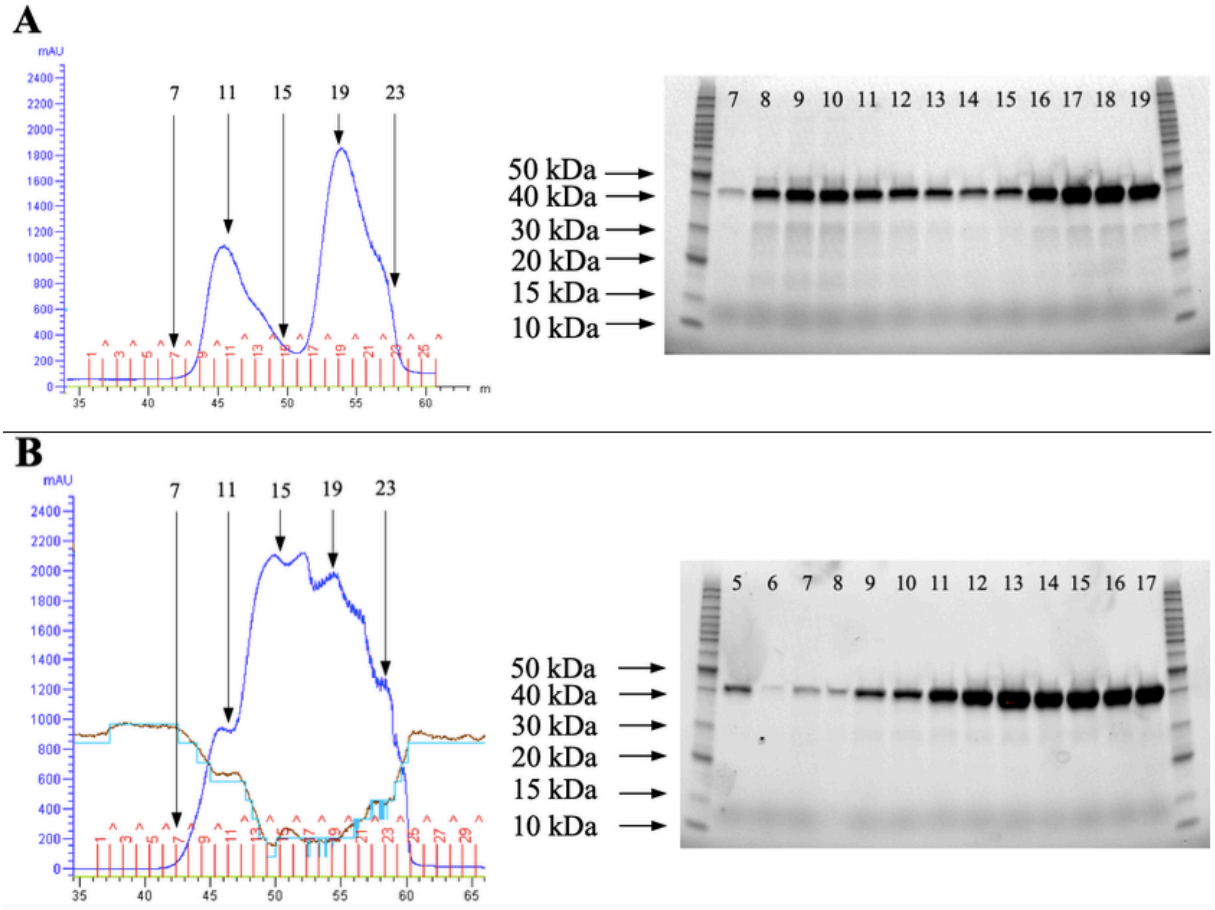


Figure 4.7. Purification of *Fp*CE2 (A) and *Fp*CEXX (B) with SEC, using a HiLoad™ 16/600 Superdex™ 75 prep grade column, and SDS-PAGE analysis. Fractions containing eluted protein (fraction 7 to 25 in A and fraction 5 to 26 for B) were collected for SDS-PAGE. In both gels, the outer lanes contain the BenchMark™ protein ladder. The lanes in both gels are marked with each collected fraction during SEC. Fractions of *Fp*CE2 (41 kDa) were collected and combined in two different stocks, lane 8 – 12 and 16 – 19 and two more fractions (data not shown). *Fp*CEXX (41 kDa) were collected from lane 11 – 17 and five more fractions (data not shown).

During buffer exchange for *Fp*CE2, the proteins from the first peak precipitated and were discarded. The second peak was fine and used for the further analyses.

4.2.2 *Ri*CEX Mutants

Cloning of *Ri*CEX_CatD and *Ri*CEX_CBM

The native *Ri*CEX is a two-domain enzyme containing a catalytic domain and a carbohydrate binding module (CBM) of 666 and 350 base pars, respectively. Truncated versions of the

protein containing single domains were generated using the specific primers in (table 3.1). The cloning and transformation resulted in three cloned CatD domain plasmids and four cloned CBM domain plasmids (Figure 4.8 A). Figure 4.8 B shows the two domains together with the native *Ri*CEX (1116 bp).

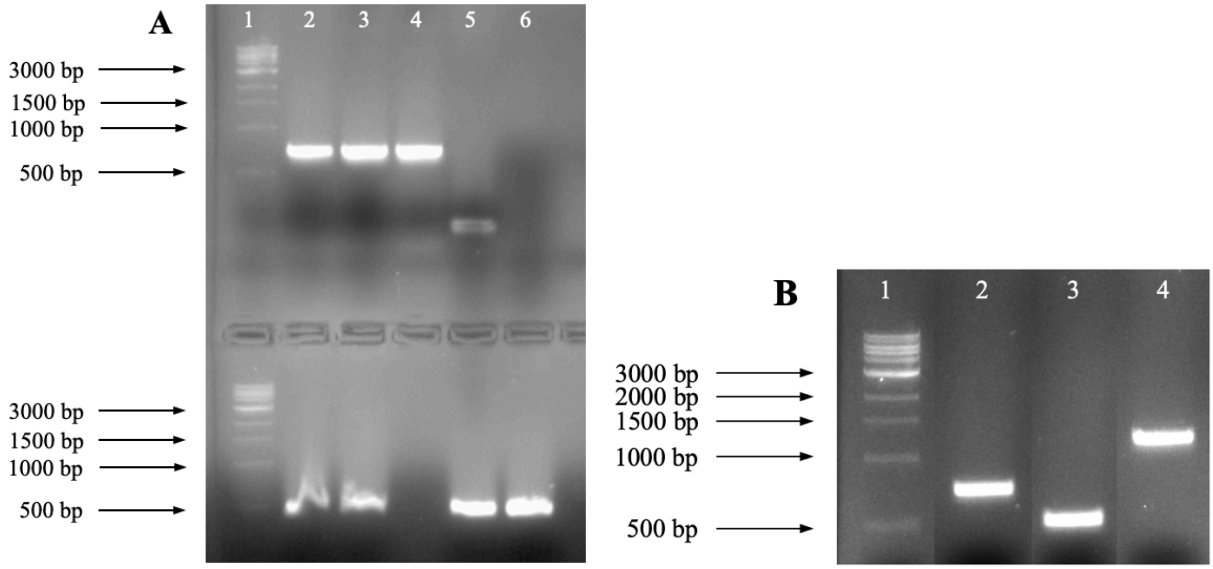


Figure 4.8. Agarose gel electrophoresis of *Ri*CEX_CatD and *Ri*CEX_CBM. Lane 1 in both gels shows the GeneRuler 1 kb DNA ladder. The upper lanes 2, 3 and 4 with *Ri*CEX_CatD (666 bp) and the lower lane 2, 3, 5 and 6 with *Ri*CEX_CBM (350 bp) in A shows, three and four respectively, successfully cloned plasmids. This was after cloning the genes into the pNIC-CH vector, transformation of the plasmids into One Shot® TOP10 cells and plasmid purification after production of the cells. B shows the *Ri*CEX_CatD (lane 2) and *Ri*CEX_CBM (lane 3) in comparison with the native *Ri*CEX (1116 bp) (lane 4).

Transformants were verified by sequencing and then transformed and grown in the Harbinger system. The proteins were purified with IMAC and appeared in two small peaks during elution (data not shown). They were expressed in smaller amounts compared with the other proteins. Figure 4.9 shows the SDS-PAGE analysis after IMAC purification. The first peaks from the fractions contained other contaminating proteins (similar to what is observed in figure 4.6), while the second peak contained the CatD domain (A) and CBM domain (B), with bands corresponding to ca 25 kDa (the theoretical M_w 25.68 kDa) and in between the 15 and 20 kDa bands on the protein ladder (theoretical M_w 17.54 kDa), respectively.

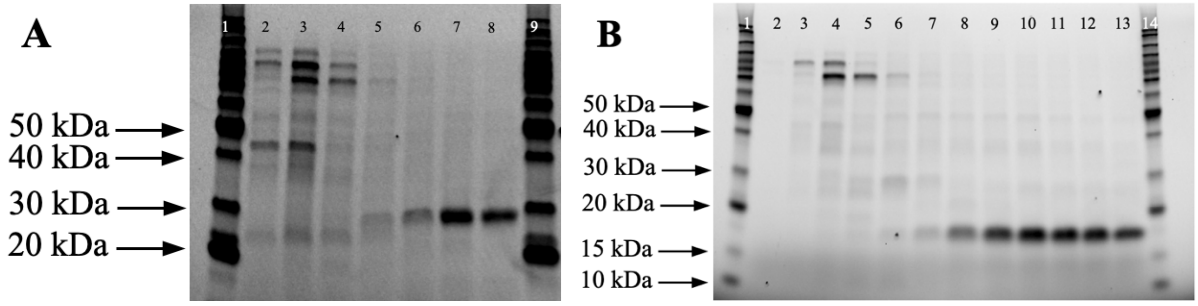


Figure 4.9. SDS-PAGE analysis of *Ri*CEX_CatD (A) and *Ri*CEX_CBM (B) after purification with IMAC. Lane 1 and 9 in A and lane 1 and 14 in B contains the BenchMark™ protein ladder. For both proteins, the chromatogram from the IMAC showed two small peaks with eluted proteins. In A, lane 2-5 contains fractions with contaminating proteins and lane 7 and 8 with the CatD domain (25.68 kDa) were collected. Gel B, with the CBM domain (17.54 kDa), lane 9-13 were collected.

*Ri*CEX Mutants #1-16

The genes for the *Ri*CEX mutants #1-16 were provided ready-made in pET28a(+) vectors by GenScript®. They were directly transformed into One Shot® BL21 Star™ (DE3) Chemically Competent *E. coli* cells for protein expression. All mutants contain a 6xHis-tag and the protein expression was induced with IPTG. They were purified using IMAC and figure 4.10 A shows an example of the purification, mutant #8, a typical result from the IMAC run and the other mutants looked very similar to this, except for mutant #3 (Figure 4.10 B).

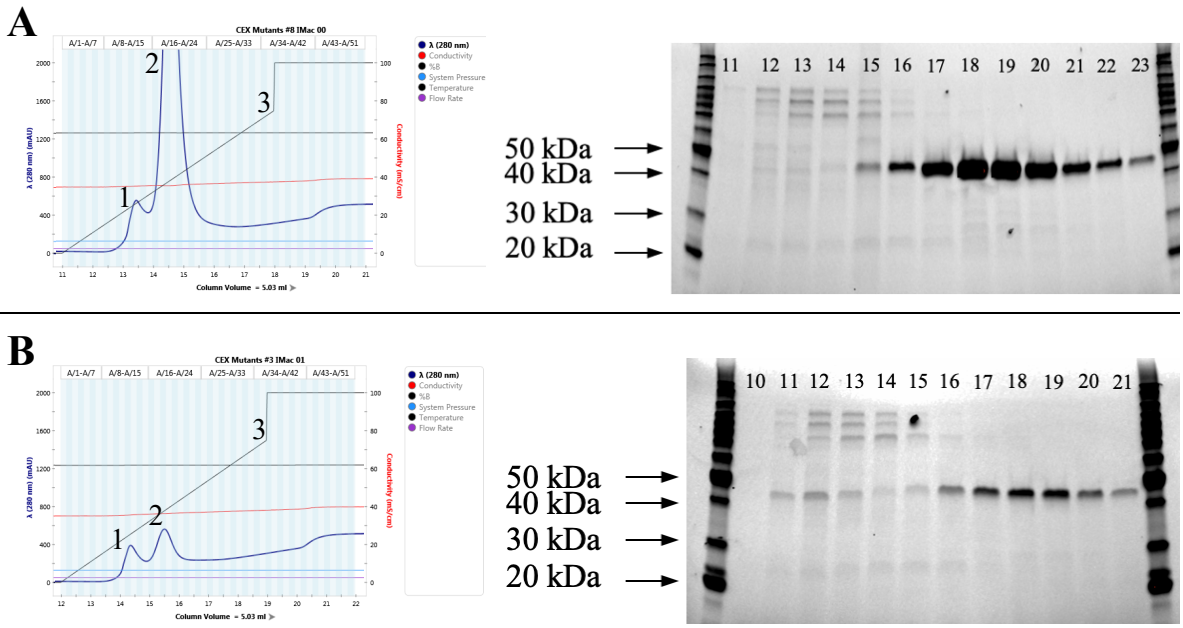


Figure 4.10. Chromatograms and SDS-PAGE analysis of *Ri*CEX mutants #8 (A) and #3 (B) after purification with IMAC. All the mutants have a molecular weight around 42.5 kDa. In both chromatograms, peak 1 contains contaminating proteins, peak 2 contains the protein of interest and the gradient line of elution buffer are shown as 3. *Ri*CEX mutants #1-16, except #3, gave similar purification chromatograms as shown for mutant #8. The *Ri*CEX mutant # 3 resulted in a very low protein yield compared with the other mutants. Fraction numbers are denoted on top of the chromatograms as intervals of 7 fractions (A1-A7 and so forth). Both gels show, with the lane number corresponding to fraction number from the chromatograms, contaminating proteins in lane 11 – 15, and the mutants in lane 17 – 23/21.

The first peak in both chromatograms contained contaminating proteins, as seen in the SDS-PAGE analysis, and the mutants were found in the second, largest peak (peak 2 in Fig 4.10). All *Ri*CEX mutants #1-16 have a molecular weight around 42.5 kDa. The numbered gel corresponds with the fraction number from IMAC. The mutants commonly eluted in fraction 16 – 23.

In general, all *F. prausnitzii* proteins and *Ri*CEX mutant proteins were produced in high amounts, except for *Ri*CEX mutant #3, *Ri*CEX_CatD and *Ri*CEX_CBM. Still, there was sufficient amounts for further characterization, also for these three low-expression proteins.

4.3 Characterization of *F. prausnitzii* enzymes

4.3.1 Activity test on mannose and non-mannose containing substrates with *Fp*CE2 and *Fp*CEXX

The activity of *Fp*CE2 and *Fp*CEXX were tested on several substrates and analysed with MALDI-ToF. The *F. prausnitzii* esterases were tested on AcGGM from Norway spruce (R5K) (figure 4.11), *Aloe vera* mannan, GH26 treated Konjac glucomannan and chemically acetylated Konjac glucomannan as mannose-based substrates. Both enzymes were active on these substrates (data not shown). The enzymes were also tested on acetylated cellulose, chitopentaose and birch xylan to test the substrate specificity for these enzymes. *Fp*CE2 and *Fp*CEXX did not show any activity on these substrates, neither when used separately or combined (data not shown).

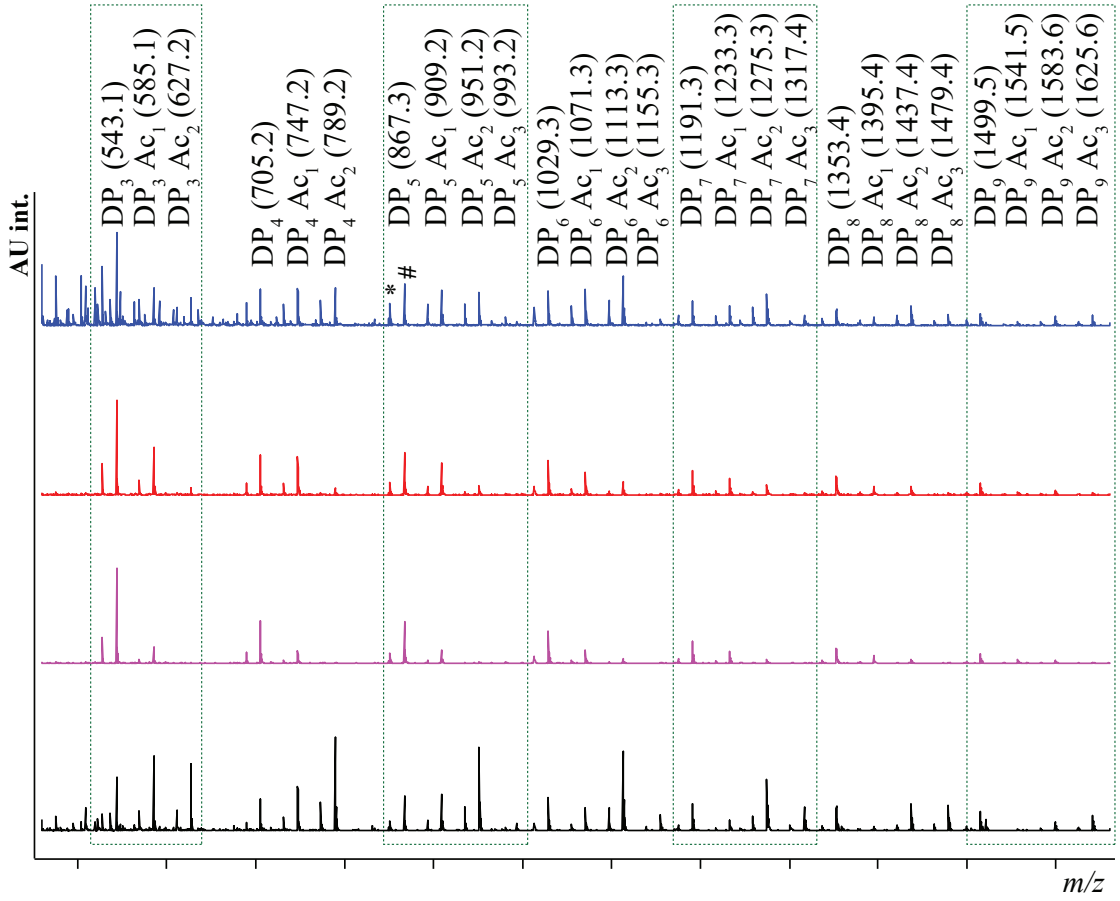


Figure 4.11. MALDI-ToF analysis of *Fp*CE2 and *Fp*CEXX on AcGGM. The negative control is shown in black, and the *Fp*CE2 (blue), *Fp*CEXX (red) and both enzymes combined (purple).

In figure 4.11, AcGGM (negative control without enzyme, black) was treated with *Fp*CE2 (in blue), *Fp*CEXX (in red) and both enzymes combined (in purple). After enzymatic treatment there is an apparent shift from peaks corresponding to acetylated to non-acetylated oligosaccharides. Several peak intensities of peaks corresponding to acetylated oligosaccharides (like e.g. *m/z* 627; DP_3Ac_2 and *m/z* 789; DP_4Ac_2) were reduced when the substrate was treated with each of the enzymes separately, and nearly completely deacetylated when used together. This pattern of partially and almost complete deacetylation, when used separately and combined, was seen for all the mannose-based substrates (data not shown). Further experiments on substrate specificity are presented in section 4.3.3.

4.3.2 Catalytic activity

Activity on pNP Acetate

The pH optimum of both *Fp*CE2 and *Fp*CEXX was 6.75, when tested on pNP-acetate in the pH range 6.0 – 8.0 (figure 4.12). *Fp*CE2 and *Fp*CEXX were used with a final concentration of 1 nM and 0.1 µM, respectively. Figure 4.12 show acetate released after 10 minutes reaction time at ambient temperature.

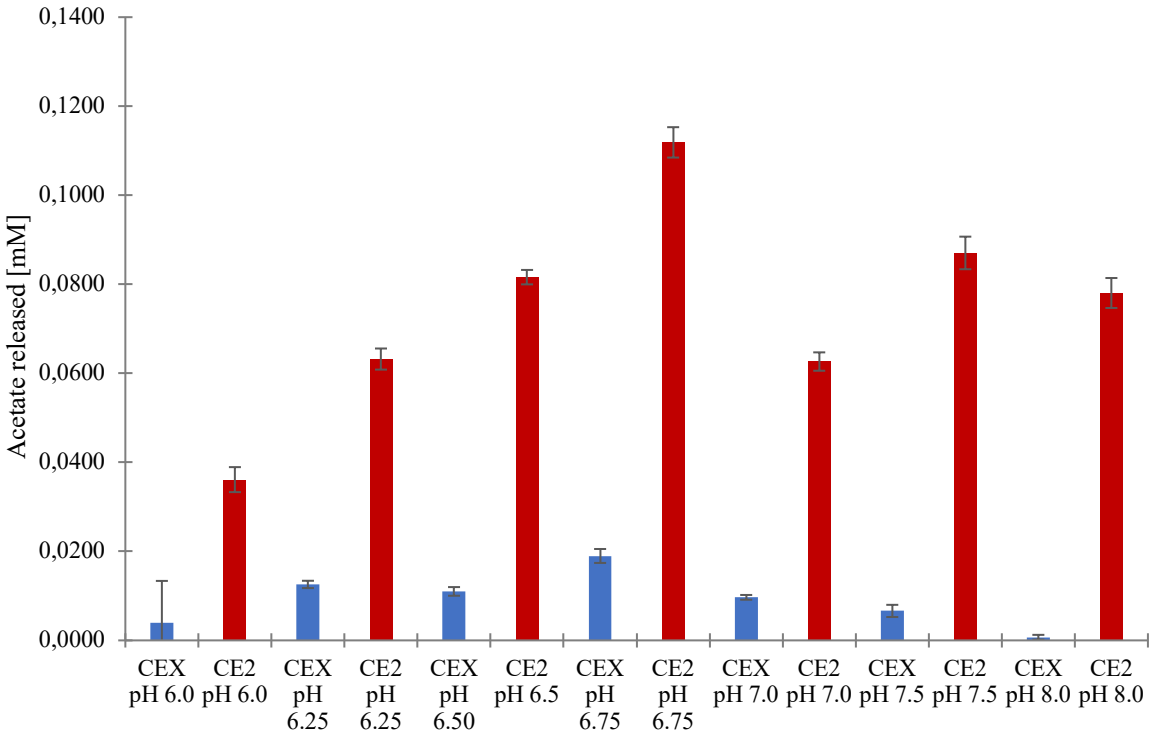


Figure 4.12. pH optimum for *Fp*CE2 and *Fp*CEXX on pNP-acetate. The esterases were tested on pNP-acetate in 50 mM sodium phosphate buffers with pH 6.0 – 8.0 at ambient temperature. Acetate release were measured after 10 minutes. The activity was also tested at 5.0 and 5.5 but showed a high variation over several tests and these results are not included.

Activity on acetylated galactoglucomannan from Norway spruce

A more in-depth characterization was conducted on AcGGM and was tested as a natural substrate. The esterases were tested at pH 6.75 and 35 °C and acetate release was measured at 5 minutes intervals. This temperature is close to physiological conditions in addition to avoid

acetyl migration. The enzymes were used in a 50 nM final concentration with 100 mg/mL GH26-treated AcGGM as substrate. The deacetylation rate was analysed using HPLC. The activity of *Fp*CE2 was linear up to 15 minutes before the reaction rate slowed down, while *Fp*CEXX was linear throughout 30 minutes and much slower than the *Fp*CE2 (figure 4.13). When combined, using 25 nM of each esterase, the deacetylation rate was close to half of the *Fp*CE2 and the double of *Fp*CEXX compared to treatments with each enzyme alone.

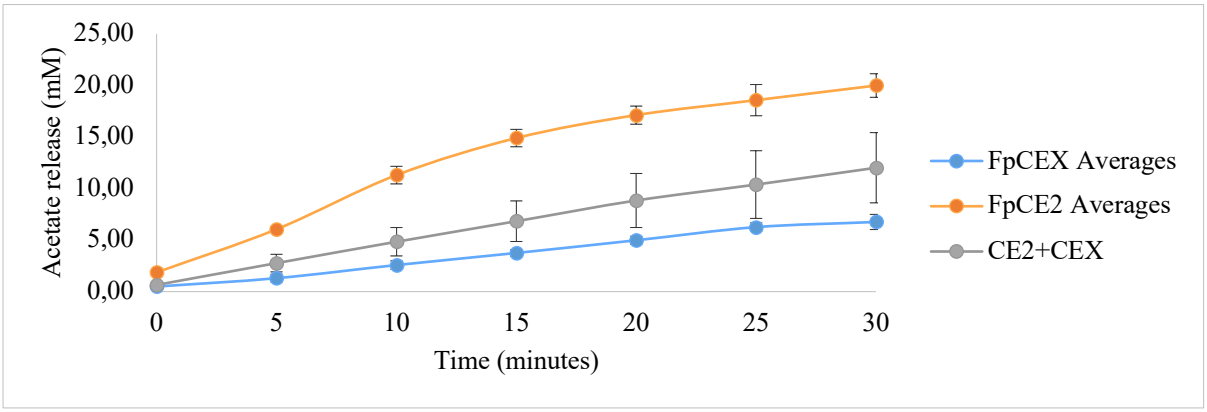


Figure 4.13. Acetate release from timecourse reactions with F. prausnitzii esterases and AcGGM GH26 treated. The reactions were loaded with 50 nM esterase, when used separate or combined, and 100 mg/mL substrate. Acetate released were measured in 5 minutes intervals.

The *Fp*CE2 released acetate approximately four times faster than *Fp*CEXX, when using the same enzyme dosages. The specific activity and turnover rate were calculated using the timepoints after 15 minutes of acetate release (Figure 4.13), when the esterases showed a linear deacetylation rate (Table 4.3).

Table 4.3. Deacetylation rate, specific activity and turnover rates of *Fp*CE2 and *Fp*CEXX on AcGGM from Norway spruce. This was calculated from the timepoints after 15 minutes.

	***Fp*CEXX**	***Fp*CE2**	**CE2+CEXX**
Deacetylation rate [nmol/s]	4166	16550	7588
k_{cat} [s^{-1}]	83	331	152
Specific activity [nmole acetate/s/µg enzyme]	2	8	

Protein thermal shift assay

The apparent melting temperature of *Fp*CE2 and *Fp*CEXX were tested in a protein thermal shift (PTS) assay (Figure 4.14). The lowest melting temperature was with pH 5.0 and was 55 °C for *Fp*CE2 and 58 °C for *Fp*CEXX, indicating they are most unstable at this pH. Both showed highest melting temperature in pH between 6.0 and 6.5.

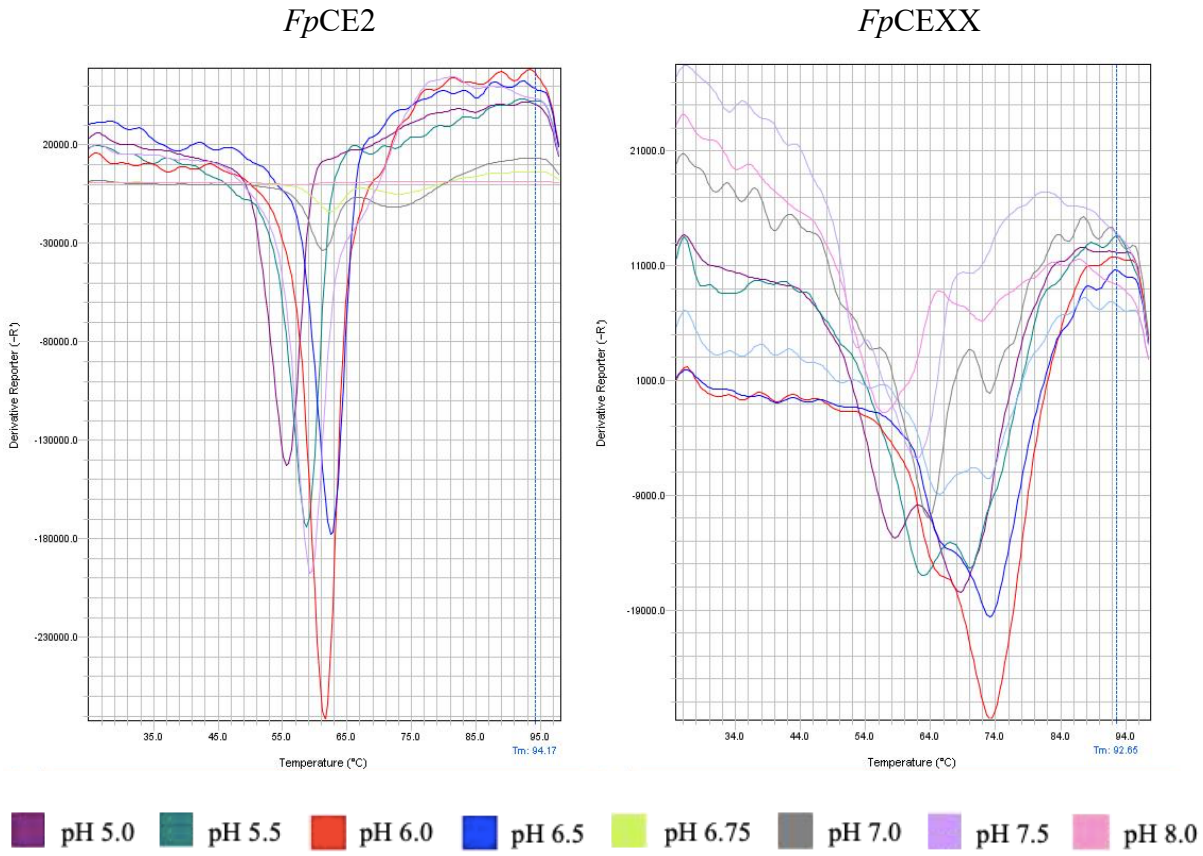


Figure 4.14. Melting curves of *Fp*CE2 and *Fp*CEX in a protein thermal shift assay. The melting temperature for the esterases were tested in 50 mM sodium phosphate buffers with pH 5.0 – 8.0.

4.3.3 Transesterification

Transacetylation and transpropylation

Transacetylation is used as a means of identifying specificities of acetylesterases (Biely, 2012). To test the possibility of the *Fp*CE2 and *Fp*CEXX to transacetylate substrates, reactions with

mannotriose and mannotetraose together with vinyl acetate were performed. Both esterases were able to transacetylate the two substrates. Figure 4.15 shows *Fp*CE2 (top spectrum) and *Fp*CEXX (bottom spectrum) transacetylation reactions, with non-acetylated mannotetraose (*m/z* 689; DP_4) and acetylated mannotetraose (*m/z* 731; DP_4Ac). The mannotriose reaction gave similar results (data not shown).

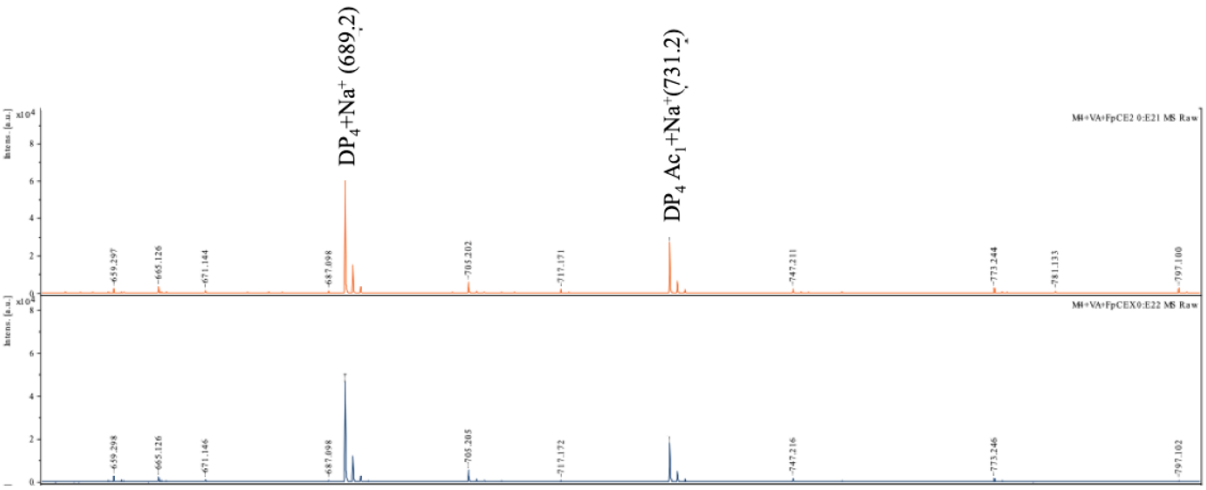


Figure 4.15. MALDI-ToF MS analysis of transacetylated mannotetraose with *Fp*CE2 (top spectrum) and *Fp*CEXX (bottom spectrum). Mannotetraose are seen as non-acetylated (*m/z* 689; DP_4) and mono-acetylated (*m/z* 731; DP_4Ac_1).

The same experiment was carried out but with vinyl propionate replacing vinyl acetate. Both esterases were able to transpropylate the substrates, but to a lesser degree than for vinyl acetate. Figure 4.16 show transpropylation of mannotriose (*m/z* 527; DP_3 non-propylated) with *Fp*CE2 (top spectrum) and *Fp*CEXX (bottom spectrum) to propylated mannotriose (*m/z* 583; DP_3Prop_1).

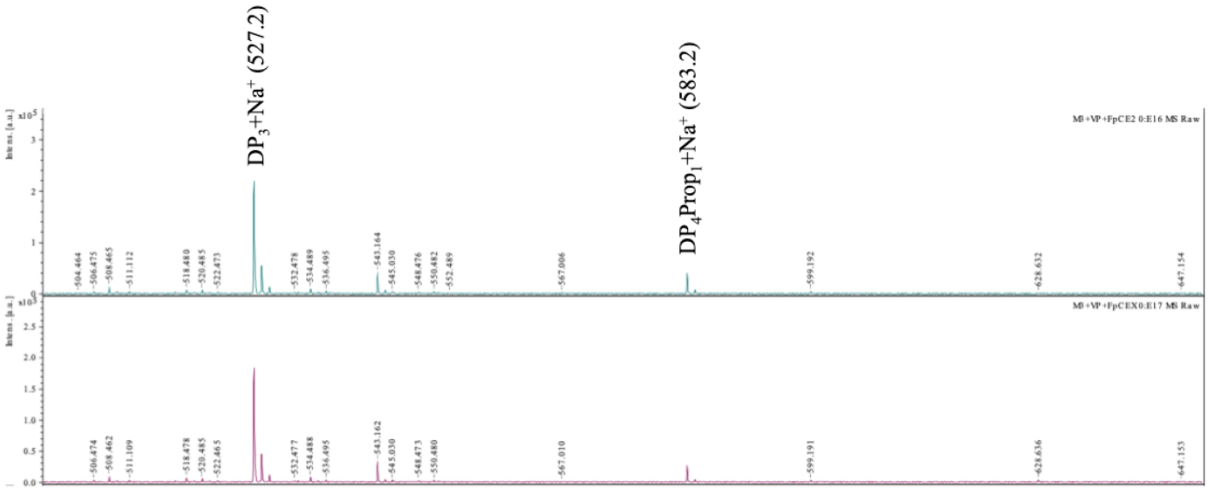


Figure 4.16. MALDI-ToF MS analysis of transpropylated mannotriose with *Fp*CE2 and *Fp*CEXX. The esterases, *Fp*CE2 (top spectrum) and *Fp*CEXX (bottom spectrum), were able, to some extent, to transpropylate mannotriose. In both diagrams, mannotriose are seen as non-propylated (*m/z* 527; DP_3) and mono-propylated (*m/z* 583; DP_3Prop_1).

Preferred substrate

An additional test to determine the preferred substrate for the esterases was conducted on a mix of mannotriose (man_3), cellotetraose (glc_4) and xylopentaose (xyl_5). These substrates were tested with vinyl acetate, vinyl propionate and vinyl butyrate on each *F. prausnitzii* esterase in transesterification reactions. Mannotriose (*m/z* 527; DP_3) and vinyl acetate were the preferred substrate for *Fp*CE2 (middle spectrum) and *Fp*CEXX (bottom spectrum), which transacetylated some of the mannotriose (*m/z* 569; DP_3Ac_1) (Figure 4.17). Cellotetraose and xylopentaose are seen as peaks at 689 m/z and 701 m/z, respectively; the top spectrum (blue trace) shows the control sample without enzyme. This was done by Dr. Leszek Michalak.

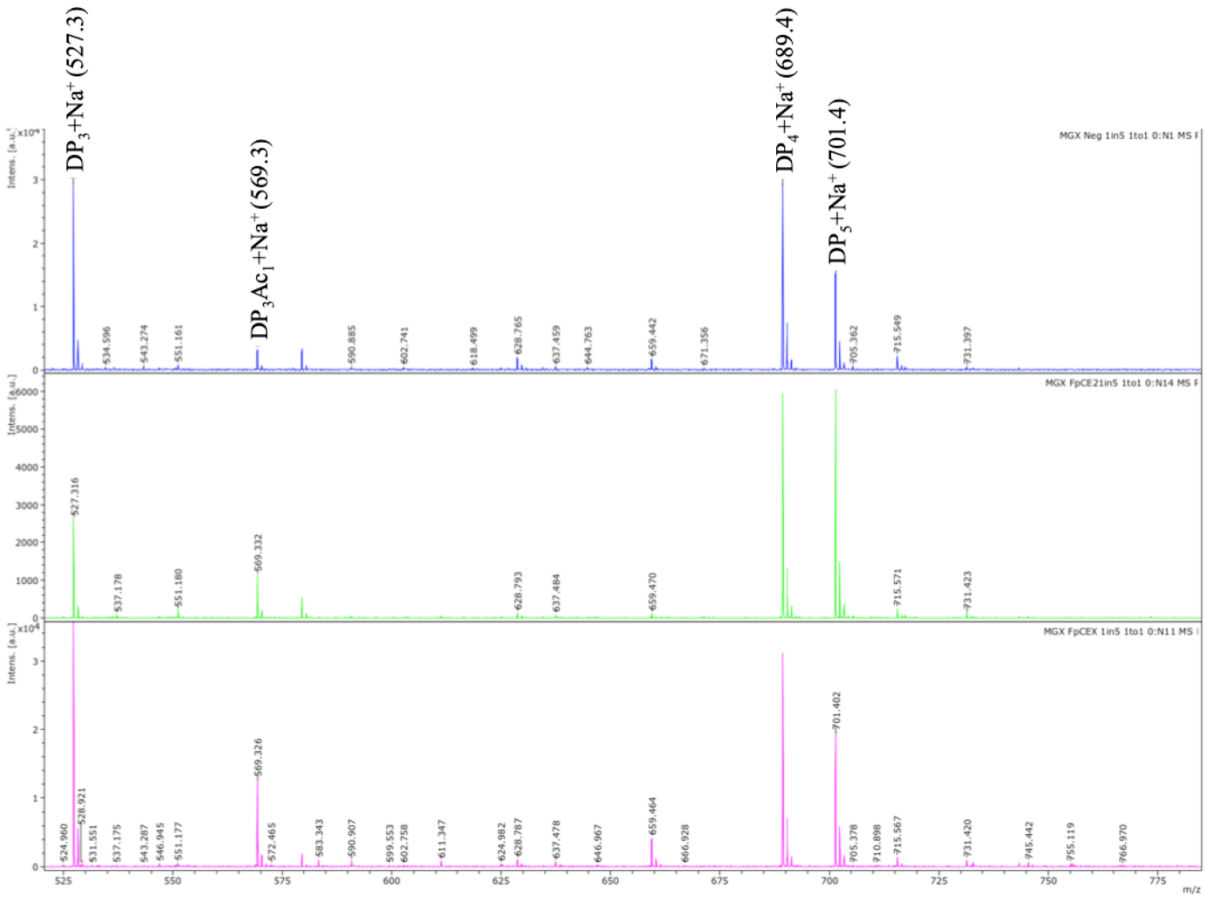


Figure 4.17. MALDI-ToF MS analysis of reactions for identification of preferred oligosaccharides for *Fp*CE2 (green spectrum) and *Fp*CEXX (magenta spectrum) in a transesterification reaction, negative control (blue spectrum). The esterases were tested on a mix with mannotriose (*m/z* 527; DP_3), cellotetraose (*m/z* 689; DP_4) and xylopentaose (*m/z* 701; DP_5) and with vinyl acetate, vinyl propionate and vinyl butyrate as ester donors. Both esterases preferred mannotriose and vinyl acetate, and transacetylated this substrate (*m/z* 569; DP_3Ac_1).

Acetyl position specificity of *Fp*CE2 and *Fp*CEXX

To investigate on which position *Fp*CE2 and *Fp*CEXX deacetylate the substrate, their activity was tested on a substrate with acetyl groups in either the 2-*O* or the 3-*O*, 4-*O* and/or 6-*O* positions. Two esterases, *Ri*CE2 and *Ri*CEX which have different specificity, were used to generate acetylated mannohexaose with transacetylation reactions. The *Ri*CE2 was used to attach acetyl groups on either the 3-*O*, 4-*O* or 6-*O* position, while *Ri*CEX on the 2-*O* position. Mannohexaose transacetylated with *Ri*CE2 and *Ri*CEX are shown in the two top and two bottom rows, respectively, in figure 4.18. Spectrum 1 and 3 were treated with *Fp*CE2 and spectrum 2 and 4 with *Fp*CEXX. The results in spectrum 1 show that *Fp*CE2 is able to deacetylate the mannohexaose, while *Fp*CEXX didn't show activity (spectrum 2). The *Fp*CEXX deacetylated the mannohexaose (spectrum 4), while *Ri*CE2 did not (spectrum 3). This

implies that the *F. prausnitzii* esterases have the same specific activity as their corresponding esterases in *R. intestinalis*.

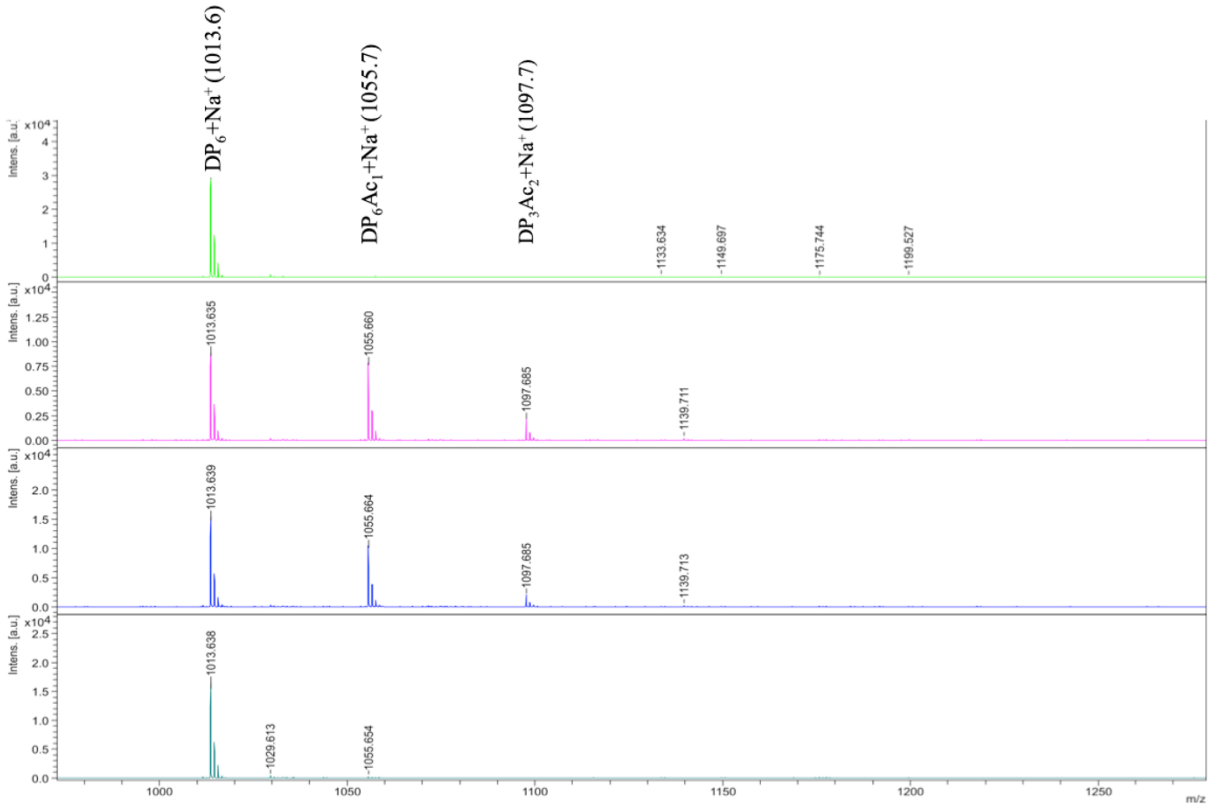


Figure 4.18. MALDI-ToF MS analysis of acetyl position specificity of *Fp*CE2 and *Fp*CEXX. The *F. prausnitzii* esterases were tested on mannohexaose with acetylgroups in either 2-*O* position (spectrum 3 and 4) or in 3-*O*, 4-*O* and/or 6-*O* position (spectrum 1 and 2). *Fp*CE2 was able to remove acetylations in the 3-*O*, 4-*O* or 6-*O* position (spectrum 1) and not the 2-*O*- acetylations (spectrum 3). *Fp*CEXX removed 2-*O*- acetylations (spectrum 4), and not the 3-*O*-, 4-*O*- or 6-*O*- acetylations (spectrum 2).

4.3.4 Crystallization

Several crystals, obtained with *Fp*CEXX, were examined at the European Synchrotron Radiation Facility in Grenoble, France. Even though several of the crystals diffracted, none of the diffraction data sets were of sufficient quality to solve the protein structure. Crystal optimization didn't result in any better crystals than from the crystallization sets.

4.4 Characterization of *R. intestinalis* mutants

4.4.1 *Ri*CEX_CatD and *Ri*CEX_CBM

*Ri*CEX_CatD did scarcely show any activity and *Ri*CEX_CBM did not have any activity when tested on GH26-treated AcGGM (R5K), R5K-2AB, GH26-treated konjac glucomannan and pNP-acetate (data not shown).

4.4.2 *Ri*CEX mutants # 1-16

All *Ri*CEX mutants (#1-16) were active when tested on GH26-treated R5K and R5K-2AB, analysed with MALDI-ToF MS. Mutants #1 and #10 are seen in figure 4.19, as an example of the activity of all mutants.

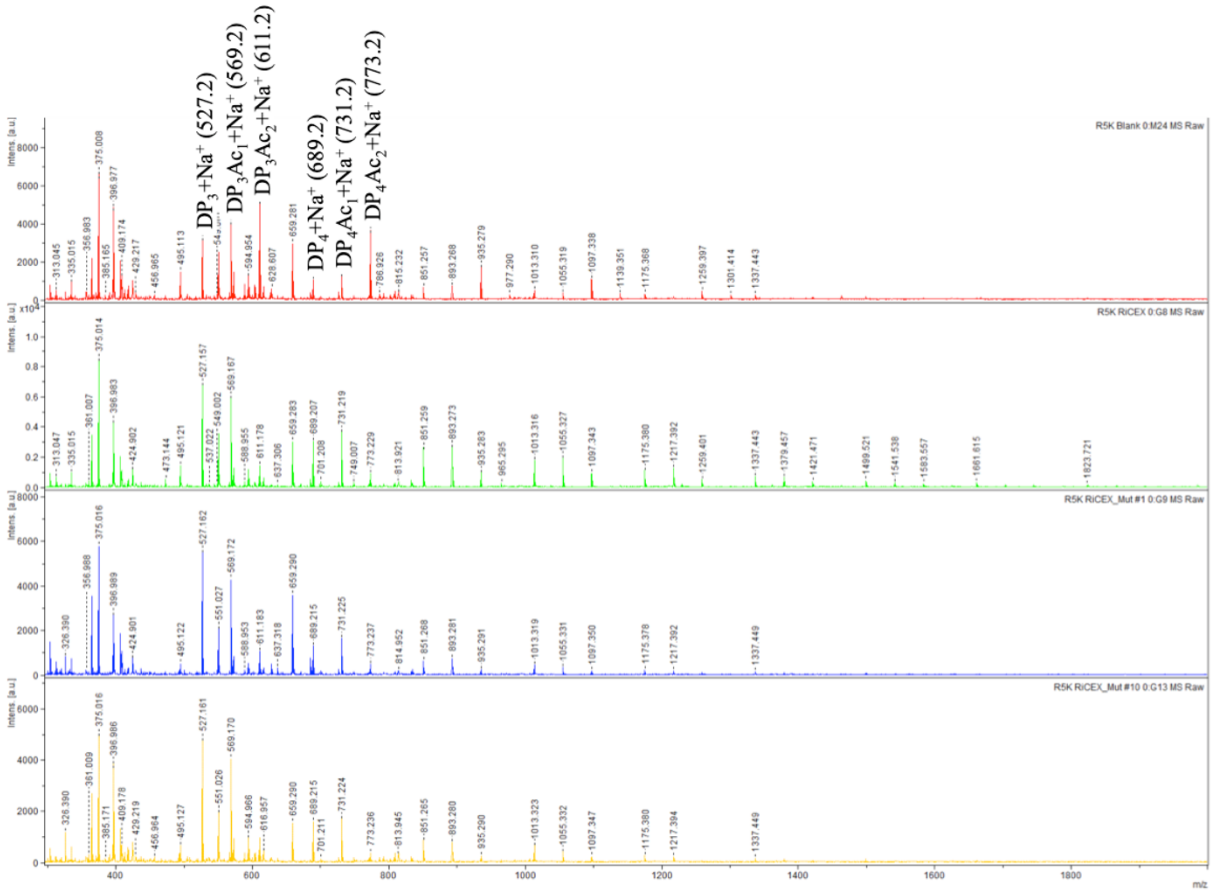


Figure 4.19. MALDI-ToF MS analysis of *Ri*CEX mutants #1 and #10 activity on R5K. Control sample, native *Ri*CEX, mutant #1 and mutant #10 are shown in spectrum top to bottom, respectively.

All mutants were active on R5K with similar spectra to the native *Ri*CEX, and mutant #1 and #10 (figure 4.19). Peaks with double acetylations, *m/z* 611.2 (DP_3Ac_2) and *m/z* 773.2 (DP_4Ac_2), are close to complete deacetylated, and mono-acetylated, *m/z* 569.2 (DP_3Ac_1) or *m/z* 731.2 (DP_4Ac_1), are partly deacetylated. Non-acetylated oligosaccharides are seen as peak *m/z* 527.2 (DP_3) and *m/z* 689.2 (DP_4). The same experiment was also conducted with labelled R5K-2AB as substrate (figure 4.20).

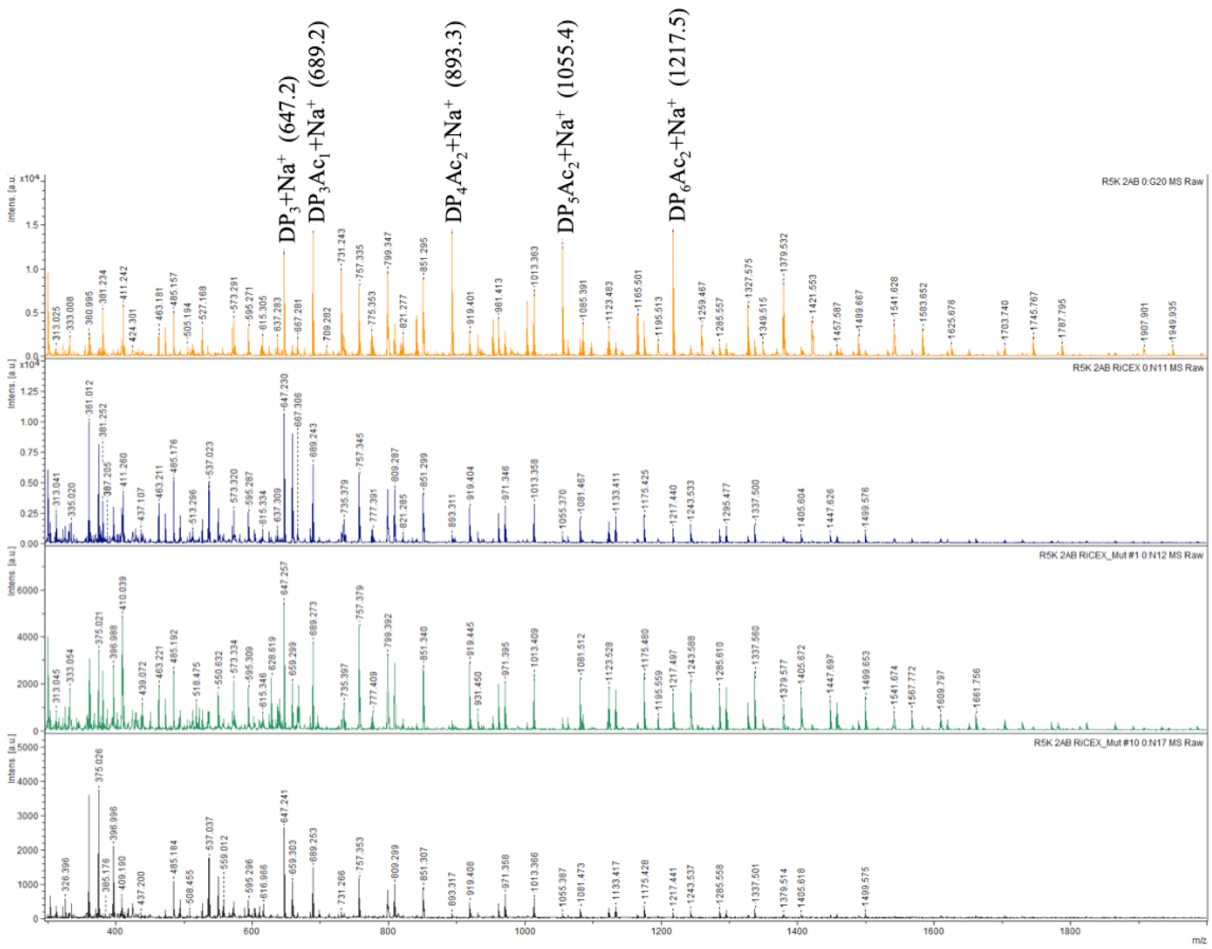


Figure 4.20. MALDI-ToF MS analysis of *Ri*CEX mutants #1 and #10 activity on 2AB-labeled AcGGM (R5K-2AB); a control containing R5K-2AB without enzyme (top spectrum), the native *Ri*CEX (the second spectrum) and mutant #1 (spectrum three) and #10 (spectrum four).

All mutants were active on the labelled R5K-2AB resulting in spectra similar to the *Ri*CEX mutant #1 and #10 (figure 4.20). Peaks corresponding to 2-AB linked double acetylated oligosaccharides, *m/z* 893.3 (DP_4Ac_2), *m/z* 1055.4 (DP_5Ac_2) and *m/z* 1217.5 (DP_6Ac_2), are almost completely removed, while peaks corresponding to mono-acetylated (m/z 689.2, 851.3, 1013.4; DP_3Ac_1, DP_4Ac_1, DP_5Ac_1) and non-acetylated oligosaccharides *m/z* 647.2 (DP_3), *m/z*

809.3 (DP_4) and *m/z* 971.4 (DP_5). The 2-AB aglycone when coupled to the oligosaccharide has an *m/z* 120.

A selection of *Ri*CEX galactose pocket mutants, #5, 6, 7, 8, 9 and 12, were tested on native R5K and R5K-2AB. The mutants were selected from whom we thought would have mutations with most effect. They were compared with the native *Ri*CEX and analysed with HILIC. With these substrates, we were not able to identify any observable difference between the native *Ri*CEX and the mutants (data not shown). Transesterification reactions were set up with all mutants and the native *Ri*CEX. Mutants were tested on mannotriose, mannohexaose, galactosylmannotriose and digalactosylmannopentaose, with vinyl acetate, vinyl propionate and vinyl butyrate as ester donors. No detectable difference was observed between any of the mutants tested (data not shown). None of the active site pocket mutants (#13-16) were able to transbutyrate, and they did not show any greater effect on transpropylation than the native *Ri*CEX. Analyses were done by Dr. Shaun Leivers and Dr. Leszek Michalak.

5 Discussion

The objective of this study was to gain more insight into some of the proteins in a gene cluster of mannan degrading enzymes in *F. prausnitzii*. Two esterases in this cluster were suspected to play a key role in the degradation of mannans since they were similar counterparts to esterases found in *R. intestinalis* (La Rosa et al., 2019b) and have been the main focus in this study. *F. prausnitzii*, has in a number of studies been linked to benefits for human and/or animal health and is thus a potential target organism for prebiotic substances. When considering prebiotic substances, a key parameter is that they provide a highly selective growth of the target species (Gibson et al., 2017). Specialised substrates are necessary to provide. Detailed insight in how they degrade complex carbohydrates are thus of high interest and highly relevant when seeking to understand aspects related to gut health.

In the interest of making more specific substrates, mutants of an esterase from *R. intestinalis* was generated. Mutations were designed with the aim at blocking substrates with 6-*O*-galactosylations or the possibility to the removal or addition of larger molecules than acetylations in the 2-*O* position on mannose residues, like propionate and butyrate. This was highly experimental and is still an ongoing project where several group members are involved.

5.1 Cloning, expression and purification

All proteins were successfully cloned into the pNIC-CH vector. The *Ri*CEX mutants and the *F. prausnitzii* epimerase were provided ready made in the pET-28a(+) vector and no cloning was thus necessary for these proteins. All proteins were produced in *E. coli* and have a 6xHis-tag that made expression and purification easy and successful, with minor obstacles and problems. *E. coli* is a well-known host for expression of proteins, but the success is not guaranteed as the proteins of interest comes from other phyla. Factors like pH, cofactors, folding mechanisms and toxicity for the host cell may affect the ability to express the target proteins (Rosano & Ceccarelli, 2014).

*Fp*CE2 and *Fp*CEXX were purified using IMAC and SEC chromatography. *Fp*CEXX was expressed and purified in large quantities, showing no problems with the attached 6xHis-tag, nor the purification methods, and was stable during long time storage in 4 °C. *Fp*CE2 behaved

well after IMAC but eluted in two peaks during SEC. SDS-PAGE analysis showed that both peaks contained proteins of correct size. The first peak may be a dimer of the protein, giving the larger size and therefore eluted first. We considered two approaches to identify if the protein dimerised in solution, first approach was to include dynamic light scattering (DLS) in the protein purification (SEC) step is a precise way, however this equipment was not set up in the lab. Secondly, it was considered to run a native gel to identify whether or not the protein dimerized (the first peak) to some extent but was not conducted since they precipitated during buffer exchange. The *Fp*CE2 was first stored in 20 mM sodium phosphate pH 5.9 buffer but samples were precipitating, and the buffer was changed to 20 mM sodium phosphate pH 6.9 to get further away from the pI, this is a common way to prevent precipitation issues. No further problems were observed after changing buffer. It is worth mentioning that the *Ri*CE2 is relatively challenging to keep in solution and needed storage at -80 °C in pH 5.9 buffer (Michalak et al., 2019). It is tempting to speculate that the CE2 fold in general is a little prone to instability in *in-vitro* conditions.

The truncated version of *Ri*CEX, *Ri*CEX_CatD and *Ri*CEX_CBM, were only purified with IMAC for initial testing of activity. Compared with the other proteins expressed and purified these proteins showed a much lower expression. This indicate that there may be problems with the folding of the proteins or with the purification. The catalytic domain (*Ri*CEX_CatD) where very weakly (negligible) active, which implies that the CBM is essential for the catalytic activity. Even though the protein is not precipitating, the abolished activity we observed could, however, simply be that problems with folding. Analysis confirming that the *Ri*CEX_CatD is not linearized should be tested by e.g. a combination of melting curves and further dynamic light scattering to ensure the protein is properly folded (Jachimska et al., 2008).

*Ri*CEX mutants #1-16 were all, except for mutant #3, expressed in high quantities and purified successfully with IMAC. This indicates that most of the mutations in this experiment have a small or no effect on the folding and stability of the proteins. Mutant #3, which has an Ala112 instead of Asp112 in the left galactose pocket, was poorly expressed and may be affected of this specific mutation. This mutation replaced the native negatively charged amino acid with a non-polar amino acid, and by this changing some of the properties of the protein. On the other hand, the other two mutations performed in this site, did not affect the expression of proteins, but they are polar amino acids.

5.2 Characterization

F. prausnitzii esterases

The *Fp*CE2 was classified as a carbohydrate esterase family 2 by analysis with InterPro. It contains homology to the N-terminal domain recognised to the CE2 family, and the SGNH esterase domain, with the catalytic residues serine and histidine that are common for the CE2 family (Till et al., 2013). *Fp*CEXX was not recognised to any of the CE families by InterPro, but an SGNH hydrolyse-type esterase domain was identified. Furthermore, since it was part of the same mannan degrading cluster in *F. prausnitzii* and had high sequence identity to the *Ri*CEX (La Rosa et al., 2019b) it was hypothesised that this was the *Ri*CEX counterpart.

The initial testing of activity with MALDI-ToF MS on different substrates indicated that they were only active on mannan-based substrates. This assumption was further strengthened during the transesterification reactions with cellulose, xylan and mannan (mannotriose) based substrates. Mannotriose was clearly the preferred substrate in this experiment for transesterification with vinyl esters. This implies that the *Fp*CE2, which is in the CE2 family where most members are classified as xylan esterases, is in fact a mannan specific esterase like some other few characterised members of this family (Biely, 2012).

The activity test also revealed that they did not deacetylate all double acetylated substrates when used separately but showed nearly complete deacetylation when used combined. This is also the case for the *R. intestinalis* esterases, *Ri*CE2 and *Ri*CEX, which was examined in an NMR study by (Michalak et al., 2019). This showed that the *Ri*CE2 has a broad specificity, removing 3-*O*-, 4-*O*- and 6-*O*-acetylations, while the *Ri*CEX is specific for 2-*O*-acetylations. For complete removal of all acetylations, the 2-*O*-acetylations needed to be removed to access the other acetylations for *Ri*CE2. Testing the *Fp*CE2 and *Fp*CEXX with transacetylated substrates made with either of these position specific *R. intestinalis* esterases revealed that these two *F. prausnitzii* esterases have the same specificity. This relate the *Fp*CEXX as a 2-*O*-specific mannose esterase and the *Fp*CE2 to be a broad specific esterase in the CE2 family. This thesis has not revealed if the *Fp*CEXX is needed to act first, with removal of the 2-*O*-acetyl position, to facilitate for the *Fp*CE2 to be able to remove all other acetylations in the 3-*O*-, 4-*O*- and 6-*O*-positions, as the *Ri*CE2 and *Ri*CEX (Michalak et al., 2019). An NMR study could examine if this also applies for the *F. prausnitzii* esterases. An experiment, which has not been conducted

but should be considered, is to run a reaction on a substrate with *Fp*CE2 first and then *Fp*CEXX on the same substrate and do the same with opposite order of the enzymes. If there is a complete deacetylation rate when the *Fp*CEXX is used first, and not the *Fp*CE2, could reveal the same pattern as for the *R. instestinalis* esterases.

The pH optimum assay with pNP-acetate showed a pH 6.75 optimum for both *F. prausnitzii* esterases. At pH's lower than 6.0, the results were very inconsistent during several reactions and with different enzyme concentrations, and the results differed highly and were sometimes negative. The drop in activity for *Fp*CE2 at pH 7.0 was not observed during all testing and may reflect experimental error. Protein thermal shift assays showed that *Fp*CE2 and *Fp*CEXX were stable up to 55 °C and 58 °C, respectively, with pH 5.0 buffer. They had the highest melting temperature with pH approximately in the range 6.0 - 6.5, indicating that they are more stable at this pH, and have a lower stability when reducing the pH. The artificial substrate pNP-acetate clearly is the source of several challenges related to the esterase characterization. This probably relates to the inherent instability of pNP-acetate at elevated pH and temperatures. This was clearly a challenge in the temperature optimum assays were also the results inconsistent during multiple attempts. During testing, *Fp*CEXX needed a 100 times higher enzyme concentration than the *Fp*CE2 to see roughly the same acetate release. This may indicate that the pNP-acetate do not fit well in the active site pocket of *Fp*CEXX, where natural substrates have their acetyl group in the axial position, and pNP-acetate has a more linear structure. This implies that this substrate is not optimal for the testing of optimum conditions for these esterases. Considering this, and the fact that the natural substrates are prone to acetyl migration during high temperatures and elevated pH (Roslund et al., 2008), 35 °C and pH 6.75 was chosen for further experiment with deacetylation rate on AcGGM from Norway spruce. This temperature is relevant in terms of the activity of the esterases in close to physiological conditions, and to keep the acetyl migration to a minimum. Preventing migration is particularly important when analysing enzyme specificity, since migration from 2-*O* to 3-*O* acetylation obviously will alter the composition of the substrate and thus easily lead to misinterpretations. The deacetylation rate on AcGGM were, when equal concentrations were used, close to four times higher for *Fp*CE2 than for *Fp*CEXX. Used in combination, they had a deacetylation rate which was close to half of the *Fp*CE2 and double of the *Fp*CEXX contrary when used separately. This indicates that they are not acting synergistically but may rather be competing about substrate, as the deacetylation rate of *Fp*CE2 is slower.

The obtained crystals with *Fp*CEXX did not result in diffraction data with sufficient quality to solve the protein structure. Further optimization may result in better quality crystals, using several steps with different pH's, salts and precipitants.

***Ri*CEX mutants**

The catalytic domain (*Ri*CEX_CatD) and the CBM domain (*Ri*CEX_CBM) of the truncated version of *Ri*CEX did hardly show any detectable activity on neither of the substrates tested. If correctly folded, this indicate that they are not functional on their own, and thus need the full enzyme to function. More characterisation on binding affinity, e.g. with isothermal titration calorimetry, can reveal if the domains are functional in terms of substrate binding (Leavitt & Freire, 2001).

Mutants, #1 - 12, were designed aiming to affect and block the galactose pockets on each side of the active site in *Ri*CEX. Mutations leading to, separate or in combinations, a change in size, charge and polarity were chosen. Mutants #13-16 were active site mutants, using amino acids with smaller sidechains to make space for larger substrates. All mutations were aimed at keeping other possible interferences with the protein to a minimum.

We were not able to identify any differences with any of the mutants compared to the wild type. Clearly this is not a straight forward task, where we have outlined several approaches that have not successfully pointed to any effects of the mutants. The way forward would be to test hydrolysates of guar gum (oligosaccharides of as sketched in figure 1.1) which are hypothesized to be a good candidate and will be subject to further research. LC-MS analysis is still a potential approach, but with the complex mixture of oligosaccharides generated from the R5K this require a considerable amount of work. It can be considered that the mutations not were sufficient to block the galactose pockets to hinder substrates from binding.

5.3 Concluding remarks and future perspectives

This study on *Fp*CE2 and *Fp*CEXX have demonstrated that they are mannan specific esterases, working in a complementary manner. The *Fp*CE2 has a broad specificity, deacetylating 3-*O*-, 4-*O*- and 6-*O*-acetylations and belongs to the CE2 family. *Fp*CEXX is deacetylating 2-*O*-

acetylations specifically and is not linked to any of the CE families but it has an SGNH hydrolyse-type esterase domain. Initial work on the other enzymes in this gene cluster have started, by several of the group members. This will gradually increase our understanding of the *F. prausnitzii* ability to break down AcGGM. Hopefully, this can reveal more about the structures of complex mannans *F. prausnitzii* can utilise and grow on, which can potentially lead to new prebiotics.

16 *Ri*CEX mutants were produced, and all initial activity testing revealed that they were active. During the work with the mutants, the primary focus in this thesis has been producing the various clones. As we could not observe any difference in activity distinct from the native *Ri*CEX, in depth characterisation of the galactose pocket mutants will be subject to further research.

6 References

Adesioye, F. A., Makhalanyane, T. P., Biely, P. & Cowan, D. A. (2016). Phylogeny, classification and metagenomic bioprospecting of microbial acetyl xylan esterases. *Enzyme and microbial technology*, 93: 79-91.

Al-Lahham, S. A. H., Peppelenbosch, M. P., Roelofsen, H., Vonk, R. J. & Venema, K. (2010). Biological effects of propionic acid in humans; metabolism, potential applications and underlying mechanisms. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*, 1801 (11): 1175-1183. doi: 10.1016/j.bbalip.2010.07.007.

Biely, P. (2012). Microbial carbohydrate esterases deacetylating plant polysaccharides. *Biotechnology advances*, 30 (6): 1575-1588.

Biely, P., Westereng, B., Puchart, V., De Maayer, P. & Cowan, D. A. (2014). Recent Progress in Understanding Mode of Action of Acetylxylan Esterases. *Journal of Applied Glycoscience*: jag. JAG-2013_018.

Boraston, A. B., Bolam, D. N., Gilbert, H. J. & Davies, G. J. (2004). Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochemical journal*, 382 (3): 769-781.

Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V. & Henrissat, B. (2008). The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. *Nucleic acids research*, 37 (suppl_1): D233-D238.

Cockburn, D. W. & Koropatkin, N. M. (2016). Polysaccharide degradation by the intestinal microbiota and its influence on human health and disease. *Journal of molecular biology*, 428 (16): 3230-3252.

Cummings, J. & Stephen, A. (2007). Carbohydrate terminology and classification. *European journal of clinical nutrition*, 61 (S1): S5.

Dave, M., Higgins, P. D., Middha, S. & Rioux, K. P. (2012). The human gut microbiome: current knowledge, challenges, and future directions. *Translational Research*, 160 (4): 246-257.

Duncan, S. H., Hold, G. L., Barcenilla, A., Stewart, C. S. & Flint, H. J. (2002). Roseburia intestinalis sp. nov., a novel saccharolytic, butyrate-producing bacterium from human faeces. *International journal of systematic and evolutionary microbiology*, 52 (5): 1615-1620.

El Kaoutari, A., Armougom, F., Gordon, J. I., Raoult, D. & Henrissat, B. (2013). The abundance and variety of carbohydrate-active enzymes in the human gut microbiota. *Nature Reviews Microbiology*, 11 (7): 497.

Flint, H. J., Bayer, E. A., Rincon, M. T., Lamed, R. & White, B. A. (2008). Polysaccharide utilization by gut bacteria: potential for new insights from genomic analysis. *Nature Reviews Microbiology*, 6 (2): 121.

Frost, G., Sleeth, M. L., Sahuri-Arisoylu, M., Lizarbe, B., Cerdan, S., Brody, L., Anastasovska, J., Ghourab, S., Hankir, M., Zhang, S., et al. (2014). The short-chain fatty acid acetate reduces appetite via a central homeostatic mechanism. *Nature Communications*, 5: 3611. doi: 10.1038/ncomms4611

https://www.nature.com/articles/ncomms4611#supplementary-information.

Gibson, G. R. & Roberfroid, M. B. (1995). Dietary Modulation of the Human Colonic Microbiota: Introducing the Concept of Prebiotics. *The Journal of Nutrition*, 125 (6): 1401-1412. doi: 10.1093/jn/125.6.1401.

Gibson, G. R., Hutkins, R., Sanders, M. E., Prescott, S. L., Reimer, R. A., Salminen, S. J., Scott, K., Stanton, C., Swanson, K. S. & Cani, P. D. (2017). Expert consensus document: The International Scientific Association for Probiotics and Prebiotics (ISAPP) consensus statement on the definition and scope of prebiotics. *Nature reviews Gastroenterology & hepatology*, 14 (8): 491.

Jachimska, B., Wasilewska, M. & Adamczyk, Z. (2008). Characterization of globular protein solutions by dynamic light scattering, electrophoretic mobility, and viscosity measurements. *Langmuir*, 24 (13): 6866-6872.

La Rosa, S. L., Kachrimanidou, V., Buffetto, F., Pope, P. B., Pudlo, N. A., Martens, E. C., Rastall, R. A., Gibson, G. R. & Westereng, B. (2019a). Wood-Derived Dietary Fibers Promote Beneficial Human Gut Microbiota. *mSphere*, 4 (1): e00554-18.

La Rosa, S. L., Leth, M. L., Michalak, L., Hansen, M. E., Pudlo, N. A., Glowacki, R., Pereira, G., Workman, C. T., Arntzen, M. Ø. & Pope, P. B. (2019b). The human gut Firmicute Roseburia intestinalis is a primary degrader of dietary β-mannans. *Nature communications*, 10.

Lassfolk, R., Rahkila, J., Johansson, M. P., Ekholm, F. S., Wärnå, J. & Leino, R. (2018). Acetyl group migration across the saccharide units in oligomannoside model compound. *Journal of the American Chemical Society*, 141 (4): 1646-1654.

Leavitt, S. & Freire, E. (2001). Direct measurement of protein binding energetics by isothermal titration calorimetry. *Current opinion in structural biology*, 11 (5): 560-566.

Leszek Michalak, J. C. G., Leidy Lagos, Sabina Leanti La Rosa, Johannes Dröge, Margareth Øverland, Phillip B. Pope, Bjørge Westereng. (2019). *Wood-derived galactoglucomannan promote butyrate-producing microbes in the wine gut microbiome*. Science, N. U. o. L. (ed.).

Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M. & Henrissat, B. (2013). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic acids research*, 42 (D1): D490-D495.

Lombard V, G. R. H., Drula E, Coutinho PM, Henrissat B. (2014). *The Carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res 42:D490–D495. [PMID : 24270786]*. Available at: http://www.cazy.org (accessed: 14.05).

Lopez-Siles, M., Khan, T. M., Duncan, S. H., Harmsen, H. J., Garcia-Gil, L. J. & Flint, H. J. (2012). Cultured representatives of two major phylogroups of human colonic Faecalibacterium prausnitzii can utilize pectin, uronic acids, and host-derived substrates for growth. *Appl. Environ. Microbiol.*, 78 (2): 420-428.

Lopez-Siles, M., Duncan, S. H., Garcia-Gil, L. J. & Martinez-Medina, M. (2017). Faecalibacterium prausnitzii: from microbiology to diagnostics and prognostics. *The ISME journal*, 11 (4): 841.

Louis, P., Scott, K. P., Duncan, S. H. & Flint, H. J. (2007). Understanding the effects of diet on bacterial metabolism in the large intestine. *Journal of applied microbiology*, 102 (5): 1197-1208.

Louis, P. & Flint, H. J. (2017). Formation of propionate and butyrate by the human colonic microbiota. *Environmental microbiology*, 19 (1): 29-41.

Lundqvist, J., Teleman, A., Junel, L., Zacchi, G., Dahlman, O., Tjerneld, F. & Stålbrand, H. (2002). Isolation and characterization of galactoglucomannan from spruce (Picea abies). *Carbohydrate Polymers*, 48 (1): 29-39.

Michalak, L., Knutsen, S. H., Aarum, I. & Westereng, B. (2018). Effects of pH on steam explosion extraction of acetylated galactoglucomannan from Norway spruce. *Biotechnology for biofuels*, 11 (1): 311.

Michalak, L., Rosa, S. L. L., Leivers, S., Lindstad, L. J., Røhr, Å. K., Aachmann, F. L. & Westereng, B. (2019). *A Pair of eterases from a commensal gut bacterium completely deacetylate highly complex mannans*. Sscience, N. U. o. L. (ed.).

Miquel, S., Martin, R., Rossi, O., Bermudez-Humaran, L., Chatel, J., Sokol, H., Thomas, M., Wells, J. & Langella, P. (2013). Faecalibacterium prausnitzii and human intestinal health. *Current opinion in microbiology*, 16 (3): 255-261.

Montanier, C., Money, V. A., Pires, V. M., Flint, J. E., Pinheiro, B. A., Goyal, A., Prates, J. A., Izumi, A., Stålbrand, H. & Morland, C. (2009). The active site of a carbohydrate esterase displays divergent catalytic and noncatalytic binding functions. *PLoS biology*, 7 (3): e1000071.

Morrison, D. J. & Preston, T. (2016). Formation of short chain fatty acids by the gut microbiota and their impact on human metabolism. *Gut microbes*, 7 (3): 189-200.

Nakamura, A. M., Nascimento, A. S. & Polikarpov, I. (2017). Structural diversity of carbohydrate esterases. *Biotechnology Research and Innovation*, 1 (1): 35-51.

Pryde, S. E., Duncan, S. H., Hold, G. L., Stewart, C. S. & Flint, H. J. (2002). The microbiology of butyrate formation in the human colon. *FEMS microbiology letters*, 217 (2): 133-139.

R. Gibson, G., Scott, K., A Rastall, R., Tuohy, K., Hotchkiss, A., Dubert-Ferrandon, A., Gareau, M., F. Murphy, E., Saulnier, D., Loh, G., et al. (2010). *Dietary prebiotics: Current status and new definition*, vol. 7.

Ramirez-Farias, C., Slezak, K., Fuller, Z., Duncan, A., Holtrop, G. & Louis, P. (2008). Effect of inulin on the human gut microbiota: stimulation of Bifidobacterium adolescentis and Faecalibacterium prausnitzii. *British Journal of Nutrition*, 101 (4): 541-550.

Rastall, R. A. & Gibson, G. R. (2015). Recent developments in prebiotics to selectively impact beneficial microbes and promote intestinal health. *Current Opinion in Biotechnology*, 32: 42-46.

Rosano, G. L. & Ceccarelli, E. A. (2014). Recombinant protein expression in Escherichia coli: advances and challenges. *Frontiers in microbiology*, 5: 172.

Roslund, M. U., Aitio, O., Wärnå, J., Maaheimo, H., Murzin, D. Y. & Leino, R. (2008). Acyl group migration and cleavage in selectively protected β-d-galactopyranosides as studied by NMR spectroscopy and kinetic calculations. *Journal of the American Chemical Society*, 130 (27): 8769-8772.

Ruhaak, L., Zauner, G., Huhn, C., Bruggink, C., Deelder, A. & Wuhrer, M. (2010). Glycan labeling strategies and their use in identification and quantification. *Analytical and bioanalytical chemistry*, 397 (8): 3457-3481.

Sankar, S. A., Lagier, J.-C., Pontarotti, P., Raoult, D. & Fournier, P.-E. (2015). The human gut microbiome, a taxonomic conundrum. *Systematic and applied microbiology*, 38 (4): 276-286.

Scheller, H. V. & Ulvskov, P. (2010). Hemicelluloses. *Annual review of plant biology*, 61.

Singh, S., Singh, G. & Arya, S. K. (2018). Mannans: An overview of properties and application in food products. *International journal of biological macromolecules*.

Tamanai-Shacoori, Z., Smida, I., Bousarghin, L., Loreal, O., Meuric, V., Fong, S. B., Bonnaure-Mallet, M. & Jolivet-Gougeon, A. (2017). Roseburia spp.: a marker of health? *Future Microbiology*, 12 (2): 157-170.

Till, M., Goldstone, D. C., Attwood, G. T., Moon, C. D., Kelly, W. J. & Arcus, V. L. (2013). Structure and function of an acetyl xylan esterase (Est2A) from the rumen bacterium Butyrivibrio proteoclasticus. *Proteins: Structure, Function, and Bioinformatics*, 81 (5): 911-917.

Timell, T. (1967). Recent progress in the chemistry of wood hemicelluloses. *Wood Science and Technology*, 1 (1): 45-70.

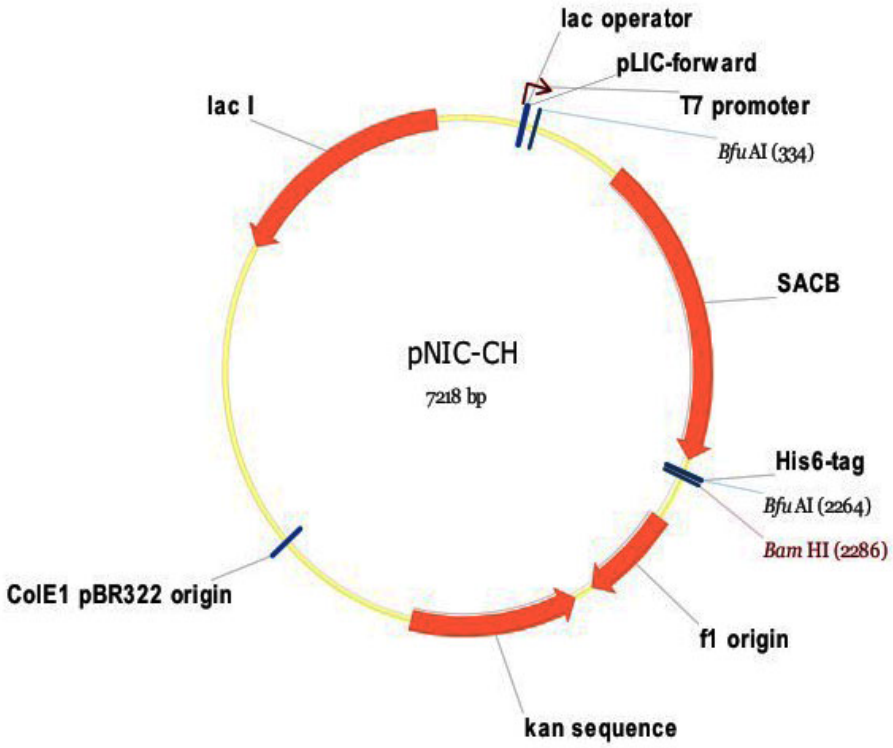
Topakas, E., Kyriakopoulos, S., Biely, P., Hirsch, J., Vafiadi, C. & Christakopoulos, P. (2010). Carbohydrate esterases of family 2 are 6-O-deacetylases. *FEBS letters*, 584 (3): 543-548.

Williamson, G., Kroon, P. A. & Faulds, C. B. (1998). Hairy plant polysaccharides: a close shave with microbial esterases. *Microbiology*, 144 (8): 2011-2023.

7 Appendix

Appendix A

An overview of the pNIC-CH plasmid used for cloning in this study. Figure from Addgene. pNIC-CH was a gift from Opher Gileadi (Addgene plasmid # 26117.



Appendix B

Gene sequences of *Fp*CE2:

```
>FP929046.1:1761869-1762981 Faecalibacterium prausnitzii SL3/3 draft genome
TTACAGGATCGTTCTCAGAAAGGCGGTCAGGACGTTTGCCGCCTGCCGGTGGTTCTCTGCGCCGGGGTGC
TGCCGGGCACCCACGGTTTCCGGCGTGGTGTTGGGCAGCTCCAGCAGATAGACAGAGCTGTCTCCGGTAA
TGGCCTTGTACTGCTCCATGCCCTGCCGCAATACCGGCAGCAGCTCGCTGCCCAGCATCCCGATGCACCA
CACCAGCTTTGCCCCGGGGTTCTTTGCCCGGAGCAGGGTCAGAAAATGCTGTACACCATCTGCCACTTTT
TGGGCGTCCGCAGGGTGGAAGTCTCCATTGGAGAGCCGCCGCAGCTGGTGGGTCTTGCCAGTGGCAGGGT
```

CTTGCCACGGCGGGTTATCAAATGCGCCCGTGTCGTTGGTGCCCAGGTTGACGATGACAGCATCCGGTTT
CCATGCAGCAAAATCGTTTTCCTGCTGTGCGCCAAGGGCAGCATTGCGCTGCCCCATCGCTACACCGCAC
ACCTGCGTATAGTAGGGCGGCATCACATGGCGCACATCGTTGTCCCAGCCGGTCACGATGCCCCAGCCAC
TTTGCGAGATGCAGCGGTATTCGGCTCCCAGTGCATCGGCGGTCAGGCGGCCGTAGTGGTTCTCTGCAGA
GAAAAAGGCTCCCACCCAGTCCTCCTCCGGCTTGGCCCCGATGGCACCCTCACCGCTGGTGATGCTGTCG
CCCACAAACTCCAGCCGATACTGCGGCTCCGGCAGGGGCAGAAATTCGCCGCCTGCATATTCCAGCCCGG
TGATCTGCAACAGATGCGCCGGGTCATCGTGCATGGCCTGCACATCCTTTAGCAGCCGGAGGTGCTTTGC
CTTGCCCGGTGTCATGCCCCGGAAAAGACAGATCCGGCTTTTGCCGGGGTTCACGGCAAACCGTGCCACC
CATGCGCCGTTCAGCTCCACGCTGACCCACGGCTCTACCGCTTCATAATCGGCGAAAAGATCCACCCACA
GCTCACTGCCGGTAAATTCCAGCTCCATGCCGCTGGCCGTCCAGAACAGGGTCAGCGGGTCACGCCCGGT
GTGCCGCCCCAGTGCTCGCACCTGCGGCAGCTGGGTCAGCGTCGCAGTTGTAAGTTGCTCCAT

Amino acid sequence of *Fp*CE2:

MEQLTTATLTQLPQVRALGRHTGRDPLTLFWTASGMELEFTGSE
LWVDLFADYEAVEPWVSVELNGAWVARFAVNPGKSRICLFRGMTPGKAKHLRLLKDVQ
AMHDDPAHLLQITGLEYAGGEFLPLPEPQYRLEFVGDSITSGEGAIGAKPEEDWVGAF
FSAENHYGRLTADALGAEYRCISQSGWGIVTGWDNDVRHVMPPYYTQVCGVAMGQRNA
ALGAQQENDFAAWKPDAVIVNLGTNDTGAFDNPPWQDPATGKTHQLRRLSNGDFHPAD
AQKVADGVQHFLTLLRAKNPGAKLVWCIGMLGSELLPVLRQGMEQYKAITGDSSVYLL
ELPNTTPETVGARQHPGAENHRQAANVLTAFLRTIL

Gene sequence of *Fp*CEXX:

>FP929046.1:1764180-1765259 Faecalibacterium prausnitzii SL3/3 draft genome
ATGAATACGTCTAATTTTGCTCGTTTGAAAGAACTGTTCCGCCGTGCCGCGGCAGGGCAGGAGTTGACCA
TCGGCTTTCTGGGCGGTTCCATCACCCAGGGCAGCCTTTCCACCCAGCCCGGTAATGCTTACGCTTTCCG
GGTGTACCAGTGGTTCGTGGATACCTTTCCGCAGTCGAAGTTTCACTATGTCAACGGCGGCATCGGCGGC
ACCAGCTCCCACTATGGCGTTGCCCGTGCCGTGACCGATGTGCTGATGTACCAGCCGGATTTTGTGGCGG
TGGATTTCAGCGTCAACGATCTGGAAGTCCCCTTCCGGCAGGAGACTTATGAGGGCGTTGTCCGCAAGCT
GCTCACATGGCCCTCCCACCCGGCGGTGGTGCTGCTGAACAACATTTATTACGATACCGGCGAGACGTCA
CAGGACGAGCACAACGCCGTGGGCGACCACTACGGTGTGCCCCACGTCAGCATCCGGGACAGCATTTACA
AGGATCTGCGTGCCGGTAAGTATGCCAGCCGCACCCTGTTGAGCCCCGATGGGCTGCACCCCAACGATTA
CGGTCACGGTCTGGTGGCAGGGGAGATCATCAAGCTGCTGGAGGAAGTCAACGCCCACCGGGAAGAACCA
GAGCAGGAGCCTGCCTTCCCGGCACCGCTCACCGCCAACGCCTACGAGAACGCCCGGCGGCTGACCATCC
GGGAGCTCTCCCCCAAGCTGGAGGGTTTCCGGGCAGACCCGGAGGAAAAGCTGGGGCATCTGGACCACTG
GAAAAACGGCTGGATCGGCCAGAAGCCCGGGGATAAGATCACCTTTGAAGTGGATGCCTCCTGCATTGCA
GTGCAGTACCGCAAGACCATCCGCCGCCCGGCGCTGCGGGCACAGCTGGTGCTGGACGGCGACACCGCCC
ACGCCGCTGTGCTGGACGGCAACTTTGATGAGGATTGGGGCGACTGCCTGTATCTCCAGCCCATTCTCCA

CCATGGTGAGCAGAAGAACCATACCGTGGAGATCACCATTCTGCCCACAGAAGATGAAACTCCCACGGAG
TTCTACCTCATGGCTCTGATTACGGCATGA

Amino acid sequence of *Fp*CEXX:

MNTSNFARLKELFRRAAAGQELTIGFLGGSITQGSLSTQPGNAY
AFRVYQWFVDTFPQSKFHYVNGGIGGTSSHYGVARAVTDVLMYQPDFVAVDFSVNDLE
VPFRQETYEGVVRKLLTWPSHPAVVLLNNIYYDTGETSQDEHNAVGDHYGVPHVSIRD
SIYKDLRAGKYASRTLLSPDGLHPNDYGHGLVAGEIIKLLEEVNAHREEPEQEPAFPA
PLTANAYENARRLTIRELSPKLEGFRADPEEKLGHLDHWKNGWIGQKPGDKITFEVDA
SCIAVQYRKTIRRPALRAQLVLDGDTAHAAVLDGNFDEDWGDCLYLQPILHHGEQKNH
TVEITILPTEDETPTEFYLMALITA

Appendix C

Amino acid sequence for *Ri*CEX:

>EEV02614.1 GDSL-like protein [Roseburia intestinalis L1-82]
MEYQIKYENGIANRGCLYRLKKVMDRAKAGEALNIAFLGGSITQGSLSSKPELCYAYHVYEWWKKTFPQA
DFTYINAGIGGTTSQFGVARAEADLLSKEPDFVIIEFSVNDDSTEHFMETYEGLVRKVYTSKTKPAVLLV
HNVFYNNGANAQLMHGRIARHYNLPAVSMQSTIYPEVVAGRIENREITPDDLHPNDAGHALVASVITYFL
DKVKTEDATEQSEPDYPAPLTKNTYEKSIRHQNSDENVVCHGFVADTSAQRDITDCFKHGWTASKKGDSI
TLDVEGCNISVQYRKSVKLPAPVAEIIVDGDAEHAVRLDANFDETWGDKLELDTILEHGENKVHKVEVRL
TETHENDAVPFYLVSVIGSSEK

Norges miljø- og biovitenskapelige universitet
Noregs miljø- og biovitskapelege universitet
Norwegian University of Life Sciences

Postboks 5003
NO-1432 Ås
Norway