Identification and characterisation of an enzyme from *Monoglobus pectinilyticus* associated with degradation of pectin from Kiwiberry

A thesis presented in partial fulfilment of the requirements for the degree of

Master of Science

in

Biological Sciences

at Massey University, Palmerston North,

New Zealand

Aymee Lewis

2022

Dedicated to my brother, Josh Lewis, for never leaving my side and to my Granma Melva Bradshaw, who rests peacefully.

Abstract

Pectin is a complex polysaccharide and a very important source of dietary soluble fibre, present in a variety of fruit and vegetables. Pectin possesses a wide range of contributions to a healthy diet and a healthy human gut, such as lowering cholesterol, protecting against intestinal inflammation and maintaining digestive health. Yet pectin is unable to be degraded by human gastrointestinal enzymes, arriving in the large intestine mostly intact. Limited research has showed the domination of Gram-negative bacteria which possess a range of secreted extracellular cell-bound and cell-free pectin degrading enzymes. These enzymes attack the pectin backbone and the accessory side chains resulting in the isolation of mono/oligosaccharides for subsequent uptake. The recent novel discovery of *Monoglobus pectinilyticus,* a Gram-positive bacterium, expanded the narrow knowledge regarding specific pectinolytic degraders. Genomic studies showed this bacterium contains an arsenal of enzymes specifically for pectin degradation. However, several putative pectin-degrading enzymes produced by *M. pectinilyticus* were unable to be identified against up-to-date databases, thus suggesting that there may be potentially novel classes of CAZyme(s) to be discovered. The proposed study will attempt to identify such enzyme(s) using New Zealand cultivar kiwiberries as the source of pectin. The initial aim of this study was to investigate the enzyme(s) involved in β-1,4-galactan by growing *M. pectinilyticus* with a mixture of pectin substrates (0.5% citrus pectin, 0.5% kiwiberry pectin and 0.5% potato galactan), in hope to identify such enzyme(s). Unfortunately, no β-galactosidase/β-galactanase was found. However a potentially novel binding enzyme, protein 0050 was isolated, cloned and expressed, which may be involved in a pectin degrading version of a cellulosome complex. Further research into isolating a variety of secreted proteins activated in presence of galactan during *M. pectinilyticus* cultivation is needed to understand the complexity of pectin degradation.

Keywords: pectin; kiwiberry; *Actinidia arguta*; *Monoglobus pectinilyticus*; gastro-intestinal tract; CAZymes; β-galactanase; β-galactosidase.

Acknowledgement

The support, help and encouragement of many individuals along the way are largely acknowledged. I extend the uttermost gratitude to all my supervisors for their guidance and support along the way. To Dr Caroline Kim for the reassurance, mentorship, academic guidance and support. To Emeritus Prof. Geoff Jameson for his kindness, encouragement and insightful feedback and patience. To Dr Erin O'Donoghue for her patience, support and wisdom. To Dr Dave Brummell for his valuable suggestions and reassurance. I also extend my appreciation to Dr Paul Blatchford, whose continued praise helped along the way. I would like to extend my sincere appreciation to the members of the Food, Nutrition and Health team at Plant & Food Research, for their constant inspiration, support, friendship, laughter and contributions to my research. A special thanks to Dr Robert Simpson, Trevor Loo and Ruby Roach for their much needed contribution to my protein work, and Sheryl Somerfield for her contribution to my pectin work.

A special thanks goes to my family and friends who provided the much the needed support, drive and determination. To my Mum, Kay, I could not have achieved this without her moral, emotional and mental support. It is entirely their love, persistence and devotion which has molded me into who I am today. To Mini moo, thank you for your croaky meows and thundering purrs to help keep me grounded. Lastly, to my partner Monique, your love, support and ability to always be there for me given me the drive to follow my passion.

Lastly, this opportunity would not be possible without the financial support awarded from the Māori Kiwiberry Scholarship provided by Zespri Ltd.

Table of Contents

List of Figures

List of Tables

List of Appendices

List of copyrighted material for which permission was obtained

Material	Source	Note
Figure 1.1	Flint *et al*., 2012. Gut Microbes. 3, 289-306.	No permissions required
Figure 1.4	Harholt *et al*., 2010. Plant Physiology 153(2): 384-395.	Permissions were granted for use
Figure 1.5	Schmitz *et al*., 2019. Applied Microbiology and Biotechnology 103: 2507-2524.	Permissions were granted for use
Figure 2.3	Labus, 2018. PLoS One 13(10): e0205532.	No permissions required

Abbreviations

AIX	Anion exchange chromatography
AS	Ammonium Sulfate
BGAL	β-galactosidase/β-galactanase
BP-EIS	Buffered-phenol alcohol insoluble solids
CAZymes	Carbohydrate active enzyme
FW	Fresh weight
GalA	Galacturonic acid
HG	Homogalacturonan
IMAC	Immobilised metal affinity chromatography
IPTG	Isopropyl-β-D-thio-galactoside
LC-MS/MS	Liquid Chromatography tandem Mass Spectrometry
LDS	Lithium dodecyl sulfate
mAb	Monoclonal antibody
MM	Mineral Media
MM	Mineral Media
MOPS	3-(Morpholin-4-yl)propane-1-sulfonic acid
MWCO	Molecular weight cut off
NAD	Nicotinamide adenine dinucleotide
NADH	Nicotinamide adenine dinucleotide phosphate
Ni-NTA	Nickel (II) nitroacetate
ONPG	Ortho-nitro phenyl-B-D-galactopyranoside
PAGE	Polyacrylamide gel electrophoresis
PGA	Polygalacturonic acid
POS	Pectin oligosaccharides
PULs	Polysaccharide utilisation locus
RG-I	Rhamnogalacturonan I
RG-II	Rhamnogalacturonan II
Rha	Rhamnose
ROS	Reactive oxygen species
SCFA	Short chain fatty acid
SDS	Sodium dodecyl sulfate
SLH	S-layer homology
UA	Uronic Acid
XGA	Xylogalacturonan

Amino acid Abbreviations

A	Alanine
R	Arginine
N	Asparagine
D	Aspartic Acid
C	Cysteine
Q	Glutamine
E	Glutamic Acid
G	Glycine
H	Histidine
I	Isoleucine
L	Leucine
K	Lysine
M	Methionine
F	Phenylalanine
P	Proline
S	Serine
T	Threonine
W	Tryptophan
Y	Tyrosine
V	Valine

Chapter 1 Introduction

The human gastrointestinal tract is inhabited by a diverse microbial community, known as the microbiome. These communities play a vital and dynamic role in human nutrition and health, from host protection against pathogenic microbes to producing a wide range of metabolites that are unable to be synthesised by the host (Thursby & Juge, 2017). The human gut microbiome consists of six major phyla – dominated by *Firmicutes* and *Bacteroidetes,* with large groups of *Actinobacteria* and *Proteobacteria*, followed by both *Verrucomicrobia* and *Euryarchaeota* (Bang *et al*., 2018). Food ingested by the host provides the gastrointestinal community with an energy source, which influences their composition and metabolic activities. They also have the capacity to convert dietary fibres into short chain fatty acids (SCFA) and other simple molecules that provide potential health benefits directly to the host (Despres *et al*., 2016; Larsen *et al*., 2019; Nakajima *et al*., 1999).

The consumption of plants, and in turn, plant cells, staples within the average human diet, provide nutrients, vitamins and minerals (Holland *et al*., 2020). Cell walls are important features of plant cells that contribute to a number of essential functions, such as providing shape, intercellular communication and plant-microbe interactions (Keegstra, 2010).

Plant cell walls make up 20-30% of daily consumed fibre (Kong & Singh, 2009) and are comprised of complex polysaccharide structures, containing a matrix of hemicellulose, pectin and cellulose (Caffall & Mohnen, 2009). This matrix provide plants with structural stability and flexibility during cellular division, as well as a barrier against environmental or pathological stress (Hofte *et al*., 2017; Lodish *et al*., 2000). Due to the variability of covalent and non-covalent cross-linkages in the plant cell wall, degradation by the human digestive system poses a challenge, because of the inability of human endogenous enzymes to cleave these carbohydrate polymers (Flint *et al*., 2012).

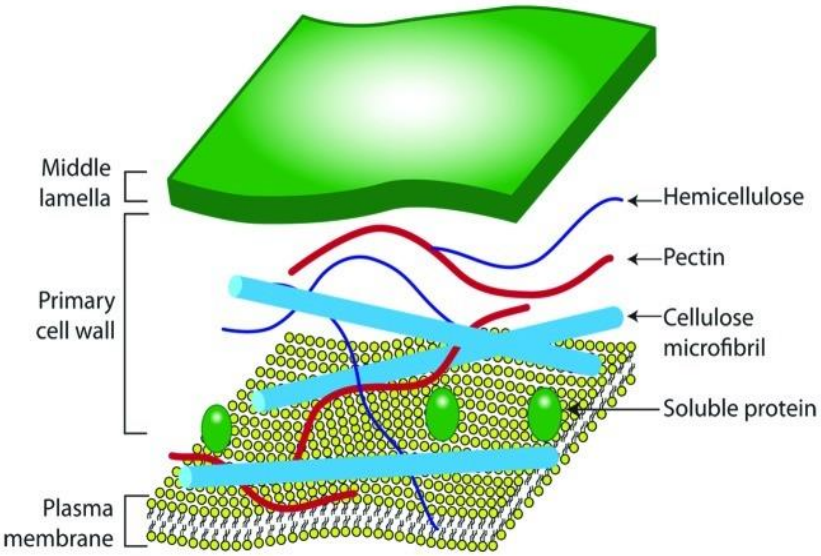


Figure 1.1 The plant cell wall primary structure – a representation of the major polysaccharide components (Flint *et al*., 2012). No copyright permission needed.

Plant cell wall degrading microbes contain genes encoding fibrolytic enzymes which are necessary for the initiation of plant cell wall degradation. The action of these initial enzymes result in the release of more accessible forms of polymers for degradation by other bacteria (Dongowski *et al*., 2000). However, a small group of specific bacteria with the capability of degrading complex polysaccharides has been isolated and identified (Carlos *et al*., 2018; Dongowski *et al*., 2000).

Pectin is a hetero-polysaccharide, rich in galacturonic acid (GalA) and makes up approximately a third of dicotyledon and non-poales monocotyledon plant cell wall dry weight (Chung *et al*., 2017). It is highly abundant, structurally and functionally complex, and contributes to physiological processes in plant tissue (Caffall & Mohnen, 2009). The health benefits of pectin have been extensively researched, establishing their contribution to reducing cholesterol levels, protecting against gastrointestinal epithelium inflammation and maintaining a healthy gut (Elshahed *et al*., 2021; Larsen *et al*., 2019; Parker *et al*., 2010; Zhu *et al*., 2019).

Actinidia arguta 'kiwiberry', is a relatively new horticultural product that is becoming increasingly popular among consumers worldwide (Cossio et al., 2015). The spike in popularity has resulted in a wide range of kiwiberry research, including fruit nutritional and

health benefits (Latocha, 2017; Latocha *et al*., 2021; Parker *et al*., 2010; Pinto *et al*., 2020; Zhu *et al*., 2019), commercial challenges (Debersaques *et al*., 2015), and postharvest challenges (Han et al., 2019; Latocha *et al*., 2021; Sutherland *et al*., 2017). Kiwifruit research has dominated the *Actinidia* field with an enriched amount of publications detailing a wide range of scientific and commercial information. Major horticultural industries are always on the hunt for other fruits that could provide similar or even better commercial success – in this case, kiwiberries. However, growers are running into various issues cultivating this vine (Debersaques *et al*., 2015; Latocha *et al*., 2021). In particular, one postharvest challenge is the short shelf and storage life of the *Actinidia arguta* fruit, due to the edible, soft skin, and the fruit's rapid decrease in firmness after harvest (Debersaques *et al*., 2015; Latocha *et al.*, 2021; Williams *et al*., 2003). These poses significant limitations and difficulties in the kiwiberry industry.

1.1 *Actinidia arguta* "kiwiberry"

Actinidia is a large genus of plants which are functionally deciduous and vigorous climbers found predominantly in China (Datson & Ferguson, 2011). Some of these plants produce berries containing seeds within a pericarp, with the majority of the berries being edible (Latocha, 2017; Latocha *et al*., 2021; Sutherland *et al*., 2017). The most common and popular species are the kiwifruits, *Actinidia deliciosa* 'Hayward' and *Actinidia chinensis* 'Hort16A', where 148.8 million trays of the fruit were exported from New Zealand, producing a total net return of NZD 2.3 billion. On the other hand, *Actinidia arguta*, commonly known as 'kiwiberry', 'baby kiwi' or 'hardy kiwi', produced a combined total net of NZD 3.7 million from both domestic and international sales (Aitken & Hewett, 2019). New Zealand commercially grows three main cultivars, *Actinidia arguta* (Sieb. Et Zucc) Planch. Ex Miq. var. arguta 'Hortgem Tahi', var. 'Takaka Green' and var. 'Marju Red' for the domestic and international market.

Figure 1.2 Kiwifruit and kiwiberry, grown from Plant and Food Research, Te Puke, New Zealand. **Left side**: *Actinidia deliciosa*, green kiwifruit. **Right side**: *Actinidia arguta* cultivar Hortgem Tahi.

Kiwiberries are a perennial vine which fruit hairless berries (Latocha *et al*., 2021). They are found widely spread in Northeast Asian countries. They display a range of polymorphic characteristics such as leaf shape, presence or absence of leaf hairs and flower attributes (Latocha *et al*., 2021; Sutherland *et al*., 2017), which may not be commercially important. In contrast, morphological variations of the kiwiberry fruit size, colour, flavour and storage life (Sutherland, *et al*., 2017; Williams, *et al*., 2003) are more favourable to the commercial grower and general consumer. Typically, kiwiberry fruit weigh between 5 – 16 grams with hairless, smooth edible skins enriched in antioxidants (Latocha, 2017; Latocha *et al*., 2021). The fruit are generally green, but may change colour, depending on the type of cultivar and fruits ripening. The colour feature can change to a darker green, bright purple or red (Ferguson, 2016; Latocha *et al*., 2021; Sutherland *et al*., 2017). The fruit shape characteristics are also dependent on the cultivar and range from globose to cylindrical ellipsoid (Ferguson, 2016; Latocha *et al*., 2021). The fruit have an aromatic smell with

intense and sweet flavours, similar to that of a ripe strawberry or blackcurrant (Latocha, 2017).

A. arguta thrives in a cold climate with reports showing its high frost hardiness (carries fruit in temperatures falling to -30°C) and its somewhat short vegetation period (Latocha, 2017; Williams *et al*., 2003). However, kiwiberries have a reputation of being difficult to grow and manage due to their ripening irregularities, short shelf and storage life, compared to other *Actinidia* species (Sutherland *et al*., 2017; Williams, *et al*., 2003). The issues arise with the delicate, smooth skin consisting of a simple epidermis. It contains a thick cuticle and a hypodermis which is only one to two cells thick. These fruits also ripen in cold storage to eating firmness within four to six weeks, making postharvest storage and shelf life difficult to overcome (Hallett & Sutherland, 2005; Han *et al*., 2019; Latocha *et al*., 2021; Sutherland *et al*., 2017). Once out of cold storage, kiwiberries take only 4 days to reach edible softness, with fruits considered overripe at day 5 (Sutherland *et al.,* 2017).

Figure 1.3 Hortgem Tahi' kiwiberries, grown from Seeka Limited, Te Puke, New Zealand. (A) 'Hortgem Tahi' fruit dangling off the vine in bundles. (B) Close up bundles of 'Hortgem Tahi'.

The fruit is considered to be an exceptional source of many vitamins, dietary fibres, actinidins and antioxidants (Latocha, 2017; Latocha *et al*., 2021; Leontowicz *et al*., 2015; Parker *et al*., 2010; Pinto *et al*., 2020). The rich nutritional profile of kiwiberry is partially due to a high degree of edibility of the fine and smooth skin, which contains up to 15 times more antioxidants than the fruit pulp alone (Latocha *et al*., 2017; Pinto *et al*., 2020). The fruits of the kiwiberry contain a large amount of essential nutrients, where on average they contain 150 mg/g fresh weight (FW), though this is dependent on which cultivar is studied (Latocha *et al*., 2021; Nishiyama *et al*., 2004). In addition, the fruit hold a significant amount of vitamin B-complex, with levels of *myo*-inositol (vitamin B_8) ~5 times higher than kiwifruit (Nishiyama *et al*., 2008). Thus, the fruit of kiwiberry contains one of the richest sources of the hexahydric sugar alcohol. The phenolic compounds of kiwiberries are thought to be very effective antioxidants (Latocha *et al.,* 2015). In comparison to Hayward kiwifruit, kiwiberries contain up to 3x more total phenolic compounds, which are dominated by flavanols (96-99% total phenolic compounds) (Latocha, 2017; Leontowicz *et al*., 2016). The higher amounts of phenolic compounds in kiwiberry may be associated to the fruit skin or peel compared to the pulp (Latocha, 2017). Within *Actinidia* species, kiwiberries contain the highest amount of lutein (0.93 mg/100 g FW) and β-carotene (0.29 mg/100g FW). Both are potent antioxidants (Latocha, 2017). While most kiwiberry are green-fleshed, those with red flesh such as Marju Red also contain a small amount of anthocyanin (Latocha, 2017). Given the number of potential antioxidants present which possess free radical scavenging abilities, kiwiberry and its nutritious composition could help reduce diseases associated with free radicals (An *et al.*, 2016; Latocha *et al*., 2021). Ingestion of a single kiwiberry provides ~2.0 – 3.8% total dietary fibre (TDF), which consists of ~2.1 – 3.1% insoluble fibre and ~0.8 – 1.0% soluble fibres (Zhu *et al*., 2019). These fibres contribute to the health of our gastrointestinal tract and consist of several polymeric plant materials associated with the plant cell wall, such as pectin (Mudgil, 2017). Kiwiberries (100g) contain between 2.17 – 3.30% pectin, though this is cultivar dependent (Latocha, 2017; Latocha *et al*., 2021). The breakdown of pectin into pectin oligosaccharides (POS) by intestinal flora have physiological activities such as prebiotic, antibacterial, anticancer and antioxidant properties (Parker *et al*., 2011; Zhu *et al*., 2019). Also, POS has antiglycation activity, providing protection against reactive oxygen species (ROS) which leads to the reduction in oxidative stress within the gastrointestinal tract (Xiangxuge *et al*., 2016; Zhu *et al*., 2019).

1.2 Plant cell wall polysaccharide - Pectin

Pectin is defined as a hetero-polysaccharide and dominantly contains repeating units of galacturonic acid (GalA) residues, in which varying amounts of acid groups are present as methoxyl esters (Cosgrove, 2022; Voragen *et al*., 2009). The GalA are in alternating chains of α-1, 4-glycosidic linkages, with side chains containing neutral sugars. Pectin contains a pattern containing 'smooth' homogalacturonic regions and ramified 'hairy' rhamnogalacturonan regions, where the latter contains the majority of the neutral sugars (Mohen, 2008; Voragen *et al*., 2009). The structure of pectin consists of four main pectic polysaccharides: homogalacturonan (HG), rhamnogalacturonan I (RG-I), xylogalacturonan (XGA) and, lastly, rhamnogalacturonan II (RG-II) (Cosgrove, 2022; Jesper *et al*., 2010; Mohen 2008).

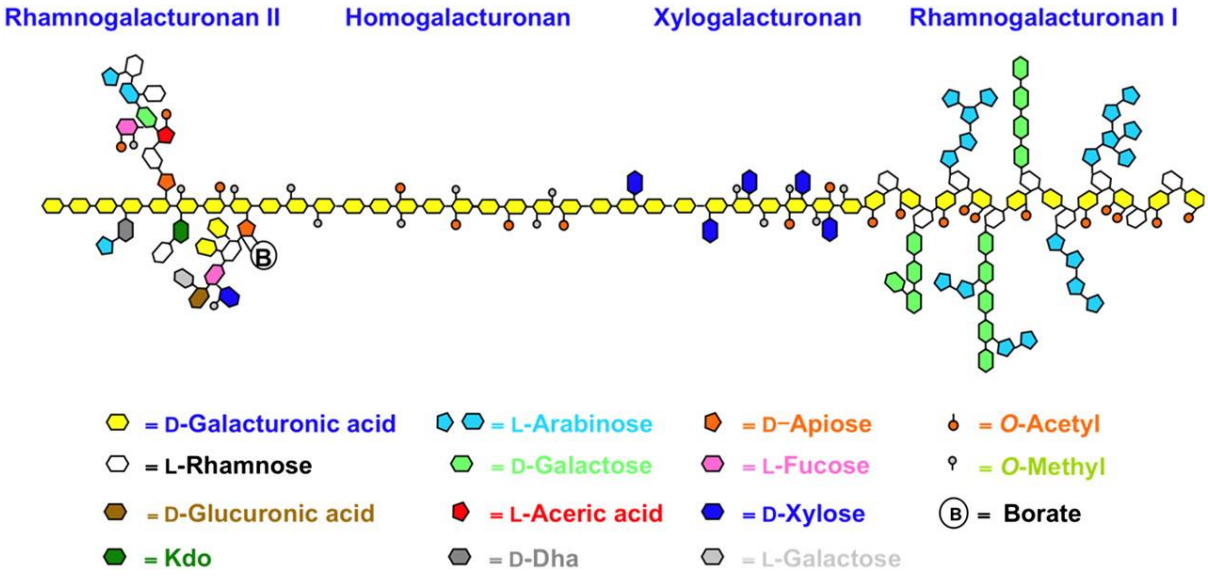


Figure 1.4 A schematic of pectin structure made up of the four types of pectic polysaccharides: HG, RG-I, RG-II and XGA (Harholt *et al.,* 2010). Permissions was granted for the use of this figure.

Homogalacturonan (HG), is the main and simplest pectic polysaccharide. HG is also known as the 'smooth' region (Cosgrove, 2022; Ropartz & Ralet, 2020). It comprises ~60% of the total pectin scaffold, containing a linear homopolymer molecule of α-1, 4-linked galacturonic acid (Caffall & Mohnen, 2009; Ropartz & Ralet, 2020). HG has a degree of polymerisation (DP) that can go up to ~100 (O'Neill *et al*., 1990). Methyl-esterification can

partially occur on the C-6 carboxyl group of the galacturonic acid which can also be *O*-acetylated on either *O*-2 or *O*-3, dependent on the type of plant (O'Neill *et al*., 1990; Ridley *et al*., 2001; Voragen *et al*., 2009). One study shows HG is generally represented as repeating units of approximately 100 galacturonic residues (Gawkowska *et al*., 2018). Other reports show it can be found as short HG regions between other pectic polysaccharides (Caffall & Mohnen, 2009). Also, the variations in the degree of acetylation (DA) or methylation (DM) on different galacturonic acid residues within homogalacturonan, occur at varying stages of plant cell growth (Mohnen, 2008). Pectin can be characterised by its degree of esterification (DE); the number of galacturonic acid residues esterified to form methyl esters to produce either high ester (DE >50%) or low ester (DE <50%) pectin (Harris & Smith, 2006; Ropartz & Ralet, 2020). The DE controls the gelling properties of pectin and can alter the structure of HG which, in turn, may affect degradation (Harris & Smith, 2006). A low methyl-esterified pectin gel forms over a wide range of pH in the presence of soluble solids, and requires the presence of divalent cations. In comparison, highly methyl-esterified pectin does form gels but only at low pH and in the presence of large amounts of soluble solids (Brejnholt, 2009). Due to these two gelling states, each type has its own function within the food industry, and also in their regulatory function on human health (Brejnholt, 2009). Galacturonic acid residues that are either methyl-esterified or acetylated carry a negative charge, which may interact with Ca^{2+} ions, enabling the formation of a characteristic gel-like texture (Caffall & Mohnen, 2009; Cosgrove, 2022). HG is able to form consecutive links to RG-I, RG-II and xylogalacturonan (Coenen *et al*., 2007; Cosgrove, 2022; Voragen *et al*., 2009).

Rhamnogalacturonan I (RG-I) is made from repeating units of GalA residues and rhamnose (Rha) residues, which are linked ($[\text{-}\alpha\text{-D-GalA-1,2-}\alpha\text{-L-Rha-1-4-}]_n$) (Figure 1.5). This polysaccharide makes up 20-35% of the pectin structure, and has a diverse number of neutral sugars, oligosaccharides and branched oligosaccharides attached to its backbone structure (Caffall & Mohnen, 2009). This variation is not fully understood and therefore, suggests that RG-1 has numerous functional roles. RG-I side chains are mostly linked to the O-4 rhamnose units and contain neutral sugars made of galactans, arabinians and potentially arabinogalactans (Cosgrove, 2022; Kaczmarska *et al*., 2020). Galactan side chains are characteristically made of unbranched polymers of 43 – 47 residues of β-D-1, 4-GalA

(Kaczmarska *et al*., 2020; Mohen, 2008). Type I arabinogalactan side chains of pectin are usually made of linear chains of β-1, 4-galactan substituted with a α-1, 5-linked arabino-oligosaccharides (Caffall & Mohnen, 2009; Voragen et al., 2009). Type II arabinogalactans consist of a β-1, 3-D-galactan backbone substituted with β-1, 6-D-galactan side chains adorned with arabinose and glucuronic acid residues (Ropartz & Ralet, 2020; Sakamoto *et al*., 2013). The decoration of neutral sugars and the amount of side chains on RG-I are dependent on plant source (Kaczmarska *et al*., 2022).

Rhamnogalacturonan II (RG-II also known as the substituted HG), though less abundant than RG-1, is the most structurally complex pectic polysaccharide (Figure 1.5). It constitutes less than 10% of the pectin structure and is made of ~13 different glycosyl resides. These residues are connected by over 21 different glycosidic linkages (Caffall & Mohnen, 2009; Silva *et al*., 2016). RG-II differs from RG-1 due to its backbone, which is composed of at least 8 linear α-1, 4-D-GalpA residues (Cosgrove, 2022; O'Neill *et al*., 2004). Many heterogeneous sugar residues attached to RG-II have been identified; however their exact location and orientation have yet to be finalised (Yapo, 2011). These identified sugars are galacturonic acid, galactose, arabinose, rhamnose, D-apiose (API), L-aceric acid (AceA) and more (Yapo, 2011). Mutations lead to a change in the degree of dimerisation, which plays a major effect in the plant cell growth. This can lead to growth defects, suggesting dimerisation of RG-II is vital for the healthy growth and developments of plants (Caffall & Mohnen, 2009; Yapo, 2011).

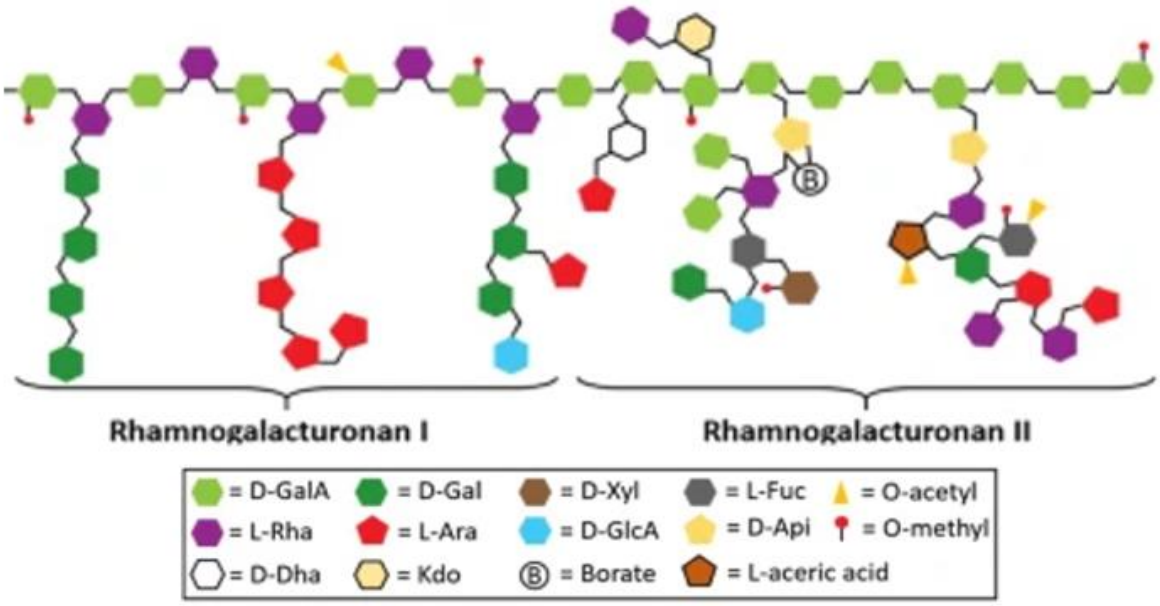


Figure 1.5 Rhamnogalacturonan I and II structure. Permissions were approved for the use of this figure; adapted from Schmitz, *et al*. 2019.

Xylogalacturonan (XGA) is another substituted HG, where a β-D-xylosyl is attached to the C-3 of the backbone α-D-galactopyranosyluronic acid (Caffall & Mohnen, 2009; Ridley *et al*., 2001). XGA has only been identified in several plant cells such as apple, watermelon, potato tuber and pea hulls (Coenen *et al*., 2007; Ropartz & Ralet, 2020). The expression of this pectic polysaccharide is restricted, although abundant in plant reproductive tissues (Caffall & Mohnen, 2009). One study observed that during pathogen response, an XGA biosynthetic gene is upregulated, suggesting that XGA could help make HG more resistant to degradation during a pathogen attack (Jensen *et al*., 2008).

1.3 Gastro-intestinal degradation of pectin by bacteria

Digestion of dietary plant material occurs in the large human intestine, where fibrolytic populations initiate the breakdown. Specifically, for pectin degradation, secreted enzymes attack the middle lamella and expose the cell wall for other bacterial attachment, allowing enzymes to attack the zone of penetration (Silva-Sanzana *et al*., 2020). The middle lamella is targeted for degradation due to its pectin-rich content (Daher & Braybrook, 2015). Although pectin is digested in the human gut, it does not add much contribution to faecal bulking,

and, therefore, pectin is extensively and almost fully digested within the human gut (Cummings *et al*, 1979). Research has shown that the Gram-negative *Bacteriodes* spp. dominate the breakdown of pectin in the gut, yet other phyla such as *Firmicutes* can contribute to pectin degradation (Comstock, 2009; Larson *et al*, 2019).

Bacteriodes is a genus containing a large arsenal of pectinolytic bacteria. As a result, our current knowledge of pectin degradation comes from studying Gram-negative models. *Bacteriodes* spp. are not very effective at binding to the cell surface as they lack the extracellular machinery to maintain proximal cell to substrate distance (Despres *et al*., 2016). As a result, this mechanism of secreting extra-cellular CAZymes is not the predominant way of producing CAZymes in *Bacteriodes* spp. for digestion (Despres *et al*. 2016). Early research and genome sequencing exposed the ability of *Bacteriodes thetaiotaomicron* to degrade and use plant polysaccharides as a carbon source. A total of 88 polysaccharide utilisation loci (PULs) were identified. These PULs encoded proteins involved in polysaccharide degradation (El Kaoutari *et al*., 2013). PULs are a set of physically linked genes that facilitate the breakdown of specific glycan. Up to 80% of genes encoded for CAZymes, which contains predicted signal peptide sequences targeting the protein to either the outer membrane or the periplasm (Dongowksi *et al*., 2000), suggests that their function is most likely due to the degradation of ingested plant material. This is supported by another publication, where no polygalacturonase enzymes or pectate lyases were identified in the extracellular supernatant during pectin degradation. This indicates the enzymes produced were not secreted but rather targeted to the outer membrane or periplasm (El Kaoutari *et al*., 2013). *Bacteriodes* spp. are known for their production of numerous glycosyl hydrolases, which hydrolyse the glycosidic linkage, and polysaccharide lysases, which cleave glycans at specifically activated glycosidic linkages. However, a simulated human gut microbiome study, found polysaccharide lyases made up less than <2 % of the glycobiome (Dongowksi, *et al*., 2000). This assumed pectin degradation by this genus is heavily focused on using glycoside hydrolases (Dongowksi, *et al*., 2000). This is in contrast to another simulated gut study that showed the main breakdown product detected was the end-product of pectin degradation, caused by the β-elimination (Chung *et al*., 2017). This reaction was catalysed by the pectate lyases of *Bacteriodes thetaiotaomicron*. Thus, the pectate lysases of the *Bacteriodes* genus plays a significant role in pectin degradation of plant-based dietary fibre

(Chung *et al*., 2017). The combination of the different *Bacterioides* species, with different glycan degradation abilities, allows maximum utilisation of plant fibres and contributes to host nutrition, by the production of short chain fatty acids (SCFA) (Comstock, 2009; Flint *et al*., 2008).

Enterobacteriaceae, a large family of Gram-negative bacteria, contains pectinolytic members such as *Escherichia coli* and *Yersinia enterocolitica,* which also encode genes for pectinases (Abbott & Boraston, 2008; Hugouvieux-Cotte-Pattat *et al*., 2014). Another study discovered *Y. enterocolitica* possesses both endo-pectinases and exo-pectinases (Anuradha *et al*., 2019). These two types are pectin lyases, known as PelY, a periplasmic-acting lyase and PelW, a cytoplasmic acting lyase (Hugouvieux-Cotte-Pattat *et al*., 2014). *Y. enterocolitica* also possesses YeCE8, which is an intracellular pectin methylesterase (Boraston, 2012).

Within the phylum *Firmicutes, Faecalibacterium prausnitzii* is one of the most abundant (~5%) bacteria within the human microbiota (Schwalm *et al*., 2016). The bacterium has a limited ability to break down pectin and can grow on pectic oligosaccharides with a Degree of Polymerisation (DP) of 4 and 5 (Chung *et al*., 2016; *Ferreira et al*., 2019; Martin *et al*., 2017; Salyers, 1977a). This is contrasted by numerous reports, stating the bacterium was unable to degrade pectin. Substrates tested were pectin from citrus peel, commercial pectin and polygalacturonate (Abbott *et al*., 2008; Salyers *et al*., 1977a). However, because of the variability in the experimental conditions, the ability of *F. prausnitzii* to contribute to pectin degradation remains unclear. Further study into whether this bacterium should be considered a true pectinolytic enzyme is needed. The genome of *Eubacterium eligens,* another Gram-positive *Firmicutes,* encodes a wide arsenal of pectinolytic enzymes, consisting of six CAZymes, including a pectate lysase (Chung *et al*., 2016; Martin *et al*., 2017). A rapid increase in cell density was reported to occur when apple pectin was added to the media in *in vitro* fermenters. This was confirmed when inulin was used as a negative control and resulted in *E. eligens* becoming undetectable (Lopesz-Stiles *et al*., 2017). A new strain of *Enterococcus mundtii,* was recently isolated from a human faecal sample, and unlike many other pectinolytic bacteria, this strain can grow solely on pectin – degrading ~35% within 13 hours into monosaccharides. Genomic insights suggest there is a variety of enzymes used to achieve this degradation (Jung *et al*., 2020).

Clostridium butyricum-beijerinckii, breaks down citrus pectin into precursor forms of short chain fatty acids under anaerobic conditions through the production of exo-/endo- pectate lyases and pectin methylesterase (Nakajima *et al*., 1999; Prade *et al*., 1999). A recent study (Kim *et al*., 2017) identified a new pectin degrading bacterium, *Monoglobus pectinilyticus,* which clusters with the family *Ruminococcaceae*. Genomic analyses showed that this bacterium possesses a highly specialised glycobiome, tailored for the digestion of pectin. It encodes a large unique set of CAZymes, where 91 genes were annotated, which encoded 108 putative CAZymes domains. This contained 47 putative pectin associated CAZymes and 9 putative hemicellulolytic enzymes. Research comparing the pool of CAZymes between *M. pectinilyticus* and other pectin-degrading bacteria such as *Bacteriodes thetaiotaomicron* highlighted the disproportionally high amounts of pectin lyase (PL) and carbohydrate esterases (CE)-dominated CAZymes (Kim *et al*., 2017). Originally, compared to *Bacteriodes thetaiotaomicron, M. pectinilyticus* lacked putative β-galactosidase/β-galactanase, required for the degradation of RG-I galactan side chains (Kim *et al*., 2017). Proteomic analysis revealed that the presence of pectin in the growth medium up-regulates the extracellular production of pectin-degrading enzymes attached to the bacterial cell surface via S-layer homology (SLH) protein domains (Kim *et al*., 2019). Proteins which contain an SLH domain (typically 50 – 60 residues long) have been proposed to bind to an anchoring structure such as a secondary cell wall polymers (SCWPs), yet the affinity and specificity of the interaction is not well established (Desvaux *et al*., 2006). Interestingly, large SLH proteins have been discovered and identified as constituents of lignocellulolytic degradation systems in both *Paenibacillus curdlanolyticus* B-6 and *Caldicellulosiruptos* spp. (Conway *et al*., 2016; Ratanakhanokchai *et al*., 2013). Thus, the purpose of *M. pectinilyticus* SLH-containing proteins may represent the constituents of a novel pectin degradation system (Kim *et al*., 2017).

1.4 Rationale of research

Interest in 'super-foods' category has significantly changed the awareness of 'modern' human diets and lifestyles. Increasingly, the world is witnessing the 'rise and demand' for these foods. *Actinidia arguta* 'kiwiberries' has joined and is competing within this food category. Our current understanding of the nutritional benefits of kiwiberries comes from a

variety of overseas cultivars, with limited resources and knowledge on New Zealand kiwiberries. Due to the current information of health benefits of the *Actinidia* genus, there are growing interests into understanding the benefits these fruits provide to humans and the human colon.

Pectin, a major non-cellulosic component of the plant cell wall, comprises of approximately one-third of the cell wall dry weight (Caffall & Mohnen, 2009). The middle lamella of the plant cells is covered with pectin layers. These layers are not digested by human enzymes, but by microbial degradation (Cheng *et al*., 1979). Despite the abundant ingestion of pectin in modern diets, very few pectin-degrading bacteria or specialised pectinolytic bacteria have been found in the human gut. The discovery of *Monoglobus pectinilyticus* (Kim *et al*., 2017), a pectinolytic bacterium, was found to play a fundamental role in pectin fermentation. Its genome encodes for a highly specialised and extensive glycobiome which is enriched with carbohydrate-active enzymes (CAZymes), specifically target various glycosidic bonds in pectin (Kim *et al*., 2019).

Proteomic analysis revealed that the presence of pectin in the growth medium up-regulates the extracellular production of pectin-degrading enzymes attached to the bacterial cell surface via S-layer homology (SLH) protein domains (Kim *et al*., 2019). *M. pectinilyticus* is unique, in a sense that it is the first pectin-specialist bacterium to be found, and its contribution to the human gut health is perhaps closely linked to the typically high pectin contents in human diet (Kim *et al*., 2017).

Interestingly, whilst *M. pectinilyticus* does not appear to utilise D-galactose and β-1, 4-galactan for its own metabolic needs, it is still capable of degrading β-1, 4-galactan into D-galactose monomers, leading to a D-galactose accumulation in the growth medium over time (Kim *et al*., 2019). The degradation of β-1, 4-galactan into D-galactose requires the enzyme β-galactanase. This prompted extensive database searches for genes encoding potential β-galactanase within the near-complete genome of *M. pectinilyticus,* but no such sequences have been identified. This suggests that *M. pectinilyticus* may produce an unknown class of β-galactanase (also referred to as β-galactosidase), which has not been identified and/or characterised before (Kim *et al*., 2019).

1.5 Aims and objectives

This project is based on the previous evidence of the vital role of *M. pectinilyticus* in the degradation of pectin. Although genomic studies clearly show the involvement by the numerous pectin-degrading CAZymes, many enzymes have not been identified, or activity characterised. Thus, the metabolic breakdown of pectin by *M. pectinilyticus* is not fully understood. Exotic kiwiberry (*Actinidia arguta*) is being commercially developed in New Zealand, gaining popularity among growers and consumers due to the fruit's rich, super food composition. However, little knowledge is known about the pectin composition and structure of these New Zealand cultivars, 'Hortgem Tahi', 'Takaka green' and 'Marju Red', and their potential health characteristics.

To give important insight to these questions, an investigation into the pectin structure of New Zealand kiwiberry cultivars and into the degradation of their pectin by CAZymes from the gut bacterium *M. pectinilyticus* is needed. The following objectives were addressed experimentally:

1. Extraction and characterisation of three kiwiberry cultivars in New Zealand: Hortgem Tahi (K2D4), Takaka Green and Marju Red.
2. Identification, isolation and characterisation of the β-1, 4-galactan degrading ability of a potentially new class of β-galactanase/β-galactosidase, from the anaerobic Gram-positive bacterium *M. pectinilyticus.*

Chapter 2 Materials and Methods

2.1 Chemicals

All reagents used in this project were of technical and/or analytical grade and were purchased from Sigma-Aldrich (Auckland, New Zealand), Merck Millipore (Auckland, New Zealand) and Thermo Fisher Scientific (Auckland, New Zealand).

All enzymes used were of high purity ordered from Sigma-Aldrich (Auckland, New Zealand), Megazyme Ltd (Bray, Ireland) and Food Tech Solutions Ltd (Auckland, New Zealand).

2.2 Media

2.2.1 Sterilisation

Media and required solutions were sterilised by autoclaving at 121 ˚C for 20 minutes followed by fast cooling using a Thermo Fisher Scientific™ front-loading autoclave. Solutions that have undergone sterilisation are referred to as “sterile” in this thesis.

2.2.2 Media and additives

Unless specified, all anaerobic media and additive were prepared using strict anaerobic technique (McSweeney *et al*., 2005). Final preparations of such media or additives were maintained under an anaerobic environment in either, 16 x 125 mm Hungate tubes sealed with an open top screw caps with rubber stops (Bellco) or 50 mL or 1 L glass serum vials plugged with butyl rubber stoppers and aluminium seals (Supelco).

Clarified rumen fluid was supplied by Caroline Kim and was prepared from hay-fed fistulated cows by AgResearch Ltd (Grasslands Research Centre, Palmerston North). Rumen fluid was bubbled with N_2 for 20 min and dispensed into 50 mL serum vials, sealed with rubber stoppers and aluminium caps. Rumen fluid was autoclaved and poured into a beaker and mixed thoroughly with 1.63 g $MgCl_2.6H_2O$ and 1.18 g $CaCl_2.2H_2O$ per 100 mL, by stirring for 30 min. Formed precipitate was removed by centrifugation at 6,000 x g for 20 minutes at 4 ˚C. The supernatant was collected, bubbled through with N_2 for 20 minutes and filter-sterilised through a 0.22 µm filter and sealed in serum vials with rubber stoppers and

aluminium caps. Vitamin solution 2% (v/v) was anaerobically prepared and added by flushed N_2-washed syringes and needles.

The vitamin solution was provided by Dr Caroline Kim. This was prepared from RPMI-1640 Vitamin Solution (R 7256) supplied by Sigma-Aldrich, which contains 0.02 g/L D-biotin, 0.3 g/L choline chloride, 0.1 g/L folic acid, 3.5 g/L myo-inositol, 0.1 g/L niacinamide, 0.1 g/L *p*-aminobenzoic acid, 0.025 g/L D-pantothenic acid·½Ca, 0.1 g/L pyridoxine·HCl, 0.02 g/L riboflavin, 0.1 g/L thiamine·HCl; 0.0005 g/L vitamin B12, 0.2 g/L KCl, 0.2 g/L KH_2PO_4 (anhyd), 8 g/L NaCl, and 1.15 g/L Na_2HPO_4 (anhyd). Vitamin solution was bubbled with N_2 for 10 minutes, and filter sterilised (0.22 µm) into sterile and N_2 washed serum vials and frozen at -20 ˚C until use.

Citrus and/or apple pectin (0.5% (w/v; Sigma-Aldrich) and potato galactan (0.5% (w/v; Megazyme) were prepared by weighing required amounts and added to mineral media under N_2 flushing.

Distilled water (900 mL) was collected in Schott bottles (Marienfeld) and autoclaved. For anaerobic water, 900 mL of distilled water was boiled and cooled on ice, under a continuous flow of N_2. L-cysteine-HCl (0.5 g) was added. Required volumes were dispensed into serum vials and sealed using caps and rubber stops. Anaerobic water was sterilised by autoclaving.

2.2.3 Mineral Medium (MM)

Routine cultivation of *M. pectinilyticus* was carried out in Hungate tubes or glass serum vials. Mineral medium per litre consisted of 1.4 g KH_2PO_4, 0.6 g $(NH_4)_2SO_4$, 1.5 g KCl, 1.0 g yeast extract, 0.02 g resazurin, 4.2 g $NaHCO_3$ and 0.5 g L-cysteine-HCl. Both $NaHCO_3$ and L-cysteine-HCl were added after boiling and cooling the media under 100% CO_2. Set volumes were dispensed into appropriate vessels under CO_2 flushing. Sealed vessels were sterilised by autoclaving. Substrates and additives were added to the medium prior to inoculation.

2.2.4 Luria-Bertani (LB) broth and agar medium

LB broth powder (20 g; Sigma) was mixed and dissolved in 950 mL of distilled water. LB broth was sterilised by autoclaving. Sterile media were cooled to 50 ˚C before adding any required additives or antibiotics. LB agar medium was prepared by the addition of 15 g of agar per litre of the medium before sterilisation. Autoclaved agar medium containing antibiotics was then poured into Petri dishes and left to set, before storing at 4 ˚C.

2.2.5 Super optimal broth (SOC)

Tryptone (20 g), yeast extract (5 g), NaCl (0.5 g) and 10 mL of 250 mM KCl were added to 950 mL of distilled water. The broth pH was adjusted to 7.0 with NaOH before sterilisation by autoclaving. Before use, 5 mL of 2 M $MgCl_2$ and 20 mL of 1 M glucose were added. Media was stored at 4 ˚C.

2.2.6 Anaerobic glycerol solution for long-term preservation of bacteria

Glycerol was mixed with distilled water (50% (v/v)), boiled and then cooled on ice under N_2. L-cysteine HCl (0.5 g) was added per litre and was dispensed into anaerobic glass vessels (e.g. Hungate tubes) under N_2 and sealed using caps and rubber stoppers. Solution was sterilised by autoclaving. Exponentially growing bacterial cultures were added to the glycerol solution at a 1:1 ratio. The mixture was stored at -80 ˚C.

2.3 Materials and methods for kiwiberry pectin isolation and characterisation

2.3.1 Collection of kiwiberries

Kiwiberry ‘Hortgem Tahi’ fruit were collected from a commercial orchard located in Te Puke, New Zealand through Seeka Limited. The fruit were stored at 0.5 °C before being

transported to Palmerston North. 'Takaka Green' and 'Marju Red' fruits were harvested from commercial orchards located in Te Puke and were packed by Kiwi Produce Limited. All fruits collected were picked during harvest maturity (>90% black seeds), with a brix score of 6.2 or higher.

2.3.2 Kiwiberry processing and firmness testing

Upon arrival, discoloured and damaged fruit were removed and the remaining berries divided into two groups. Group One was labelled harvest ripeness and was processed by cutting into quarters and were immediately frozen into liquid nitrogen and stored at -80°C. Group Two kiwiberries were transferred to a post-harvest room for 4 days at 20°C (Sutherland, *et al*., 2017). Subsequently, all post-harvest fruits were processed identically to Group One.

For compression firmness testing, a Texture Analyser TA.XT plusC (Stable Micro Systems, USA) was used according to manufacturer's instructions. Firmness measurement was performed using a 100 mm aluminum compression platen (P/100) at a speed of 1 mm/s. The same five kiwiberries were tested each day over 4 days at 20 ˚C. The compression test was performed on three random sides of the kiwiberry, with the average taken.

2.3.3 Buffered phenol-ethanol insoluble method (BP-EIS)

Frozen whole kiwiberries were ground in liquid nitrogen using a blender (MultiBender™ Pro, Sunbeam). Ground frozen tissue (11 g) was weighed in 50 mL tubes (Eppendorf, Germany). Tubes were stored at -80 ˚C until further use.

Before extracting cell walls, 30 mL of absolute ethanol was added to the frozen tissue, mixed and left at 1 ˚C overnight. Tissues were centrifuged (Eppendorf centrifuge 5920 R, Germany) at 12, 000 x *g* for 20 minutes. Supernatants were discarded. Buffered phenol was made by adding 25 mL 100mM Tris-HCl, pH 7.5 to 50 mL melted phenol, mixed via magnetic stirrer and left to settle at 4 ˚C overnight. The upper aqueous layer was removed before use.

Preparation of EIS was according to O'Donoghue *et al*, 2009, using Tris-buffered phenol (Huber, 1991) as prepared above. Buffered phenol (10 mL) was added to the kiwiberry tissue, mixed and placed on ice for 30 minutes. Absolute ethanol (40 mL) was added, and the mixture vortexed and left to precipitate before centrifugation at 4,000 x *g* for 10 minutes (Eppendorf centrifuge 5920 R, Germany). Supernatant was discarded. The pellet was washed 3x with 25 mL of 80% ethanol, centrifuged at 4,000 x *g* for 10 minutes and the supernatant discarded. This was followed by 2x 25 mL acetone washes. A 25 mL 1:1 chloroform: methanol wash was left at room temperature for 30 minutes, followed by centrifugation and discarding of the supernatant. Lastly, after 3x washes with 30 mL acetone, the sample was centrifuged and the supernatant discarded. Tubes were then left to dry then transferred to a heat cupboard for further drying before weighing. Samples were stored at -20 ˚C until further use.

2.3.4 Pectin fractionation of BP-EIS

The following solutions were prepared and used accordingly to O'Donoghue *et al*, 2009.

CDTA solution

The CDTA solution was made of *trans*-1, 2-diaminocyclohexane-*N*,*N*,*N*',*N*'-tetraacetic acid monohydrate (3.463 g; CDTA) and 0.8204 g; 50 mM NaOAc was dissolved in 150 mL of water. The pH was adjusted to 6.5 using 2 M KOH, allowing CDTA to completely dissolve, then readjusted to a pH of 6. The final volume of the solution was made up to 200 mL with distilled water.

Sodium Carbonate (Na_2CO_3) solution

To 200 mL of distilled water, 1.0599 g Na_2CO_3 and 0.1510 g $NaBH_4$ were dissolved to give respectively 50 mM and 20 mM concentrations. This solution was made fresh each time.

Dried BP-EIS (100 mg) was sequentially extracted with of 10 mL of d-H_2O, CDTA (in 50 mM NaOH, pH 6.5) and 50 mM Na_2CO_3 (containing 20 mM $NaBH_4$). All extractions were made at room temperature for 16 hours at constant shaking. Extracts were centrifuged with supernatant retained and stored at -20 ˚C

2.3.5 Total pectin content of cell walls

Dried BP-EIS (5 mg) was extracted in 50 mL boiling test tubes. Chilled concentrated H_2SO_4 (2.5 mL) was added and stirred for 5 min. Sterile water (1 mL) was added very slowly and allowed to stir for a further 2.5 min. Water addition was repeated 4x and left to mix for 1 hour. The sample was poured into a measuring cylinder with sample topped up with sterile water to 10 mL, and then, transferred into a 15 ml tube. Aliquots (1.5 mL) was taken for uronic acid analysis, with stock stored at -20 ˚C.

2.3.6 Uronic acid microplate assay

Acid reagent

Di-sodium tetraborate (4.767g $Na_2B_4O_7.10H_2O$) was dissolved in 1 L of concentrated sulphuric acid (H_2SO_4) in a fume hood. The solution was mixed overnight and stored in a freezer.

Di-phenyl reagent

Reagent was made from sodium hydroxide (0.5 g NaOH) and 3-phenylphenol (50 mg $C_{12}H_{10}O$) was weighed into a 100 mL bottle containing 100 mL sterile water. Solution was kept at room temperature.

Galacturonic acid standards

Stock solution of 200 µg/mL GalA was prepared by dissolving 1.92 mg into 50 mL sterile water. Standards ranging from 200 µg/mL to 0 µg/mL were made and stored at room temperature.

All pectin fraction samples and dilutions were assayed for determining the amount of uronic acids present. The method based on Blumenkranktz and Asboe-Hansen (1973), modified to 96 well-microplate by O'Donoghue (2007) was used. Pectin fractions (40 µL) were added to a 96-well plate and 200 µL aliquots of acid reagent were dispensed rapidly to mix the sample cohesively with the acid. The plate was covered and incubated at 95 ˚C for 1 hour. A pre-read was performed at 520 nm. The plate was then cooled on ice for 20 minutes, before the addition of 10 µL di-phenyl reagent, after which the plate was sealed, vigorously mixed and read after 5 minutes. A standard of galacturonic acid was included, and the concentration of uronic acid from the pectin fractions was calculated using the standard. To calculate the concentration as pectin, the concentration of galacturonic acid determined is multiplied by (176.1/194.1 = 0.9072).

2.3.7 Glycan profiling using microarrays

Phosphate buffered saline (PBS)

The 1 L of 10x PBS was made from 1.78 g Na_2HPO_4, 2.40 g KH_2PO_4, 80 g NaCl and 2 g KCl to give concentrations of, respectively, 100 mM, 18 mM, 1.37 M and 27 mM. Buffer was sterilised by autoclaving and cooled before use and was kept at room temperature.

Skim milk phosphate buffered saline (MPBS)

Non-fat skimmed milk powder (5% w/v) was dissolved in 1 x PBS and kept chilled until use.

Monoclonal antibodies

A selection of antibodies supplied by Plant Probes (Leeds, UK) was used to identify specific for cell-wall glycan epitopes (Table 2.1). Antibodies used will be referred to by their antibody name.

Printing microarrays

Water-, CDTA- and Na_2CO_3-soluble extracts stock solutions were used undiluted and as 1:5, 1:10 and 1:20 dilutions. The plate layout and robot parameters (such as pin height, collection and dwell time) were designed by Erin O'Donoghue. The robot printed each dilution for each extract onto nitrocellulose membrane (0.2 µm pore, Bio-Rad). Approximately, 1 nL was used for each printed spot. Each sample was spotted in triplicate. This resulted in a grid of spots from each extract of each cultivar on a membrane. Eight separate membranes were used for glycan array analyses.

Table 2.1 Antibodies used to detect specific cell wall glycan epitopes.

Antibody	Recognition	References
LM5	Region of linear (1,4)-linked galactose, at least 4 residues	Jones *et al*., 1997
LM6	Region of (1, 5)-linked arabinose, backbone and/or branches	Verhertbruggen *et al*., 2009
LM8	Xylogalacturonan, recognises regions highly substituted with xylose.	Willats *et al*., 2004; Mort *et al.,* 2008.
LM13	Long stretch of (1, 5)-linked arabinose, no branches	Verhertbruggen *et al*., 2009
LM18	Partially methyl esterified pectin. Low affinity for low esterified pectin < ~45% including PGA	Verhertbruggen *et al*., 2009
LM19	Partially methyl esterified pectin. High affinity for PGA but also, can bind to pectin up to 66% esterified (block wise)	Verhertbruggen *et al*., 2009
LM20	Highly esterified pectin. Does not bind to PGA but can bind to 18% DE pectin.	Verhertbruggen *et al*., 2009

Probing of glycan microarrays

In order to develop the microarrays, each membrane was blocked with 5 mL of MPBS overnight at room temperature with shaking. The microarray was washed twice with 5 mL 1x PBS, for 10 minutes each. Each microarray was then individually probed with 5 mL of a single monoclonal antibody diluted 1/200 with MPBS (Table 2.1) for 2 hours. Each array was washed 3x with PBS for 5 minutes to remove any species binding non-specifically then treated within anti-mouse antibody conjugated with alkaline phosphate in MPBS for 2 hours. Membranes were washed a minimum of 5x with PBS for 5 minutes. To develop the array, chromogenic substances (NBT/BCIP at a dilution of 1:50) overlaid the membranes and

gently rocked until spots developed a purple colour (times varied for each primary antibody).

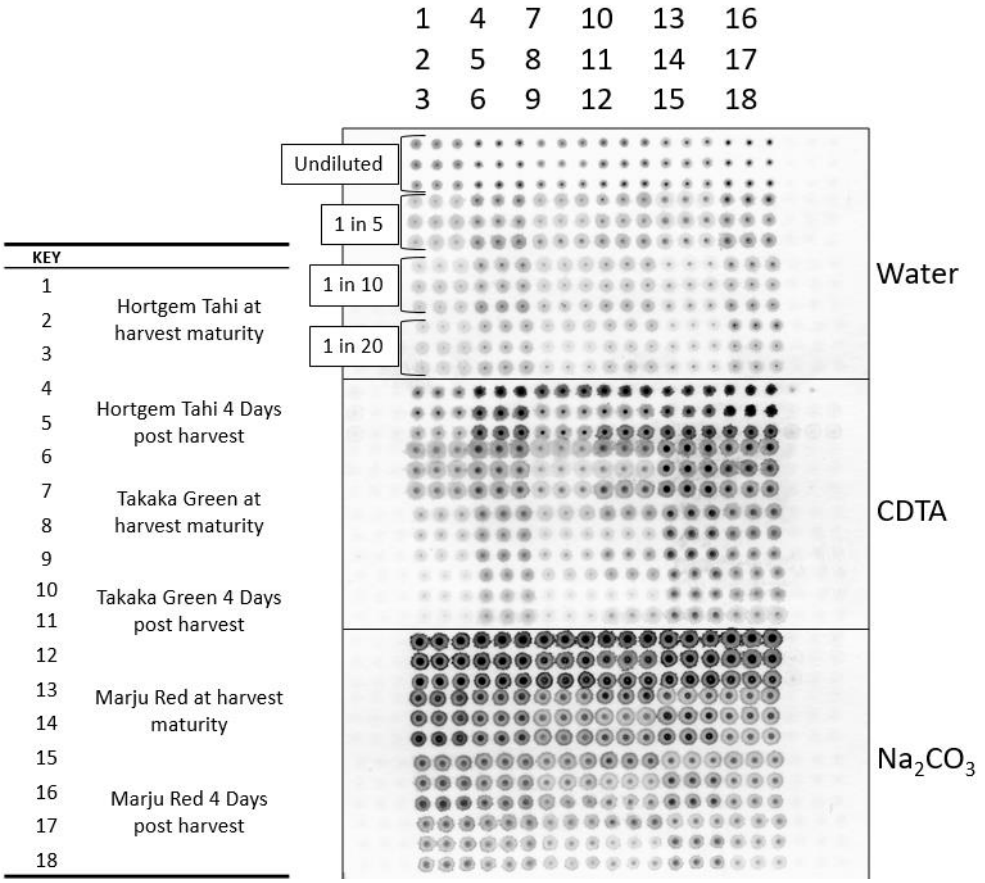


Figure 2.1 Layout of each nitrocellulose membrane for glycan array. Each membrane contained all soluble extracted pectin in order of isolation – water, CDTA and Na_2CO_3.

Quantification of microarrays

Each microarray was scanned using a high-resolution desktop scanner and the density of the colour for each spot was estimated using Xplore Image Processing Software (LabNEXT). This integral spot intensity is derived from the sum of pixels in the grid area surrounding each spot densities and were averaged over each triplicate spot. Original extractions were made from BP-EIS, therefore the density of each spot was converted to units on a fresh weight basis for the different recoveries of BP-EIS from the fresh tissues of each cultivar. The trends for each dilution of each extracts were checked but only the 1/5 dilutions responses were used for preparation of graphs.

2.4 Materials and methods for the isolation and identification of protein(s) associated with pectin degradation

2.4.1 Microbiology

Monoglobus pectinilyticus was grown under anaerobic conditions as previously described (Kim, et al., 2017). Final media made were maintained under anaerobic conditions in either 1 L glass serum vials, 50 mL serum vials (Supelco), which can be plugged with a butyl rubber stopper and crimped with aluminum seals (Supelco) or 125 mm Hungate tubes (Bellco) sealed with rubber stoppers and open top screw caps (Bellco). The liquid medium was then sterilised by autoclaving at 121 ˚C for 20 minutes (Tuttnauer Autoclave 5075 EL, Netherlands). Additional substrates/additives were sterilised and added to the media by inserting an N2-flushed syringe and needle through the stopper.

2.4.2 Gram staining

Gram staining was carried out by using the method described by Bartholomew & Mittwer (1952). All required solutions were provided by the CMN laboratory (Plant and Food Research, Palmerston North, New Zealand). A small volume of bacterial suspension was spread onto a glass microscope slide and allowed to air dry under a Bunsen burner. The suspension was then heat fixed by passing through a flame. The slide was flooded with crystal violent solution for 1 minute, rinsed gently with water, followed by Gram iodine for 1 minute and washed with water. Cells were briefly washed with decolourising agent and drained before safranin solution was used to counterstain for 30 secs. The slide was washed and drained using a paper towel. The cells were observed at 1,000 x power magnification with oil immersion using Olympus CX21 microscope.

2.4.3 Crude enzyme extraction

In order to filter and concentrate enough bacteria grown in MM, OD_{600} of samples was recorded daily with the addition of Gram staining to validate bacteria growth and to check for any contamination. Media with an OD_{600} between 0.3 – 0.4 and not contaminated were

subjected to ultra-filtration through 0.1 μm pore size (Sartorius Sartocon 200 PESU, Germany). Retentate was kept at 4 ˚C and permeate was further concentrated through a 50 kDa pore. Both the retentate and permeate were collected. The 50 kDa retentate was pipetted into a 100 kDa amplicon ultra-15 centrifugal tube (Amicon® Ultra-15 Centrifugal Filter Units, Merck, Germany) and spun at 12,000 x *g* for 20 minutes. Again, both retentate and permeate was collected. This process is repeated until no 50 kDa retentate was left. Protein Inhibitor Cocktail (P8465, Sigma-Aldrich, Germany) was added to each solution to stop protein degradation. Retentate and permeate of all molecular weight cut-offs were stored at 4 ˚C until further use.

2.4.4 Protein concentration

Protein concentration was measured by the Bradford assay method (Bradford, M., 1976). Bradford assay, using reagent supplied by Sigma-Aldrich, was performed in a 96 well plate following manufacturer's instructions. Unknown protein (5 µL) was added, followed by 250 µL of Bradford reagent. Samples were diluted if required. Standards (Figure 2.2) were included using endo-β-galactanase (E-GALCT, Megazyme, Ireland) from a range of 0.0 – 2.5 mg/mL. After an incubation time of 30 minutes at room temperature, the absorbance of samples was measured at 595 nm. Protein concentration was calculated from the standard curve included in every Bradford assay.

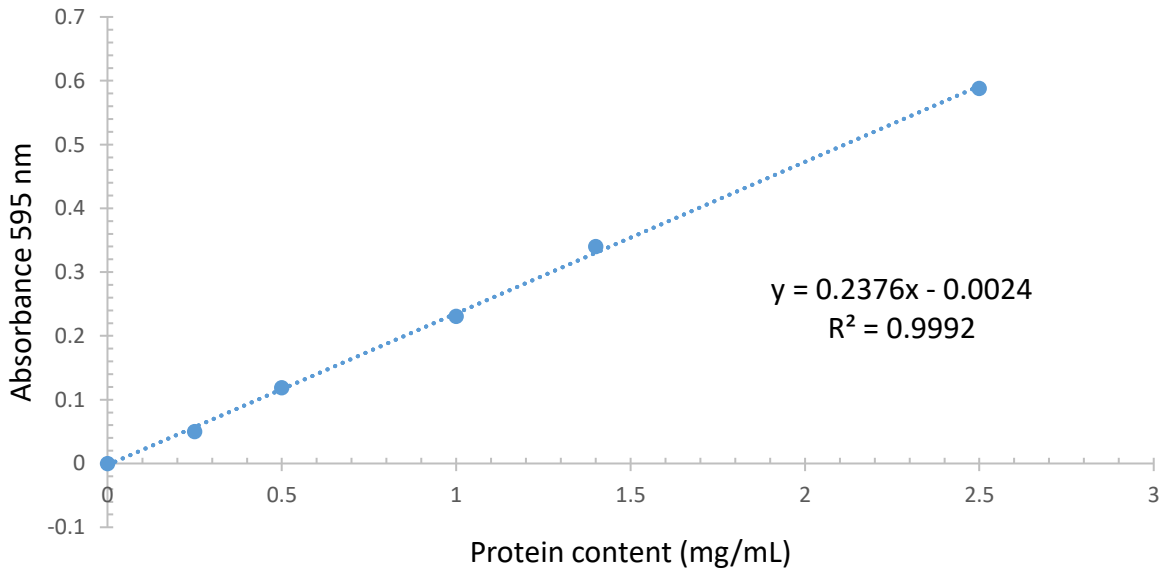


Figure 2.2 Protein standard curve. The absorbance vs mg/mL protein was prepared using endo-β-galactanase (E-GALCT, Megazyme, Ireland) from a range of 0.0 – 2.5 mg/mL

2.4.5 Enzyme identification

To identify any β-galactosidase and/or any β-galactanase activity of the enzyme(s), two assays were used.

Ortho-nitro phenyl-β-D-galactopyranoside (ONPG)

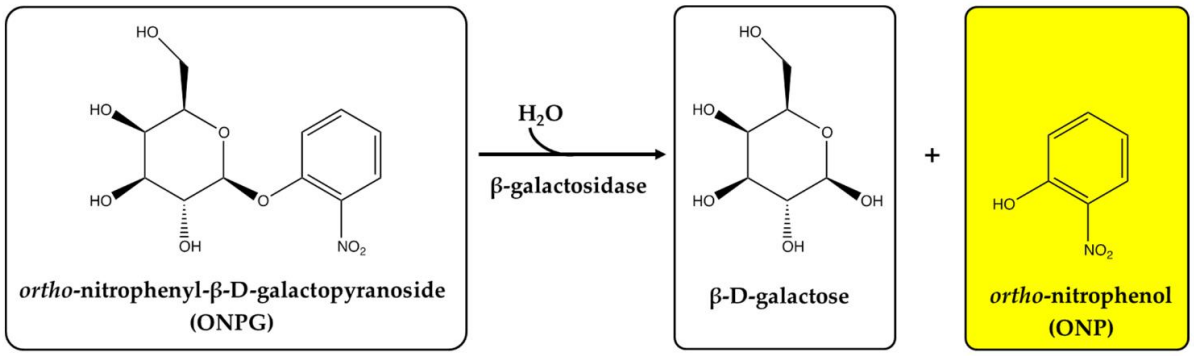


Figure 2.3 Enzymatic reaction of ONPG. The β-galactosidase hydrolyses ONPG to produce β-D-galactose and o-nitrophenol (which provides the yellow colour), allowing β-galactosidase activity to be measured by absorbance at 420 nm, Labus, K. (2018). No copyright permissions needed.

The ONPG test adapted for a microplate reader was routinely used (O'Donoghue, 2009). A 96 well microplate incorporates 15 µL of crude protein sample or 15 µL of suspected anion exchange fraction, 15 µL 200 mM NaOAc, pH 7.5 and 15 µL 5 mM ONPG. The plate is then shaken and incubated for 30 minutes, after which 45 µL of 200 mM Na_2CO_3 was added to stop the reaction and the absorbance was read at 420 nm. Nitro-phenol (Sigma-Aldrich) standards were made from 0 – 2.5 mM. Positive, negative and additional controls were used. Concentrations of nitrophenol were calculated by the standard curves included in each ONPG assay (Figure 2.4).

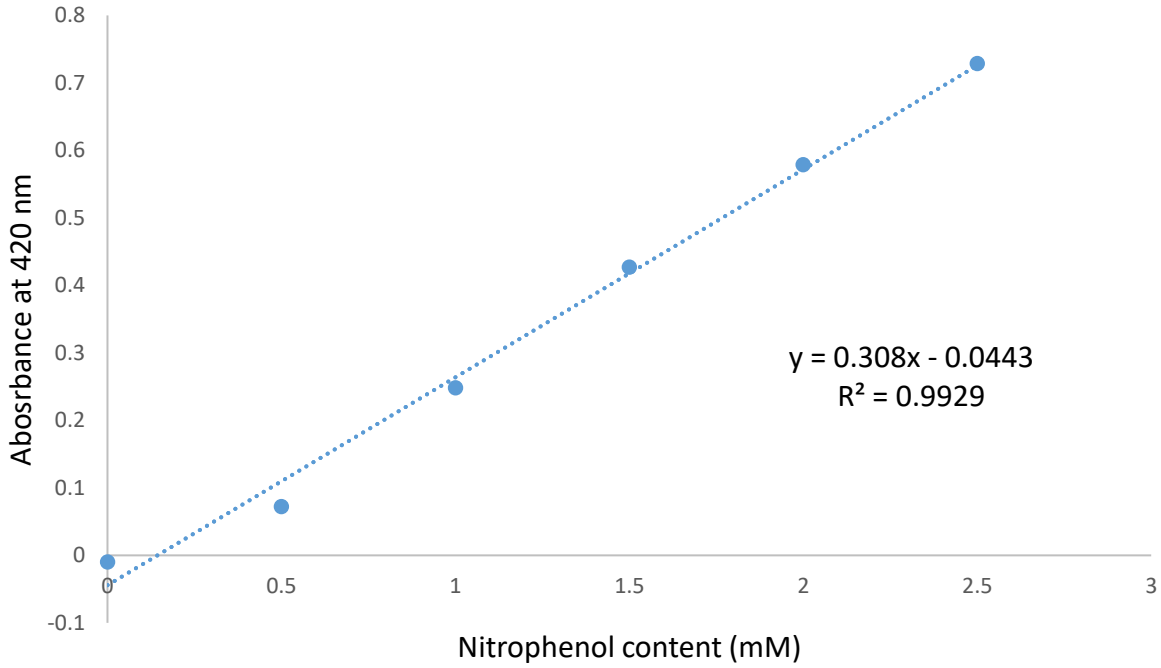


Figure 2.4 Standard curve of absorbance vs nitrophenol content (mM). The nitrophenol content (mM) is measured against the standard curve.

Enzyme activity (µmol product formed per min) was calculated using the standard curve for each assay at 21 ˚C. Specific enzyme activity was calculated as enzyme activity (µmol min^{-1})/protein content (mg).

Modified L-arabinose/D-galactose Assay

A modified L-arabinose/D-galactose assay kit (K-ARGA, Megazyme, Ireland) provides information on exo-galactanase activity, using potato galactan. The assay followed manufacturer's procedure with variations (Figure 2.5). The substrate sample solution used was potato galactan (1% w/v), enzyme suspension was the crude or pure enzyme extracted and β-galactose dehydrogenase was the final enzyme solution added. Blanks and standards were not altered from the kit. Positive and negative controls were incorporated.

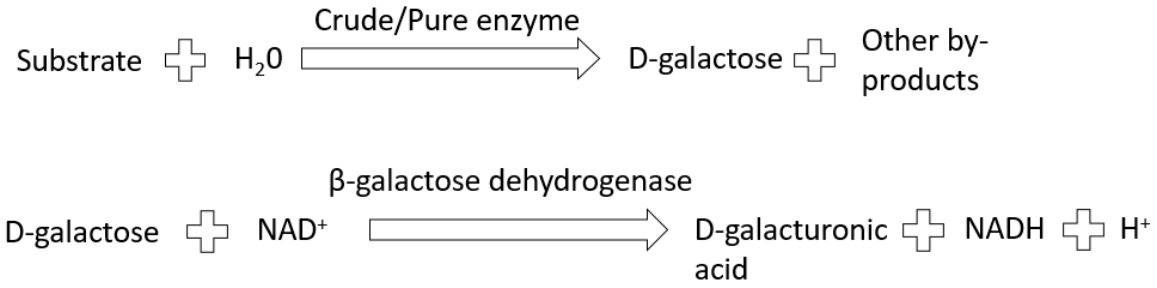


Figure 2.5 Modified D-galactose assay enzyme reaction. The increase in NADH is measured by absorbance at, 340 nm.

Before assaying, testing substrate and unknown enzyme were incubated at 1:1 ratio for 1 hour at room temperature. To this, solution 1: Buffer pH 8.6 (20 µL), solution 2: NAD^+ (10 µL) and distilled water (200 µL) was added. The plate was mixed and left for 3 minutes before reading absorbance at 340 nm (A_1), after which 2 µL of solution 3: β-galactose dehydrogenase is added. The plate was remixed and left for 6 minutes before a final absorbance (A_2) was read. Blank controls and standards were run with each plate. To determine L-arabinose, the same method was applied, with an extended incubation after the addition of solution 3, from 6 minutes to 30 minutes.

The amount of NADH formed in Figure 2.4 is stoichiometric with the amount of L-arabinose or D-galactose released. It is the NADH which is measured by the increase in absorbance at 340 nm. The calculation for the amount of L-arabinose or D-galactose follows $c = \left(\frac{V\ x\ MW}{\varepsilon\ x\ d\ x\ v}\right) x\ \Delta A$. Where: ΔA is the difference between $A_2 - A_1$; V: final volume of curvette; MW: molecular weight of L-arabinose or D-galactose (g/mol); ε: extinction coefficient of NADH at 340 nm, 6300 (l x mol^{-1} x cm^{-1}); d: light path (1 cm) and lastly, v: sample volume (mL). Concentration is in units of g/L. For any dilutions incorporated, results must be multiplied by the dilution factor.

2.4.6 Protein purification

Crude filtered fractions were subjected to ammonium sulfate precipitation (Nakano *et al*., 1990). Fractions were subjected to 50% and 80% saturation of solid ammonium sulfate. After 16 hours, samples were centrifuged at 22,000 x *g* with the pellet collected and dissolved in two buffers: (1) 20 mM acetate, pH 6.0 and dialysed against the same buffer (Nakano *et al*., 1990) and (2) 100 mM Tris-HCl, pH 8.0. An aliquot of each sample and ammonium sulfate fraction was kept for SDS-PAGE visualisation. Samples were kept at 4 ˚C until further use.

Anion exchange chromatography (AIX) was performed to further purify the samples showing BGAL enzyme activity. GE AKTA Explorer (Amersham Pharmacia, UK) system containing a Mono Q 5/5 anion exchange column was chosen (Sigma-Aldrich, USA). The column is prepacked with polystyrene/divinyl benzene matrix, with bead particle size of 10 µm and a bed volume of 1 mL, allowing an average of 50 mg protein load. Injections of sample (5 x 1 mL) were loaded onto the column at a flow rate of 0.3 mL/min, maintaining a pressure of 1.40 MPa. The AKTA explorer system (Amersham Pharmacia, UK) and the column (Sigma-Aldrich, USA) were pumped with start buffer of 100 mM Tris-HCl, pH 8.0. The sample was injected once column reached equilibrium with protein being eluted on a routine 0 – 100% salt gradient using elution buffer of 100 mM Tris-HCl, pH 8.0 and 1 M NaCl. Samples were prepared in 100 mM Tris-HCl, pH 8.0. If eluted proteins were re-run through anion exchange, the salt gradient was altered from 0 – 100% gradient to a segmented salt gradient based on previous anion results. Eluted fractions (200 – 300 µL) were collected into 96 well microplates, and immediately taken for enzyme activity assay, SDS-PAGE and stored at 4 ˚C.

Samples subjected to prior ammonium sulfate treatment were loaded into the same anion exchange system as mentioned above. However, the Mono Q 5/5 column was equilibrated with the dialysis buffer used for the sample: (1) 20 mM Na-acetate buffer, pH 5.0) or (2) 100 mM Tris-HCl, pH 8.0. The bound proteins were then eluted following the above protocol, with every collected fraction assayed for enzyme activity and protein content (mg/mL). Plates were stored at 4 ˚C.

2.4.7 Protein visualisation

Precast NuPAGE (Invitrogen, USA) was used to identify any potential proteins or enzymes within the retentate or permeate. SDS-PAGE was performed using a 10% Bis-Tris 1.0 mm mini protein gel with 10 wells. Each well contained 5 µL of LDS, 5 µL of water and 10 µL of sample that were subjected to 90 ˚C incubation for 10 minutes. 1x MOPS buffer was used. PageRuler Unstained Broad Range was used as a protein ladder and Mark12 (Thermo Fisher, USA) or endo-1, 4-β-galactanase (Megazyme, Ireland) was used as protein controls. Gels were run at 200 V for 45 minutes in an XCell SureLock™ Mini-Cell (ThermoFisher, USA). For visualisation gels were stained with SimplyBlue Stain overnight, followed by washing of the gel with distilled water for 2 hours.

Table 2.2 Preparations of gels for SDS-PAGE.

Gel %	ddH_20	40% acrylamide	1.5 M Tris-HCl, pH 8.8	0.5M Tris-HCl, pH 6.8	10% SDS	10% APS[1]	TEMED[2]
Resolving 10%	3.80 mL	4.30 mL	2.60 mL	N/A	0.10 mL	0.10 mL	0.01 mL
Stacking 5%	5.86 mL	1.34 mL	N/A	2.60 mL	0.10 mL	0.10 mL	0.01 mL

Note: [1]- Ammonium Persulfate (APS), was made fresh each time. [2]- Tetramethylethylenediamine (TEMED) was added just before pouring the gel.

Handmade SDS-PAGE acrylamide/bis-acrylamide gels were also used for identifying any filtered and crudely purified protein (Table 2.2). Gels were casted with a 10% resolving layer and a 5% stacking layer. SureCast™ Gel Handcast (Invitrogen, USA) hardware was used for casting, following manufacturer's protocol. The gels were run at 180 V for 60 minutes. Overnight SimplyBlue Stain was used for gel visualisation and analysed after a 2-hour water wash. All gels were photographed by Amersham Imager 600 (GE Healthcare, UK).

2.4.8 Protein sequencing

Sample preparation for protein sequencing by service providers, Proteomics International Laboratories Ltd, was as follows. Bands were cut out from SDS-PAGE, air dried in Eppendorf tubes and dispatched. Protein samples were trypsin digested and peptides extracted according to standard techniques (Bringans *et al*., 2008). Peptides were analysed by electrospray ionisation mass spectrometry using the Shimadzu Prominence nano HPLC system (Shimadzu) coupled to a 5600 TripleTOF mass spectrometer (Sciex). Tryptic peptides were loaded onto an Agilent Zorbax 300SB-C18, 3.5 μm (Agilent Technologies) and separated with a linear gradient of water/acetonitrile/0.1% formic acid (v/v). Spectra were analysed to identify proteins of interest using Mascot sequence matching software (Matrix Science) with SwissProt and UniProt database. (Bringans *et al*., 2008).

2.5 Material and methods for the cloning and attempted characterisation of candidate protein(s)

2.5.1 Genomic DNA extraction for amplification of candidate protein(s)

TES Buffer (pH 8.0)

TES buffer (0.254 g Tris-HCl, 0.048 g Tris, 0.116 g NaCl and 0.068 g EDTA dissolved into 400 mL of distilled water) was sterilised by autoclaving.

Lysozyme solution

Lysozyme (200 mg) from chicken egg white (≥ 40,000 U/mg, Sigma) was added to 10 mL of autoclaved water and dispensed into 1.5 mL micro centrifuge tubes in 250 µL aliquots, and kept at -20 ˚C until use.

Proteinase K solution

Proteinase K (200 mg) from *Tritirachium album* (41 U/mg, Sigma) was dissolved into 10 mL of autoclaved water and dispensed in 250 µL aliquots into 1.5 mL micro centrifuge tubes and stored at -20 ˚C until use.

RNase solution

Ribonuclease A (200 mg) from bovine pancreas (≥ 50 Kunitz U/mg, Sigma) was dissolved into 10 mL of autoclaved water and dispensed into 250 µL aliquots of the solution and kept at −20˚C until further use.

TNS solution (pH 8.0)

TNS solution (500 mM Tris-HCl, 100 mM NaCl and 10% (w/v) SDS dissolved in distilled water) was autoclaved.

PEG solution

PEG solution was prepared by dissolving 0.92 g NaCl in 10 mL (1.6 M) and polyethylene glycol 6000 was added to make a 30% (w/v) solution that was autoclaved.

Elution buffer (EB)

EB buffer was prepared in 10 mL distilled water with 0.012 g; 10 mM Tris. The pH was adjusted with HCl to 8.5. Buffer was sterilised by autoclaving.

Sodium acetate solution

Sodium acetate was prepared in 10 mL distilled water with 2.46 g; 3 M NaOAc. The solution was adjusted pH to 5.2 with glacial acetic acid and was autoclaved.

Phenol-chloroform extraction of genomic DNA

Genomic DNA (gDNA) was extracted from frozen bacterial cell pellets of *M. pectinilyticus.* Cells were re-suspended in 0.5 mL of TES buffer and vortexed vigorously. Lysozyme solution (10 µL) was added, mixed and incubated in a heat block at 37 ˚C for 2 hours. Immediately after, samples were transferred to ice where 10 µL of proteinase K solution and 10 µL of RNase solution were added. The mixtures were vortexed and incubated in a water bath at 65 ˚C for 1 hour. TNS solution (100 µL) was added, and samples were further incubated at 65 ˚C for 30 min. Samples were transferred to a clean 2 mL Eppendorf tube, where 700 µL of phenol: chloroform: *iso*-amylalcohol (25:24:1, pH 8.0, Sigma) was added. The mixture was inverted gently then centrifuged at 20,000 x *g* for 20 minutes at 4 ˚C. A volume of 500 µL of the upper layer was carefully pipetted and transferred into a fresh 2 mL Eppendorf tube and 500 µL of chloroform: *iso*-amylalcohol (24:1, Sigma) was added. The sample was mixed by gentle inversion, followed by centrifugation at 20,000 x *g* for 5 minutes at 4 ˚C. A volume of 400 µL of the upper layer was carefully pipetted and transferred into a clean 2 mL Eppendorf tube. PEG solution (800 µL) was added to the sample and mixed by gentle inverting 5x. Precipitated DNA was centrifuged at 20,000 x *g* for 1 hour at 4 ˚C. Liquids were carefully discarded and 500 µL of ice cold 70% (v/v) ethanol was added to wash the pellet. Precipitates were centrifuged at 20,000 x *g* for 10 minutes at 4 ˚C. The supernatant was carefully poured out and the ethanol-wash step was repeated. Ethanol was discarded and tubes were kept inverted on a paper towel to air-dry the pellet. DNA precipitates were eluted in 30 µL EB buffer and left to dissolve in the fridge overnight. DNA purity and concentration were checked by QIAxpert System (Qiagen, Germany).

Ethanol precipitation of DNA

Extracted DNA samples were pooled with total volume recorded. A 1/10 of sample volume of sodium acetate solution was added to the DNA samples and mixed carefully. A 2.5 volume of ice cold 100% (v/v) ethanol was added and mixed. The mixture was stored at -80˚ C overnight. The mixture was centrifuged at 20,000 x *g* for 15 min at 4 ˚C and the supernatant decanted. The pellet was then washed with 1 mL of 70% (v/v) ice cold ethanol. Precipitated DNA was centrifuged at 20,000 x *g* for 10 min at 4 ˚C, with the ethanol carefully decanted. The ethanol-wash step was repeated. Tubes were kept inverted on a paper towel to air dry to the pellets. Pellets were then eluted in 20 µL EB buffer and left to dissolve over night at 4 ˚C.

DNA quantification

DNA quantification was carried out using the QIAxpert system, using the DNA QIAamp setting, which measures absorbance at 260 nm.

TE buffer

TE buffer (10 mM Tris-HCl and 1 mM EDTA mixed in distilled water with 2 M NaOH used to adjust the pH to 8.0) was sterilised by autoclaving. Primers from Integrated DNA Technologies (IDT®) were reconstituted in TE buffer to make up 100 µM stocks.

2.5.2 PCR amplification

PCR amplification of numerous gene regions (Table 2.3) of *M. pectinilyticus* was performed using 25 mM solutions of lyophilised primers purchased from IDT®.

Table 2.3 Primers created for gene cloning.

Primer	Sequence (5' -3')	T_m (˚C)	PCR product (kb)	Notes
FP_0050HisT_XhoI	CTCGAGCATCATCACCATCACCACTTGAAAAAATATATTTCGCTATTTAT	48	1323	His tagged
RP_0050HisT_EcoRI	GAATTCCATCATCACCATCACCACTTTAGATTTTTTTACTCGGT	47		Stop codon removed for His Tagged
FP_00398HisT_Xho1	CTCGAGATGCATCATCACCATCACCACATTATACGAA AAGTATTATCTGCAATCACCGC	60	6084	His Tagged
RP_00398_NheI	GCTAGCTTACTTAAGATTATCAATACTGCCTAAGTCATATATTTTTGTA	57		
FP_02288HisT_Xho1	CTCGAGCATCATCACCATCACCACATGAAAGAAAATGGAATAA AAAATAATGAT	52	268	His Tagged
RP_02288_EcoRI	GAATTCACTTTATATTATTATTTTTTATATACGTATTATACTCGGGTTTCC	54		
FP_01597HisT_Xho1	CTCGAGATGCATCATCACCATCACCACAGAACCTTTAAGAGAAGTTTGGCC	58	9555	His Tagged
RP_01597_EcoRI	GAATTCTTAAATAATTGAATTAACAAAAGTTGCATTCACTGTAACG	59		

Note*:* FP: Forward primer; RP: Reverse primer; HisT: His-Tagged.

KOD Hot Start Master Mix (Sigma-Aldrich) was used for long target templates (up to 12 kbp), with high fidelity and creates blunt-ended PCR products. A 50 µL reaction consists of 25 µL KOD HotStart Master Mix (0.04 U/µL), 1.5 µL each of forward and reverse primer (10 µM), 20 µL PCR-grade water and 2 µL template DNA. Optimised PCR cycling conditions recommended by the manufacturer were used: polymerase activation: 95 ˚C for 2 minutes; denaturation of 95 ˚C for 20 seconds; annealing of primer at chosen primer lowest Tm ˚C for 20 seconds; lastly, extension at 70 ˚C for 20 seconds per kb size of target gene; and storage at 4 ˚C until further processing. Denaturation and annealing steps repeated for 20 - 40 cycles. KOD requires primers to have a T_m of over 60 ˚C.

Phusion High-Fidelity DNA polymerase (Thermo Scientific) was used for products up to 7.5 kb and has high fidelity and generates blunt-ended PCR products. A 50 µL reaction consists of: 25µL of 2x Phusion Master Mix, 2.5 µL each of 10 µM forward and reverse primer, 2 µL template DNA and 18 µL PCR-grade water. Optimised PCR cycle conditions recommended by manufacturer were used: initial denaturation: 98 ˚C for 30 seconds; denaturation: 98 ˚C for

10 seconds; annealing: lowest primer T_m ˚C for 20 seconds; extension at 72 ˚C for 20 seconds per kb of target gene. Final extension at 72 ˚C for 5 minutes and storage at 4 ˚C until further processing. Denaturation, annealing and extension was cycled 30x, before the final extension.

PCR was run using Applied Biosystems GeneAmp® PCR system 9700. PCR products were cleaned up using the QIAquick® PCR purification Kit (Qiagen, Germany) according to manufacturer's instructions. DNA purity and quantity were checked by QIAxpert system, following manufacturer's instructions. DNA quality was visualised by gel electrophoresis.

2.5.3 Agarose gel electrophoresis

TAE buffer

TAE stock solution (50x) was prepared by mixing 2 M Tris, 50 mM EDTA and 60 mL of glacial acetic acid into 940 mL distilled water. The pH was adjusted to 8.0 by 2 M NaOH. Stock solution was diluted to a 1x working solution of TAE.

Gel electrophoresis

Agarose gel (1.5% (w/v)) was prepared with 1x TAE buffer. DNA was stained using ethidium bromide solution (10 mg/mL; Sigma-Aldrich, USA). DNA samples were mixed with BlueJuice™ Gel Loading Buffer (10x) (Invitrogen, USA) before being loaded onto the gel. A 1 kb plus DNA ladder standard (Invitrogen) was loaded alongside. 1x TAE buffer was used as running buffer. Electrophoresis was carried out at 120 V for 30 minutes at ambient temperature using electrophoresis subsystem 70 and/or 150 (Labnet, USA). Bands were visualised by UV transilluminator at 254 nm, imaged on an Amersham Imager 600 (GE Healthcare, UK).

2.5.4 Candidate protein cloning and expression

Restriction enzyme digestion

Chosen restriction enzymes for this thesis were NdeI, XhoI, EcoRI and NotI. All restriction enzymes were supplied by New England BioLabs, USA. DNA used for digestion was purified product and/or pET28a (+) DNA (Novagen, South Africa). Restriction enzyme digests were completed according to CutSmart® buffer (New England BioLabs, USA) instructions. A 50 µL reaction consisted of 1 µL restriction enzyme (two restriction enzymes used); 1 µL of 1 µg/µL DNA, 5 µL CutSmart NE buffer and 42 µL sterile water. The reaction was incubated at 37 ˚C for 1 hour and the digestion stopped by heat inactivation at 65 ˚C for 15 minutes. Digested DNA was stored at -20 ˚C until use.

Insert and vector purification

Both digested pET28a (+) and PCR product were run on a 1% (w/v) agarose gel prepared with 1x TAE buffer. A variety of controls was included to assess efficiency of digestion and ligation. DNA was stained using ethidium bromide solution (10mg/mL; Sigma-Aldrich, USA). DNA samples were mixed with BlueJuice™ Gel Loading Buffer (10x) (Invitrogen) before being loaded onto the gel. A 1 kb plus DNA ladder standard (Invitrogen) was loaded alongside. 1x TAE buffer was used as running buffer. Electrophoresis was carried out at 100 V for 30 minutes at ambient temperature using electrophoresis subsystem 70 and/or 150 (Labnet). Bands were visualised by UV transilluminator at 254 nm, imaged on an Amersham Imager 600 (GE Healthcare). Positively digested vector bands were sliced out and purified following manufacturer's instructions of the Monarch® DNA gel extraction kit (New England BioLabs). DNA quality and quantity was checked by QIAxpert system, following manufacturer's instructions. Treated insert and vector were stored at -20 ˚C until use. DNA clean and concentrate kit (Zymo Research) was also used following manufacturer's instructions. Positively digested insert was purified using QIAquick PCR purification kit (Qiagen), following manufacturer's instructions. Samples were eluted in 25 µL TE buffer and stored at -20 ˚C.

Ligation

Purified vector and insert were ligated in a 20 µL reaction consisting of 2 µL T4 DNA ligase buffer (10x) (New England BioLabs), vector DNA and insert DNA using a molar ratio of 1:3 or 1:4, 1 µL T4 DNA ligase (New England BioLabs) and up to 20 µL nuclease-free water. The reaction was gently mixed and incubated at room temperature for 10 minutes or 37 ˚C for 2 hours, followed by heat inactivation at 65 ˚C for 10 minutes. Controls must be included to check the efficiency of the digestion and ligation. Ligated reactions were stored long term at -20 ˚C or used right away by being chilled on ice. A 1 – 5 µL ligated reaction was used for transformation.

NovaBlue competent cells transformation

NovaBlue competent cells (Novagen) were used as the host for initial cloning, due to its high transformation efficiency and high yields of quality DNA. Inactivation of ligase is not required before transformation. For transformation into NovaBlue, 1 µL ligation reaction was added to 20 µL chilled NovaBlue competent cells. Controls were included. The cells were then incubated on ice for 5 min and then subjected to heat-shock step in a 42 ˚C water bath for exactly 30 seconds, with no shaking. Tubes were then left on ice for 2 minutes before the addition of 80 µL of SOC medium. This thesis used pET28a (+) vectors, which contain resistance to antibiotic kanamycin. Before inoculation onto LB agar containing 50 µg/mL kanamycin, cells were incubated at 37 ˚C at 250 rpm for 30 min – 1 hour. Aliquots (25 µL) of each transformation was spread on LB agar plates and incubated overnight at 37 ˚C. Plates were checked for appropriate growth and controls analysed for successful transformation.

Storage of strains

Positive transformed colonies were picked and inoculated into 50 mL LB broth containing 50 µg/mL kanamycin and left to incubate with vigorous shaking at 37 ˚C until OD_{600} reached 0.6

- 0.8, at which stage 0.9 mL was removed and added to a cryovial with 0.1 mL of 80% glycerol (v/v). The tube was mixed well and stored at -80 ˚C.

Inoculation of strain from storage

To grow up strains from storage stocks, contents of cryovial tubes were either slightly melted or scraped from the surface and streaked onto appropriate agar plates and incubated at 37 ˚C overnight. Remaining cryovial tubes were returned to -80 ˚C storage.

Direct colony PCR

The analysis of pET recombinants was performed to understand the presence and orientation of the appropriate insert. Direct colony PCR was performed using the same colony picked for storage, directly transferred to 0.5 mL micro centrifuge tube containing 50 µL sterile water. Tubes were vortexed and placed on a heat block at 99 ˚C for 5 minutes to lyse the cells. The tubes were centrifuged at 12,000 x *g* for 1 minute and the supernatant taken for PCR.

To determine the insert orientation and size, PCR was performed using primers (Table 2.4). Phusion High-Fidelity DNA polymerase (Thermo Scientific) was used for products up to 7.5 kb, has high fidelity and generates blunt-ended PCR product. A 50 µL reaction consisted of: 25 µL of 2x Phusion Master Mix, 2.5 µL each of 10 µM forward and reverse primer, 2 µL template DNA and 18 µL PCR-grade water. Optimised PCR cycle conditions recommended by the manufacturer were used: initial denaturation: 98 ˚C for 30 seconds; denaturation: 98 ˚C for 10 seconds; annealing: lowest primer T_m ˚C for 20 seconds; extension at 72 ˚C for 20 seconds per kb of target gene. Final extension at 72 ˚C for 5 minutes and storage at 4 ˚C until further processing. Denaturation, annealing and extension were cycled 30 times, before the final extension.

Table 2.4 Primers used for pET28a (+) amplification.

Primers	Sequence (5' – 3')	T_m ˚C
T7 Forward	TAATACGACTCACTATAGGG	47.5
T7 terminator Reverse	GCTAGTTATTGCTCAGCGG	53.0

Note*:* T7: T7 RNA polymerase. T_m ˚C: Melting point of primers.

Extracting plasmid DNA and DNA sequencing

To identify the correct insert into the vector and verification of the reading frame, plasmid DNA was extracted using the Monarch® Plasmid Miniprep kit (New England BioLabs). Plasmid DNA was extracted from 3 mL pelleted bacteria following manufacturer's instructions, eluted in 20 µL of DNA elution buffer.

BL21 (DE3) competent cells transformation

Target protein can be expressed from pET recombinants by the induction of the T7 or T7*lac* promoter. For transformation into BL21 (DE3), a 50 µL aliquot of BL21 (DE3) cells was thawed on ice with 2 µL DNA. The cells were incubated on ice for 30 minutes before being heat-shocked for exactly 30 seconds at 42 ˚C in a water bath, followed by another incubation on ice for 2 minutes. SOC medium (250 µL) was added and incubated on a shaking incubator at 37 ˚C, 250 rpm for 1 hour. Two volumes were spread (50 µL and 200 µL) onto selective LB agar containing 50 µg/mL kanamycin. Plates were incubated overnight at 37 ˚C.

$CaCl_2$ competent cells for transformation

The calcium chloride method was used to generate chemically competent cells. The Ca^{2+} ions bind to the phosphate of the DNA and minimise the electrostatic repulsion. Existing competent cells were inoculated into a starter culture, made of 5 mL of SOC media with no

antibiotics, and incubated at 30 ˚C overnight. A 3 mL sample of the starter culture was added to 300 mL of SOB (1:100) and shaken at 37 ˚C until the OD_{600} reached 0.4 – 0.5. Once this OD_{600} was reached, cells were chilled on ice for 20 minutes, with the occasional swirl to ensure even cooling. The cell culture was then spun at 2,000 x *g* for 15 minutes at 4 ˚C. Supernatant was discarded and the cell pellet resuspended in 20 mL of ice cold 100 mM $MgCl_2$. Cells were centrifuged again at 2,000 x *g* for 15 minutes at 4 ˚C. Supernatant was decanted and the cell pellet resuspended in 30 mL of 100 mM $CaCl_2$ and cooled for 20 minutes on ice. Cells were spun as before and resuspended in 1 mL of 85 mM $CaCl_2$ in 15% glycerol (v/v). Aliquots (50 µL) were placed into microfuge tubes, snap frozen in liquid nitrogen and stored at -80 ˚C.

Rosetta (DE3) pLysS competent cells transformation

To compensate for any rare codons contained in the target protein, the plasmid was transformed into Rosetta DE3 cells, which are designed to enhance the expression of proteins that contain rare codons that are rarely used in *E. coli*. Laboratory made Rosetta cells were prepared using the $CaCl_2$ method above. For transformation, a 50 µL aliquot of Rosetta (DE3) was thawed on ice and 2 µL of DNA was added. The cells were incubated on ice for 30 minutes before being heat-shocked for exactly 30 seconds at 42 ˚C in a water bath, followed by another incubation on ice for 2 minutes. SOC medium (250 µL) was added and the mixture incubated on a shaking incubator at 37 ˚C, 250 rpm for 1 hour. Serial dilutions of 1/10 and 1/50 were made and 100 µL was plated on LB containing 50 µg/mL kanamycin and 30 µg/mL chloramphenicol. Plates were incubated overnight at 37 ˚C.

Protein expression

Target protein expression from the plasmid established within BL21 (DE3) or Rosetta (DE3) cells was induced by the addition of IPTG to a growing culture. Following manufacturer's instructions, four transformants were picked and inoculated to 5 mL of LB broth containing 50 µg/mL kanamycin (BL21 (DE3)) or 50 µg/mL kanamycin and 25 µg/mL chloramphenicol (Rosetta (DE3)). These broths were grown overnight at 37 ˚C with vigorous shaking until

they reached an OD_{600} ≥ 2. Using overnight culture, a fresh LB medium containing the appropriate antibiotics was inoculated and cells grown to an OD_{600} of 0.05 – 0.1, a 1:50 dilution. Glycerol stocks were made of remaining overnight culture. The fresh cultures were incubated until an OD_{600} reached 0.6 at which point, expression was induced by the addition of IPTG to a final concentration of 1 mM. Incubation at 37 ˚C was continued for 8 hours and then moved to 25 ˚C for overnight incubation. To find optimal protein expression, time point collections of: (Tb) before induction; (T0) at induction; (T30) 30 minutes induction; (T1) 1 hour induction; (T2) 2 hours induction and (T4) 4 hours induction. Controls were included alongside.

2.5.5 Protein verification

To assess the target gene expression, various fractions were assessed for the recovery of the target protein.

1x PBS solution

One tablet of phosphate buffered saline (Sigma-Aldrich) was dissolved in 200 mL of distilled water to create a 1x PBS solution that was then sterilised by autoclaving.

Total cell protein (TCP) fraction

Target cell protein was collected by taking a 1 mL sample and collecting the cell pellet by centrifugation at 10,000x g for 1 min and discarding the supernatant. The pellet was re-suspended in 100 µL of 1x phosphate buffered saline (PBS). 100 µL of 4x SDS Sample buffer was added and the whole solution passed through a 27-gauge needle several times to reduce sample viscosity. The sample was heated at 85 ˚C for 3 minutes before storage at -20 ˚C until SDS-PAGE analysis.

Medium fraction

Target protein may have been leaked rather than undergoing true secretion (Stader *et al*., 1990). A volume of 40mL of culture was collected, centrifuged at 10,000 x *g* for 10 minutes at 4 ˚C. The supernatant was decanted and the cell pellet kept at 4 ˚C. Supernatant (1 mL) was taken and concentrated by using Amicon ultra centrifugal filter, with a MWCO of 10 kDa. The filtrate was discarded and retentate collected and volume recorded. The filter membrane was rinsed with hot 2x SDS sample buffer and pooled with the concentrated retentate. The sample was heated for 3 minutes at 85 ˚C before storage at -20 ˚C until SDS-PAGE analysis.

Soluble cytoplasmic fraction

Target proteins have histidine-tags, where when expressed in *E. coli*, they can accumulate either as biologically functional soluble protein or as inclusion bodies. Bacterial cell cultures underwent protein extraction using BugBuster® Protein Extraction (Novagen), where two samples were collected for protein purification under native and denaturing conditions.

In the absence of BugBuster® Protein Extraction (Novagen), cell pellets collected after IPTG induction were subjected to cell lysis by sonication. The supernatant collected was subjected to protein purification, while the pellet was frozen at -20 ˚C for insoluble cytoplasmic fraction verification.

Gel staining of pure protein extracts

Colloidal blue Coomassie stain was routinely used for visualisation of SDS-PAGE gels loaded with induced samples and pure protein from purification. Coomassie solution consisted of 0.02% w/v CBB G-250, 5% w/v aluminum sulfate (14 – 18) hydrate, 10% (v/v) 96% ethanol and 2% (v/v) 85% orthophosphoric acid into 2 L of milli-Q water. This solution was stored at room temperature, wrapped in foil. Gels were fixed with 40% MeOH/10% acetic acid for 30

minutes and stained overnight with four parts colloidal Coomassie stain to one part methanol. Following this, the Coomassie solution was removed and the gel rinsed with Milli-Q water. Gels were destained in Milli-Q water for 1 – 2 hours before analysing.

Gel digestion for mass spectrometry

For verification that the correct protein was purified, gel bands with correct size (kDa) were excised from the colloidal stained gel and processed for mass spectrometry. All solvents and water were of LC-MS grade.

1 x ABC solution

Ammonium bicarbonate (0.04 g) was dissolved in 10 mL of Milli-Q water.

Reducing solution

A solution of 10 mM DTT (0.00154 g) dissolved in 1 x ABC solution.

Alkylation solution

A solution of 20 mM iodoacetamide (0.0037 g) dissolved in 1 x ABC solution

Digestion solution

Sigma SOLu-Trypsin (Ems0004) was diluted 50-fold in 1 x ABC and 1 mM $CaCl_2$.

Gel bands were excised with a sterile scalpel, chopped into tiny pieces and placed into a 1.5 mL micro centrifuge tube where the gel was destained with 1 x ABC/50% MeOH solution

(200 µL) at 45 ˚C until clear. Gels were dehydrated with 80% MeCN/water for 1 min. Once gels have become opaque and clear, the fluid was removed and tubes placed into a centrifugal evaporator (SpeedVac Concentrator, ThermoFisher) for 10 minutes. Reducing agent (30 µL) was added to the tubes and incubated at 45 ˚C for 1 hour. Supernatant is removed and gel pieces washed with 1 x ABC (100 µL). Gel is dehydrated with 80% MeCN (300 µL) for 5 min, fluid removed and dried via SpeedVac. Alkylation solution (30 µL) was added and gel samples left to incubate for 30 minutes in the dark. Fluid was removed and gel pieces washed with 100 µL 1 x ABC and fluid discarded. The gel pieces were dehydrated with 80% MeCN (300 µL) for 5 min and fluid removed. The rehydrating and dehydrating process was repeated, with gel pieces (fluid removed) dried by SpeedVac for 10 minutes. Digestion solution (30 µL) submerged the gel pieces and incubated on ice for 10 minutes. Excess solution was removed and 30 µL 1 x ABC was added and left to incubate overnight at 37 ˚C. Tubes were centrifuged briefly and subjected to sonication for 2 minutes and re-spun. Supernatant was collected into a new LoBind Eppendorf tube and pooled with the following washes. Formic acid 5%/40% MeCN (50 µL) was added, sonicated and centrifuge with supernatant pooled. Formic acid 0.5%/80% MeCN (50 µL) was added, sonicated and centrifuged with supernatant pooled. Samples volumes were reduced by SpeedVac concentrating down from 130 µL to 30 µL. Reduced samples are stored at -80 ˚C until analysed by mass spectrometry.

2.5.6 Tagged protein purification, proposed structure and ligand binding

Ni-NTA protein purification

A three-step purification method was used to obtain protein of high purity. Firstly, BugBuster® Ni-NTA His-Bind Purification kit was used, following manufacturer's instructions with some modifications for smaller samples.

Table 2.5 The concentrations of Ni-NTA purification buffers.

	NaH_2PO_4	NaCl	Imidazole
1 x Ni-NTA bind buffer	50 mM	300 mM	10 mM
1 x Ni-NTA wash buffer	50 mM	300 mM	20 mM
1x diluted 1 x Ni-NTA elution buffer	50 mM	300 mM	125 mM
1 x Ni-NTA elution buffer	50 mM	300 mM	250 mM

Ni-NTA resin was prepared by addition 1 mL of 1x Ni-NTA Bind buffer to 250 µL 50% Ni-NTA His-bind slurry. Solution was mixed and resin allowed to settle by gravity and 4 mL of supernatant was removed. The cleared 1 mL lysate from protein verified sample was pipetted into the slurry and mixed gently at 4 ˚C for 1 hour. The whole sample was loaded onto a capped column. Flow through was collected and saved for SDS-PAGE and protein activity. Next, 2 x 1 mL of 1x Ni-NTA wash buffer was added, and flow through collected. Then, 2 x 0.5 mL of diluted 1x Ni-NTA elution buffer was added and eluate collected. Finally, 2 x 0.5 mL of 1x Ni-NTA elution buffer was added and eluate collected. All fractions were analysed by SDS-PAGE for protein profile, protein content (mg/mL) and enzyme activity. Eluates were stored at -20 ˚C.

PsiPred/pGenThreader and AlphaFold analysis

To predict the structure of protein 0050, two analysis servers were used to further understand the possible folding of the protein. PSIRED (McGuffin *et al*., 2000) is a protein structure prediction server where users submit a sequence and choose the prediction method and receive results to analyse. There are three methods of prediction: PSIPRED: highly accurate secondary structure prediction; MEMSAT: transmembrane topology prediction method and, lastly, pGenTHREADER: a method which looks at the sequence profile base fold (McGuffin *et al.,* 2000; Lobley *et al*., 2009). Sequences (Figure 5.2) were uploaded to “Data Input” on the site http://bioinf.cs.ucl.ac.uk/psipred/. Analyses PSIPRED

4.0 and pGenTHREADER were selected for prediction. Data were presented online and emailed. Exported data were analysed.

AlphaFold (Jumper *et al*., 2021) is a computational method used to generate protein structures at an atomic accuracy even where no homologous protein structure is available. Protein sequence is analysed through a neural network-based model, a deep learning algorithm, which incorporates the physical and biological knowledge about protein structure (Jumper *et al*., 2021). Using AlphaFold requires bioinformatics knowledge, specifically Python and following instructions located on GitHub: https://github.com/deepmind/alphafold. Data produced by deep learning were stored in a subdirectory folder created by the user on the C: drive (Jumper *et al*., 2021).

Ligand – Protein binding using BLItz™ Software

Baseline Buffer

The buffer was prepared in 40 mL of Milli-Q water with 0.060 g Trizma base. The pH was adjusted to 8.0 by 1 M HCl. Volume was topped up to 50 mL. Buffer was sterilised by filtration and stored at room temperature.

Regeneration Buffer

The buffer was prepared in 40 mL Milli-Q water with 0.037 g glycine with the pH adjusted to 1.7 by 1 M HCl. Volume was topped up to 50 mL and sterilised by filtration. Stored at room temperature.

Recharging Buffer

The buffer was prepared in 50 mL of 0.064 g nickel (II) chloride. Buffer was sterilised by filtration. Bottle wrapped in tinfoil to minimise exposure to light and stored within a dark cupboard.

The preliminary ligand binding of protein were performed on BLItz™ instrument. This followed the BLItz user manual using BLItz Pro™ software, investigating a binding yes or no experiment. Anti-Penta-HIS biosensors were used to bind his-tagged proteins for analysis. One concentration of pectin substrate was used (100 ng/µL) against construct one (P1) and construct two (P2). Samples were run through a five-step experiment: (1) Initial baseline with baseline buffer; (2) 4 µL drop of protein loaded; (3) Re-baseline with baseline buffer(4): 4 µL load of pectin substrate to measure protein-ligand association; (5) Dissociation with baseline buffer, to measure breaking between protein-ligand. After each round of experiment, the biosensor probe was cleaned with 10mM glycine pH 1.7 and recharged with 10mM $NiCl_2$. Raw data were exported and corrected for potential substrate binding to the probe and interference between probe interactions with solution at each step.

Advanced kinetics was run similarly to the above using the advanced kinetics BLItz Pro™ software, with two concentrations of substrate used: (1) 100 ng/µL and (2) 100 µg/mL. Data analysis was performed by the BLItz Pro™ software, where the raw data was corrected for each step interaction with the probe and a curve plotted to the raw data graph with enzymatic kinetic constants presented. Analysis was exported and tabulated.

Chapter 3 Ripening in kiwiberry '*Actinidia arguta*' fruit – structural and compositional changes in the cell wall

3.1 Introduction

The fruit of *Actinidia arguta*, commonly known as kiwiberry, originating in Eastern Asia, have been cultivated around the world (Latocha, 2017). There are three main cultivars grown in New Zealand, *Actinidia arguta* (Sieb. Et Zucc) Planch. Ex Miq. var. arguta 'Hortgem Tahi', var. 'Takaka Green' and var. 'Marju Red' (as described in section 1.1). We have included the preceding three cultivars in this study.

This chapter aims to quantify and characterise pectin on whole New Zealand kiwiberry fruit of three cultivars: Hortgem Tahi, Takaka Green and Marju Red. This work was derived from the missing structural information surrounding all three of the commercially available New Zealand kiwiberries (as described in section 1.1). So far, the recent characterisation of the Hortgem Tahi cell wall structure and composition was restricted to only kiwiberry core tissue (Sutherland *et al*., 2017). It was hypothesised that there were structural changes during the softening of the kiwiberries. Firmness analysis by compression (as described in section 2.3.2) was considered ideal for monitoring the fruit's softening, as it is widely used in post-harvest assessments (Schotsmans & Mawson, 2004). Thus, samples of kiwiberries were stored under post-harvest conditions (20 ˚C) for four days, with daily compression tests. Since naturally-occurring whole fruit softening is the result of pectin solubilisation and pectin degradation, analysis of cell wall from whole kiwiberries was investigated. The methodology of alkaline fractionation was reported to extract up to 82% of polygalacturonic acid from fruit cell walls (Redgwell *et al*., 1997). Water-soluble pectin typically contains HG content, where RG-I may be co-extracted (O'Donoghue *et al*., 2017). Pectin molecules can be cross-linked in various ways (Schols *et al*., 2003; Wang *et al*., 2018). Two HG chains can be linked by the insertion of Ca^{2+} ions interacting between the unesterified carboxyl groups of galacturonosyl residues. The use of CDTA removes the Ca^{2+} ion present, hence releasing stretches of unesterified HG and any intervening stretches of RG-I (O'Donoghue *et al*., 2017; Schols *et al*., 2003). Aqueous sodium carbonate (Na_2CO_3) typically extracts most of the branched pectins from plant cell walls (Kaczmarksa *et al*., 2022; O'Donoghue *et al*., 2017),

although in doing so will de-esterify any remaining pectin. Pectin solubilisation from kiwiberry cell walls (as described in section 2.3.6) was explored by using increasingly alkaline conditions (water, CDTA and Na_2CO_3). Afterwards, soluble uronic acid content was quantified based on a colorimetric assay of Blumenkranktz and Asboe-Hansen (1973), modified to 96 well-microplate by O'Donoghue (2017). Glycan profiling was chosen to characterise the pectin composition within the soluble-uronic acid fractions, as this method is used to screen and track a diverse range of glycan structures rapidly, and provides an overview of cell wall composition (Moller *et al*., 2012). Therefore, soluble uronic acid fractions were printed onto nitrocellulose and probed with specific monoclonal antibodies, with label-specific glycan epitopes (as described in section 2.3.6).

Analysis of the structural and compositional changes of the three kiwiberry cultivars reveal significant differences. The fruit softened at various rates and this softening could be related to the diversity in the stage of their development at harvest. The solubility of the pectin (determined by uronic acid content) and the types of pectin present (as shown by glycan arrays) also changed over time and differed between cultivars. This shows that the fruit may be in different stages of fruit ripening. The work here provides key data into the postharvest and pectin composition of New Zealand kiwiberries and also, offers starting knowledge to support future work on bacterial digestion processes, simulated by kiwiberry consumption.

3.2 Results and Discussion

3.2.1 Firmness analysis of New Zealand kiwiberries

Five kiwiberries from each cultivar were tested for firmness by compression. Each test was performed on three random sides of the kiwiberry, at harvest maturity and, subsequently daily, during postharvest (2.3.2). Daily assessments of the kiwiberries showed a loss of firmness over four days at 20 ˚C (Figure 3.1), as well as significant differences in firmness at time of harvest (Day 0).

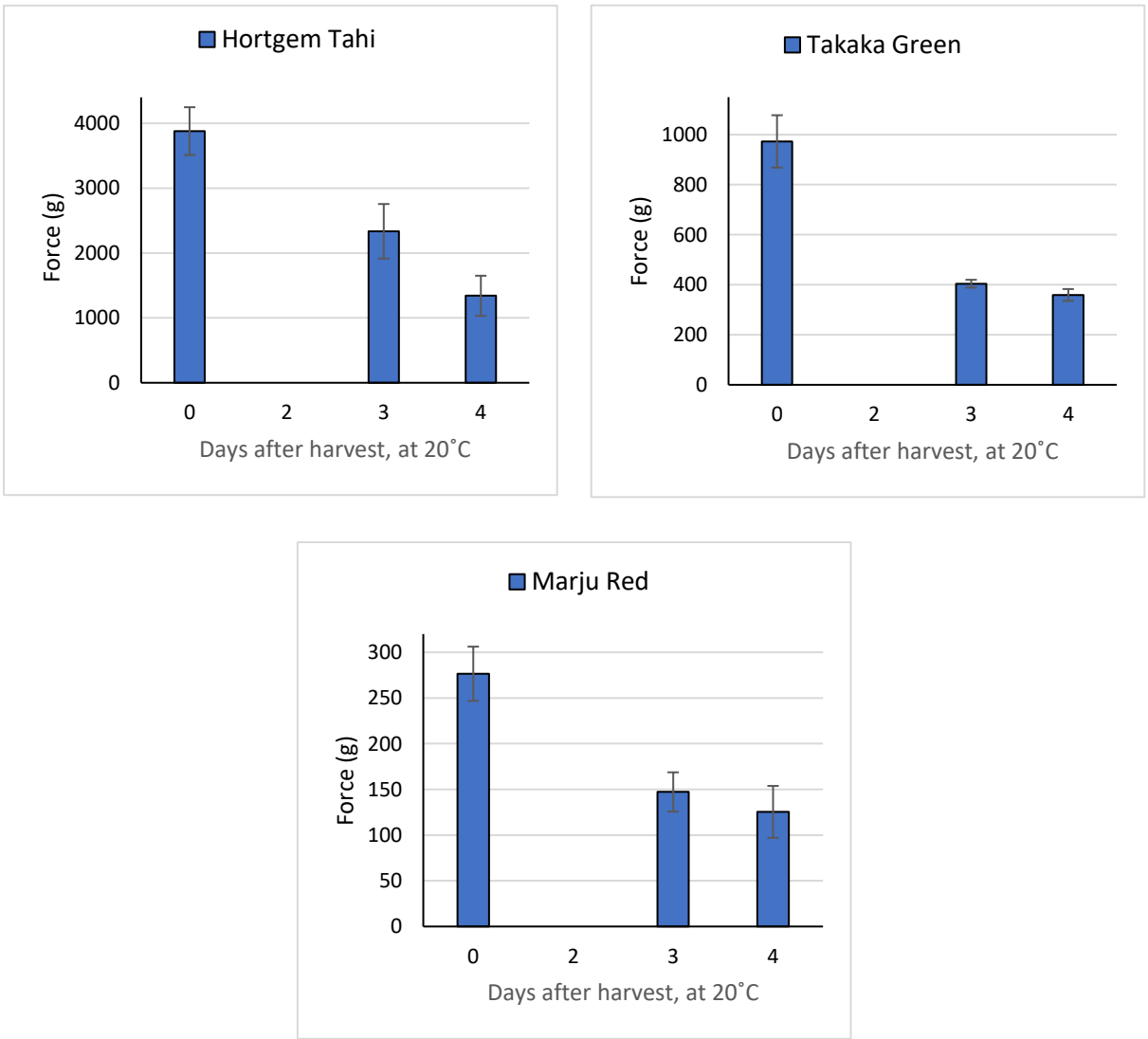


Figure 3.1 Compression firmness of Hortgem Tahi, Takaka Green and Marju Red kiwiberries after harvest and up to 4 days at 20 ˚C. Compression force (*g*) values are the average of 5 fruit tested on 3 sides, +/- standard error.

Considerable variations in firmness of all three cultivars were observed over the 4 days at 20 ˚C. Hortgem Tahi had a high initial firmness of ~3900 force (g), decreasing as the fruit was kept over time at 20 ˚C. The rate of the decrease in fruit firmness did not taper off on Day 4 at 20 ˚C, suggesting the fruit was still ripening. Takaka Green and Marju Red fruit firmness decreased over the 4 days at 20 ˚C. Noticeably, both Takaka Green and Marju Red softening rates slowed between Day 3 and Day 4 at 20 ˚C, signalling an almost, or near-complete, fruit ripening/fruit softening (Sutherland, 2017). However, the size of the kiwiberry fruit does influence the compression firmness measurement (Han *et al*., 2019). In support of these results, Sutherland *et al*. (2017) also showed a significant decrease in firmness of Hortgem Tahi from harvest until five days after harvest at 20°C, and they considered the fruit overripe on Day 5.

The results presented here indicate, by the change in firmness, that the fruits were softening during ripening (Paniagua *et al*., 2014), but the rate at which the fruits softened was different between the cultivars. Softening of fruits during ripening is generally accompanied by the dissolution of the middle lamella of the cell walls of the tissue, and the modification of the composition of polymers, such as pectin and its neutral side chains, present in the primary cell wall (Paniagua *et al*., 2014; Pose *et al*., 2018). The softening of the fruits is mainly due to the action of pectin-metabolising enzymes (Paniagua *et al*., 2014; Wang *et al*., 2018). There is also evidence to suggest that the presence of hydroxyl radicals can also non-enzymatically contribute to the degradation of pectin (Airianah *et al*., 2016). Measurement of firmness by compression is only one method of studying fruit softening and cell wall chemistry. Further investigation of ethylene and carbon dioxide production and soluble solids contents (measured by Brix) would provide a deeper insight into the cell wall chemistry changes during fruit ripening (Sutherland *et al*., 2017).

3.2.2 The uronic acid content of Hortgem Tahi, Takaka Green and Marju Red

Pectin was extracted from buffered-phenol alcohol insoluble solids (BP-EIS) prepared from kiwiberry whole tissue (2.3.3). Extracts were analysed for uronic acid (UA), and the results were expressed on a fresh tissue basis (Figure 3.2). The uronic acid content was calculated per fresh weight of kiwiberry to incorporate seeds.

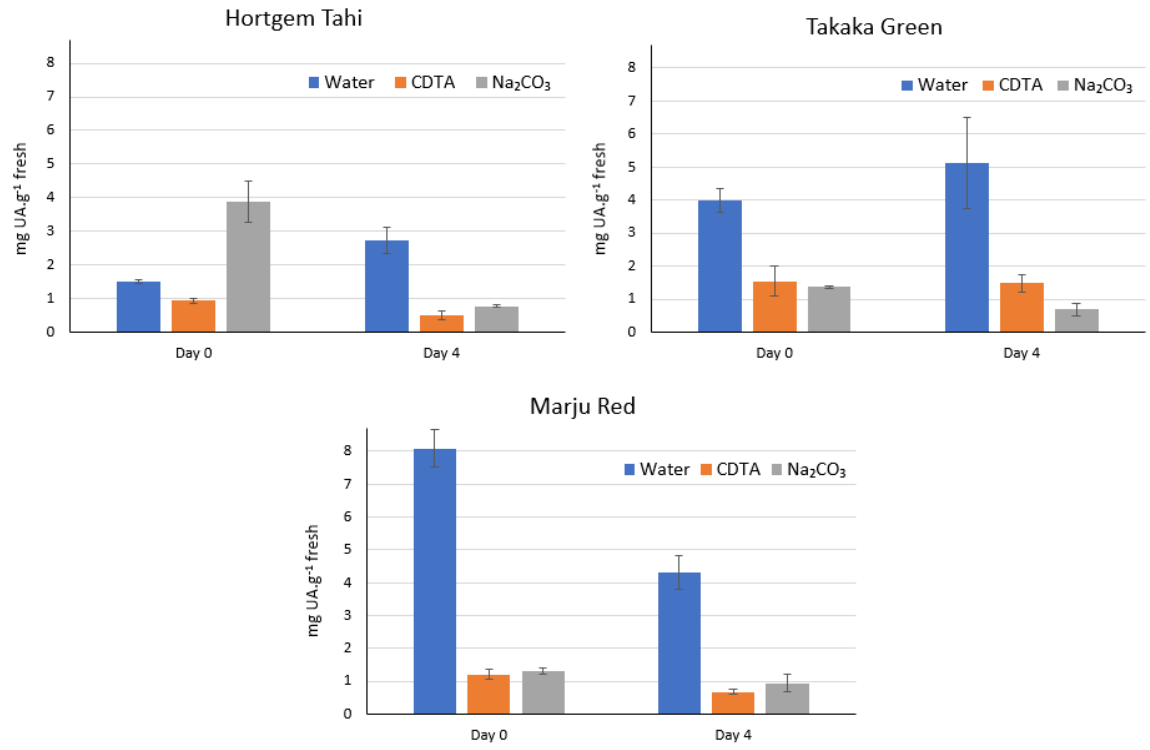


Figure 3.2. Analysis of uronic acid (UA) content (mg/g fresh) extracted from kiwiberry by water, CDTA and Na_2CO_3 at Day 0 and Day 4 at 20 ˚C. (A) Hortgem Tahi. (B) Takaka Green. (C) Marju Red. Results are the average of 3 separate extractions at each time point (+/- standard error).

Figure 3.2 shows that each cultivar has a range of uronic acid content, with changes (increase or decrease) in pectin content on Day 4 at 20 ˚C. Hortgem Tahi (Figure 3.2A) and Takaka Green (Figure 3.2B) show that uronic acid content in the water-soluble fraction increases from Day 0 to Day 4 at 20 ˚C, while that of both CDTA- and Na_2CO_3-soluble fractions lowered from Day 0 to Day 4 at 20 ˚C. In Hortgem Tahi, the uronic acid content of Na_2CO_3-soluble fraction is significantly higher compared to water- and CDTA-soluble pectin

on Day 0 at 20 ˚C. On Day 4 at 20 ˚C, the uronic acid content in water-soluble fraction was higher by 45.1% while that of CDTA- and Na_2CO_3-soluble fractions was lower (Figure 3.2A). The uronic acid content of Takaka Green in water-soluble fraction is higher, in comparison to the uronic acid contents of CDTA- and Na_2CO_3-soluble fractions at Day 0 and Day 4, in contrast to Hortgem Tahi. In Marju Red (Figure 3.2C), the uronic acid content in the water-soluble fraction is more than 7 times the amounts in the CDTA- or Na_2CO_3-soluble fractions at Day 0, dropping to almost half this by Day 44. Each cultivar had a different pattern of pectin solubilisation (Figure 3.2), which changed between Day 0 (at harvest) and Day 4 (ripening at 20 ˚C). The differences in uronic acid profiles are likely to be due to the structural changes occurring in the cell wall of the kiwiberries over the four days at 20 ˚C (Pose *et al*., 2018; Sutherland *et al*., 2017; Wang *et al*., 2018). The increase in the uronic acid content in water-soluble fractions (Figure 3.2A & B) indicates that pectin is undergoing solubilisation (Sutherland *et al*., 2017). Polysaccharide degrading enzymes would contribute to the changes in the fruit cell wall due to the fruit ripening (Tan *et al*., 2018; Wang *et al*., 2018) and these might differ between each cultivar. Some examples of such enzymes are endo-/exo-polygalacturonase, β-galactosidase, pectin methylesterase and pectate lyase (Brummel, 2006). No studies have been published on cell wall enzymes during kiwiberry ripening so far.

3.2.3 Glycan Array of Hortgem Tahi, Takaka Green and Marju Red

Kiwiberry pectin fractions (water, CDTA and Na_2CO_3) were probed with specific glycan-binding monoclonal antibodies (mAb) using glycan array technology. Details of pectin extraction are in Section 2.3.7. For the glycan array, dilutions of each extract from each cultivar were spotted on nitrocellulose membrane and treated with a range of antibodies directed at polysaccharide epitopes found in the plant cell wall. A secondary anti-rat antibody conjugated to alkaline phosphatase was used to detect the binding of the glycan antibodies. The visualisation was the result of the presence of chromogenic substances (Moller *et al*., 2012). The membranes are shown in Figure 3.3. Raw density of each spot was converted to units on a fresh weight basis. As each antibody has a unique affinity for its epitope (Moller *et al*., 2012), direct comparisons between antibodies cannot be made,

although comparisons of their trends are permissible. The data represented here are of semi-quantitative nature and does not imply absolute amounts.

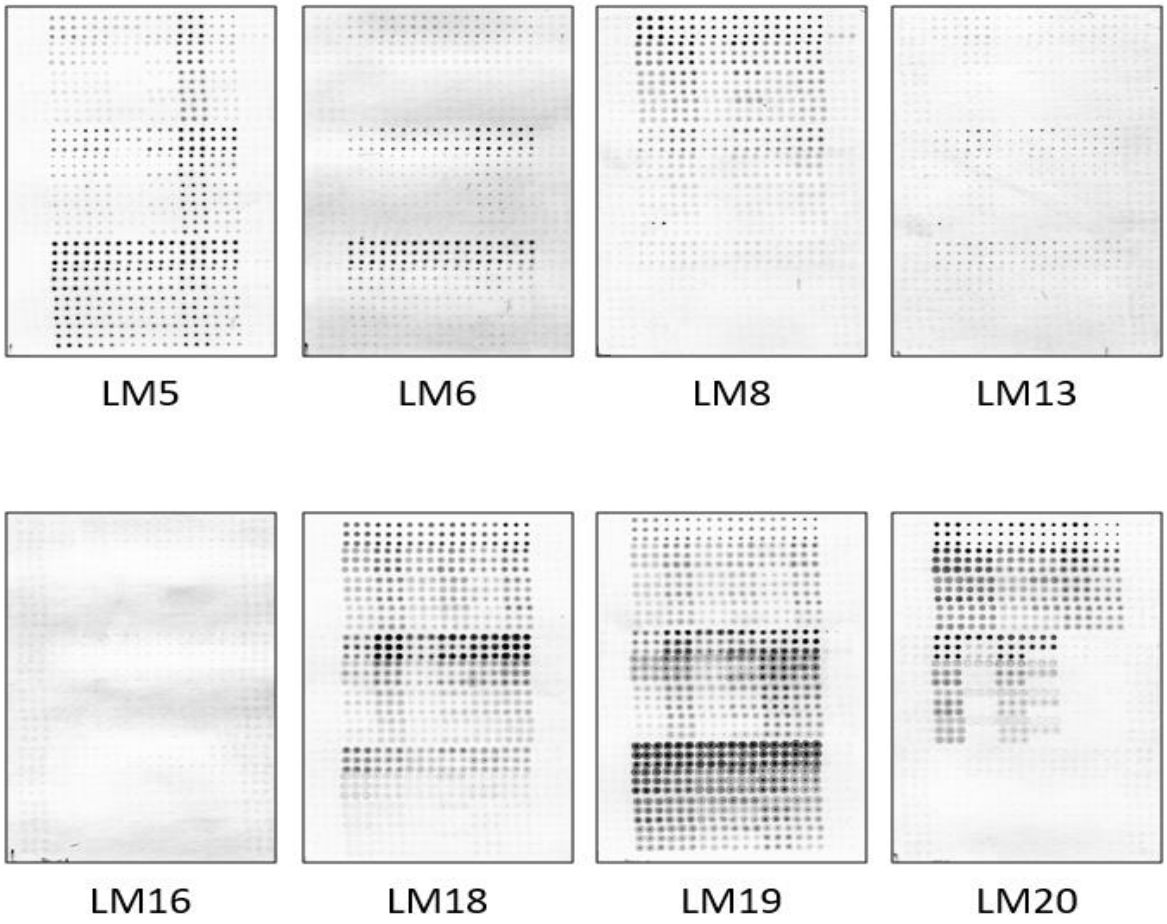


Figure 3.3 Nitrocellulose membranes dotted with soluble uronic acid content from three Kiwiberry cultivars, probed with monoclonal antibodies (labelled by LM) specific for cell-wall glycan epitope**.** Each LM is described in section 2.3.7. In summary, each nitrocellulose membrane is divided into the three soluble-uronic acid samples: top: water-soluble; middle: CDTA-soluble and bottom: Na_2CO_3. Within each third: across is each kiwiberry (at harvest maturity and post-harvest) sample dotted in repeats. While, down a column is a serial dilution of the kiwiberry sample. Refer to Figure 2.1, for a more detailed explanation of the array.

The binding of antibodies LM18, LM19 and LM20 (Verhertbruggen *et al*., 2009b) allows the visualisation of the esterification of pectin.

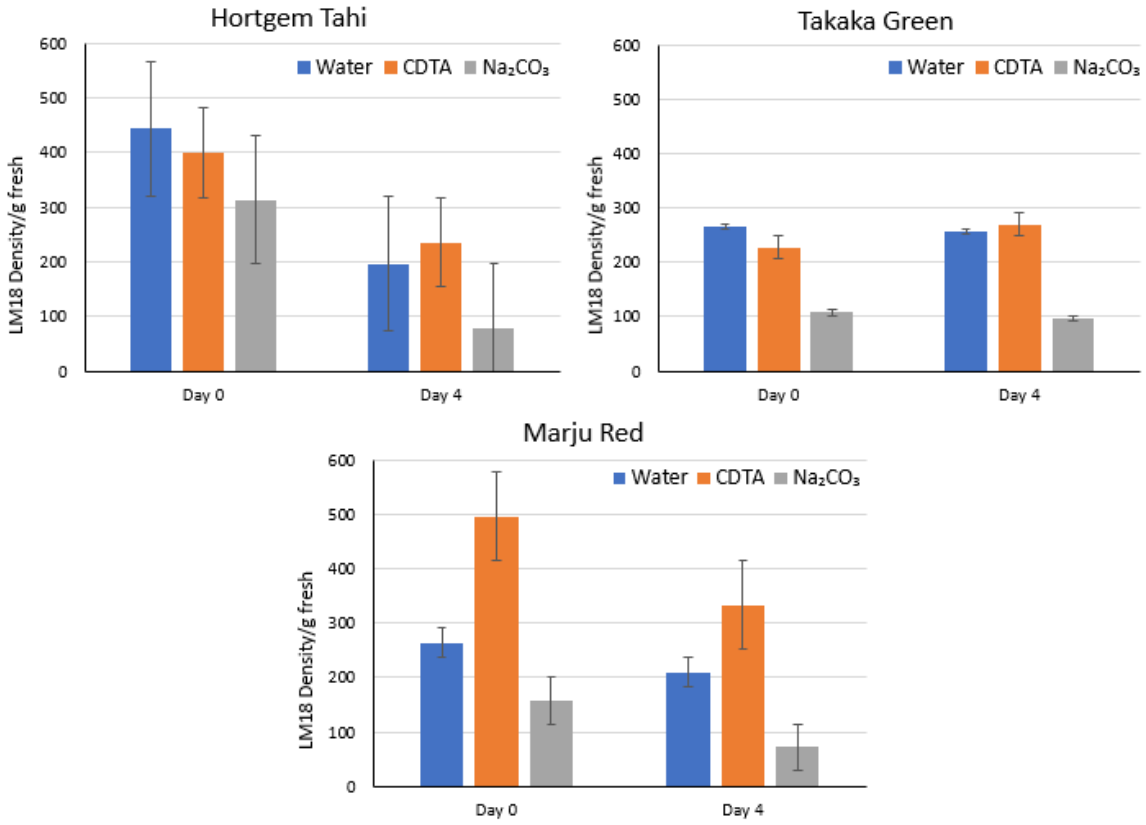


Figure 3.4 mAb LM18 binding (density/g fresh) in water-soluble, CDTA-soluble and Na_2CO_3-soluble extracts of Hortgem Tahi, Takaka Green and Marju Red kiwiberry.

LM18 antibody binds to partially methyl-esterified homogalacturonan (HG). It has a low affinity for low esterified pectin (<45%) and polygalacturonic acid (Verhertbruggen *et al*., 2009b) so the nature of this esterification epitope is complex. Figure 3.4 shows the labelling of LM18 in all soluble-pectin fractions of the three cultivars and for the most part labelling decreased between Day 0 and Day 4. A substantial amount of LM18 labelling was present at Day 0 for Hortgem Tahi in all three pectin extracts and binding decreased by almost half in all of them between Day 0 and Day 4. LM18 labelling of Hortgem Tahi water- and Na_2CO_3-soluble extract decreases over 50% after 4 days. On the other hand, CDTA-soluble extract of Hortgem Tahi does not have such a big decrease in LM18 labelling. The labelling of LM18 in pectin extracts of Takaka Green displayed a similar profile to Hortgem Tahi, although Takaka Green profiles of all fractions show smaller decreases from Day 0 to Day 4. In fact, LM18

labelling of esterified pectin increases slightly in the CDTA-soluble fraction. Marju Red had substantially more LM18 labelling in the CDTA-soluble extracts than in water-soluble extracts, different to the other two cultivars. All fractions in Marju Red showed a decrease in LM18 labelling over 4 days.

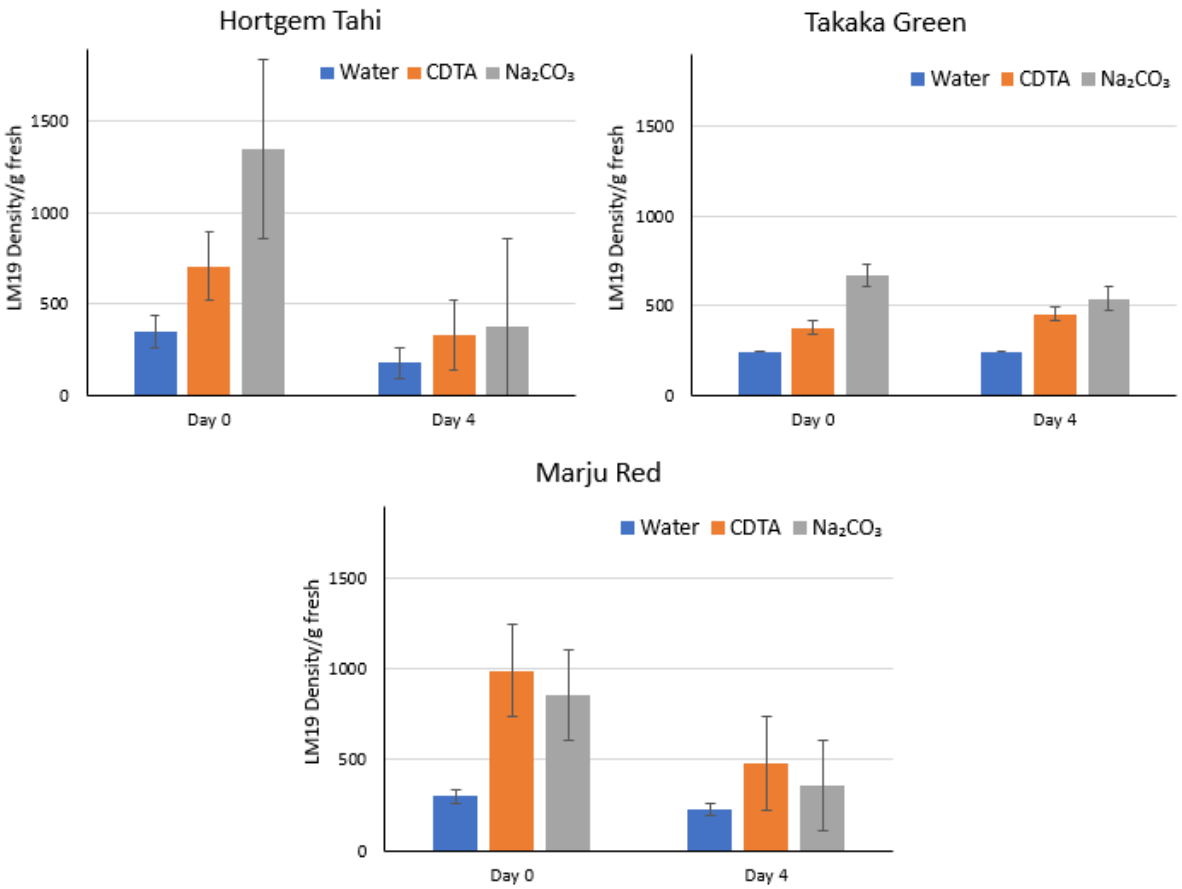


Figure 3.5 mAb LM19 binding (density/g fresh) in water-soluble, CDTA-soluble and Na_2CO_3-soluble extracts of Hortgem Tahi, Takaka Green and Marju Red kiwiberry.

LM19 typically binds to pectin of low esterification content, with no discrimination of extent and pattern of the methyl-esterification. It has, therefore, a high affinity for polygalacturonic acid (Verhertbruggen *et al.*, 2009b). Figure 3.5 shows Hortgem Tahi and Takaka Green LM19 labelling in Na_2CO_3 extracts in large amounts at Day 0, followed by CDTA- and water-soluble pectin fractions. For both cultivars, in all fractions, there was a decrease in the labelling of LM19 between Day 0 and Day 4, although this decrease was much more marked for Hortgem Tahi. LM19 labelling of Marju Red soluble fractions showed a different LM19 binding profile. There was low LM19 labelling in the water-soluble fraction. Greater labelling

of LM19 was identified in the CDTA-soluble fraction, followed by Na_2CO_3, with both fractions showing decreased labelling from Day 0 to Day 4.

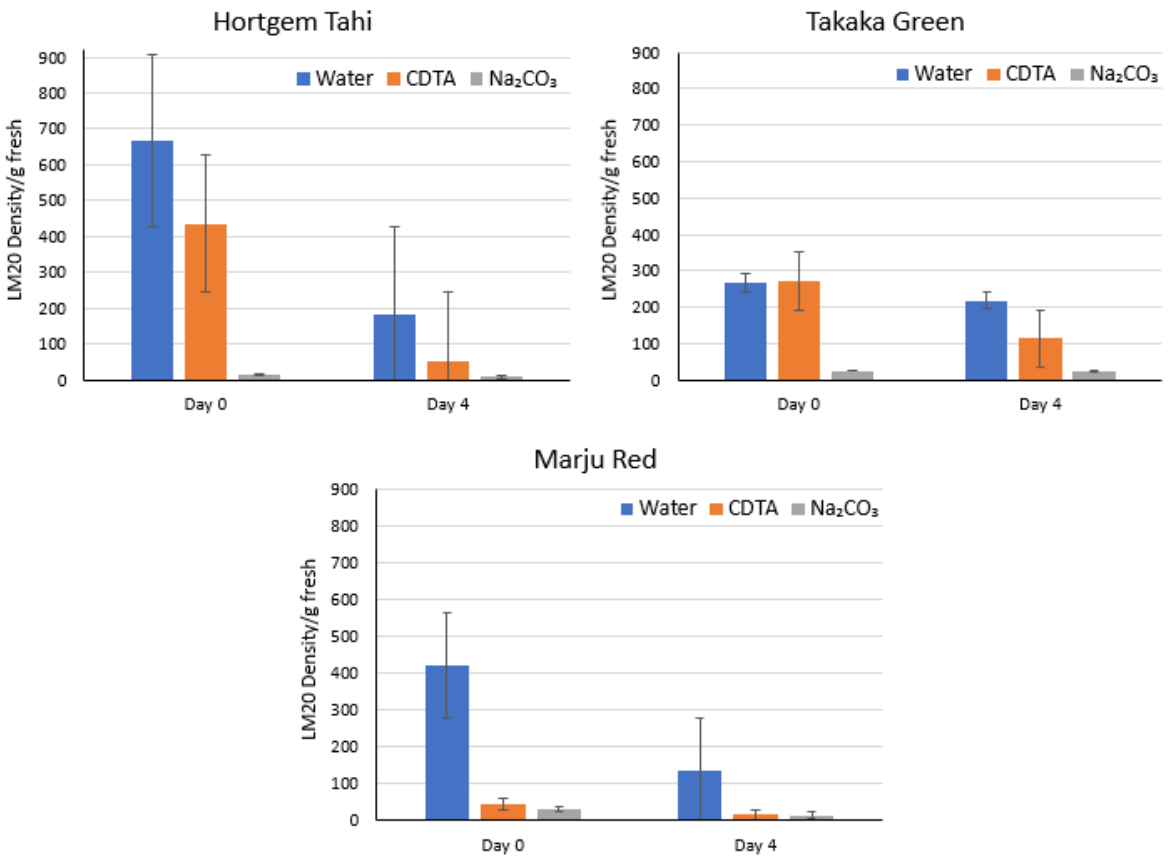


Figure 3.6 mAb LM20 binding (density/g fresh) in water-soluble, CDTA-soluble and Na_2CO_3-soluble extracts of Hortgem Tahi, Takaka Green and Marju Red kiwiberry.

The LM20 antibody recognises highly esterified (>66%) pectin (Verhertbruggen *et al*., 2009b). Figure 3.6 shows substantial labelling of LM20 in water-soluble fractions, followed by CDTA-soluble fractions in all three cultivars. The labelling of LM20 to Na_2CO_3-soluble pectin is very low in all three cultivars. Between Day 0 and Day 4, there was a substantial decrease in the labelling of LM20 in water- and CDTA-soluble fractions in all three cultivars, especially for Hortgem Tahi and Marju Red cultivars. The reduction of LM20 binding after four days is possibly due to the activity of pectin methylesterase, an enzyme which deesterifies pectin during fruit ripening (Verhertbruggen *et al*., 2009b). There was little highly esterified pectin (no LM20 labelling) in the Na_2CO_3-soluble fractions because the use of Na_2CO_3 saponifies the ester group on pectin during solubilisation (de-esterifies pectin).

The binding of the antibody to galactan side chain: LM5 (Jones *et al*., 1997) and antibodies to arabinan side chains; LM6 and LM13 (Verhertbruggen *et al*., 2009a) give insight into the composition of RG-I.

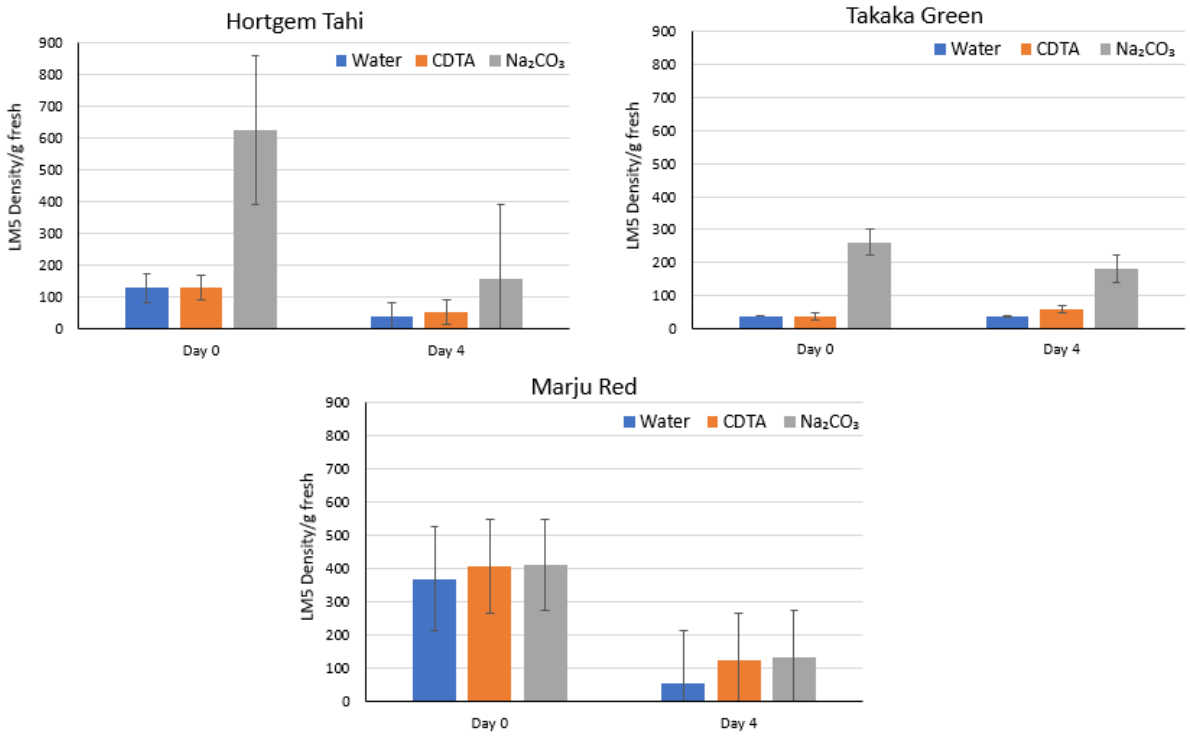


Figure 3.7 mAb LM5 binding (density/g fresh) in water-soluble, CDTA-soluble and Na_2CO_3-soluble extracts of Hortgem Tahi, Takaka Green and Marju Red kiwiberry

LM5 antibody binds to regions of linear (1, 4)-linked galactose which have more than three contiguous units of 1, 4-β-galactosyl residues (Jones *et al*., 1997). Figure 3.7 shows the labelling of LM5 in all extracts (water-, CDTA- and Na_2CO_3-soluble pectin) for all three cultivars. The binding of LM5 is highest in Na_2CO_3-soluble fractions from Hortgem Tahi and Marju Red. There was a substantial decrease in the labelling of Na_2CO_3-soluble fractions between Day 0 and Day 4 for Hortgem Tahi (especially) and Marju Red cultivars. Takaka Green Na_2CO_3-soluble fractions also decreased from Day 0 to Day 4, but not as dramatic. In Marju Red, LM5-positive epitopes were present in all soluble extracts at Day 0, and all decreased by Day 4. The loss of mAb LM5 binding indicates the loss of pectic galactan, most likely due from the degradation of RG-I side chains (Sutherland *et al.,* 2017).

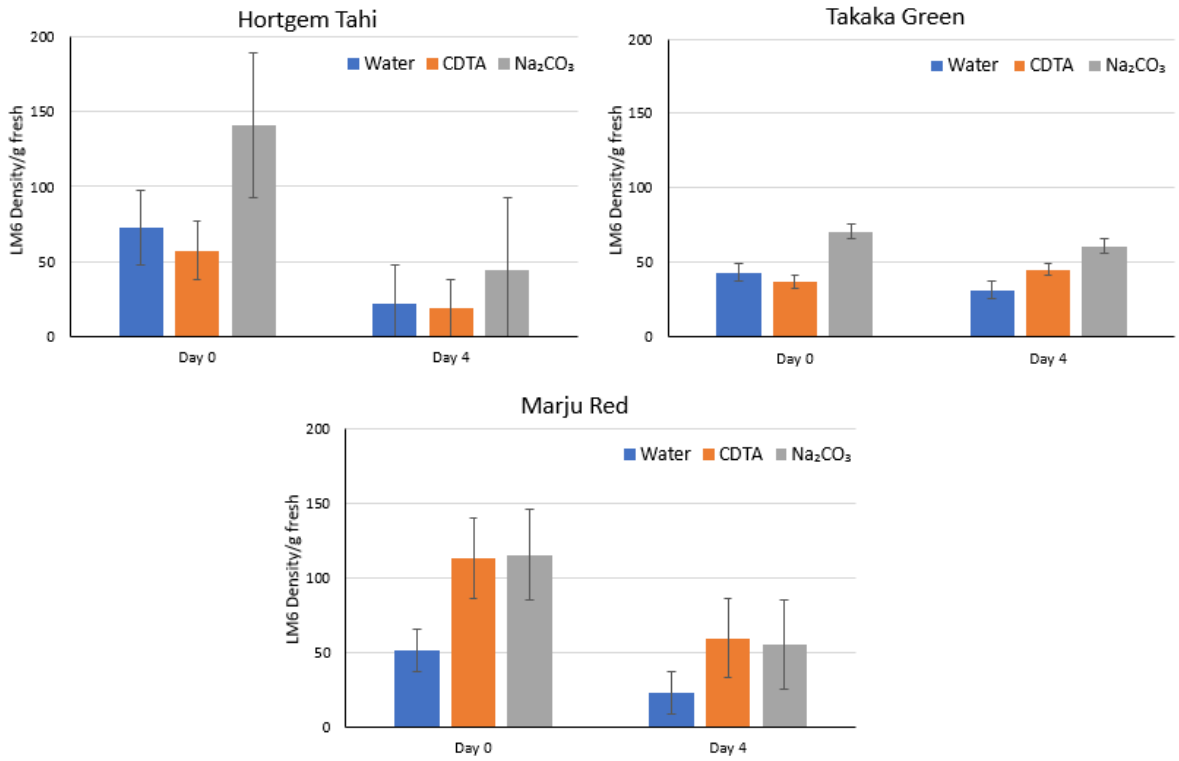


Figure 3.8 mAb LM6 binding (density/g fresh) in water-soluble, CDTA-soluble and Na_2CO_3-soluble extracts of Hortgem Tahi, Takaka Green and Marju Red kiwiberry.

The binding of LM6 is specific to a region of 1, 5-linked arabinose with typically 5-6 arabinose residues within an extended polymer backbone or branches (Verhertbruggen *et al*., 2009a; Willats *et al*., 1998). Like galactans (LM5) these epitopes tend to be present as pectin side chains. Figure 3.8 shows the labelling of LM6 in all extracts for all kiwiberry cultivars. In Hortgem Tahi and Takaka Green, the labelling of LM6 was highest in Na_2CO_3-soluble fractions, followed by water- and CDTA-soluble fractions. There was a decrease in LM6 labelling of all cultivars and all fractions between Day 0 and Day 4, although for Takaka Green these decreases were slight. Labelling of LM6 in Marju Red fractions presents a different LM6 profile compared to Hortgem Tahi and Takaka Green. LM6 binding in Marju Red is well presented in both CDTA- and Na_2CO_3-soluble fractions, followed by substantially less to water-soluble fractions (although note the relatively high errors associated with Marju Red samples). The binding of LM6 decreased from Day 0 to Day 4. The shortening of the 1, 5-linked arabinose side chain by arabinanases may reduce the binding affinity of LM6. This may contribute to the lower labelling of LM6 on Day 4 (Verhertbruggen *et al*., 2009a).

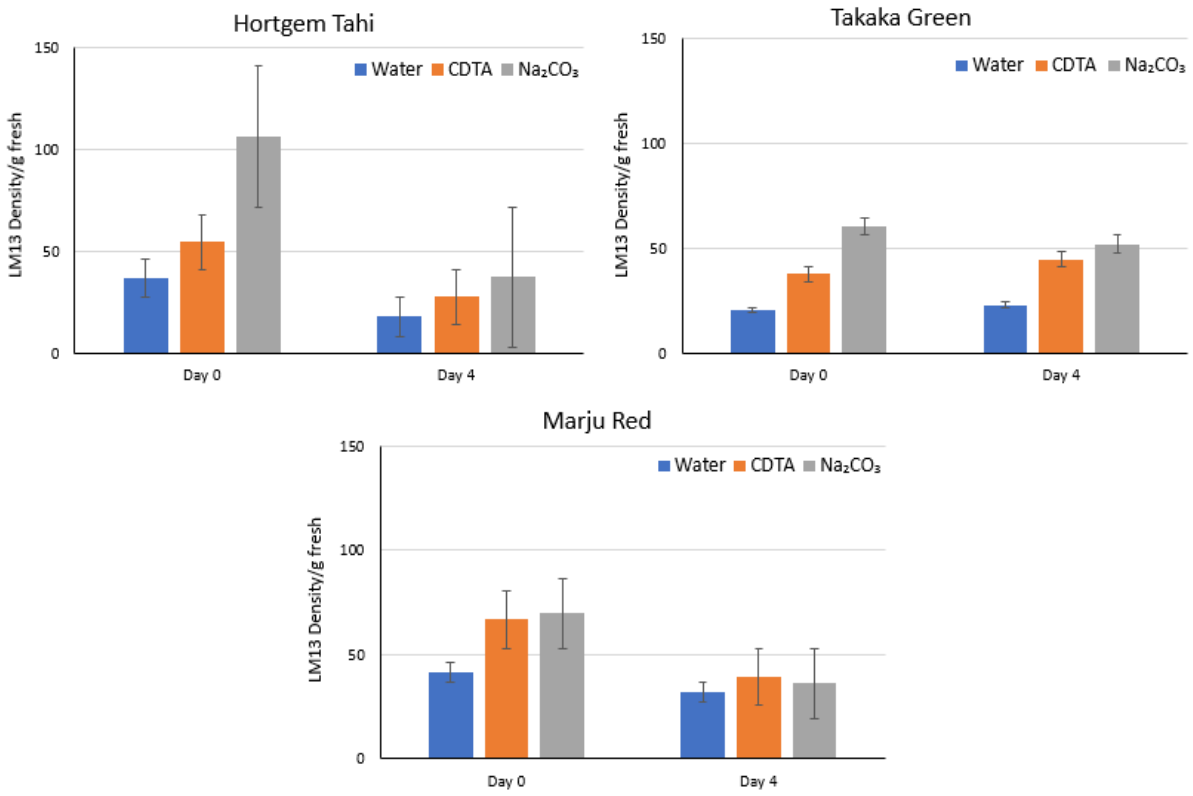


Figure 3.9 mAb LM13 binding (density/g fresh) in water-soluble, CDTA-soluble and Na_2CO_3-soluble extracts of Hortgem Tahi, Takaka Green and Marju Red kiwiberry.

LM13 binds to long stretches of 1, 5-linked-α-arabinosyl residues, such as linear arabinan (Verhertbruggen *et al*., 2009a). LM13 binds to a specific subset of pectic arabinans and to longer stretches of 1, 5-linked arabinosyl residues that are more likely to be abundant in unbranched arabinans. LM13 is also highly sensitive to the presence of arabinanase action (Verhertbruggen *et al*., 2009a) that would reduce the chain length. Figure 3.9 shows the presence of LM13 labelling in all soluble extracts of each kiwiberry cultivar but labelling tended to be highest in the Na_2CO_3-soluble extracts, decreasing between Day 0 and Day 4. The level of LM13 labelling in CDTA-soluble fraction in Marju Red was similar to that in Na_2CO_3-soluble fraction. The decrease in the labelling of LM13 may be due to the presence of kiwiberry pectin degrading enzyme, arabinanase, processing the longer chains (Verhertbruggen *et al*., 2009a; Wefers *et al*., 2018). The extent of decrease was most marked for Hortgem Tahi cultivar.

The binding of LM8 to pectin fractions gives insight into the presence of xylogalacturonan and highly substituted xylose (Mort *et al*., 2008; Willats *et al*., 2003).

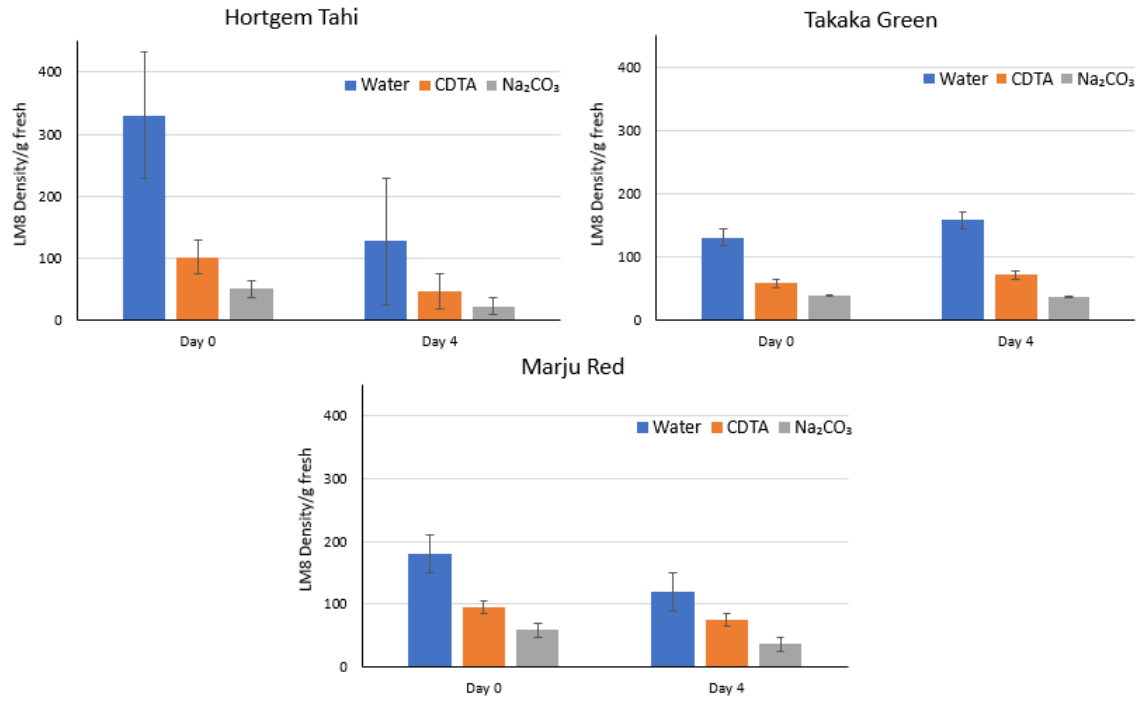


Figure 3.10 mAb LM8 binding (density/g fresh) in water-soluble, CDTA-soluble and Na_2CO_3-soluble extracts of Hortgem Tahi, Takaka Green and Marju Red kiwiberry.

Xylogalacturonan is associated with cell detachment and separation (Willats *et al*., 2003). Xylogalacturonan has a homogalacturonan backbone with ~50% of the galacturonic acid present possessing at least one xylose as a side group (Coenen *et al*., 2007). Figure 3.10 shows that there is LM8-positive xylogalacturonan in all pectin fractions, with the most found in water-soluble fraction, especially for Hortgem Tahi at Day 0. The labelling of LM8 decreased in all pectin fractions on Day 4 for Hortgem Tahi and Marju Red. For Takaka Green, however, the situation was slightly different: water-soluble and CDTA-soluble pectins increased in LM8 binding from Day 0 to Day 4. The solvents used for pectin extraction do not specifically target solubilisation of xylogalacturonan. Thus, xylogalacturonan might have been covalently bound to either a homogalacturonan or a rhamnogalacturonan I structural element in the three kiwiberry cultivars and co-extracted (Coenen *et al*., 2007).

Each New Zealand kiwiberry cultivar had different rates of firmness change, and different uronic acid contents, and pectin structures present. Changes in firmness (measured by compression tests) of the fruits over four days showed each cultivar softened between Day 0 and Day 4 at 20 ˚C, with Hortgem Tahi still steadily softening on Day 4 at 20 ˚C (Figure 3.1). This indicated there were differences in fruit ripening stages among the cultivars. This difference was also consistent with the report by Sutherland *et al*., 2017, which showed variability of whole fruit and core firmness of Hortgem Tahi, likely due to the different maturing range the fruits on the vine. The fruit softening during the ripening, results from the cell wall undergoing modifications, particularly pectin solubilisation (Paniagua *et al*., 2014). These changes are caused by the action of enzymes. (Tan *et al*., 2018; Wang *et al*., 2018). For example: endo- and exo-polygalacturonase cleaving HG; pectin methylesterase, removal of methyl ester groups from HG and β-galactosidase, removal of galactan side chains of RG-I (Brummell, 2006). Extracting solubilised pectin from kiwiberry plant cell wall extracts (Figure 3.2) provides information of the changes occurring to the pectin structure during fruit ripening. In addition, the semi-quantitative glycan array data (Figure 3.4 – 3.10) gives further information about the presence of heterogeneous glycans present within the soluble pectin fractions (Moller *et al*., 2012). The pectin solubilisation and glycan array results also show changes over time, and differences between, the three cultivars.

Initially, Hortgem Tahi contained low amounts of water-soluble pectin, which increased at Day 4. Takaka Green data for water-soluble pectin was similar to Hortgem Tahi. In contrast, the amount of Marju Red water-soluble pectin decreased by Day 4. This suggests that Marju Red was further along the lines of breaking down, or has ripened earlier than the other two cultivars. There are also noticeable differences in the relative presence of pectin esterification epitopes (LM19 and LM20; Figure 3.4 and 3.5) and the presence of galactan (LM5, Figure 3.6) and arabinan (LM6, LM13, Figure 3.7 and Figure 3.8) side-chain epitopes. These show that the cultivars have different pectin compositions before and during ripening, though this is not unusual, as several genotypes of kiwifruit possess a spectrum of softening behaviours (Fullerton *et al*., 2020; Wang *et al*., 2012; White *et al*., 2005). In addition, the presence of xylogalacturonan (LM8, Figure 3.9) in all soluble-pectin fractions for all cultivars is notable, as this pectic structure is only present in a low level in fruit. It is also not usually found in cell wall of ripening fruits (Thibault & Ralet, 2001; Tucker *et al*.,

2017), as well as peas, soybeans, baobab, apples, pears and onion (Dimopoulou *et al*., 2021; Zandleven *et al*., 2007).

Collectively, the data shows differences between cultivars Hortgem Tahi, Takaka Green and Marju Red – in their fruit firmness, pectin solubilisation and their glycan composition over time. The data suggests that Hortgem Tahi may have been less ripe than the other two cultivars at harvest, while Marju Red ripening was more advanced. The variability among the ripening stages between the three kiwiberry cultivars is consistent with the overall findings by Sutherland *et al*., 2017, despite their research focusing on the kiwiberry core. Sutherland *et al*., 2017 identified that the general softening of kiwiberry Hortgem Tahi differs from *A. chinensis* var. *deliciosa* 'Hayward' (green kiwifruit) in more than just the rate of ripening. Further studies into neutral sugar analysis of uronic acid samples, immunolabelling of plant cell wall structures of Takaka Green and Marju Red fruits and molecular weight analysis would confirm the similarities and differences between Hortgem Tahi, Takaka Green and Marju Red. In addition, there appears to be marked differences among the three kiwiberry cultivars in the levels of various pectin-processing enzymes expressed during ripening. For example, the binding of LM6, which is specific to a extend regions of 1, 5-linked arabinose, shows different patterns among the different cultivars with respect to time and soluble fraction. Similarly, the binding of LM18, which binds to highly methyl-esterified homogalacturonan, shows different patterns over time among the cultivars and soluble fractions

The pectin structures identified here have previously been found to contribute to human gut health (Larsen *et al*., 2019; Parker *et al*., 2010; Wu *et al*., 2021). For human gut health, the data here represent a preliminary framework around the types of pectic fractions (including xylogalacturonan) of three kiwiberry cultivars that would be involved in digestion and fermentation in the human gut.

3.3 Conclusion

All three kiwiberry cultivars displayed different firmness, uronic acid contents and glycan epitope profiles at harvest and when ripe. While the softening trajectory is somewhat similar among the three cultivars, the rate of the softening varied among the cultivars. Pectin solubility (determined by uronic acid content) and the types of pectin structures present (shown by glycan arrays) also changed over time and between the three cultivars. Therefore, Hortgem Tahi, Takaka Green and Marju Red may have different pectin compositions, possibly the result of different levels of expression of pectin-processing enzymes, during ripening. Additional studies would provide the reasons why the fruits have these differences.

Chapter 4 Isolation and purification of protein(s) associated with pectin degradation

4.1 Introduction

Monoglobus pectinilyticus, described in Kim *et al*., 2017, is the only human gut bacterium so far to possess a glycobiome that specifically targets pectin for degradation and utilisation. Although *M. pectinilyticus* does not utilise D-galactose and β-1, 4-galactan, it is still capable of degrading β-1, 4-galactan during its growth (Kim *et al*., 2019). Up-to-date databases were used to search for genes encoding potential β-galactanase within the near-complete genome of *M. pectinilyticus*, but no such sequences were identified (Kim *et al*., 2019). This suggests that *M. pectinilyticus* may produce a novel class of β-galactanase and/or β-galactosidase that has never been identified and/or characterised before. Previous study has identified that *M. pectinilyticus* encodes numerous proteins containing S-layer homology (SLH) modules and 108 putative CAZymes (Kim *et al*., 2017). SLH proteins are studied in their variety of functions, such as bacterial protection from pathogens (Martinez *et al*., 2012), fusion with other enzymes for specific interactions (Ilk *et al*., 2011), and constituents of novel lignocellulolytic degradation systems in both *Paenibacillus curdlanolyticus* B-6 and *Caldicellulosiruptor spp*. (Conway *et al*, 2016; Ratanakhanokchai *et al*., 2013). Due to the SLH involvement in lignocellulolytic degradation, analogous *M. pectinilyticus* SLH proteins may represent constituents of a novel pectin degradation system (Kim *et al*., 2017). The CAZymes found in *M. pectinilyticus* are spread widely across 34 CAZyme families, with 47 of these CAZyme domains predicted to have pectinolytic functions (Kim *et al*., 2017). The enzymes produced are putatively anchored via SLH domains or secreted into the environment (Kim *et al*., 2017). However, the exact role of SLH proteins and their involvement in pectin degradation have yet to be explored (Kim *et al*., 2017). Based on the assumption that proteins bearing SLH-domains are also involved in the degradation of pectin, *M. pectinilyticus* was grown in a medium containing complex pectin and other substrates (Kim *et al*., 2017).

This chapter presents the results investigating the possible protein(s) produced by *M. pectinilyticus* that possess β-galactosidase/β-galactanase (abbreviated BGAL) activity. We hypothesised that adding these complex pectin structures and additional substrates, such as potato galactan, will enable *M. pectinilyticus* to produce galactan-degrading enzymes in a greater quantity than cells growing in simple sugars (e.g. fructose as sole carbon source). Thus, *M. pectinilyticus* was grown in MM (as described in section 2.2.3) with added substrates (0.5% w/v citrus pectin and/or 0.5% w/v potato galactan). The study of proteins typically includes explorations into the physical and chemical protein properties that determine their function (Simpson, 2004). Therefore, a series of methods was used to extract and purify *M. pectinilyticus* proteins from growth media. Ultra-filtration, SDS-PAGE, ammonium sulfate precipitation and anion exchange were used to separate proteins based on their size (kDa), protein solubility and protein net surface charge (Simpson, 2004). To identify the numerous uncharacterised proteins expressed during pectin degradation, protein sequencing by LC-MS/MS was ideal. Figure 4.1 provides a workflow of the sample processing, purification and sequencing analyses.

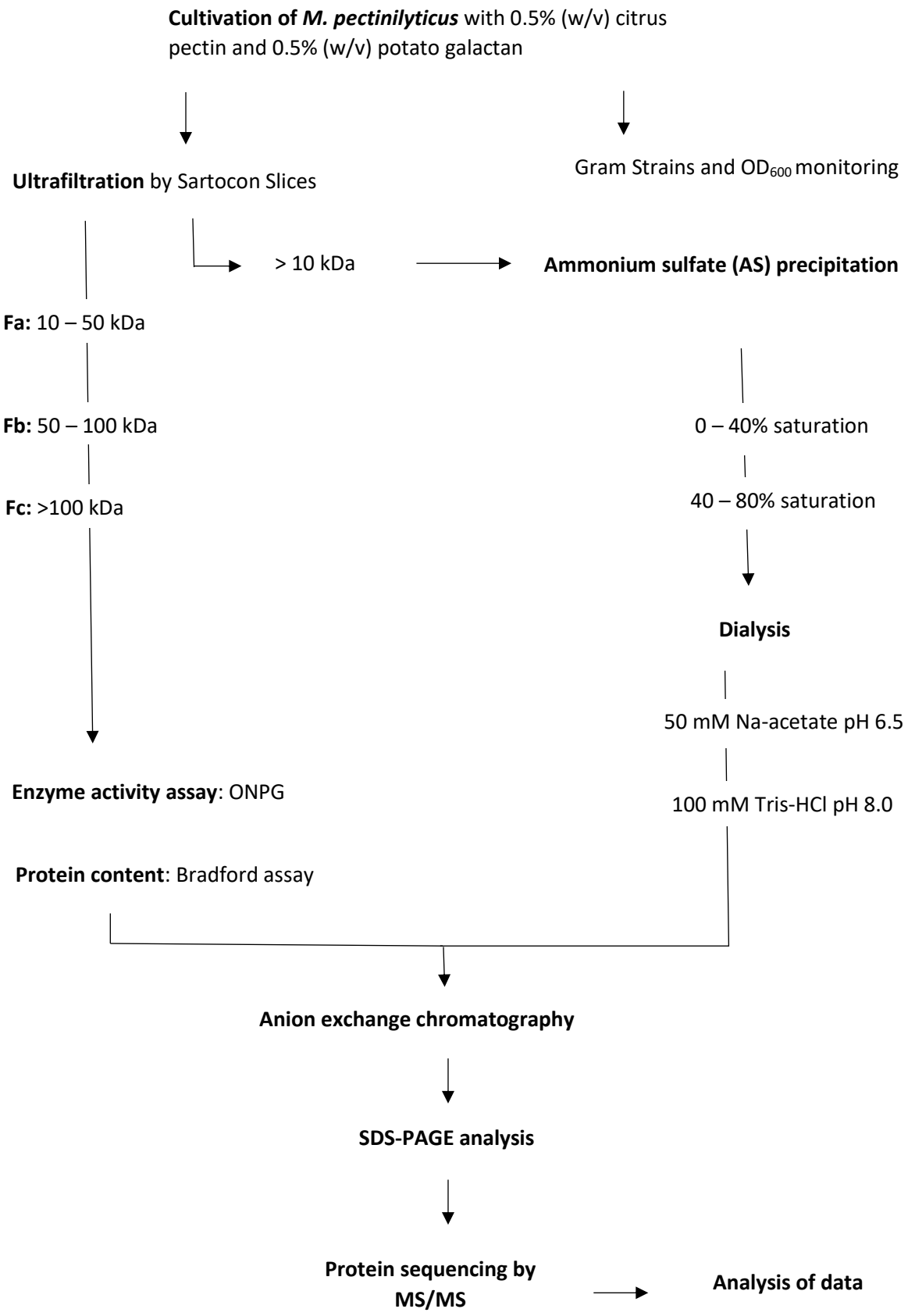


Figure 4.1 Workflow diagram of sample processing, purification and sequencing analysis.

4.2 Results and Discussion

4.2.1 Microbiology

Cultivation of *M. pectinilyticus* within MM containing added substrates, in both Hungate tubes and serum vials (50 mL and 1 L) under strictly anaerobic conditions, had a slow growth time (Kim *et al*., 2017). The addition of clarified rumen fluid, vitamin K solution and specific substrates were required to help the growth of *M. pectinilyticus*. According to Kim *et al*. 2017, maximum optical density ranged from 0.2 – 0.3. In this study, we observed a maximum OD_{600} of 0.68 after two weeks of growth.

The strain grew relatively well at 37 ˚C (pH 7.0), with constant shaking. Gram staining and OD_{600} readings were routinely performed to check for culture contamination and bacterial growth (Figure 4.2A). Gram-positive cocci are found either singly or in pairs. The bacteria, though non-motile, adhere to the pectin-rich substrate for degradation (Figure 4.2B).

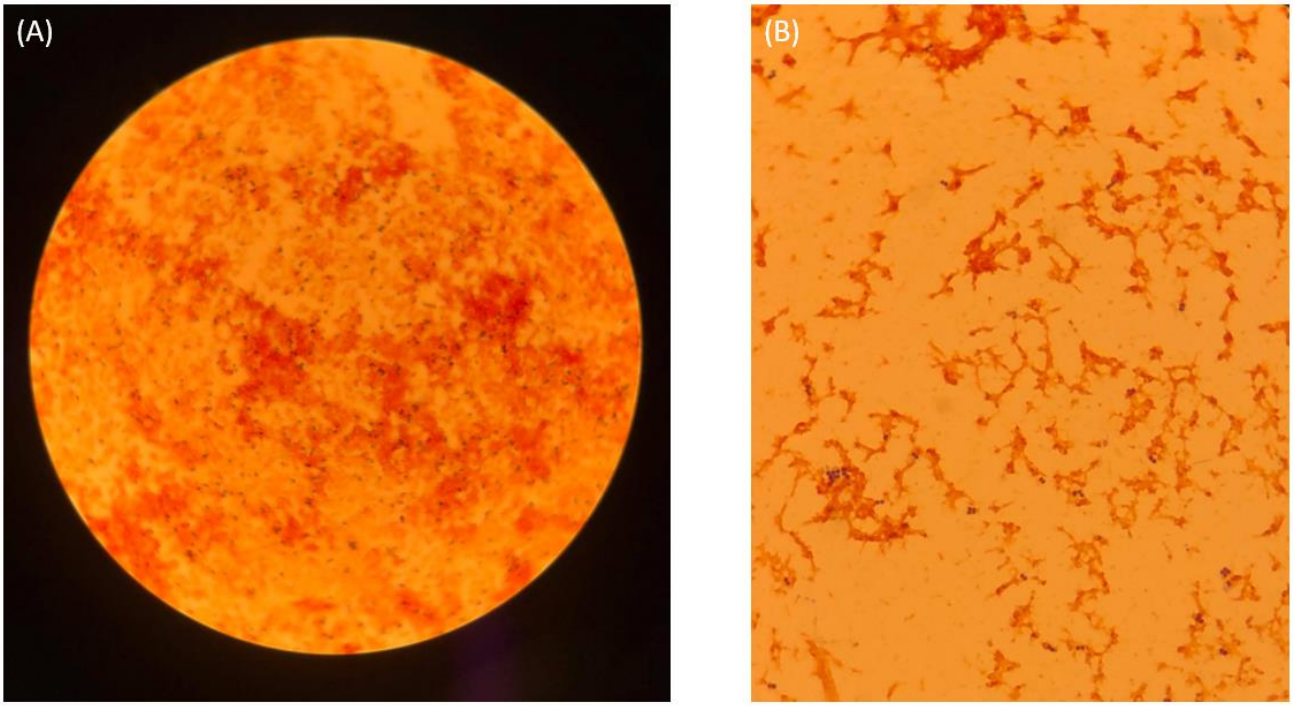


Figure 4.2. Gram stains of Gram-positive *M. pectinilyticus*. (A) 100x magnification. (B) 400x.

4.2.2 Protein content, enzyme activity and ion-exchange chromatography

The results for protein content and enzyme activity are summarised in Table 4.1. Ultrafiltration successfully separated proteins based on their size (kDa) into three different fractions, checked by SDS-PAGE. Predominantly, fraction Fc/Fc1 displayed BGAL activity, with lower molecular-weight fractions Fb and Fa displaying very little to no activity (section 2.4.5).

Table 4.1 Summary of results for ultrafiltration, protein content and enzyme activity.

	Culture	**Retained**	**Fa**	**Fb**	***Fc**	**Fc1**	**S/N**	**MM**
MWCO	-	>0.1 μm	10-50 kDa	50-100 kDa	>100 kDa	>100 kDa	<10 kDa	-
OD_{600}	0.46	-	-	-	-	-	-	0.00
Enzyme activity (μmol min^{-1} mg^{-1})	-	0.039	0.00	0.006	0.026	0.220	0.00	0.00
Protein content (mg/mL)	-	1.090	0.649	0.570	7.710	0.590	0.289	0.491

Note: S/N: supernatant; Blank control: MM: mineral media. MWCO: molecular weight cut-off. OD_{600}: Optical density range set at 600 for bacterial growth; * - Fc sample was destroyed during anion exchange chromatography, thus repeated and labelled Fc1. Fc will not be discussed, only used for reference.

The BGAL activity (as described in section 2.4.5) was highest in Fc1 compared to the other three fractions (Table 4.1). In contrast, Fb with an MWCO range of 50 – 100 kDa produced low enzyme activity. Therefore, it may be that either the protein or proteins responsible for β-galactosidase/β-galactanase (BGAL) activity are moderately large 50 – 100 kDa, or perhaps small multi-subunit proteins combine to mediate the enzyme activity. Fraction Fa with an MWCO of 10 – 50 kDa, although having high protein content, did not engage in BGAL

activity. This is consistent with the literature, as most BGAL enzymes described by bacteria are ~110 kDa (Matthews, 2005). Therefore, fraction Fa was not taken for further analysis.

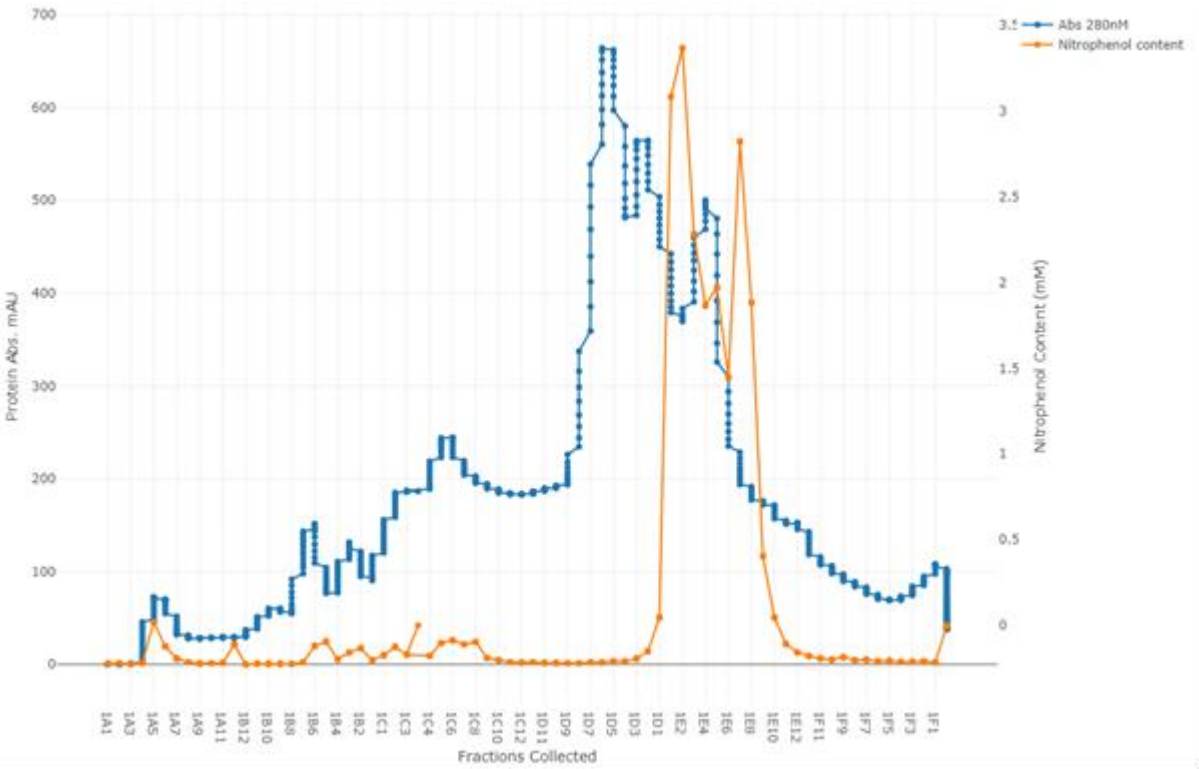


Figure 4.3 Analysis of Fb sample by anion exchange chromatography and eluted fractions tested for enzyme activity**.** A total of 5 mL sample was pumped through a Mono Q 5/5 column, with protein absorbance measured in mAU and subsequent fractions collected checked for BGAL activity by measuring the absorbance of by-product nitrophenol (mM).

Samples Fb and Fc1 were analysed by anion exchange chromatography as described in 2.4.5. Anion exchange of Fb observed poorly resolved peaks (Figure 4.3), with two distinct peaks covering a wide area. Injections occurred at fraction collection points of FV (before fraction collection), 1B11, 1B7, 1B4 and 1B1. Many unbound proteins ran straight through the column and eluted before applying a salt gradient. Salt elution started from fraction 1C4 onwards, where within one column volume (CV), NaCl concentration increased from 0 to 0.25 M NaCl (25%). An increase from 0.25 M to 0.5 M NaCl over 5 CV started from 1C6 to 1E10, whereas the last 50% increase to 1 M NaCl was applied over 1 CV. Majority of the bound proteins eluted over the range of 0.25 M NaCl to 0.50 M NaCl. Those unbound proteins may not have bound to the column due to the possibility of incorrect pH, incorrect buffer, or overloading of the column.

Enzyme activity measured in all collected fractions was assayed as described in 2.4.4. Only fractions 1D9 to 1E19 were able to hydrolyse *o*-nitro phenyl-β-D-galactopyranoside and were taken for further analysis.

Figure 4.4 Analysis of Fc1 sample by anion exchange chromatography and eluted fractions tested for enzyme activity**.** A total of 1 mL of sample was pumped through a Mono Q 5/5 column, with protein absorbance measured in mAU and subsequent fractions collected checked for BGAL activity by measuring the absorbance of by-product, nitrophenol (mM).

Anion exchange of sample Fc1 (Figure 4.4) observed significantly different results from sample Fb (Figure 4.3). Multiple prominent peaks were observed throughout the run. Though Fc1 was run identically as sample Fb, we observed unbound protein eluting almost instantly after injection. Unbound protein may result from an overloaded column or previously clogged column, incorrect buffer and pH used or large protein complex/aggregate. At the beginning, a salt gradient of 0 – 0.25 M was applied over 1 CV from 1B7 to 1B3, with low elution of proteins. An increase of salt, 0.25 M – 0.5 M NaCl, was applied over 5 CV from 1B2 – 1D1, with a significant amount of proteins unbinding. The remaining 0.5 – 1 M NaCl eluted off the remaining unbound proteins from 1D2 to 1F11.

Each fraction collected was assayed for BGAL enzyme activity (2.4.2). One significant peak was observed at 1A9 – 1A11, and one less resolved peak over a broader range, 1C4 – 1D8. The second less resolved peak contains one more prominent and two smaller peaks. In contrast, 1A9 – 1A11 contained proteins that did not bind to the column. These proteins may be leftover proteins from previous runs, resulting in contamination or, more likely, the column was already overloaded (Simpson, J. 2004).

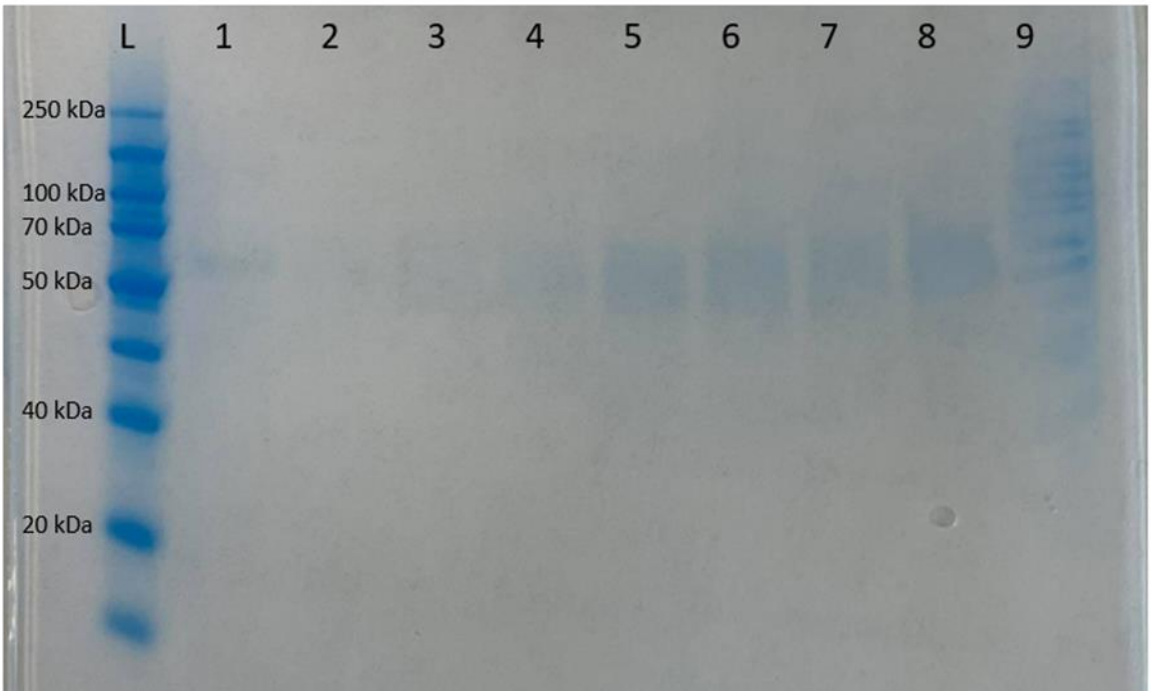


Figure 4.5 SDS-PAGE analysis of Fb purified by anion exchange chromatography identified with BGAL activity. Lane L was loaded with PageRuler unstained Broad Range protein ladder (ThermoFisher). Lanes 1 – 7 were loaded from fractions collected between E1 – E7 (Figure 4.3). Lane 8 was loaded with the original sample, Fb. Lane 9 was loaded with Mark12 standard. Minor visual differences among fractions.

A selection of fractions showing positive enzyme reactions from Fb and Fc1 was analysed by SDS-PAGE as described in 2.4.6. Fb gel shows smeary, very faint and potentially singular bands increasing in intensity from E1 – E7, with band sizes ranging from 50 - 70 kDa. The Fc1 gel shows multiple proteins eluted out within the fractions, with similar profiles and varying protein intensities.

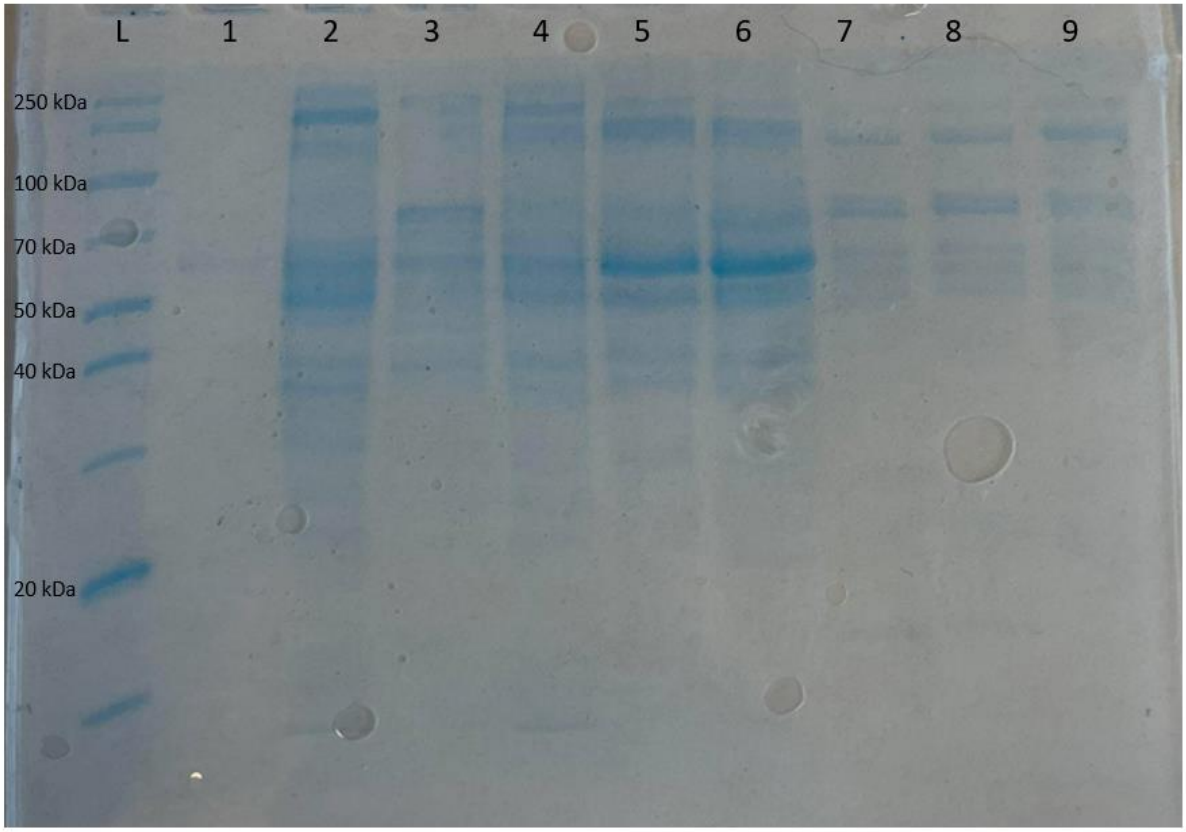


Figure 4.6 SDS-PAGE analysis of Fc1 purified by anion exchange chromatography identified BGAL activity. Lane L was loaded with PageRuler unstained Broad Range protein ladder (ThermoFisher). Lane 1 loaded with Fb E1 for comparison. Lanes 2 – 9 were loaded from fractions collected from the two prominent peaks displaying BGAL activity (Figure 4.4). Visible differences among Fc1 collected fractions were observed by electrophoresis.

Typically, β-galactosidases/β-galactanases are large proteins expected at or above the 100 kDa marker (Matthews, 2005). Noticeably, Figure 4.6 shows varying protein sizes (10 – 250 kDa), despite being ultra-filtrated by MWCO >100 kDa. Several factors may explain this discrepancy; (1) Protein modifications, such as glycosylation or phosphorylation; (2) Presence of protease, breaking down retained proteins; (3) Aberrant migration of protein due to the SDS interaction with the protein, (4) Fragmentation of bigger proteins into smaller units; and (5) disassembly of large protein complex into constituent proteins The denaturing ability of the detergent removes the secondary, tertiary and quaternary dimensional structure, transforming the protein to be uniformly negatively charged for migrating through the gel towards a positively charged electrode.

SDS-PAGE showed a variety of proteins that "may or may not" be the enzyme. Initially, Fc1 resulted in the highest enzyme activity after ultra-filtration but lost significant enzyme activity after anion exchange chromatography. However, Fb initially had lower enzyme activity after ultra-filtration, the fraction produced stronger enzyme activity after anion exchange. SDS-PAGE visualised a smeary single protein that is predicted to be responsible.

4.2.3 Protein purification

Based on previous results, culturing of *M. pectinilyticus* and ultra-filtration were repeated with modifications. Purification method, "salting out" by ammonium sulfate precipitation, was incorporated to clarify the fractionated samples (Figure 4.7). As previously, each stage was checked for protein content and enzyme activity, determining which samples were taken for subsequent analysis.

Table 4.2 Summary of protein purification.

	Culture	**Permeate**	**<10 kDa**	**A**	**B**	**C**	**D**	**MM**
MWCO	-	<0.2 µm	<10 kDa	8 –10 kDa	8–10 kDa	8–10 kDa	8–10 kDa	-
OD_{600}	0.680	-	-	-	-	-	-	0.00
Enzyme activity (µmol min^{-1} mg^{-1})	-	0.008	0.020	0.012	0.012	0.003	0.031	0.00
Protein Content (mg/mL)	-	0.341	0.589	0.119	0.261	0.208	0.459	0.00

Note: AS: ammonium sulfate; A: AS 0-40% Na-acetate pH 6.5; B: AS 40-80% Na-acetate pH 6.5; C: AS 0-40% Tris-HCl, pH 8.0; D: AS 40-80% Tris-HCl, pH 8.0. Blank control: MM: mineral media. MWCO: molecular weight cut-off. OD600: Optical density range set at 600 for bacterial concentration.

Only two ultra-filtration MWCO were used: 0.2 µm pore with permeate concentrated down to 200 mL using 10 kDa MWCO filters. Aliquots of MM, permeate and <10 kDa were analysed as controls (Table 4.2). The filtrate was processed as described in 2.4.2 and 2.4.3. Most samples produced relatively low enzyme activity whereas sample D (40-80% ammonium sulfate) produced the most activity and protein content.

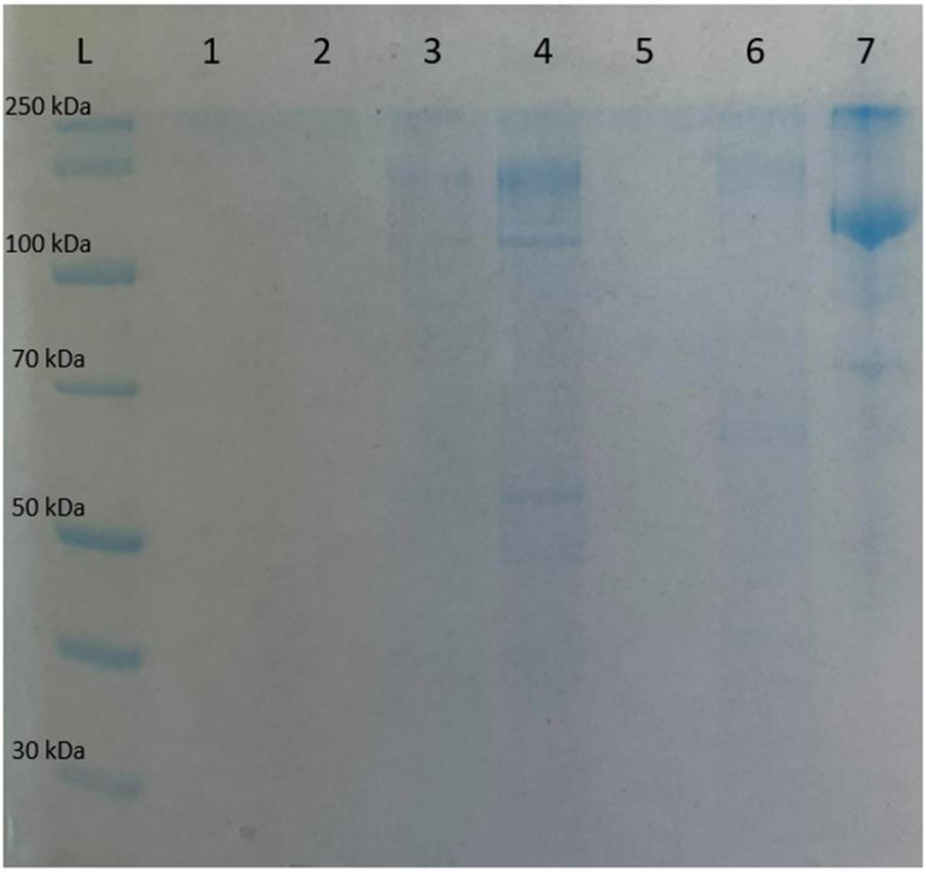


Figure 4.7 SDS-PAGE analysis of ammonium sulfate samples dialysed in respective buffers before being applied to anion exchange chromatography**.** Lane L was loaded with PageRuler -unstained Broad Range protein ladder (ThermoFisher). Lanes 1 -2 were loaded with 20 mM Na-acetate dialysate samples, 0 – 40% AS and 40 – 80% ammonium sulfate. Lane 3 -4 were with 100mM Tris-HCl dialysate samples and, respectively, 0 – 40% and 40 – 80% ammonium sulfate. Lane 5 was loaded with 100 mM Tris-HCl as a control. Lane 6 was loaded with Mark12, and Lane 7 was loaded with β-galactosidase from *E*. *coli* (E-ECBGAL, Megazyme). Visible differences were observed in AS fractions by electrophoresis.

SDS-PAGE analysis of ammonium sulfate samples of both buffers visualised two different results. Both saturation ranges for Na-acetate showed very low protein content, low enzyme activity and no protein band was visualised (Figure 4.7). The difference in saturation and dialysis did not retain or purify any significant protein or activity. In contrast, 40 – 80% ammonium sulfate, dialysed with Tris-HCl, resulted in positive enzyme activity with low protein content (mg/mL). Three definite bands and smeary lower (kDa) bands were visualised with a ~100 kDa band showing similarity to lane 7, *E. coli* β-galactosidase (E-ECBGAL, Megazyme).

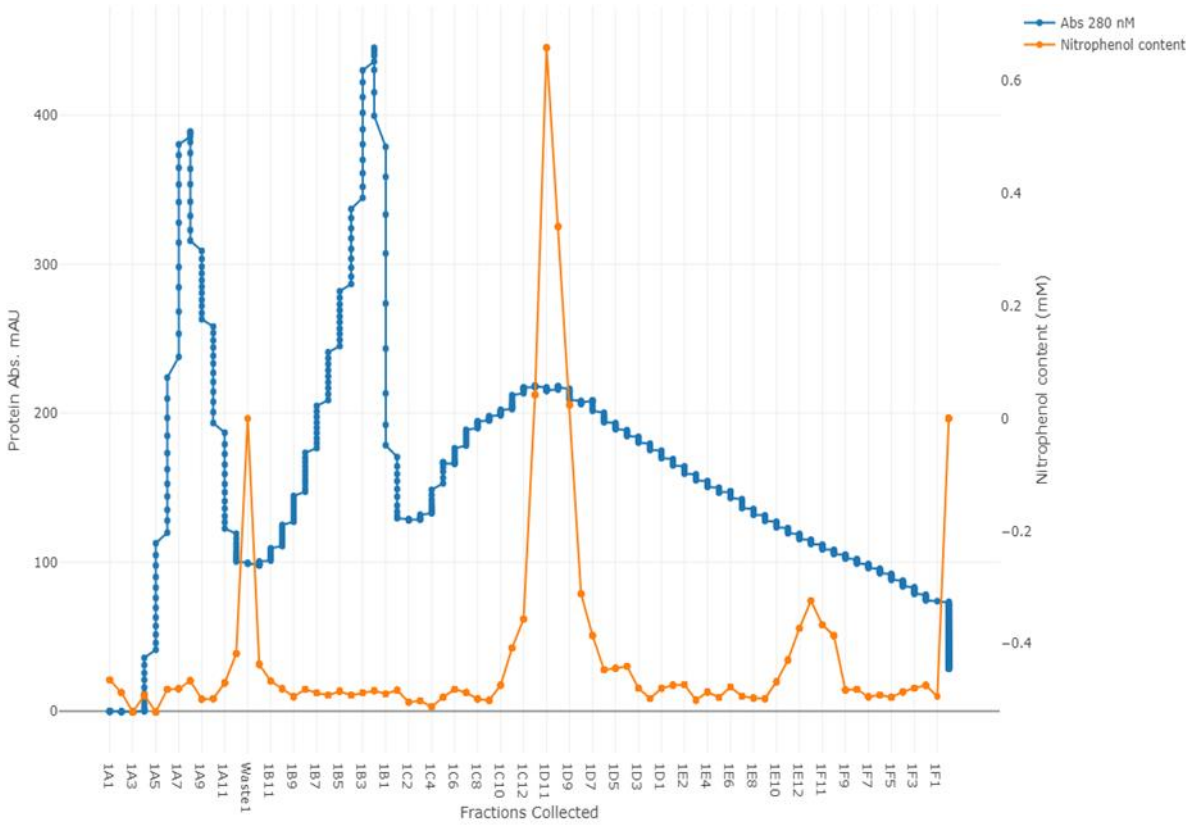


Figure 4.8 Analysis of ammonium sulfate sample dialysed in 100 mM Tris-HCl further purified by anion exchange chromatography and eluted fractions tested for enzyme activity. A total of 2 mL of sample was pumped through a Mono Q 5/5 column, with protein absorbance measured in mAU and subsequent fractions collected checked for BGAL activity by measuring the absorbance of by-product, nitrophenol (mM).

Subsequently, 40 – 80% Tris-HCl pH 8.0 was analysed by anion exchange chromatography. Three prominent peaks were observed (Figure 4.8): Two well-defined peaks and one less defined broader peak. Unlike Fb (Figure 4.3) and Fc1 (Figure 4.4), anion exchange was run with a steady increase of NaCl concentration from 0 – 100% over 10 column volumes, collecting in 200 µL fractions. The conditions here were altered due to the incorporation of ammonium sulfate in the protein purification method and not knowing where the protein may elute. Thus, a salt gradient was applied from 1B12 until the end of the run with 1 M NaCl achieved at 1F11. Similarly to Fb and Fc1, Figure 6.8 shows unbound protein passing through the column before applying a salt gradient. All fractions were tested for BGAL activity, with only one very distinct peak from fractions C10 – D8 (Figure 4.8). The single peaks for 1E10 – 1F10 and waste are at 0.00 mM NaCl.

4.2.4 Mass spectrometry (MS/MS) sequencing by electrospray

Due to the variety of protein bands present in collected fractions (Figure 4.5, 4.6 and 4.7) which contained BGAL activity (Figure 4.3, 4.4 and 4.8), ten samples were sent for protein sequencing (Figure 4.9 and 4.10). Majority of proteins sent for sequencing were from Fc1 samples, due to the variety. Sequencing was performed by tandem MS (LC-MS/MS) after enzymatically digested gel bands to produce fragmented peptides (Proteomics International, 2022). SwissProt is a curated protein sequence database that strives to provide a high level of annotated sequences. The database distinguishes itself by three criteria: annotation, minimal redundancy and integration with other databases (Bairoch & Apweiler, 1996). UniProt database is a freely accessible database of protein sequences and is largely un-reviewed (The UniProt Consortium, 2017).

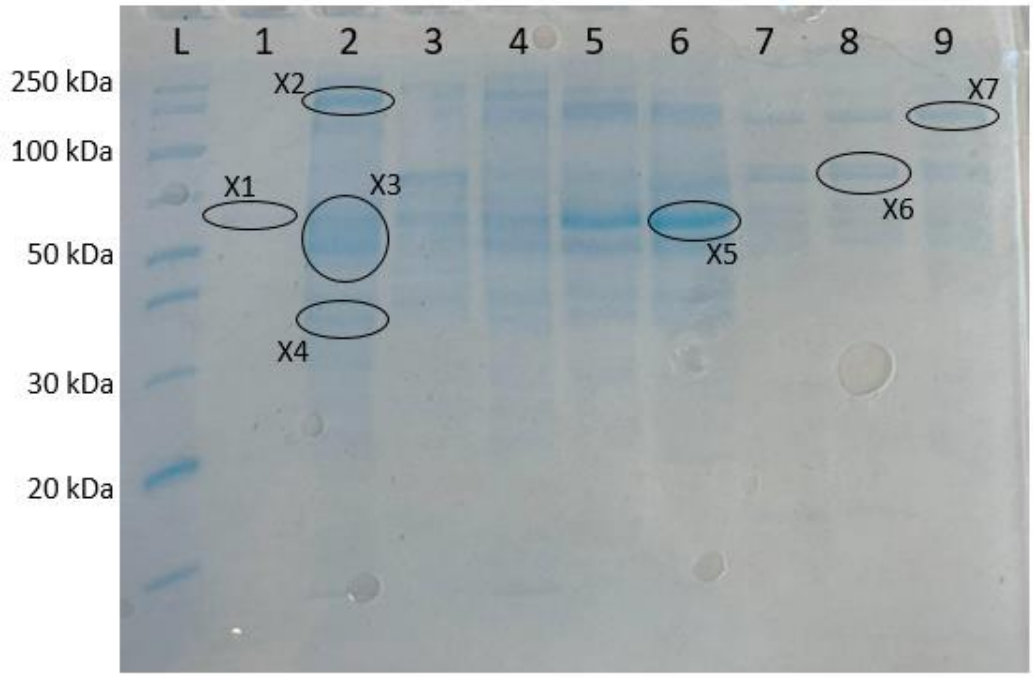


Figure 4.9 SDS-PAGE bands excised for protein LC-MS/MS sequencing**.** Lane L was loaded with ladder. All samples loaded from Lane 2 – 9 were positive for BGAL activity. Lane 1 loaded with Fb E1 fraction. Lane 2 loaded with Fc1 A10. Lane 3 loaded with Fc1 D5. Lane 4 loaded with Fc1 A11. Lane 5 loaded with Fc1 C4. Lane 6 loaded with Fc1 C5. Lane 7 loaded with Fc1 C10. Lane 8 loaded with Fc1 C11 and Lane 9 loaded with Fc1 D12.

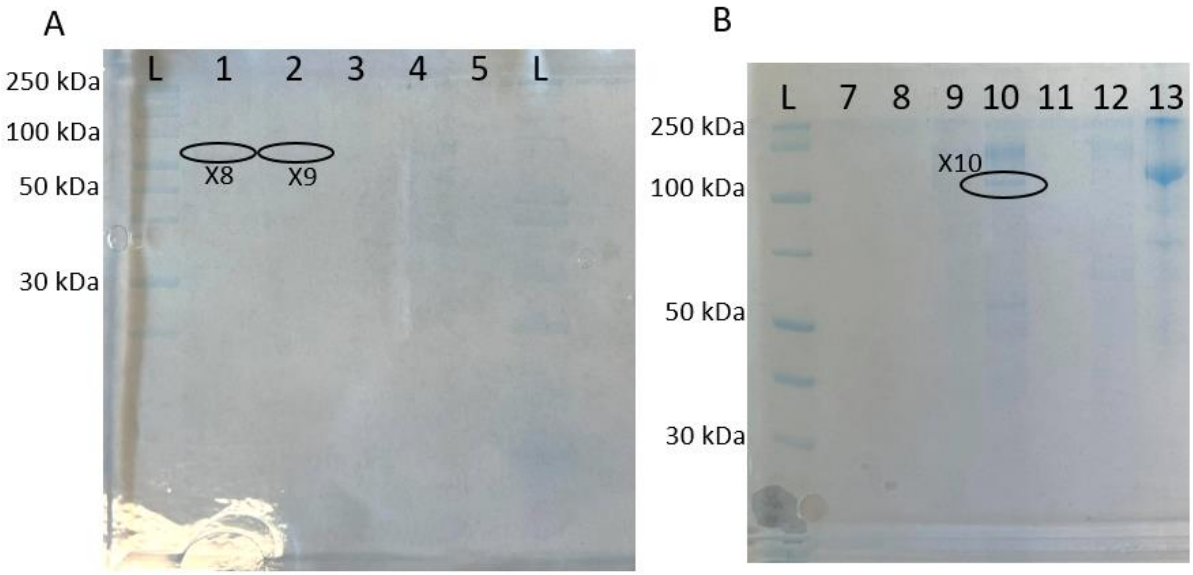


Figure 4.10 SDS-PAGE bands excised for protein LC-MS/MS sequencing**.** Lane L was loaded with ladder. All samples loaded from Lane 2 – 13 are from AIX, loaded with 40 – 80% AS, dialysed in 100 mM Tris-HCl, pH 8.0 and were positive for BGAL activity. Lanes 1 and 7 were D9. Lanes 2 and 8 were loaded with D10. Lanes 3 and 9 were loaded with D11. Lanes 4 and 10 were loaded D12. Lanes 4 and 11 were with 100 mM Tris-HCl, pH 8.0. Lane 5 and 12 were incorrectly loaded with Mark12. Lane 13 were loaded with *E. coli* β-galactosidase/

Table 4.3 summarises the protein fragments identified from MS/MS that were queried against two databases; (1) SwissProt filtered to bacteria only and (2) Universal Protein Database (UniProt), specifically against *M. pectinilyticus*. Top hits were ordered by the top score, which acknowledges the amount of alignment between queries and protein sequence. Several queries match, and b/y ions were also recognised and incorporated into finalising the results. Most of the samples found top matches that helped identify the suspected protein. Only sample X1 resulted in inconclusive results. Two samples, X5 and X10, matched to enzyme β-galactosidase, though of different domains. Sample X3 and X6 gave consistent results between SwissProt and UniProt databases, solidifying the identification.

Table 4.3 Summary of proteins sent for sequencing by Proteomics International laboratories LTD.

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10
SDS-Page MW	60 kDa	150 kDa	50 kDa	30 kDa	60 kDa	70 kDa	150 kDa	60 kDa	60 kDa	120 kDa
Sprot_all bacteria top hits	No hits	No hits	Uronate isomerase, *B. licheniformis*	1. L-Lactate dehydrogenase, *B. licheniformis* 2. Malate dehydrogenase, *B. velezensis*	β-galactosidase, *Planococcus sp.*	60 kDa chaperonin, *E. sibiricum*	60 kDa chaperonin, *B. thuringiensis*	No hits	Elongation factor G, *Bacillus sp*.	β-galactosidase *E. coli* O6:K15:H31
UniProt_*M. pectinilyticus* top hits	No hits	1. DNA helicase B9019_01364 2: S-Layer domain protein, B9019_01597	1. Uronate isomerase 2. S-layer domain-containing protein, B9019_01597	1. Uncharacterised protein, B9019_02288 2: Pectate Lysase Fam 1/S-layer domain-containing protein, B9019_00909	1. Transcriptional regulator, B9019_00547 2. S-layer domain-containing protein, B9019_01597	60 kDa chaperonin	1. Uncharacterised protein, B9019_02288 2. Uncharacterised protein, B9019_00805	1. Glucose-6-phosphate isomerase, B9019_00362	1. Propeptide PepSY and peptidase M4 2. Glucose-6-phosphate isomerase	1. S-layer domain-containing protein, locus_01597 2. pectin esterase family protein, locus_01626

Notes*:* top two hits presented are labelled 1 and 2, respectively.

While Table 4.3 shows a variety of protein results, few stood out as potential candidates for possessing BGAL catalytic capabilities. S-layer domain-containing protein B9019_01597 appeared in many protein hits, a possibility from dominated protein fragments present in many of the samples. Protein 01597 is large (3,174 amino acids) with only an S-layer homology domain identified between 1 – 152 amino acid residues and the rest largely unknown.

The top protein match for sample X10 (β-galactosidase OS=Escherichia coli O6:K15:H31) amino acid sequence was taken and blasted against *Monoglobus pectinilyticus* specifically. This resulted in only one protein: AUO18237.1 S-layer domain-containing protein 0050, with a query cover of 2%. The query cover blasted against the whole NCBI non-redundant (nr) database and matched 100% query and identity to numerous β-galactosidase enzymes and hypothetical proteins from bacteria. The overlapping query peptide was DLPLSDMYTPYVFPSEN-GLRCGTRELNYGP. Therefore, protein 0050 corresponds to sample X10 (Figure 4.10 and Table 4.3).

4.2.5 *de novo* sequencing

Interestingly, even though sample X1 was the only band present in Fb that produced BGAL activity (Figure 4.4, 4.5), the sample failed to match any proteins available on the SwissProt database and against *M. pectinilyticus* (Table 4.3). Therefore, sample X1 was subjected to *de novo* sequencing by Proteomics International.

Analysis of Table 4.4 identifies a potential BGAL candidate – glycosyl hydrolase family 3 (GH3) 00398. However, the second hit for peptide 1, glycosyl hydrolase family three (exo-acting enzymes), are a widespread group that hydrolyses the glycosidic bond between two or more carbohydrates (Macdonald *et al*., 2015). GH3 catalytic site found between 1 – 360 amino acid residues, larger sections of amino acid of 00398 are mainly unknown. Thus, it is possible that protein could be a dual-functioning enzyme with galactanase/arabinose or xylosidase activity.

Table 4.4 *de novo* sequencing of sample X1 blasted results against *M. pectinilyticus*.

Peptide Sample	Peptide fragment	Confidence (%)	Blastp Results
1	LASYLDK	99.0%	1. Histidine-tRNA ligase 00129 2. Glycosyl hydrolase family 3 00398
2	YALPLEYGLLK	99.0%	1. Maltodextrin glucosidase 00229 1. Hypothetical protein 01257
3	LAADQTLR	98.2%	1. Hypothetical protein 000733 2. type-II-A-CRISPR 00444
4	TLLAEGLK	89.2%	1. Chaperonin GroEL 00602 2. DNA-directed RNA-polymerase 01445
5	RPLKKSLSLSPGK	70.3%	1. SafA/ExsA family spore 00076 2. von Willebrand factor type A domain/S-layer domain-containing protein 00543
Combine 1 - 2	LASYLDKYALPLEYGLLK	N/A	No results
Combine 1 - 5	1 - 5	N/A	No results

Note: fragments were blasted against the NCBI nr database with the top two hits recorded, specifically screening against *M. pectinilyticus*. Confidence (%) indicates confidence in each amino acid assignment. Confidences below 80% shown are used with caution.

4.3 Conclusion

Proteins involved in pectin degradation were isolated and purified. β-Galactosidase/β-galactanase activity was identified in two samples. These samples differed by MWCO and enzyme activity. Numerous proteins were extracted and purified in samples with BGAL activity. Attempts were made to narrow these proteins to no prevail. Thus, a variety of proteins was sequenced and analysed against two databases for identification and function. Four candidate proteins were identified. Due to time restraints, only protein 0050 (X10) was taken for further analysis. The work here provides a starting framework to isolate out the proteins involved in *M. pectinilyticus* pectin

degradation. In addition, the methods here can be used to tailor the expression of other potentially novel proteins.

Chapter 5 Investigations into the characterisation of wild-type 0050 and its variants

5.1 Introduction

Expressing target protein identified from initial screenings in recombinant systems, allows the pure protein to potentially be obtained in large quantities. Typically, recombinant proteins are expressed with a fusion tail or 'tag', such as a poly-histidine. The tag chosen determines the purification method and subsequently, characterisation methods. Affinity chromatography, X-ray crystallography, NMR spectroscopy, enzyme activity, ligand binding assays, and computational analysis are some methods to characterise an unknown protein.

This Chapter present attempts to characterise of protein 0050. The work is derived from the possibility of protein 0050 possessing BGAL enzyme capabilities. It was hypothesised that protein 0050 contains the catalytic ability of β-galactosidase/β-galactanase and is involved in pectin degradation. The use of computational protein structure analysis was considered ideal due to the capability of the programming to compare similar structures and produce a highly confident evaluation. Therefore, three computational tools, PsiPred/pGenThreader (Lobley *et al*., 2009), domain association by PISA analysis (Krissinel & Henrick, 2007; Krissinel, 2010) and neural networking scripts, AlphaFold (Jumper, 2021) were used. Cloning and expressing a target protein is the standard method to obtain large quantities of pure protein. Histidine-tagged wild-type protein 0050 and two truncated variants P1 and P2 which were synthesised and expressed in *E. coli* systems, with immobilised metal affinity (IMAC) used for protein purification. Numerous protein characterisation methods were incorporated together to investigate the physiological and biochemical attributes of proteins (Simpson, 2004). Here, BGAL enzymatic activity using a colorimetric assay (as described in 2.4.5) and ligand-binding assays (as described in 2.5.5) were used to identify some characteristics of protein 0050.

PsiPred and AlphaFold show that protein 0050 consists of two main domains, a beta-sheet N-terminal domain and an alpha-helix C-terminal domain, with an unusually long 90 amino acid loop connecting the two domains, with a repeating motif of 12 glycine. Two variants of protein 0050 were constructed, in one the hydrophobic N-terminal helix was excised and in the other this helix and the unusual glycine-containing loop were excised to express soluble protein from which

protein function might be determined. Enzymatic assays were repeated to test for β-galactosidase/β-galactanase activity, with no identification for either activity. Binding assays using bilayer interferometry (BLItz Pro software (Forte Bio, Inc.)) determined the interaction of the truncated 0050 variants binding to various pectin fragments. Initial results identified a change in binding (nm) with substrates citrus pectin, linear 1, 5-α-L-arabinan, D-(+)-xylose and di-galacturonic acid. Preliminary enzyme kinetics data identified protein association and dissociation with citrus pectin, linear 1, 5-L-α-arabinan and potato galactan.

5.2 Results and Discussion

5.2.1 Preparation of protein 0050 for expression

Genomic DNA (gDNA) from *M. pectinilyticus* was extracted and concentrated by ethanol precipitation from bacterial cell pellets collected from MM broths (2.5.1). DNA was extracted in two sets of duplicates and pooled, with DNA concentration checked by QIAxpert (Table 5.1). Previously extracted *M. pectinilyticus* gDNA (provided by Caroline Kim) was used as a positive control.

Table 5.1 DNA concentrations of *M. pectinilyticus.*

	ng/µL	**A260/A280**	**A260/A230**
DNA 1	151.8	1.78	1.73
DNA 2	158.8	1.73	1.75
Positive gDNA	186.4	1.95	2.39
Elution Buffer	0.000	0	0

Both DNA 1 and 2 were of high purity (A260/A280) with a ratio close to the expected 1.8. The measurement of nucleic acid purity (A260/A230) was lower than the expected range of 2.0. Thus, there may be the presence of contaminants that absorbed at 230 nm. When compared to positive gDNA, both DNA 1 and 2 extracted lower concentrations. However, the purity (A260/A280 ratio) of the DNA 1 and 2 was better.

Table 5.2 Primers for protein 0050 amplification.

	Sequence 5' – 3'	**PCR Size**	**Mer size**	**GC (%)**	**T_m ˚C**
FP	CTCGAGCATCATCACCATCACCACTTGAAAAAATATATTTCGCTATTTAT	1323 bp	50	34	48
RP	GAATTCCATCATCACCATCACCACTTTAGATTTTTTTACTCGGT		44	36	47

Note: FP: Forward primer; RP: Reverse primer; GC: Nucleotide G and C content; Mer: length of oligonucleotide.

Primers were designed using SnapGene software (from Insightful Science; available at snapgene.com) to amplify the 0050 gene (Table 5.2). Additional motifs were included in the

primer design to allow for his-tagging and restriction enzyme digestion. Protein 0050 was amplified from newly extracted gDNA and positive control *M. pectinilyticus* gDNA, producing a band size between 1300 – 1350 bp (Figure 5.1). Smeary bands are due to degraded DNA and/or overloading of the wells. Lane 5 is PCR produced from gDNA, purified by QIAquick PCR purification kit (Qiagen).

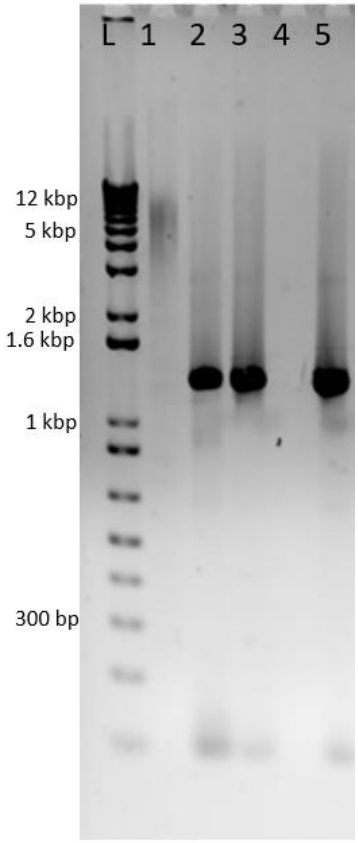


Figure 5.1 Gel electrophoresis of DNA for 0050**.** Lane L loaded with 1 kb+ ladder (Invitrogen); Lane 1 was gDNA of *M. pectinilyticus*. Lane 2 was loaded with 0050 PCR product from gDNA. Lane 3 was loaded with 0050 PCR product from positive gDNA. Lane 4 is a negative control. Lane 5 was loaded with PCR 0050 purified sample.

5.2.2 Computationally predicted structure of protein 0050

The amino-acid sequence of protein 0050 (Figure 5.2) was passed onto Geoff Jameson, who submitted for protein structure analysis using computational tools, PsiPred/pGenThreader and AlphaFold, two highly recognised prediction models, the former for secondary structure and threading onto known structures, the latter for *de novo* structure prediction.

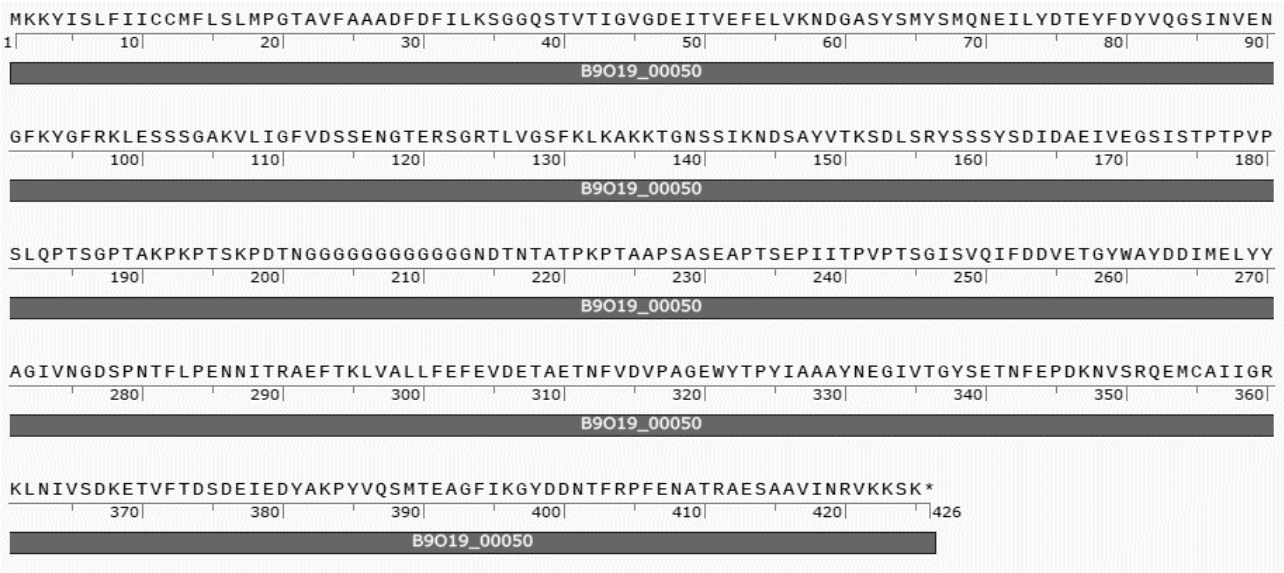


Figure 5.2 Amino sequence of protein 0050**.** Protein 0050 is 426 amino acid residues long. * - a stop codon. Genbank ID: AUO18237.1.

The predicted secondary structure of protein 0050 consists of a helical N-terminus followed by 10 β-strands (predicted with fair confidence), which is considered the first domain of the protein (Figure 5.3). From residue 210 onwards, a 6 α-helical structure forms the second domain of the protein. Between the two domains is a ~90 amino acid loop from residues ~170 to ~250. Within this unusual loop is a repeating motif of 12 glycine residues (positioned at amino acids 202 – 213). The repeating glycine$_{12}$ residues may give the protein high flexibility for domain binding, where the loop can be easily disrupted by connections with partner molecules, such as protein – protein interactions. Once the process or interaction is completed, it is assumed based on literature, that the original conformation of the protein is restored (Sachetto-Martins *et al*., 2000).

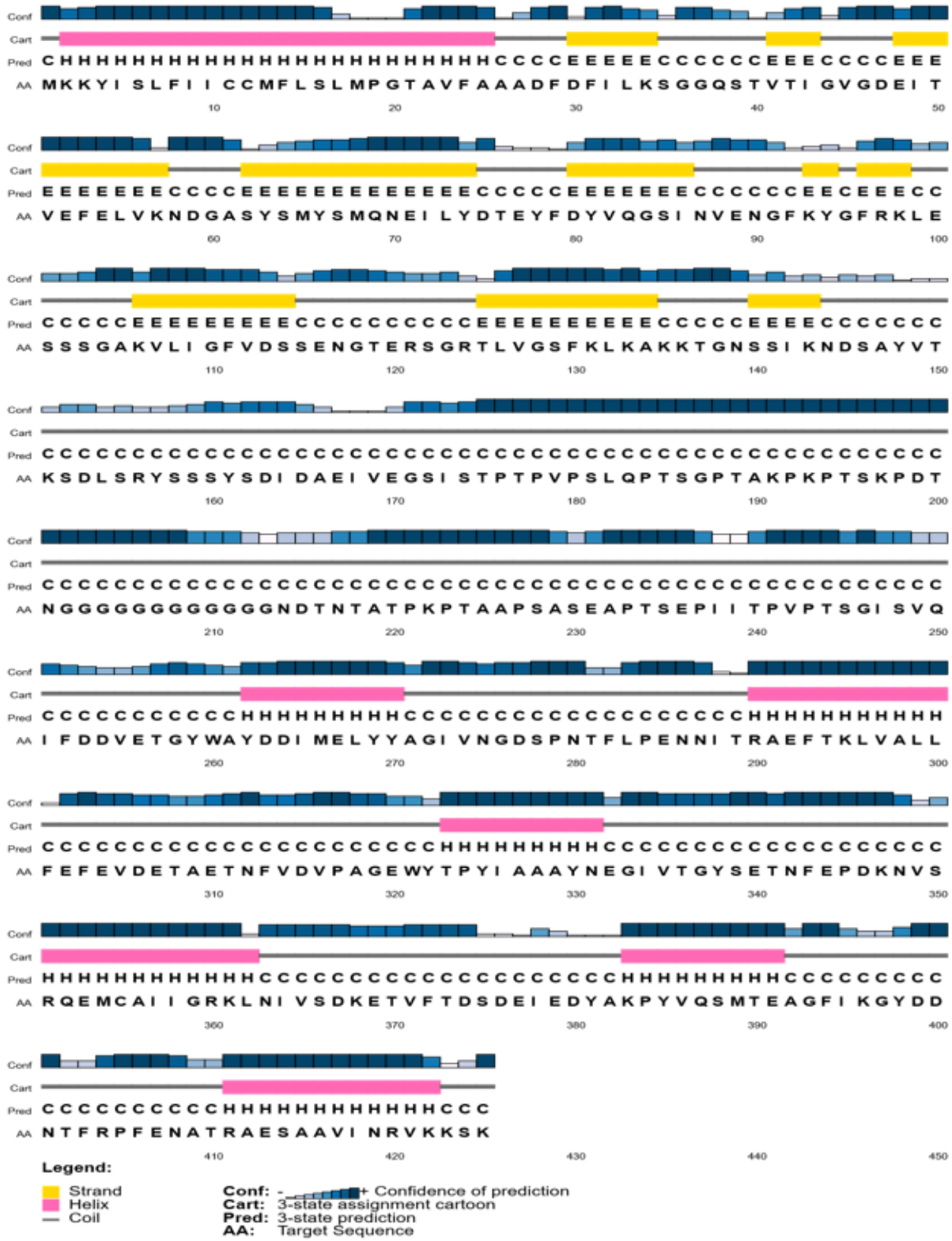


Figure 5.3 PsiPred analysis of protein 0050**.** Note the extended region of the coil from residues 170 to 260. Pink sections: α-helix structure. Yellow sections: β-sheet strands. Grey line: coil. Confidence in the predicted structure is indicated by the white (low confidence) to dark blue (high confidence).

Sequence alignment with pGenThreader using predicted secondary structure elements for protein 0050 yielded two hits to proteins of known structure with certainty (Table 5.3). The alignment of both these proteins covers only the C-terminal domain. Protein with PDB code 6cwm aligned with 0050 from residues 258 to 425 showed with 22.2% identity. Protein 3pyw matched the C-terminal region of protein 0050 from 251-425 with 28.7% identity. It also matched a small stretch of the loop region of protein 0050 from amino acids 175 to 251, but with only 6.7% identity. There was no similarity between these two proteins (6cwm and 3pyw) with respect to the repeating glycine motif.

Table 5.3 Alignment of protein 0050 to proteins of known structure.

*PDB		Alignment length	E-value	% identity	Organism	Classification
6cwm	CERT	171[a]	$1x10^{-8}$	22.2%	*Paenibacillus alvei*	Sugar binding protein by SLH domain
3pyw	CERT	178[a]	$3x10^{-7}$	28.7%	*Bacillus anthracis*	SLH domain
3bwz	HIGH	171[b]	$6x10^{-4}$	10.5%	*Acetivibrio cellulolyticus*	Type II cohesion module from cellulosome
5nrm	HIGH	141[b]	$6x10^{-4}$	12.8%	*Acetivibrio cellulolyticus*	Sixth cohesion in scaffoldin B in complex with Cel5 dockerin
2ccl	HIGH	149[b]	$8x10^{-4}$	11.4%	*Acetivibrio cellulolyticus*	Mutant of Type I cohesion-dockerin from cellulosome

Note: *PDB: protein data bank, https://www.rcsb.org/. [a] Alignment to C-terminal α-helical domain of protein 0050 (residues ~251 to 425). [b] Alignment to N-terminal □-sheet domain of protein 0050 (residues ~27 to ~150).

The structure of protein 0050 predicted by AlphaFold showed a two-domain protein with areas of high and low certainty (Figure 5.4). The positioning of the helix predicted at the N-terminus, which leads into the β-sheet domain, is uncertain. Except for the first two residues following the initial methionine (a pair of lysine), residues 3-27 are hydrophobic. Thus, this N-terminal tail may function as a membrane anchor. The threading of the β-strands in the N-terminal domain (residues ~50 to ~170) is predicted with fairly good confidence, giving a degree of certainty to the

β-sheet tertiary structure visualised here. AlphaFold could not model the ~90 amino acid loop containing the glycine$_{12}$ motif. The tertiary structure of the α-helical C-terminal domain is predicted with high confidence by AlphaFold. In addition, AlphaFold predicts a close association of the N-terminal β-sheet domain with the C-terminal α-helical domain.

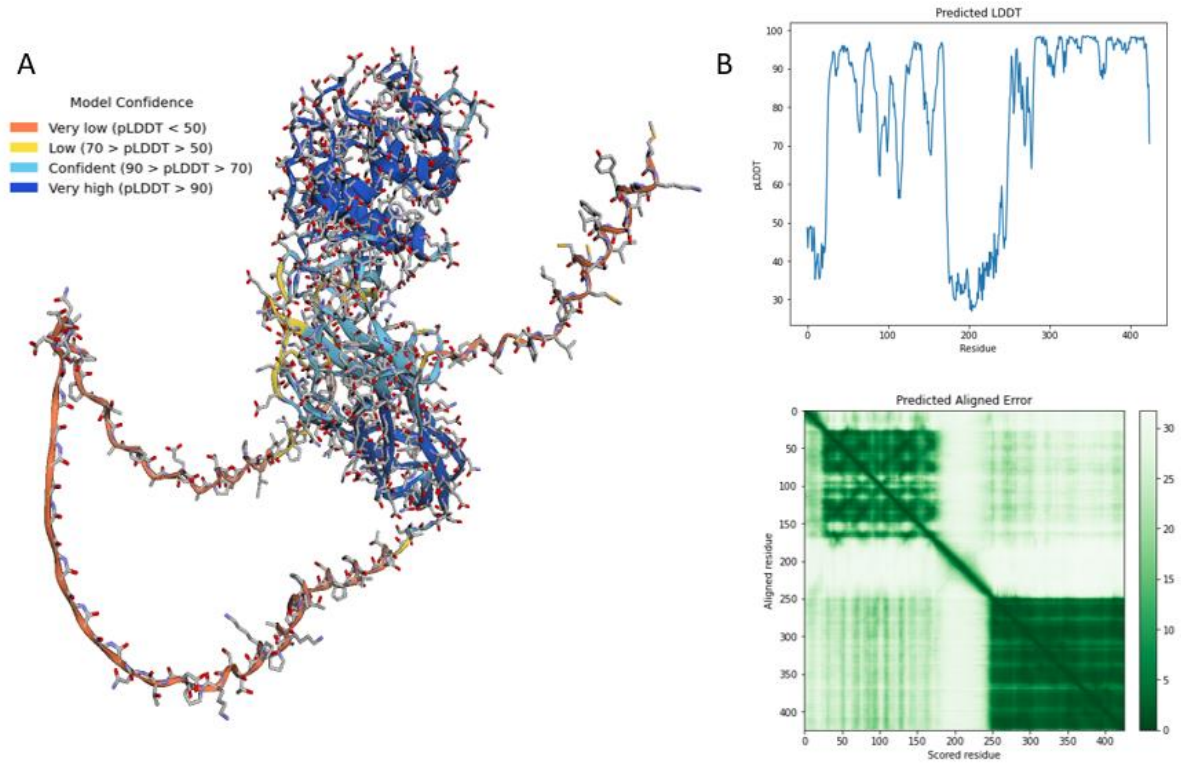


Figure 5.4 AlphaFold analysis of protein 0050. (A) Predicted structure of protein 0050. (B) Model per-residue estimate of Alpha Fold's confidence in its prediction on a scale of 0 – 100, where a pLDDT of 70 and below are low confidence, pLDDT 70 – 90 are expected to be modelled well, with anything over 90 considered high accuracy of structure modelling.

Protein 0050 was blasted against the UniProt database with only the top 3 hits presented. The top-scoring protein, QO6848, annotated as cellulosome-anchoring protein, and is the most similar to protein 0050 in terms of length, similarity over amino acid overlap and the presence of the poly-glycine motif.

Table 5.4 Results of BLAST of protein 0050 against the UniProt database.

UniProt	Length (aa)	% identity	% similarity	Z-Score	Organism	Classification
QO6848	447	28.7%	54.4%	309.4	*Acetivibrio thermocellus*	Cellulosome-anchoring protein
P38536	1861	29.9%	56.4%	296.0	*Thermoanaerobacterium thermosulfurigenes*	Amylopullulanase
C6CRV0	1462	26.0%	57.1%	282.4	*Paenibacillus sp.*	Endo-1,4-beta-xylanase A

The cellulosome is a multi-modular extracellular complex of structural and enzymatic proteins produced to process polysaccharides. The cellulosome, in particular, focuses on the degradation of cellulose and hemicellulose (Bule *et al*., 2018). The homology of protein 0050 to a protein annotated as a cellulosome-anchoring protein suggests that protein 0050 may be involved in a multi-modular pectin-degrading complex. Both putative 0050 and cellulosome-anchoring protein QO6848 share similarities with the glycine$_{12}$ motif, where protein QO6818 has a glycine$_{7}$ motif, supporting the idea that the loop may be essential for protein flexibility for the interaction of both domains.

The association of protein 0050's domains (α-helical and β-sheet domain) predicted by AlphaFold was investigated using PISA (protein interface and surface analysis) (Krissinel, & Henrick, 2007; Krissinel, 2010). Both domains may potentially interact to create a binding site or a catalytic site to interact with pectic substrates. To do the PISA analysis, protein 0050 was split into two at the glycine$_{12}$ loop, producing the β-sheet domain and the α-helical domain. Computational results show that the total surface of the domains is ~17,056 A^2. When the two domains come together, a total of ~871 A^2 is buried, which constitutes approximately 5% of the total surface area. The solvation free energy gain for the interface between the β-sheet domain and the α-helical domain is -6.4 kcal mol^{-1} (-27 kJ mol^{-1}). Thus, there is a significant hydrophobic interaction between the two domains within the AlphaFold prediction association. While computationally significant, there is no certainty that this association between the two domains occurs in solution or is biologically relevant.

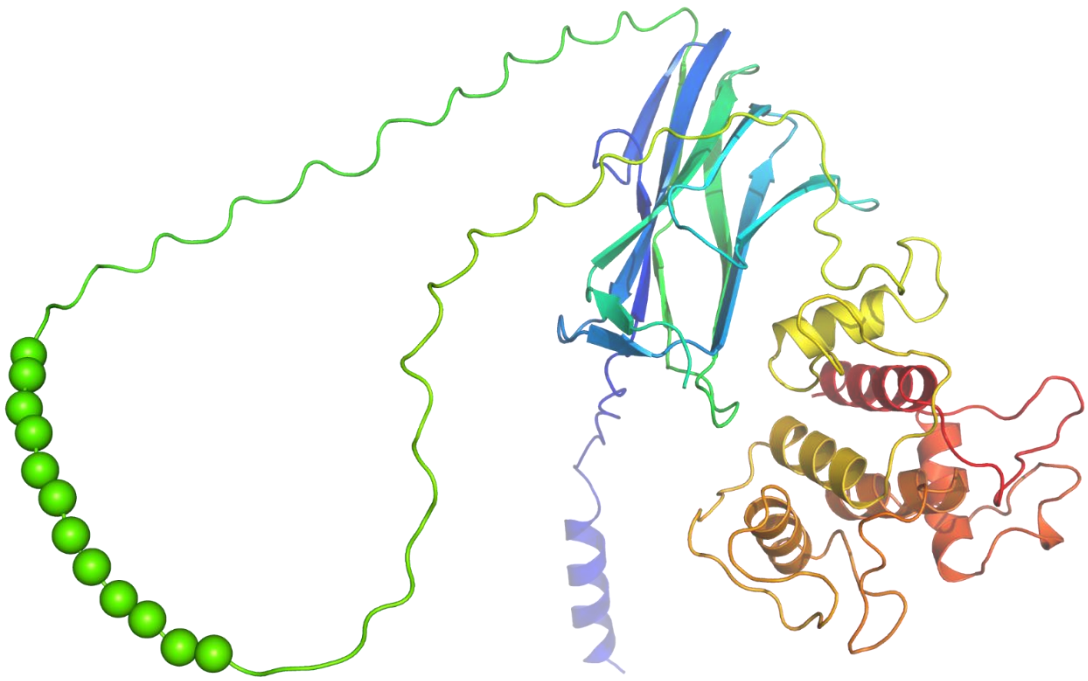


Figure 5.5 Predicted structure of protein 0050. Green spheres denote the glycine$_{12}$ motif within the loop. The diagram is rainbow coloured, with the N-terminus dark blue and the C-terminus red. Diagram prepared with PyMol.

AlphaFold is reasonably confident at modelling the two major domains of protein 0050 structure, where a β-sheet domain linked by a glycine$_{12}$ loop to an α-helical domain. The glycine$_{12}$ loop, though poorly modelled, may provide a flexible link facilitating the interaction between the two domains and/or the domains to other proteins – (protein-protein interactions).

Due to the unusual ~90-residue loop and low confidence for positioning the first 27 hydrophobic amino acid residues, two variations of protein 0050 were constructed alongside the original for protein expression. Construct one, labelled P1, has the first 25 amino acid residues removed. Construct two, labelled P2, has the first 27 amino acids and the ~90 amino acid loop truncated to ~25 amino acids, effectively removing the glycine$_{12}$ motif.

5.2.3 Vector construct

A FASTA file of protein 0050 (Figure 5.2) was taken from NCBI and uploaded into SnapGene (Figure 5.6). The sequences of the three constructs of protein 0050 were aligned into vector pET28a (+). Full vector and inserts 0050 were sent to GenScript® for synthetic construction optimised for expression in *E. coli* BL21 cells. As stated above, two alternative protein variants of 0050 were constructed (Figure 5.6).

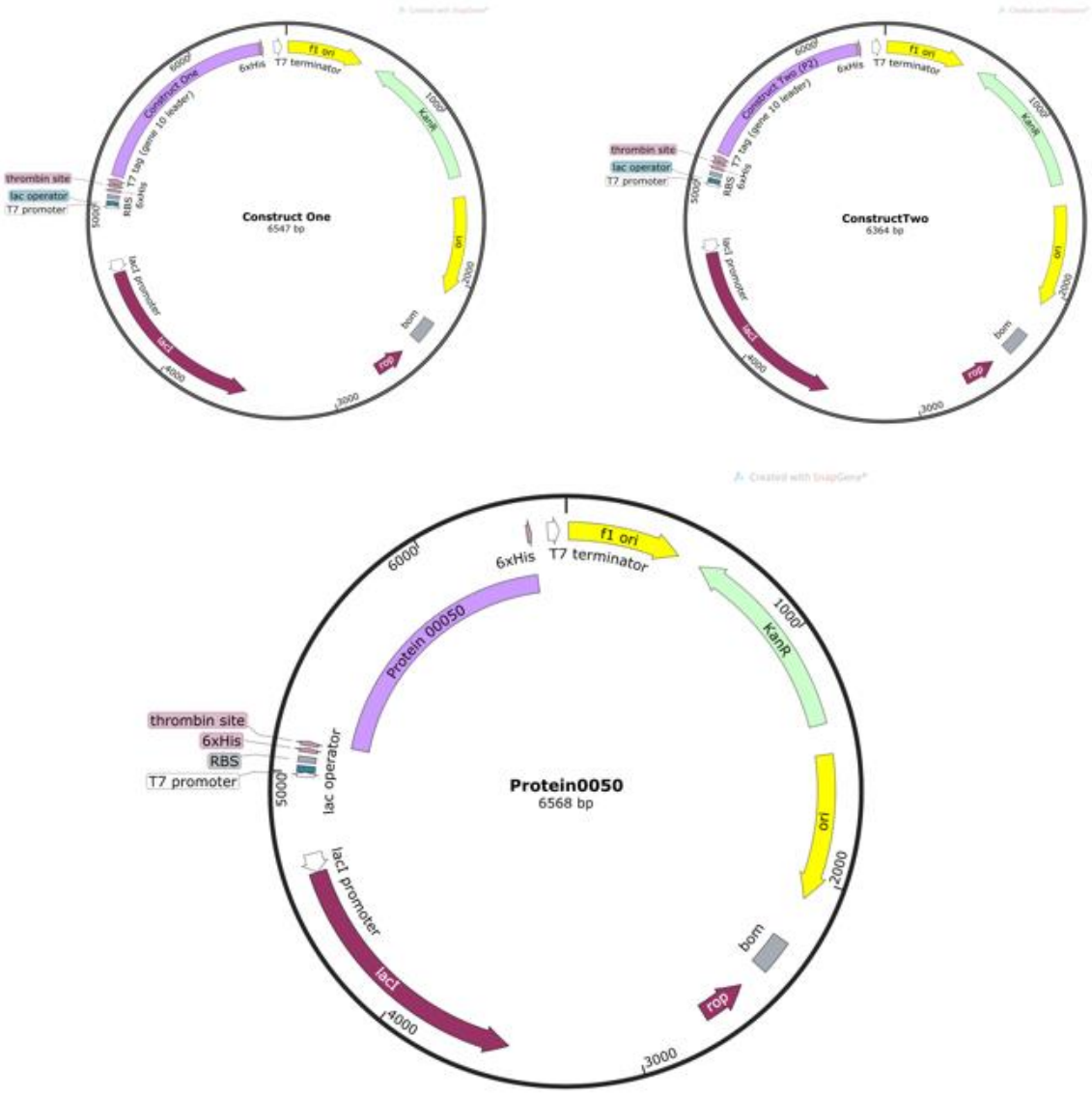


Figure 5.6 Construct of pET28a (+) vectors to express protein 00050 and its two variants. The light purple colour annotates protein 0050 and its variants.

5.2.4 Expression and purification of His-tagged 0050

Expression of wild type protein 0050 was attempted in the *E. coli BL21* system under normal conditions. Extensive assistance was provided by Trevor Loo in the troubleshooting of protein 0050. Figure 5.7 shows that while 0050 is present (Lane 1), there was very little increase in band intensity over time. Collection times at 8 hours and 16 hours showed the most significant increase in band intensity. Either the expression system is having difficulties expressing soluble 0050, or there is little to no expression occurring. Repeated expressions produced consistent results – little expression or none at all. Codon analysis identified that the numerous rare codons present within the sequence of 0050 had not been optimised for ideal expression in *E. coli* BL21 cells, despite an explicit request to GenScript for codon optimisation (Appendix 1). There are double AUA negative repeat elements near the sequence's middle, which could contribute to expression issues.

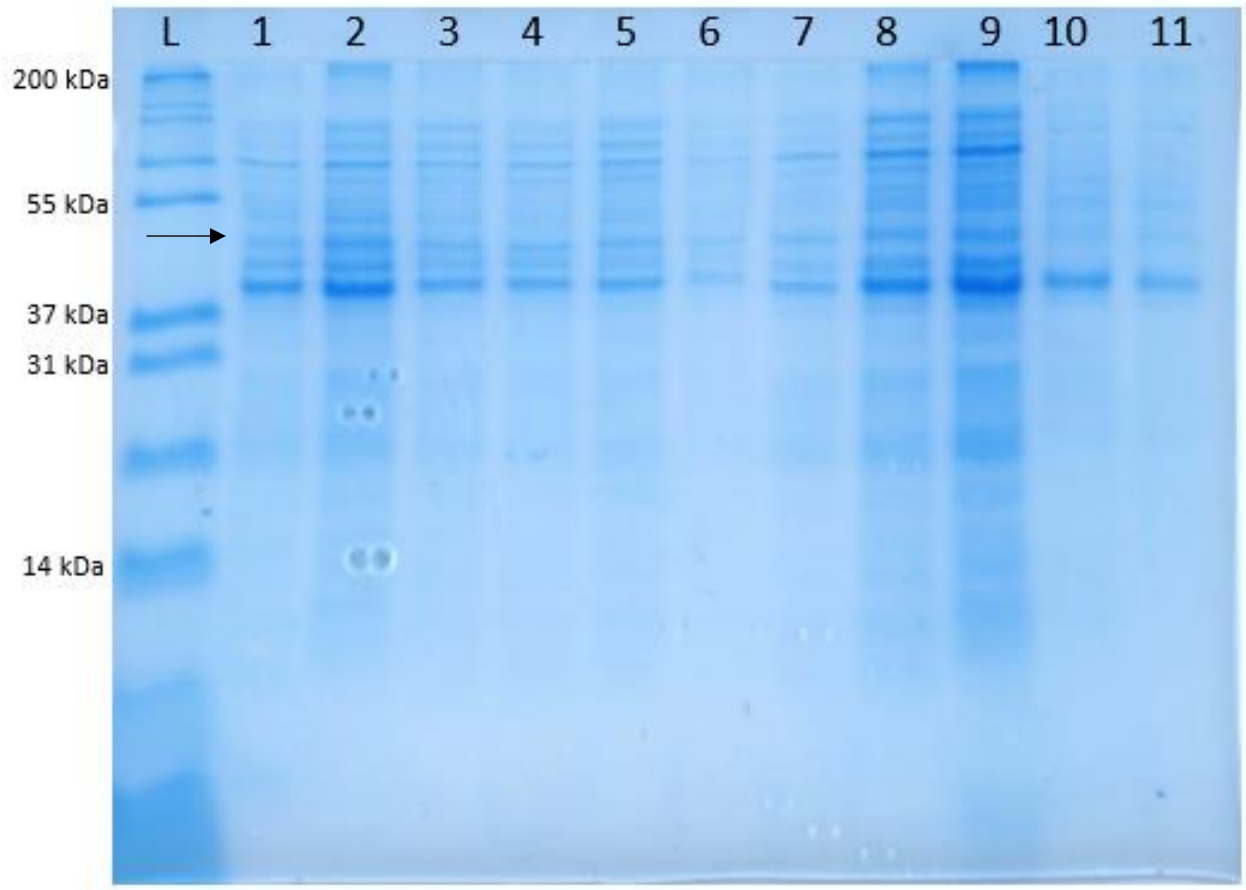


Figure 5.7 Expression of 0050 in *E. coli* BL21 cells induced by 1 mM IPTG. Lane L was loaded with Mark12 ladder. Lane 1 loaded with sample before induction (Tb). Lanes 2 – 9 loaded with time-collected samples: respectively, 0 minutes, 30 minutes, 1 hour, 2 hours, 3 hours, 4 hours, 8 hours and 16 hours. Lane 10 was loaded with uncut pET28a (+) induced with IPTG for 4 hours. Lane 11 is similar to Lane 10 but induced for 16 hours.

Rosetta cells, a BL21 derivative enhanced to deal with proteins that contain rare codons, were used instead of BL21 E. coli cells

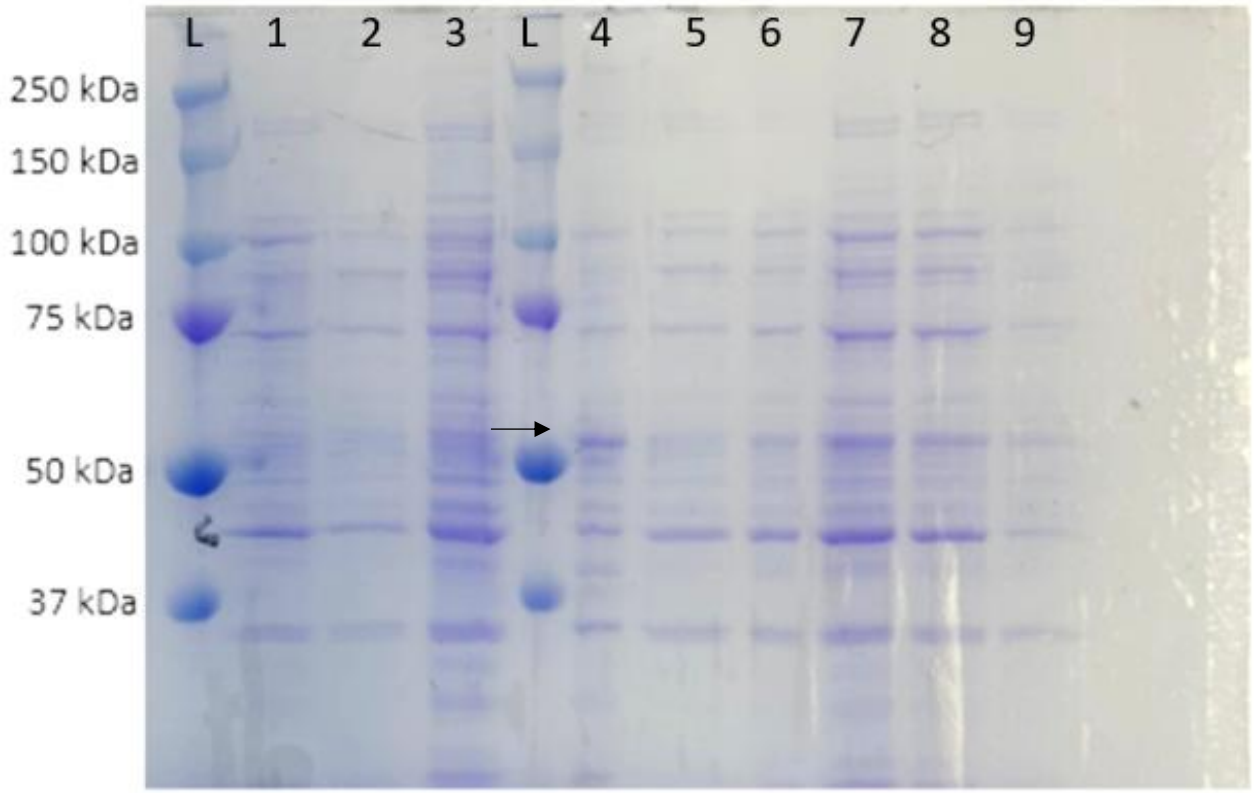


Figure 5.8 Expression of 0050 using Rosetta cell system induced by 1 mM IPTG. Lane L was loaded with Mark12 ladder. Lanes 1 – 3 loaded with uncut pET28a (+) transformed into Rosetta cells, with time-collected samples: 0 minutes, 1 hour and overnight. Lanes 4 – 9 loaded with time-collected samples: respectively, 0 minutes, 1 hour, 2 hours, 4 hours, 8 hours and overnight at 25 ˚C

A dark band is present at just above 50 kDa, indicated by a black arrow in Figure 5.8, represents the presence of protein 0050. The uncut pET28a vector, with no 0050 insert transformed in Rosetta cells, showed no expression of protein 0050 (Lanes 1 – 3). Lanes 4 – 9 show a faint presence of the protein with a slight increase in intensity as the time over which the protein is expressed increased. The expressed protein after 4 hours is not as intense as lane 4, suggesting leakiness of the protein before induction. Rosetta cells struggled to express the protein despite Rosetta correcting for rare codons. Overnight expression at 25 ˚C did not express large quantities of protein 0050. Despite the error in codon optimisation, *E. coli* systems failed to express in significant quantities soluble protein 0050 from Gram-positive *M. pectinilyticus*. The protein may have also not been expressed in soluble form or, less likely, is toxic in *E. coli*.

5.2.5 Expression and purification of His-tagged variants

Protein 0050 variants (construct one and two) from GenScript were successfully expressed using the BL21 (DE) system (Figure 5.9 and Figure 5.10). Construct one and construct two protein expressions are labelled P1 and P2, respectively. It should be noted, the extensive help and trouble-shooting from Trevor Loo, in the expression and purification of variants, P1 and P2.

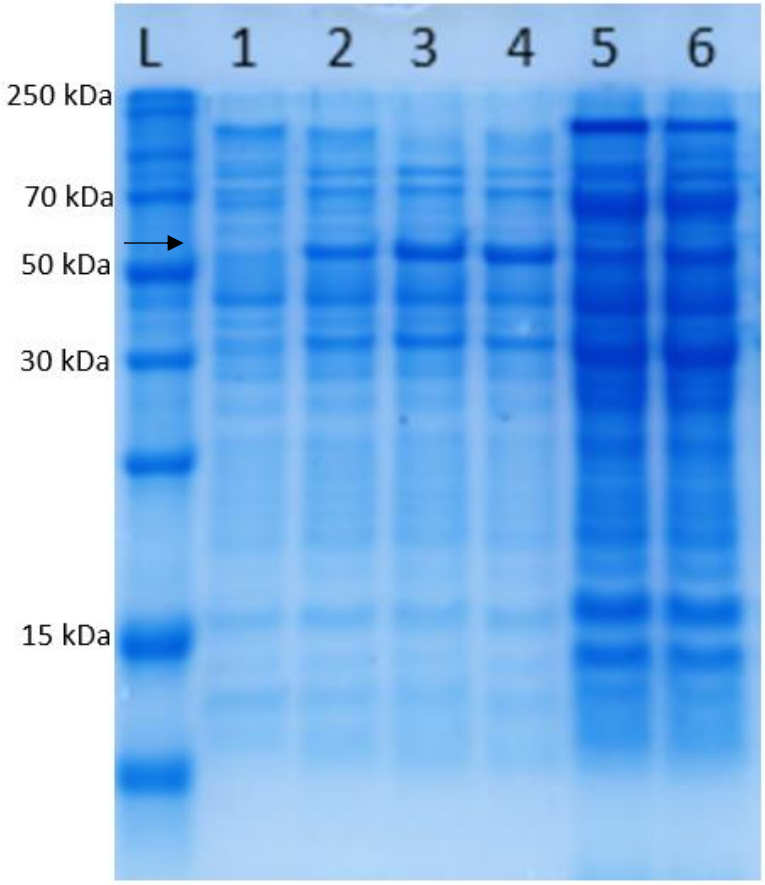


Figure 5.9 Construct one (P1) protein 0050 using the *E. coli* BL21 expression system. Lane L was loaded with ladder. Lane 1 loaded with uncut pET28a (+) vector with no added protein. Lanes 1 – 6 were loaded with time-collected samples: respectively, before induction, 1 hour, 2 hours, 4 hours, overnight at 37 ˚C and overnight at 25 ˚C.

Construct P1 of 0050, identified by the black arrow, was successfully expressed. Despite a bigger than the expected size of 48.8 kDa, there is an apparent increase of protein intensity over the period for induction, with the optimal time being between 2 – 4 hours. Similarly to Figure 5.8, expression levels are not as high as expected, though significantly more dramatic than wild-type 0050. The glycine$_{12}$ loop could be involved in the poor expression levels due to its high flexibility and relatively high hydrophobicity.

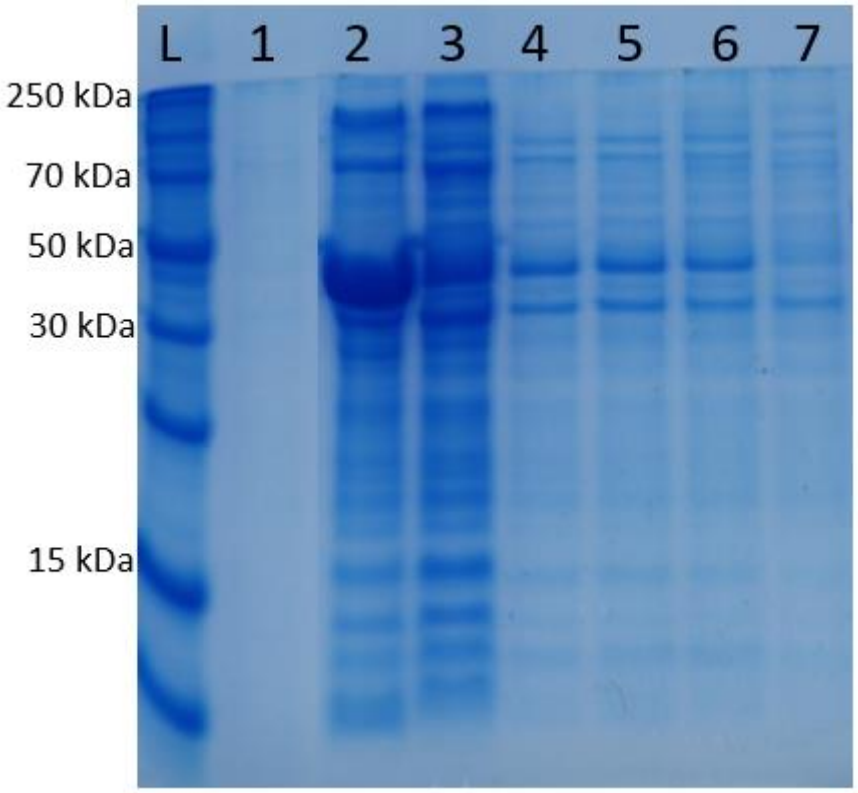


Figure 5.10 Construct two (P2) protein 0050 using the *E. coli* BL21 expression system. Lane L was loaded with ladder. Lane 1 loaded with uncut pET28a (+) vector with no insert. Lanes 2 – 7 loaded with time-collected samples: respectively, before induction, 1 hour, 2 hours, 4 hours, overnight at 37 ˚C and overnight at 25 ˚C.

The expression of construct P2 was substantially more apparent compared to wild-type and construct P1 (Figures 5.8 and 5.9). Lane 3 contains an enormous band at ~45 kDa before any IPTG was added to the culture. This large band is the result of a leaky system. The expression of construct P2 over 2 hours to overnight at 37 ˚C did not change much, while induction overnight at 25 ˚C resulted in little to no induction of construct P2.

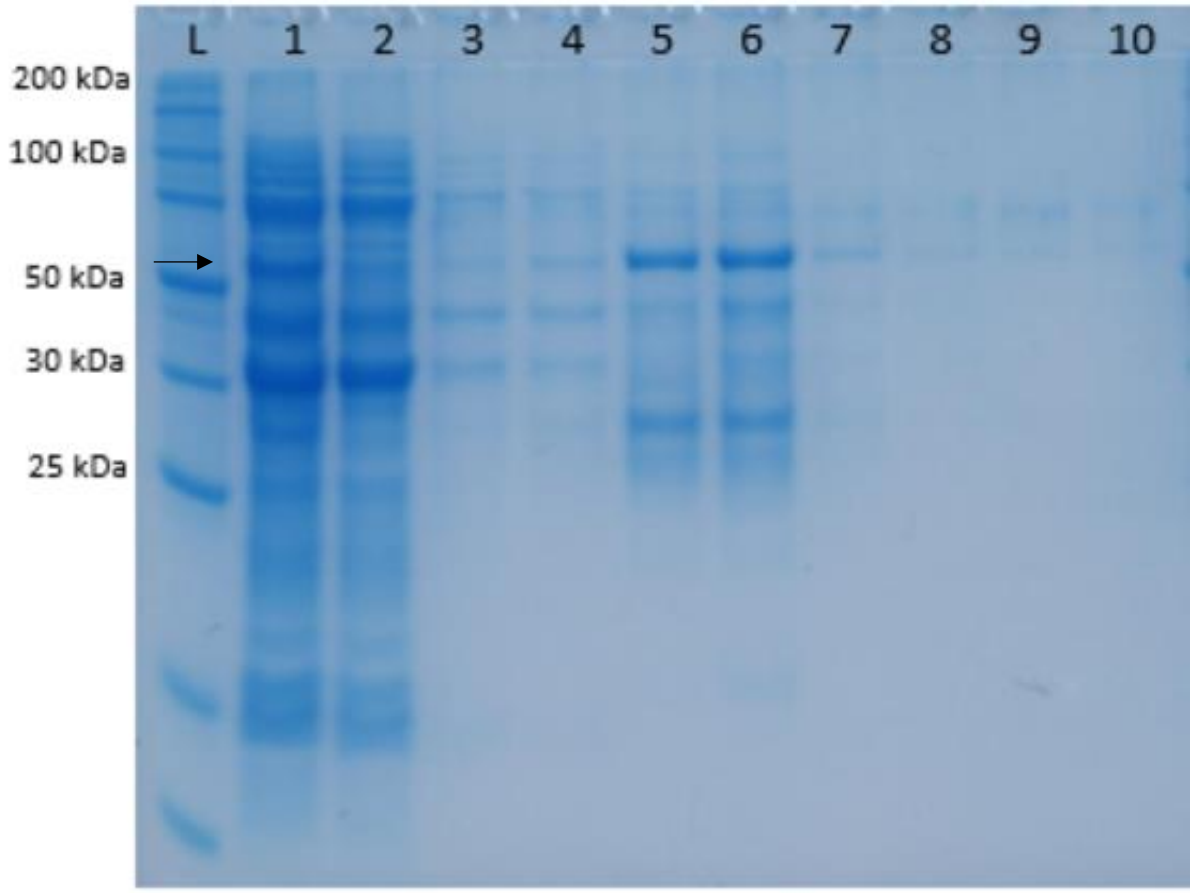


Figure 5.11 Purification of construct P1 by IMAC. Lane L was loaded with ladder. Lane 1 loaded with the soluble cytoplasmic fraction of construct one. Lane 2 loaded with flow through. Lanes 3 – 4 are loaded with wash elutions. Lanes 5 – 10 were loaded with elution buffer at different concentrations of imidazole. Lanes 5 – 6 contained 125 mM imidazole samples. Lanes 7 - 10 contained 250 mM imidazole samples.

Construct P1 was present firmly in the soluble cytoplasmic fraction and bound to the Ni-NTA resin, allowing strong elution in the presence of 125 mM imidazole. The intensity of the dark bands in lanes 5 and 6 is consistent with Figure 5.9.

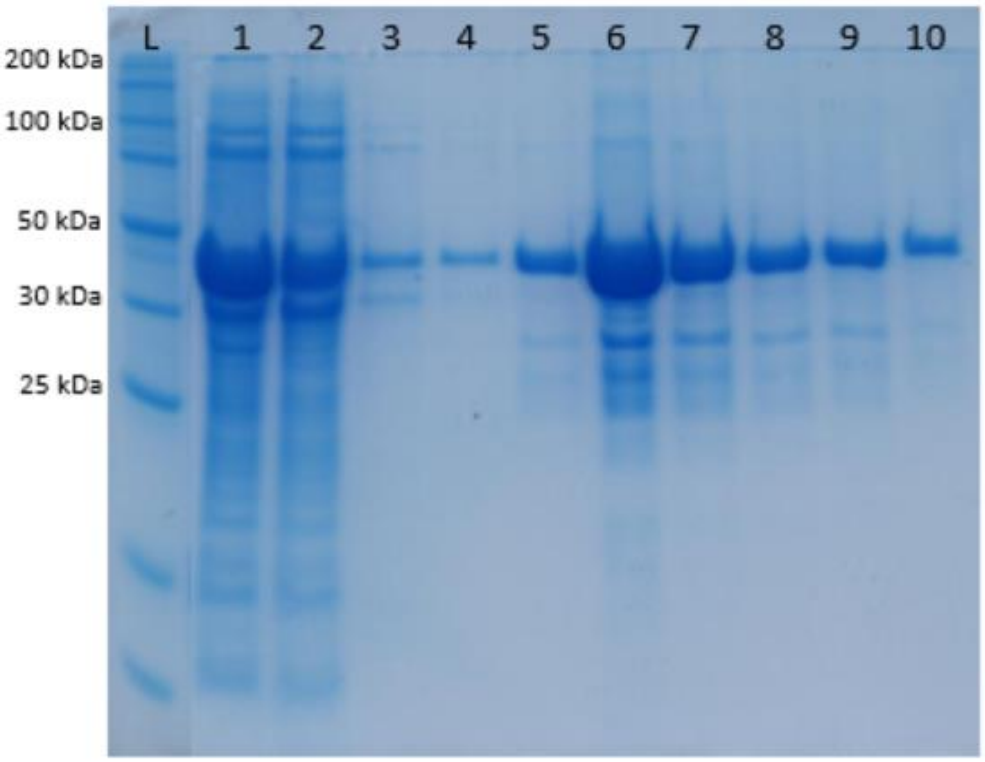


Figure 5.12 Purification of construct P2 by IMAC. Lane L was loaded with ladder. Lane 1 loaded with construct P2 soluble cytoplasmic fraction. Lane 2 loaded with flow through. Lanes 3 – 4 are loaded with wash elutions. Lanes 5 – 10 are loaded with elution buffer at different concentrations of imidazole, with 125 mM imidazole used in lanes 5 – 6 and 250 mM imidazole in lanes 7 – 10.

Construct P2 was successfully purified by IMAC in large quantities as indicated by lanes 5 – 10 in Figure 5.12, with the purest fractions eluting with 125 mM imidazole. Construct P2 is the same size (~46 kDa) as identified in Figure 5.10. Both constructs, P1 and P2, were excised from the gel for validation by mass spectrometry (Appendix 3).

The expression of both variants differed from each other. Variant P2 was expressed and purified in much larger quantities than variant P1. The only difference between the two is the ~90 amino acid loop. The highly flexible loops may affect the protein's solubility and /or structural stability resulting in poor expression and purification.

5.2.5 β-galactosidase/β-galactanase activity

Wild-type and variants P1 and P2 were tested for enzymatic activities associated with pectin degradation (2.4.5). All three proteins did not have specific β-galactosidase and β-galactanase activity (Table 5.5). The positive control included was *E. coli* β-galactosidase. Negative controls were distilled water, and elution buffer from Ni-NTA purification was included to validate the results.

Table 5.5 Enzymatic results of all three proteins.

	Protein 0050	**P1**	**P2**	***dH$_2$0**	***Ni-NTA elution buffer**	**^β-galactosidase (*E. coli*)**
Enzyme activity (µmol min^{-1} mg^{-1})	-	0.00	0.00	0.00	0.00	0.026
D-galactose/L-arabinose content	-	0.00	0.00	0.00	0.00	0.00
Protein content (mg/mL)	-	1.090	0.649	0.00	0.00	7.710

Note: *: negative controls; ^: positive control. P1: Construct one protein. P2: Construct two protein.

5.2.6 Ligand binding assay

Due to the computational structural prediction identifying two domains with similarity to sugar-binding proteins, structural proteins (Table 5.3) and cellulosome anchoring proteins (Table 5.4), the binding capabilities of wild-type, P1 and P2 were investigated. The incorporation of His-tag to all proteins allowed for the ligand-binding assays to be studied, using bi-layer interferometry (BLItz) (2.5.6). The assistance of Ruby Roach aided in the BLItz sample set up and the data processing for sample analysis. Only variant P1 and P2 were included due to unsuccessful validation of the original full-length protein 0050 by mass spectrometry. A qualitative initial yes-or-no assay was incorporated to narrow down candidates for advanced kinetic assays (Appendix 3). A quantitative assay followed this to identify the binding affinity of the protein to the ligand (Table 5.6). The analysis of the BLItz data was provided by Ruby Roach.

Pectin consists of a variety of saccharides; thus, thirteen saccharides present in pectin were tested (Table 5.6). Protein P1 bound to 11 of the 13 substrates, whereas construct P2 is only bound to 9 of the 13 substrates. Construct P1 bound strongest to citrus pectin, D-(+)-xylose, di-galacturonic acid, arabinan and the mixed water-soluble pectin from Kiwiberry. On the other hand, construct P2 bound firmly to di-galacturonic acid. The decrease in substrate binding between the two proteins may result from the $glycine_{12}$ loop. The flexibility of construct P1 may allow the protein to bind more efficiently to the desired substrate. However, the numerous substrates to which both constructs P1 and P2 bind tell a complicated story. The initial binding results indicate that constructs P1 and P2 are attracted to the monomer of xylose, galacturonic acid and arabinan.

Table 5.6 Summary of initial screening of ligand binding to protein 0050.

Substrate	Composition	Purity	Binding	
			P1	P2
Citrus pectin	Galacturonic acid ≥ 74%	>74%	++	+/-
D-(+)-Xylose	Monomers of D-(+)-xylose	99%	++	+/-
Di-galacturonic Acid	-	>95%	++	++
Tetra-galacturonic Acid	-	>90%	+	+/-
Linear 1,5-α-L-arabinan	Arabinose: Galactose: Rhamnose: Galacturonic acid = 85.2: 7.6: 1.5: 5.7	95%	+	+
Arabinan	Arabinose: Galactose: Rhamnose: Galacturonic acid: Other sugars = 69: 18.7: 1.4: 10.2: 0.7	95%	++	-
Rhamnogalacturonan I	Galacturonic acid: Rhamnose: Arabinose: Xylose: Galactose: Other Sugars = 61.0: 6.2: 2.5: 0.5: 23.1: 6.7	>90%	+/-	-
Polygalacturonic Acid	Galacturonic acid: Galactose: Arabinose: Rhamnose: Xylose: Glucose = 94: 1.0: 0.2: 1.0: 1.0: 0.3	96%	+/-	-
Galactan	Galactose: Arabinose: Rhamnose: Galacturonic acid = 87: 3: 4: 6	>85%	+	+/-
Mixed pectin	Kiwiberry Hortgem Tahi – water soluble extract	N/A	++	++
Mixed pectin	Kiwiberry Takaka Green – water soluble extract	N/A	++	++
Mixed pectin	Kiwiberry Marju Red – water soluble extract	N/A	++	++
Cellulose	-	N/A	-	+/-

Note: N/A: not applicable due to mixed composition of pectin or data not available. +/-: increase in binding ≤0.1nm, +: increase in binding 0.1 - 0.2 nm, ++: increase in binding >0.2 nm, -: no increase in binding (nm).

Advanced kinetics was performed by BLItz to observe the strength of association and dissociation of the ligand binding to protein (Appendix 3). Table 5.7 summarises the data from ligand binding. There is no fixed threshold for binding affinity. Measurement of association rate constant (k_a (1/M s) and dissociation rate constant (k_d (1/s) tell a story of the interactions between the binding of the protein and ligand. The k_a measures the rate at which free ligand interacts with protein, while k_d measures the rate at which a bound ligand dissociates from a protein. The equilibrium constant K_D (M), the ratio of rate constants k_d/k_a, reveals the affinity of the interaction between the ligand and protein. The higher the K_D value, the weaker the ligand and protein are attracted to one another, whereas the lower *the* K_D value, the greater the binding affinity between the ligand and protein. A K_D range of $x10^{-5}$ – $x10^{-9}$ are considered moderate to high binding affinity.

Table 5.7 Kinetics of ligand binding to protein 0050.

Substrate	$^{*}K_D$ (1/M)		$^{\wedge}k_a$ (1/M s) ± s.d		$^{\circ}k_d$ (1/s) ± s.d		R^2	
	P1	P2	P1	P2	P1	P2	P1	P2
Citrus pectin	$4.08x10^{-4}$	$1.26x10^{-5}$	(1.5 ± 0.8) $x10^{2}$	(2.1 ± 3.9) $x10^{3}$	(6.2 ± 2.2) $x10^{-2}$	(2.64 ± 0.08) $x10^{-2}$	0.92	0.92
D-(+)-Xylose	$1.27x10^{-2}$	1.084	7 ± 5	1.60 ± 5.1	(8.80 ± 0.18) $x10^{-2}$	1.73 ± 0.23) $x10^{-1}$	0.97	0.76
Di-galacturonic acid	$1.43x10^{-1}$	$2.70x10^{-3}$	$4.71x10^{-1}$ ± 7.36	(1.5 ± 0.8) $x10^{1}$	(6.78 ± 0.12) $x10^{-2}$	(4.09 ± 0.09) $x10^{-2}$	0.97	0.96
Tetra-galacturonic acid	2.05	$4.44x10^{-1}$	$3.16x10^{1}$	(0.05 ± 6) $x10^{2}$	$6.48x10^{1}$	(1.99 ± 0.6) $x10^{-1}$	0.00	0.50
Linear 1,5-α-L-Arabinan	$7.26x10^{-6}$	$2.31x10^{-5}$	(5.4 ± 0.2) $x10^{3}$	(1.3 ± 3.1) $x10^{2}$	(3.90 ± 0.06) $x10^{-2}$	(3.04 ± 0.07) $x10^{-2}$	0.97	0.95
Arabinan	$2.54x10^{-5}$	$1.33x10^{-2}$	(2.3 ± 0.7) $x10^{3}$	0.0004 ± 2) $x10^{5}$	(5.8 ± 0.2) $x10^{-2}$	0.05 ± 1.1	0.90	0.00
Rhamno-galacturonan I	$9.04x10^{-6}$	$<1x10^{-12}$	(6.3 ± 2.3) $x10^{3}$	$5.69x10^{2}$	(5.7 ± 0.2) $x10^{-2}$	$<1x10^{-7}$	0.92	0.00
Polygalacturonic acid	$7.32x10^{-5}$	$3.46x10^{-4}$	(7.0 ± 1.0) $x10^{2}$	(3.16 ± 2.9) $x10^{5}$	(5.1 ± 1.6) $x10^{-1}$	1.09 ± 5.34	0.92	0.00
Galactan (Potato)	$6.82x10^{-4}$	$1.20x10^{-6}$	(0.07 ± 2) $x10^{3}$	(4.62 ± 0.5) $x10^{4}$	(4.72 ± 0.14) $x10^{-2}$	5.6 ± 0.2) $x10^{-2}$	0.93	0.89
Mixed pectin – Hortgem Tahi	$2.07x10^{-6}$	$2.10x10^{-7}$	(1.67 ± 0.09) $x10^{4}$	(2.56 ± 0.05) $x10^{4}$	(3.46 ± 0.06) $x10^{-4}$	(5.40 ± 0.18) $x10^{-3}$	0.98	0.92
Mixed pectin – Takaka Green	$3.12x10^{-7}$	$1.55x10^{-7}$	(9.33 ± 0.21) $x10^{4}$	(2.48 ± 0.07) $x10^{5}$	2.92 ± 0.04) $x10^{-2}$	(3.85 ± 0.12) $x10^{-2}$	0.98	0.96
Mixed pectin – Marju Red	$6.74x10^{-7}$	$4.11x10^{-6}$	(2.48 ± 0.03) $x10^{4}$	(2.9 ± 1.1) $x10^{3}$	(1.67 ± 0.01) $x10^{-2}$	(1.21 ± 0.02) $x10^{-2}$	0.99	0.97
Cellulose	$6.86x10^{2}$	$4.16x10^{-1}$	$2.33x10^{3}$	(0.006 ± 1.7) $x10^{2}$	$1.60x10^{6}$	(2.70 ± 0.7) $x10^{-2}$	0.00	0.58

Note: *: measured interaction affinity in molar units; ^: association rate constant; ˚: dissociation rate constant; ± s.d: ± standard deviation. R^2: correlation coefficient for the fit of the model to data. Estimated molecular weights of the substrates used for binding calculations: Citrus pectin, $2x10^4$ – $2x10^5$ g/mol; D-(+)-Xylose, 150.13 g/mol; Di-galacturonic acid, 379.26 g/mol; Tetra-galacturonic acid, 546.4 g/mol; Linear 1,5-α-L-Arabinan, $1.5x10^4$ g/mol; Arabinan, $1.5x10^4$ g/mol; Rhamno-galacturonan I, $6.4x10^4$ g/mol; Polygalacturonic acid, ~$3.75x10^5$ g/mol; Galactan (Potato), $1.0x10^5$ g/mol; Mixed pectin of Hortgem Tahi, Takaka Green and Marju Red, ~$5.00x10^5$ g/mol; and Cellulose, 162.14 g/mol.

Overall, out of the substrates tested, cellulose, D-(+)-xylose and tetra-galacturonic acid are the primary substrates that have a high K_D value from both proteins (constructs P1 and P2). While the remaining substrates had somewhat lower K_D values. The best binding affinity of both proteins (P1 and P2) was with the kiwiberry mixed pectin samples of all three cultivars: 'Hortgem Tahi', 'Takaka

Green' and 'Marju Red'. These samples represent mixed pectic-saccharides, which simulate part of a realistic pectin structure found in authentic plant samples. Intriguing, for the two kiwiberry substrates and several other substrates, the removal of glycine$_{12}$ increased the binding of the protein to the ligand. On the other hand, for some substrates, removing the loop decreased the binding affinity.

Though a complex story due to the numerous binding to pectin, shows that protein 0050 identified and extracted in Chapter 4 is involved in pectin degradation. Other than the pectic-saccharide kiwiberry samples – the strongest binding affinity with both protein constructs was observed for substrates citrus pectin, linear-1, 5-α-L-arabinan and potato galactan. These three substrates contained the same monosaccharide compositions of arabinose, galactose, rhamnose and galacturonic acid, though in different ratios. The presence of xylose in rhamnogalacturonan I and polygalacturonic acid may deter the protein from binding. Given the poor binding of di- and tetra-galacturonic acid, constructs P1 and P2 appear to show a stronger attraction towards monomers of galacturonic acid (Table 5.7). Arabinose, galactose, rhamnose and galacturonic acid are monosaccharides made up of pyranose – a six-membered ring of five carbon atoms with one oxygen atom. Interference within the substrates may weaken the binding affinity of the proteins to the substrate, such as arabinan, which contains other unknown sugars.

5.3 Conclusions

Protein 0050 faced complications with expression, and the original protein and its variants failed to possess any catalytic activity identified in Chapter 4. Predicted structure analysis identified an unusual structure – an extremely hydrophobic N-terminal followed by a β-sheet domain linked to a α-helix domain by a long, poorly defined loop with an unusual $glycine_{12}$ motif. Only a tiny portion of the protein, the C-terminal domain, was identified as containing an SLH domain. Wild-type protein, P1 and P2 failed to hydrolyse ONPG. Thus, protein 0050 and its variants are not by themselves a β-galactosidase/β-galactanase enzyme. Qualitative kinetics and preliminary quantitative kinetics identified successful binding between protein variations (P1 and P2) and various pectin substrates. Interestingly, the most successful binding affinity was observed for complex substrates, in particular kiwiberry pectin. However, there was significant binding with samples containing only monosaccharides of pyranose sugars. Preliminary binding affinities suggest the need for further investigation into the binding of the monosaccharides: arabinose, rhamnose and galactose. These results may identify a new family of carbohydrate-binding modules or the possibility of a protein containing multiple CBMs domains.

Chapter 6 General Discussions and Conclusions

This Project provides three overall conclusions: (1) New Zealand kiwiberries have a different pectin composition during fruit softening; (2) *M. pectinilyticus* does possess β-galactosidase/β-galactanase activity, though no singular protein was identified; and (3) Protein 0050, thought to be a β-galactosidase/β-galactanase, is involved in pectin degradation as a putative anchoring protein, which may be involved in a multi-protein complex.

6.1 Softening differences between New Zealand Kiwiberry cultivars

The first objective of this research was to characterise New Zealand kiwiberry pectin. We incorporated the characterisation of three cultivars, Hortgem Tahi, Takaka Green and Marju Red, to reflect what is commercially grown and sold (Aitken & Hewett, 2019). Summarising Chapter 3, the data shows differences between the three cultivars in their firmness, pectin solubilisation and glycan composition. It is important to reaffirm, the glycan array composition is of semi-quantitative analysis and does not mean absolute pectin compositions (Moller *et al.,* 2012).

The texture of fruit is an important quality that contributes to consumer acceptability. These textures are manipulated by the biochemical and/or chemical changes accompanied by the dissolution of the middle lamella, as well as by the modification of cell wall polymers, in particular, pectin and its neutral side chains (Brummell, 2006; Pose *et al*., 2018; Sutherland *et al*., 2017). Measurement of firmness by compression, is a physical test to observe fruit softening. The firmness of three kiwiberry fruits ripened rapidly between Day 0 and Day 4. Sutherland *et al*. (2017) also observed a rapid reduction in firmness (N) over the course of 5 days. Interestingly, core tissue firmness studied in Sutherland *et al*., (2017) was initially more firm than the whole kiwiberry, suggesting *Actinidia arguta* fruit tissues start to ripen at different times (Boyd *et al*., 2007; White *et al*, 2005; Jackson & Harker, 1997). The differences in fruit maturity is also reflected at the time of harvest (Sutherland *et al*., 2017; Latocha, 2021). The softening of fruits arises from cell wall modifications, in particular pectin solubilisation (Brummell, 2006; Paniague *et al*., 2014). Extracting soluble pectin allows for the investigation of cell wall changes over softening. The amount of uronic acid within each cultivar at Day 0 and Day 4, coincides with the variation of fruit firmness. When the fruit firmness decreases, the amount of pectin water-soluble increases, thus indicating pectin solubilisation. The decrease in uronic acid content indicates pectin degradation,

and results in the weakening of plant cell walls. This relates to low fruit firmness (Sutherland *et al*., 2017; White *et al*., 2005). Noticeably, Hortgem Tahi and Takaka Green after four days are still undergoing pectin solubilisation while Marju Red has reached its 'peak' (Figure 3.2). This is consistent with the reoccurring theme of ripening differences. The changes of firmness and uronic acid content could be the action of pectin-metabolising enzymes (Wang *et al.,* 2018) but also, the presence of hydroxyl radicals (Airianah *et al*., 2016), changing the cell turgor and primary cell wall structures (Brummell, 2006). Sutherland *et al*. (2017) suggests the fruit ripening rate may be the result of the early reduction of galactose content which can be found in the side chains of RG-I. The reduction of galactose would increase the cell wall porosity and thereby assists in pectin solubilisation (Latocha, 2021; Sutherland *et al*., 2017). Antibody labelling of glycan epitopes suggests pectin degradation between Day 0 to Day 4, where each cultivar has its own pectin composition. While this difference is not unusual, Hortgem Tahi and Takaka Green pectin composition patterns mimic each other, with different densities of antibody labelling. Marju Red labelling of pectin is quite different. Homogalacturonan-related epitopes (identified by antibodies: LM18, LM19 and LM20) have shown that there is a loss of esterified pectin and polygalacturonic acid content between Day 0 and Day 4 as the fruit softens. Even the loss in LM19 – 20 binding, a significant amount of labelling was retained (especially in LM18 and LM19). This suggests that, even though the decrease indicates pectin degrading, the depolymerisation of the HG chains may not be fully progressed. Therefore, pectin may be still trapped within the wall matrix (Sutherland *et a*l., 2017; Verhertbruggen *et al*., 2009). Antibody LM5 (specific for galactan) and LM6 (specific for arabinan) identifies side chains attached to RG-I. The significant decrease in LM5 and LM6 labelling indicates a substantial breakdown of RG-I side chains during ripening. The presence of xylogalacturonan (LM8, Figure 3.9) in all soluble-pectin fractions for all cultivars was intriguing as this pectic structure is only present in small proportions and is not usually found in cell wall studies of ripening fruits (Tucker *et al*., 2017; Thibault & Ralet, 2001). Xylogalacturonan has been identified in a small number of fruits and vegetables (Willats et al., 2003). The presence of xylogalacturonan in Day 4, suggested that the polymer persists in kiwiberry and is consumed.

For gut microbiota and health benefits, the results reflect the composition of kiwiberry pectin that would be ingested. Degradation of bigger pectin polymers into pectin oligosaccharides (POS) are selectively fermented by intestinal flora (Beukema *et al*., 2020; Wicker *et al*., 2014; Zhu *et al*., 2019). In addition, POS have been proposed to provide beneficial physiological activities (Parker *et*

al., 2010; Zhu *et al*., 2019). Interestingly, the presence of xylogalacturonan retained at Day 4, gives the assumption that XGA is consumed. The purpose of xylogalacturonan within the human gut and its potential interaction with the gut microbiota has yet to be extensively researched (Wicker *et al*., 2014). One study shows that xylogalacturonan possessed an anti-depressant like effect in mice (Patova *et al*., 2021). However, the molecular mechanisms at a gastro-intestinal level, which regulates the physiological reaction to decrease depression, remain unclear. In addition, little is known about pectinases that can degrade xylogalacturonan in the human gut (Zandleven *et al*., 2005).

6.2 *Monoglobus pectinilyticus* β-galactosidase/β-galactanase activity detected

The second objective of this research was to identify, isolate and characterise a potentially new class of β-galactanase/β-galactosidase (BGAL), from *M. pectinilyticus*. In Chapter 4, by using potato galactan (β-1, 4-galactan), I detected activity of a β-galactanase/β-galactosidase enzyme expressed by *M. pectinilyticus*. Manipulated cultures were grouped into different molecular weight cut-offs (MWCOs), with activity dominantly found in extracts with protein molecular weights ≥100 kDa, which is consistent with the general monomeric size of β-galactosidase (Schmidt & Stougaard, 2010; Bartesaghi *et al*., 2014). In contrast, BGAL activity was also detected in the smaller protein MWCO. The majority of the galactanases (or novel β-galactosidase) currently discovered are of different sizes, ranging from 36 kDa to 120kDa (Böger *et al*., 2019; Bonnin *et al.*, 1995; Buckeridge & Reid, 1994; Hirofumi *et al*., 2014; Nakano *et al*., 1990; Ogasaware *et al.*, 2007; Sakamoto *et al*., 2013). Despite the difference in sizes, these galactanases were able to catalyse the hydrolysis of *o*-nitro phenyl-beta-D-galactopyranoside. This has shown that β-galactanase/β-galactosidase enzymes can be structurally different from one another, despite possessing a similar catalytic function.

Inconsistent with literature (Buckeridge & Reid, 1994; Nakano *et al*, 1990; Ogasaware *et al*, 2007; Sakamoto *et al*., 2013), the use of several protein purification methods (Simpson, 2004) did not identify an individual β-galactanase/β-galactosidase. Numerous proteins persisted throughout the process, with samples still detecting BGAL activity.

Protein sequencing analyses of samples possessing BGAL activity also provided mostly inconclusive findings. Collectively, while BGAL activity was successfully distinguished, extracting the probable protein candidate proved difficult. The presence of numerous proteins with uncharacterised functions questions the mechanism as to how BGAL activity is performed. In attempt to identify the enzyme responsible for BGAL activity, protein 0050 was cloned and successfully expressed. However, this protein did not possess the catalytic function needed to hydrolyse ONPG and β-1, 4-galactan. With these results in mind, it is not far-fetched to suggest that a multi-complex protein may be involved that facilitates the degradation of potato galactan and ONPG. Multi-protein systems such as cellulosomes, are known for complex polysaccharide degradation (e.g. cellulose). The combination of multiple enzymes may facilitate the breakdown of the complex polymers (Artzi *et al*., 2016).

6.3 Preliminary characterisation of protein 0050

Due to the observation that no one protein exhibited β-galactanase/β-galactosidase activity, the nature of protein 0050 and its involvement with pectin was investigated. . Summarising Chapter 5, protein 0050 possesses an unusual structure, which binds to various pectin substrates with high affinity to complex pectin structures and monosaccharides of pyranose sugars. Protein 0050 structure is predicted to have two main domains connected by a glycine$_{12}$ motif linker/loop. The modelling predicted with high confidence that the N-terminal domain is made of β-sheets, whereas the C-terminal domain is α-helical. The N-terminal domain is highly predicted to contain 10 β-sheets separated by short loops, whereas C-terminal domain contains 6 α-helices (Figure 5.3 & 5.4). The N-terminal domain had weak similarity to cell adhesion and cohesion type proteins (Table 5.3). This together with the hydrophobic anchor (amino acid residues 3 – 27; Figure 5.4) tells us that the N-terminal domain of protein 0050 may be used for some sort of cell adhesion and/or a cohesion domain. On the other hand, the C-terminal had similarity to sugar binding by SLH domain (Table 5.3; Blacker *et al*., 2018) and is unlikely to possess catalytic function due to inconsistent catalytic results (Table 5.4). However, SLH domains are typically anchoring domains into the cell surface (Kosugu *et al*., 2002; Sára, 2001), not ligand- or protein-binding domains. Studies by Ozdemir *et al*. (2012) identified two SLH domain-containing proteins from *C. saccharolyticus* that are distinguished from the typical cell wall anchoring SLH-domain proteins

due to the presence of putative binding domains. Thus, the SLH present in protein 0050 may have an alternative purpose, such as pectin binding function rather than a cell anchor

The glycine$_{12}$ loop was poorly modelled due to the unusual repeating glycine amino acid. These repeating glycine motifs are not uncommon in the literature of plant proteins, where substantial research has explored the plant glycine-rich proteins and categorised into families (Czolpinska & Rurek, 2018; Sachetto-Matins *et al*., 2000). Although protein 0050 is not dominated by glycine residues, the placement of the glycine motif seems to act as a linker between the N-terminal and C-terminal domain. Interestingly, these glycine-rich linkers are important components for generating fusion proteins, due to their high flexibility and ability to not interfere with interacting domains (Chichili *et al*., 2013). Some applications of protein linkers between domains allow the independent folding of the domains. Alternatively, the properties of the linker may directly affect the functional properties of the protein (Rosmalen *et al*., 2017). Typically, functional linkers are repeats of glycine and serine residues, combining not only the flexibility of glycine, but also the solubility of serine (Rosmalen *et al*., 2017). These linkers also hold a coil structure, which is not a true 'regular' secondary structure (Chichili *et al*., 2012; Rosmalen *et al*., 2017) to allow for domain interactions.

Intriguingly, the majority of the proteins that share a similarity to protein 0050 matched to only one of the two domains, except a cellulosome-anchoring protein (UniProt: Q06848) produced by *Acetivibrio thermocellus*. This protein is anchored into the cell wall by SLH domains which then allow the cohesion domain to be exposed, and available for binding to the C-terminal end of a CipA (cellulosome-interacting protein A) (Brás *et al*., 2012; Fujino *et al*., 1993).

Ligand binding tests performed on protein 0050 showed there were interactions with pectin-saccharides. Though binding affinity is dependent on substrate composition, there were strong interactions between protein 0050 and complex pectin structures. Testing out individual subunits of pectin showed successful binding affinity for monosaccharides containing the pyranose sugar structure. The removal of the glycine$_{12}$ motif in protein 0050, did not clarify the binding of the domain(s). Contrasting results showed that for some substrates binding affinity to P2 was stronger than to P1, whereas for other substrates binding affinity was weaker. However, for variant P2 protein the glycine$_{12}$ linker loop was reduced from ~90 aa to ~29 residues, and so the shortening may have affected the folding of the domains and/or the interactions of the domain(s) to substrates. Binding between protein and ligand are typically identified as carbohydrate binding

modules (CBM) (Boraston *et al*., 2004). However, no singular CBM family was identified to be able to class the substrate interactions we had observed. Thus, there is a possibility that the pectin binding domain of protein 0050 may be a new class of CBM which incorporates a SLH-domain and/or may contain multiple CBMs sites. This may explain why there were numerous binding affinities to various pectin substrates.

Based on the computational analyses and ligand binding of protein 0050, I propose a putative model for its involvement in pectin degrading (Figure 6.1). This proposed function is similar to some enzymes involved in the multi-protein complex such as cellulosomes (Artzi *et al*., 2017). The model would attempt to address the function of protein 0050 and why it was found in positive BGAL detected samples. I propose that protein 0050 is an anchoring protein, involved in both the binding of pectin and the binding of a multi-complex scaffold, to facilitate pectin breakdown. It should be noted that this model would remain hypothetical until there is further research to solidify its purpose.

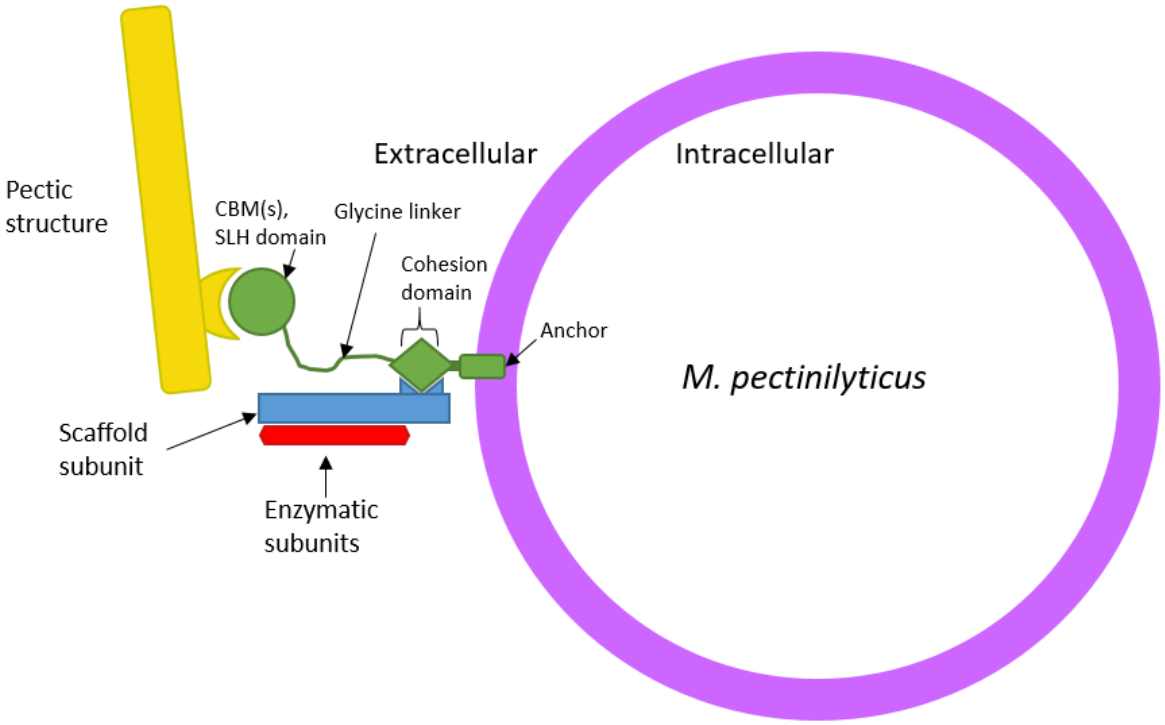


Figure 6.1 A proposed model for protein 0050 involvement in pectin degradation. Purple: *M. pectinilyticus*; Green: protein 0050; Yellow: Pectic structure; Blue: primary scaffold subunit; Red: Enzymatic subunits.

Collectively, the computational analysis of the protein 0050 structure and the ligand binding has highlighted how protein 0050 may be involved in pectin degradation. These results are preliminary but provide a strong basis of the function of protein 0050. It can be hypothetically-assumed that protein 0050 acts as a cohesion-type protein, which binds to pectic-saccharides, where the binding favours complex pectin structures with a strong affinity to monosaccharides that have a pyranose structure. This interaction would hypothetically propose that protein 0050 could act as part of the scaffolding subunit of a multi-protein complex (similar to a cellulosome) that binds the pectin substrate, though further investigation is needed. Typically, the scaffold subunit would carry enzymatic subunits. The glycine$_{12}$ motif linker provides the flexibility to allow the two domains to interact.

6.4 Conclusions

In conclusion, the findings from this thesis sheds light on the pectin composition and fruit softening changes of three New Zealand kiwiberry cultivars, and the pectin degrading capabilities that *M. pectinilyticus* holds. The investigation into the kiwiberries illustrate Hortgem Tahi, Takaka Green and Marju Red possess pectic-saccharides, xylogalacturonan, and that each cultivar possesses a different pectin composition during ripening. The preliminary data suggest a framework of the kiwiberry pectin composition and what types of pectic fractions would be involved in digestion and fermentation. The proteomics analysis of *M. pectinilyticus* reveals that the bacterium does possess β-galactosidase/β-galactanase activity and expresses a protein which is involved in a potentially primitive pectin degrading strategy. Protein 0050 is proposed as an anchoring protein with the potential capability to bind pectin and form part of a multi-complex scaffold. The two domains would work together to facilitate the degradation of pectin. This study provides a novel pectin composition of Takaka Green and Marju Red. It also showcases the possibly of a different pectin degradation system utilised by *M. pectinilyticus,* which extends similarities to the multi-complex protein, cellulosome.

6.5 Future Directions

This thesis reports on the findings of three kiwiberry pectin compositions during fruit softening and the identification of *Monoglobus pectinilyticus* β-galactosidase/β-galactanase activity and a pectin-ligand binding protein. While, composition of pectin is extensively researched, insight to New Zealand kiwiberry pectin is relatively new. Neutral sugar composition by Gas Chromatography, biochemical post-harvest analysis (e.g., ethylene production and soluble solids content), microscopy staining and molecular weight by Size Exclusion Chromatography together would provide an added in-depth analysis of the kiwiberry pectin. Investigations into the potential of New Zealand kiwiberry pectin interactions with human health (in particular gut health) have yet to be extensively researched. This can be facilitated by research into simulated gastrointestinal digestion and immunology cell culture.

The nature of β-galactosidase/β-galactanase activity by *M. pectinilyticus* still remains unknown. Typically, characterisation by cloning and expressing recombinant proteins is followed by enzymatic assay (e.g. colorimetric and/or spectrophotometric) and/or ligand/glycan binding assays (e.g. antibody-mediated glycan array). Incorporation of screening a large number of proteins and additional purification methods (e.g. affinity chromatography and size exclusion chromatography) needs to be highly considered. Due to the potential of a multi-complex protein to be involved in pectin degradation, additional investigations are needed. Creating mutants from wild-type proteins needs to be considered for the investigation into the function and purpose of the proteins involved. Expressed and purified proteins should be tested for both catalytic function, ligand binding activities and degradation of end-product for analysis. For the interest of protein – protein interaction, numerous biochemical applications are available including Western-blot analysis, cross-linkage analysis, protein – protein binding kinetics, thin-layer chromatography and co-immunoprecipitation methods. In addition, crystallisation of proteins can be used to characterise the proteins 3D structure, providing insight to architecture of sugar-binding domains and/or catalytic sites. Any additional microbiological and proteomic characterisation of the *M. pectinilyticus* pectin degradation is required to provide better understanding of the mechanism in place.

6.6 Research affected

It is important to recognise the impact of world and research events that have affected the research investigated here.

6.6.1 Worldwide pandemic

The world has been greatly affected by the sudden, rapid emergence of SARS-CoV-2 in 2019. Since then, we have seen the world rapidly closing its borders and restriction to national travel. New Zealand went into lockdown level 4 on the 25th March 2020 – during the first semester of my studies. This lockdown period continued till the move to level 3 on the 27th April 2020, and to level 2 on the 13th May 2020. During this time, there was no access to either Massey University campus or the Plant and Food Research, Palmerston North site. I attended back at Plant and Food Research campus at the start of June, 2020. Due to the Covid-19 outbreak and lockdown, this resulted in a delay of laboratory work. On top of this, due to the out-sourcing for consumables, ordering and delivery times were extended. Similarly, on the 17th Aug 2021, due to the Delta-variant outbreak, the whole of New Zealand was again moved into lockdown level 4. The Manawatū region transitioned into level 3, the 1st September and to level 2 on the 7th September. Again, as above for Covid-19, the same rules and paperwork were required. In total, about 5 months of laboratory work was affected reducing the scope of this thesis.

6.6.2 GenScript Failing to codon optimise protein

Due to the limited time and affordability of ordering vector and gene, protein 0050 and vector pET28a (+) were ordered from GenScript. Unfortunately, GenScript did not codon optimise the protein sequence despite explicit instructions to do so and thus led to 5 – 6 weeks of protein expression troubleshooting (data not shown). By checking the codons and running protein 0050 through a codon optimisation software, we were able to identify the issue. Protein 0050 contains rare codons and thus, the E. coli system BL21 (DE3) was having issues expressing the protein in large quantities to purify it. Fortunately, GenScript was able to provide a substantial discount on ordering new synthetic vector and protein 0050 – one without the N-terminus and another without the N-terminus and the glycine motif. However, due to this loss of time, characterisation of these proteins was limited to one protein.

Chapter 7 References

Abbott, D. W. and A. B. Boraston. (2008). Structural biology of pectin degradation by *Enterobacteriaceae*. *Microbiology and Molecular Biology Reviews, 72*(2), 301–316. https://doi.org/10.1128/MMBR.00038-07.

Airianah, O. B., Vreeburg, R. A. M., & Fry, S.C. (2016). Pectic polysaccharides are attacked by hydroxyl radicals in ripening fruit: Evidence from a fluorescent fingerprinting method. *Annals of Botany, 117*(3), 441–455. https://doi.org/10.1093/aob/mcv192.

Aitken, A & Hewett, E. (2019). *New Zealand Horticulture*. www.freshfacts.co.nz.

Anuradha, K., & Bhawni, B. P. (2019). Structural insights into the molecular mechanisms of pectinolytic enzymes. *Journal of Proteins and Proteomics, 10*(4), 325–344. https://doi.org/10.1007/s42485-019-00027-5.

Bairoch, A., & Apweiler, R. (1996). The SWISS-PROT protein sequence data bank and its new supplement TREMBL. *Nucleic Acids Research, 24*(1), 21–25. https://doi.org/10.1093/nar/24.1.21.

Bang, S., Kim, G., Lim, M. Y., Song, E. –J., Jung, D. –H., Kum, J. –S., Nam, Y. –D., Park, C. –S., & Seo, D. –H. (2018). The influence of in vitro pectin fermentation on the human fecal microbiome. *AMB Express, 8*(98). https://doi.org/10.1186/s13568-018-0629-9.

Blumenkranktz N, & Absore-Hansen, G. (1973). New method for quantitative determination of uronic acids. *Analytical Biochemistry, 54*(2), 484–489. https://doi.org/10.1016/0003-2697(73)90377-1.

Böger, M., Hekelaar, J., Leeuwen, S. S., Dijkhuizen, L., & Lammerts van Bueren, A. (2019). Structural and functional characterization of a family GH53 β-1, 4-galactanase from *Bacteroides thetaiotaomicron* that facilitates degradation of prebiotic galactooligosaccharides. *Journal of Structural, Biology 205*(1), 1–10. https://doi.org/10.1016/j.jsb.2018.12.002.

Boyd, L., Petley, M., Martin, P., Jackman, R., Requejo, C., & Williams, M. (2007). The effects of harvest maturity on fruit quality in Actinidia arguta. *N.Z Kiwifruit Journal,* (179), 36–39.

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry, 7*(72), 248–254. https://doi.org/10.1016/0003-2697(76)90527-3.

Brejnholt, S. M. (2009). Pectin. *Food stabilisers, thickeners and gelling agents,* 237–265. https://doi.org/10.1002/9781444314724.ch13.

Bringans, S. D., Kendrick, T., Lui, J., & Lipscombe, R. J. (2008) A comparative study of the accuracy of several *de novo* sequencing software packages for datasets derived by matrix-assisted laser desorption/ionisation and electrospray. *Rapid Communications Mass Spectrometry, 22*(21), 3450–3454. https://doi.org/10.1002/rcm.3752.

Brummell, D. A. (2006). Cell wall disassembly in ripening fruit. *Functional Plant Biology, 33*(2), 103–119. https://doi.org/10.1071/FP05234.

Buckeridge, M. S., & Reid, J. S. G. (1994). Purification and properties of a novel β-galactosidase or exo-(1 → 4)-β-D-galactanase from the cotyledons of germinated *Lupinus angustifolius* L. seeds. Planta, 192(4), 502–511. https://doi.org/10.1007/BF00203588.

Bule, P., Pires, V. M. R., Fontes, C. M. G. A., & Alves, V. D. (2018). Cellulosome assembly: Paradigms are meant to be broken. *Current Opinion in Structural Biology*, *49,* 154–161. https://doi.org/10.1016/j.sbi.2018.03.012.

Caffall, K. H., & Mohnen, D. (2009). The structure, function, and biosynthesis of plant cell wall pectic polysaccharides. Carbohydrate Research, *344*(14), 1879–1900. https://doi.org/10.1016/j.carres.2009.05.021.

Chung, W. S. F., Walker, A. W., Louis P., Parkhill, J., Vermeiren, J., Bosscher, D., Duncan, S. H., & Flint, H. J. (2016). Modulation of the human gut microbiota by dietary fibres occurs at the species level. *BMC Biology, 14*(3). https://doi.org/10.1186/s12915-015-0224-3.

Chung, W. S. F., Meijerink, M., Zeuner, B., Holck, J., Louis, P., Meyer, A. S., Wells, J. M., Flint, H. J., Duncan, S. H. (2017). Prebiotic potential of pectin and pectic oligosaccharides to promote anti-inflammatory commensal bacteria in the human colon. *FEMS Microbiology ecology, 93*(11). https://doi.org/10.1093/femsec/fix127.

Coenen, G. J., Bakx, E. J., Verhoef, R. P., Schols, H. A., & Voragen, A. G. J. (2007). Identification of the connecting linkage between homo- or xylogalacturonan and rhamnogalacturonan type I. *Carbohydrate Polymers, 70*(2), 224–235. https://doi.org/1016/j.carbpol.2007.04.007.

Comstock, L. E. (2009). Importance of glycans to the host-*Bacteroides* mutualism in the mammalian intestine. *Cell Host & Microbe, 5*(6), 522–526. https://doi.org/10.1016/j.chom.2009.05.010.

Conway, J. M., Pierce, W. S., Le, J. H., Harper G. W., Wright, J. H., Tucker, A. L., Zurawski, J. V., Lee, L. L., Blumer-Schuette, S. E., & Kelly, R. M. (2016). Multidomain, surface layer-associated glycoside hydrolases contribute to plant polysaccharide degradation by *Caldicellulosiruptor* species. *Journal of Biological Chemistry*, *291*(13), 6732–6747. https://doi.org/10.1074/jbc.m115.707810.

Cossio, F., Debersaques, F., & Latocha, P. (2015). Kiwiberry (*Actinidia arguta*): New perspectives for a great future. *Acta Horticulturae, 1096,* 423–434. https://doi.org/10.17660/ActaHortic.2015.1096.51.

Datson P. M., & Ferguson A. R. (2011). *Actinidia.* In: Kole C. (eds) Wild Crop Relatives: Genomic and Breeding Resources. Springer, Berlin, Heidelberg. https://link.springer.com/chapter/10.1007/978-3-642-20447-0_1.

Debersaques, F., Mekers, O., Decorte, J., Labeke, M. V., Schoedl-Hummel, K., & Latocha, P. (2015). Challenges faced by commercial kiwiberry (*Actinidia arguta* Planch.) Production. *Acta Horticulturae, 1096*, 435–442. https://doi.org/10.17660/ActaHortic.2015.1096.52.

Despres, J., Forano, E., Lepercq, P., Comtet-Marre, S., Jubelin, G., Yeoman, C. J., Miller, M. E. B., Fields, C. J., Terrapon, N., Bourvellec, C. L., Renard, C. M. G. C., Henrissat, B., White, B. A., & Mosoni, P. (2016). Unraveling the pectinolytic function of *Bacteroides xylanisolvens* using a RNA-seq approach and mutagenesis. *BMC Genomics, 17*(147). https://doi.org/10.1186/s12864-016-2472-1.

Dimopoulou, M., Alba, K., Sims, I. M., & Kontogiogos, V. (2021). Structure and rheology of pectic polysaccharides from baobab fruit and leaves. *Carbohydrate Polymers, 273*, 118540. https://doi.org/10.1016/j.carbpol.2021.118540.

Dongowski, G., Lorenz, A., & Anger, H. (2000). Degradation of pectins with different degrees of esterification by *Bacteroides thetaiotaomicron* isolated from human gut flora. *Applied and Environmental Microbiology, 66(*4), 1321–1327. https://doi.org/10.1128/aem.66.4.1321-1327.2000.

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), 497–504. https://doi.org/10.1038/nrmicro3050.

Ferreira, A., Moreno, F. J., Cueva, C., Gil-Sánchez, I., & Villamiel, M. (2019). Behaviour of citrus pectin during its gastrointestinal digestion and fermentation in a dynamic simulator (simgi®). *Carbohydrate Polymers, 207,* 382–390. https://doi.org/10.1016/j.carbpol.2018.11.088.

Flint, H. J., & Bayer, E. A. (2008). Plant cell wall breakdown by anaerobic microorganisms from the mammalian digestive tract. *Annals of the New York Academy of Science, 1125*, 280–288. https://doi.org/10.1196/annals.1419.022.

Flint, H. J., Scott, K. P., Duncan, S. H. Louis, P., & Forano, E. (2012). Microbial degradation of complex carbohydrates in the gut. *Gut Microbes, 3*(4), 289–306. https://doi.org/10.4161/gmic.19897.

Gawkowska, D. Cybulska, J. & Zdunek, A. (2018). Structure-related gelling of pectins and linking with other natural compounds: A review. *Polymers, 10*(7), 762. https://doi.org/10.3390/polym10070762.

Hallett, I & Sutherland, P. (2005). Structure and development of kiwifruit skins. *International Journal of Plant Science, 166*, 693–704. https://doi.org/10.1086/431232.

Han, N., Park, H., Kim, C. –W., Kim, M. –S., & Lee, U. (2019). Physiochemical quality of hardy kiwifruit (*Actinidia arguta* L. cv. Cheongsan) during ripening is influences by harvest maturity. *Forest Science and Technology, 15*(4), 187–191. https://doi.org/10.1080/21580103.2019.1658646.

Harholt, J., Suttangkakul, A., Scheller, H. V. (2010). Biosynthesis of pectin. *Plant Physiology, 153*(2), 384–395. https://doi.org/10.1104/pp.110.156588.

Harris, P. J. & B. G. Smith. (2006). Plant cell walls and cell-wall polysaccharides: structures, properties and uses in food products. *International Journal of Food Science and Technology 41*, 129–143. https://doi.org/10.1111/j.1365-2621.2006.01470.x.

Hirofumi, N., Takenishi, S., & Watanabe, Y. (2014). Purification and properties of two galactanases from *Penicillium citrinum*. *Agricultural and Biological Chemistry, 49*(12), 3445–3454. https://doi.org/10.1080/00021369.1985.10867285.

Huber, D. (1992). The inactivation of pectin depolymerase associated with isolated tomato fruit cell wall: implications for the analysis of pectin solubility and molecular weight. *Physiologia Plantarum, 86*(1), 25–32. https://doi.org/10.1111/j.1399-3054.1992.tb01307.x.

Hugouvieux-Cotte-Pattat, N., Condemine, G., & Shevchik, V. E. (2014). Bacterial pectate lyases, structural and functional diversity. *Environmental Microbiology Reports, 6(*5), 427–440. https://doi.org/10.1111/1758-2229.12166.

Ilk, N., Egelseer, E., & Sleytr, U. B. (2011). S-layer fusion proteins - Construction principles and applications. *Current Opinion in Biotechnology*, *22*(6), 824–831. https://doi.org/10.1016/j.copbio.2011.05.510.

Jensen, J. K., Sorensen, S. O., Harholt, J., Geshi, N., Sakuragi, Y., Moller, I., Zandleven, J., Bernal, A. J., Jensen, N. B., Soresen, C., Pauly, M., Beldman, G., Willats, W. G. T., & Scheller, H. V. (2008). Identification of a xylogalacturonan xylosyltransferase involved in pectin biosynthesis in Arabidopsis. *Plant Cell, 20(*5), 1289–1302. https://doi.org/10.1105/tpc.107.050906.

Jones, L., Seymour, G. B., Knox, J. P. (1997). Localisation of pectic galactan in tomato cell walls using a monoclonal antibody specific to (1[->]4)-B-D-Galactan. *Plant Physiology, 113*(4), 1405–1412. https://doi.org/10.1104/pp.113.4.1405.

Jumper, J., Evans R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredas, B., Nikolov, S., Jain, R., Adler, J., …Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature, 596*, 583–596. https://doi.org/10.1038/s41586-021-03819-2.

Jung D. -H., Seo, D. –H., Shin, J. –H., Park, C. –S., & Chung, W. –H. (2020). Genome analysis of Enterococcus mundtii Pe103, a Human gut originated pectinolytic bacterium. *Current Microbiology, 77,* 1839–1847. https://doi.org/10.1007/s00284-020-01932-5.

Kaczmarksa, A., Pieczywek, P. M., Cybulska, J., & Zdunek, A. (2022). Structure and functionality of Rhamnogalacturonan I in the cell wall and solution: A review. *Carbohydrate Polymers 278*, 118909. https://doi.org/10.1016/j.carbpol.2021.118909.

Kim, C. C., Healey, G. R., Kelly, W. J., Patchett, M. L., Jordens, Z., Tannok, G. W., Sims, I. M., Bell, T. J., Hedderley, D., Henrissat, B., & Rosendale, D. I. (2019). Genomic Insights from *Monoglobus pectinilyticus*: A pectin-degrading specialist bacterium in the human colon. *ISME Journal, 13*(6), 1437–1456. https://doi.org/10.1038/s41396-019-0363-6.

Kim, C. C., Kelly, W. J., Patchett, M. L., Tannock, G. W., Jordens, Z., Stoklosinki, H. M., Taylor, J. W., Sims, I. M., Bell, T. J., & Rosendale, D. I. (2017). *Monoglobus pectinilyticus* gen. nov., sp. nov., a pectinolytic bacterium isolated from human faeces. *International Journal of Systematic and Evolutionary Microbiology, 67*(12), 4992–4998. https://doi.org/10.1099/ijsem.0.002395.

Kong, F. B. & Singh, R. P. (2009). Modes of disintegration of solid foods in simulated gastric environment. *Food Biophysics, 4*(3), 180–190. https://doi.org/10.1007/s11483-009-9116-9.

Krissinel, E. (2010). Crystal contacts as nature's docking solutions. *Journal of computational chemistry, 31(*1), 133–143. https://doi.org/10.1002/jcc.21303.

Krissinel, E., & Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. *Journal of Molecular Biology*, *372*(3), 774–797. https://doi.org/10.1016/j.jmb.2007.05.022.

Labus, K. (2018). Effective detection of biocatalysts with specified activity by using a hydrogel-based colourimetric assay – β-galactosidase case study. *PLOS ONE, 13*(10), e0205532. https://doi.org/10.1371/journal.pone.0205532.

Larsen, N., Bussolo de Souza, C., Krych, L., Cahú, T. B., Wiese, M., Kot, W., Hansen, K. M., Blennow, A., Venema, K., & Jesperson, L. (2019). Potential of Pectins to Beneficially Modulate the Gut Microbiota depends on their Structural Properties. *Frontiers in Microbiology 10,* 223. https://doi.org/10.3389/fmicb.2019.00223.

Latocha, P. (2017). The nutritional and health benefits of kiwiberry (*Actinidia arguta*) – A review. *Plants Foods Human Nutrition, 72*(4), 325–334. https://doi.org/10.1007/s11130-017-0637-y.

Latocha, P., Debersaques, F., & Hale, I. (2021). *Actinidia arguta* (kiwiberry): Botany, production, genetics, nutritional value, and postharvest handling. In Horticultural Reviews, I. Warrington (Ed.). https://doi.org/10.1002/9781119750802.ch2

Leontowicz, H., Leontowicz, M., Latocha, P., Jesion, I., Park, Y. –S. Katrich, E., Barasch, D., Nemirovski, A., Gorinstein, S. (2016). Bioactivity and nutritional properties of hardy kiwi fruit *Actinidia arguta* in comparison with *Actinidia deliciosa* 'Hayward' and *Actinidia eriantha* 'Bidan'. *Food Chemistry, 196,* 281–291. https://doi.org/10.1016/j.foodchem.2015.08.127.

Lobley, A., Sadowski, M. I., & Jones, D. T. (2009). pGenTHREADER and pDomTHREADER: new methods for improved protein fold recognition and superfamily discrimination. *Bioinformatics, 25(*14), 1761–1767. https://doi.org/10.1093/bioinformatics/btp302.

Lodish H., Berk, A., Zipursky, S. L., Matsudaira, P., Baltimore, D., & Darnell, J. (2000). The Dynamic Plant Cell Wall. Molecular Cell Biology. 4th edition. New York: W. H. Freeman; Section 22.5.

Lopesz-Stiles, M., Duncan, S. H., Garcia-Gil, L. J., & Martinez-Medina, M. (2017). *Faecalibacterium prausnitzii:* from microbiology to diagnostics and prognostics. *ISME Journal, 11,* 841–885. https://doi.org/10.1038/ismej.2016.176.

Macdonald, S. S., Blaukopf, M., & Withers, S. G. (2014). N-acetylglucosaminidases from CAZy family GH3 are really glycoside phosphorylases, thereby explaining their use of histidine as an acid/base catalyst in place of glutamic acid. *Journal of Biological Chemistry, 290(*8), 4887–4895. https://doi.org/10.1074/jbc.M114.621110.

Martin, R., Miquel, S., Benevides, L., Bridonneau, C., Robert, V., Hudault, S., Chain, F., Berteau, O., Azevedo, V., Chatel, J. M., Sokol, H., Bermudez-Humaran, L. G., Thomas, M., Langella, P. (2017). Functional characterization of novel *Faecalibacterium prausnitzii* strains isolated from healthy volunteers: A step forward in the use of *f. prausnitzii* as a next-generation probiotic. *Frontiers in Microbiology*, *8*, 1226. https://doi.org/10.3389/fmicb.2017.01226.

Martínez, M. G, Acosta, M. P., Candurra, N. A., & Ruzal, S. M. (2012). S-layer proteins of *Lactobacillus acidophilus* inhibits JUNV infection. *Biochemical and Biophysical Research Communications, 422*(4), 590–595. https://doi.org/10.1016/j.bbrc.2012.05.031.

Matthews, B. W. (2005). The structure of *E. coli* β-galactosidase. *Comptes Rendus Biologies, 328*(6), 549–556. https://doi.org/10.1016/j.crvi.2005.03.006.

McGuffin, L. J., Bryson, K., & Jones, D. T. (2000). The PSIPRED Protein Structure Prediction Server. *Bioinformatics, 16*(4), 404–405. https://doi.org/10.1093/bioinformatics/16.4.404.

Desvaux, M., Dumas, E., Chafsey, I., & Hébraud, M. (2006). Protein cell surface display in Gram-positive bacteria: from single protein to macromolecular protein structure. *FEMS Microbiology Letters, 256*(1), 1–15. https://doi.org/10.1111/j.1574-6968.2006.00122.x.

Mohnen, D. (2008). Pectin structure and biosynthesis. *Current Opinion in Plant Biology, 11*(3), 266-277. https://doi.org/10.1016/j.pbi.2008.03.006.

Elshahed, M. S., Miron, A., Aprotosoaie, A. C., & Farag, M. A. (2021). Pectin in diet: Interactions with the human microbiome, role in gut homeostasis, and nutrient-drug interactions. *Carbohydrate Polymers, 255*. https://doi.org/10.1016/j.carbpol.2020.117388.

Nakajima, N., Ishihara, K., Matsubara, K., Matsuura, Y. (1999). Degradation of pectic substances by two pectate lyases from a human intestinal bacterium, *Clostridium butyricum-beijerinckii* group. *Journal of Bioscience and Bioengineering, 88*(3), 331–333. https://doi.org/10.1016/s1389-1723(00)80020-1.

Nakano, H., Takenishi, S., Kitahata, S., Kingugasa, H., & Watanabe, Y. (1990). Purification and characterisation of an exo-1, 4-B-galactanase from a strain of *Bacillus subtilis*. *European Journal of Biochemistry, 193*(1), 61–67. https://doi.org/10.1111/j.1432-1033.1990.tb19304.x.

O'Donoghue, E. M., Somerfield, S. D., Deroles, S. C., Sutherland, P. W., Hallett, I. C., Erridge, Z. A., Brummell, D. A., & Hunter, D. A. (2017). Simultaneous knock-down of six β-galactosidase genes in

petunia petals prevents loss of pectic galactan but decreases petal strength. *Plant Physiology and Biochemistry, 113*, 208–221. https://doi.org/10.1016/j.plaphy.2017.02.005.

O'Donoghue, E. M., Somerfield, S. D., Watson, L. M., Brummell, D. A., & Hunter, D. A. (2009). Galactose Metabolism in Cell Walls of Opening and Senescing Petunia Petals. *Planta, 229*(709). https://doi.org/10.1007/s00425-008-0862-6.

O'Neill, M., Albersheim, P., & Darvill, A. (1990). The pectic polysaccharides of primary cell walls. *Methods in Plant Biochemistry, 2*, 415–441. https://doi.org/10.1016/B978-0-12-461012-5.50018-5.

O'Neill, M., Ishii, T., Albersheim, P., Darvill, A. G. (2004). Rhamnogalacturonan II: Structure and function of a borate cross-linked cell wall pectic polysaccharide. *Annual Review of Plant Biology, 55,* 109–139. https://doi.org/10.1146/annurev.arplant.55.031903.141750.

Paniague, C., Posé, S., Morris, V. J., Kirby, A. R., Quesada, M. A., & Mercado, J. A. (2014). Fruit softening and pectin disassembly: an overview of nanostructural pectin modifications assessed by atomic force microscopy. *Annals of Botany, 114(*6), 1365–1383. https://doi.org/10.1093/aob/mcu149.

Parker, S., Redgate, E., Wibisono, R., Luo, X., Koh, E. T. H., & Schröder, R. (2010). Gut health benefits of kiwiberry pectins: Comparison with commercial functional polysaccharides. *Journal of Functional Foods, 2*(3), 210–218. https://doi.org/10.1016/j.jff.2010.04.009.

Pinto, D., Delerue-Matos, C., & Rodrigues, F. (2020). Bioactivity, phytochemical profile and pro-healthy properties of *Actinidia arguta*: A review. *Food Research International, 136*(3), 109449. https://doi.org/10.1016/j.foodres.2020.109449.

Redgwell, R. J., Fischer, M., Kendal, E., & MacRaw, E. A. (1997). Galactose loss and fruit ripening: high-molecular-weight arabinogalactans in the pectic polysaccharides of the fruit cell walls. Planta, 203(2), 174–181. https://www.jstor.org/stable/23385028.

Sachetto-Martins, G., Franco, L. O., & de Oliveira, D. E. (2000). Plant glycine-rich proteins: a family or just proteins with a common motif? *Biochimica et Biophysica Acta – Gene Structure and Expression, 1492*(1), 1–14. https://doi.org/10.1016/s0167-4781(00)00064-6.

Salyers, A. A., West, S. E., Vercellotti, J. R., & Wilkins, T. D. (1977). Fermentation of mucins and plant polysaccharides by anaerobic bacteria from the human colon. *Applied Environmental Microbiology*, 34(5), 529–533. https://doi.org/10.1128/aem.34.5.529-533.1977.

Schmitz, K., Protzko, R., Zhang, L., & Benz, P. (2019). Spotlight on fungal pectin utilization—from phytopathogenicity to molecular recognition and industrial applications. *Applied Microbiology and Biotechnology, 103,* 2507–2524. https://doi.org/10.1007/s00253-019-09622-4

Schwalm N. D., Townsend G. E., & Groisman E. A. (2016) Multiple signals govern utilization of a polysaccharide in the gut bacterium *Bacteroides thetaiotaomicron*. *American Society for Microbiology, 7(*5). https://doi.org/10.1128/mBio.01342-16.

Scotsmans, W. C., & Mawson, A. J. (2004). Non-destructive firmness measurement of Zespri™ Gold using acoustic impulse response technique and compression tests. *ISHS Acta Horticulturae, 687,* 107-112. https://doi.org/10.17660/ActaHortic.2005.687.12.

Simpson, R. J. (2004). Purifying proteins for proteomics: A laboratory manual. Cold Spring Harbor, N.Y: Cold Spring Harbor Laboratory Press.

Sutherland, P., Fullerton, C. G., Schröder, R., & Hallett, I. C. (2017). Cell wall changes in *Actinidia arguta* during softening. *Scientia Horticulturae, 226*, 173–183. https://doi.org/10.1016/j.scienta.2017.08.027.

The UniProt Consortium. (2018). UniProt: the universal protein knowledgebase. *Nucleic Acids Research, 46(*5), 2699. https://doi.org/10.1093/nar/gky092.

Thibault, J. F., & Ralet, M. C. (2001). Pectins, their origin, structure and functions. In: McCleary, B. V., Prosky, L. (eds.) *Advanced Dietary Fibre Technology.* Blackwell Science, Oxford, pp. 369–378. https://doi.org/10.1023/B:VERC.0000014152.80334.86.

Thursby, E., & Juge, N. (2017). Introduction to the human gut microbiota. *The Biochemical journal, 474*(11), 1823–1836. https://doi.org/10.1042/BCJ20160510.

Verhertbruggen, Y., Marcus, S. E., Haeger A, Verhoef, R., Schols, H. A., McCleary, B. V., McKee, L., Gilbert, H. J., & Knox, J. P. (2009a). Developmental complexity of arabinan polysaccharides and their processing in plant cell walls. *The Plant Journal, 59,* 413–425. https://doi.org/10.1111/j.1365-313X.2009.03876.x.

Verhertbruggen, Y., Marcus, S. E., Haeger, A., Ordaz-Ortiz, J. J., & Know, P. (2009b). An extended set of monoclonal antibodies to pectic homogalacturonan. *Carbohydrate Research, 344(*14), 1858–1862. https://doi.org/10.1016/j.carres.2008.11.010.

Wang, D., Yeats, T. H., Uluisik, S., Rose, J. K. C., & Seymour, G. B. (2018). Fruit softening: Revisiting the role of pectin. *Trends in Plant Science, 23*(4), 302–310. https://doi.org/10.1016/j.tplants.2018.01.006.

Wang, H., Chen, F., Yang, H., Chen, Y., & Zhang, L. (2012). Effects of ripening stage and cultivar on physiochemical properties and pectin. *Carbohydrate Polymers, 89*(4), 1180–1188. https://doi.org/10.1016/j.carbpol.2012.03.092.

Wefers, D., Flörchinger, R., & Bunzel, M. (2018). Detailed structural characterisation of arabinans and galactanase of 14 apple cultivars before and after cold storage. *Frontier Plant Science, 9*, 1451. https://doi.org/10.3389/fpls.2018.01451.

White, A., Nihal de Silva, N., Requeho-Tapia, C., & Harker, R. F. (2005). Evaluation of softening characteristics of fruit from 14 species of *Actinidia. Postharvest Biology and Technology, 35(*2), 143–151. https://doi.org/10.1016/j.postharvbio.2004.08.004.

White, B., Lamed, R., Bayer, E. A., & Flint, H. J. (2014). Biomass utilisation by gut microbiomes. *Annual Review of Microbiology,* 68, 279–296. https://doi.org/10.1146/annurev-micro-092412-155618.

Willats, W. G., Marcus, S. E., & Knox, J. P. (1998). Generation of monoclonal antibody specific to (1-->5)-alpha-L-arabinan. *Carbohydrate Research, 308*(1-2), 149–152. https://doi.org/10.1016/s0008-6215(98)00070-6.

Willats, W. G., McCartney, L., Steele-King, C. G., Marcus, S. E., Mort, A., Huisman, M., Jan van Alebeek, G., Schols, H. A., Voragen, G. J. Le Goff, A., Bonnin, E., Thibault, J., & Knox, J. P. (2003). A xylogalacturonan epitope is specifically associated with plant cell detachment. *Planta, 218,* 673–681. https://doi.org/10.1007/s00425-003-1147-8.

Williams, M., Boyd, L. M., McNeilage, M. A., MacRae, E A., Ferguson, A. R., Beatson, R. A., & Martin, P. J. (2003). Development and commercialisation of 'Baby Kiwi' (*Actinidia arguta* planch.). *Acta Horticulturae, 610,* 81–86. https://doi.org/10.17660/ActaHortic.2003.610.8.

Wu, D., Ye, X., Linhardt, R. J., Liu, X., Zhu, K., Yu, C., Ding, T., Liu D., He, Q., & Chen, S. (2021). Dietary pectic substances enhance gut health by its poly component: A review. *Comprehensive Reviews in Food Science and Food Safety, 20*(2), 2015–2039. https://doi.org/10.1111/1541-4337.12723.

Xiangxue, A., Lee, S. G., Kang, H., Heo, H. J., Cho, Y. – S., Kim, D. –O. (2016). Antioxidant and anti-inflammatory effects of various cultivars of kiwi berry (*Actinidia arguta*) on lipopolysaccharide-stimulate raw 264.7 cells. *Journal of Microbiology Biotechnology, 26*(8), 1367–1374. https://doi.org/10.4014/jmb.1603.03009

Yapo, M. B. (2011). Pectic substances: From simple pectic polysaccharides to complex pectins—A new hypothetical model. *Carbohydrate Polymers, 86*(2), 373–385. https://doi.org/10.1016/j.carbpol.2011.05.065.

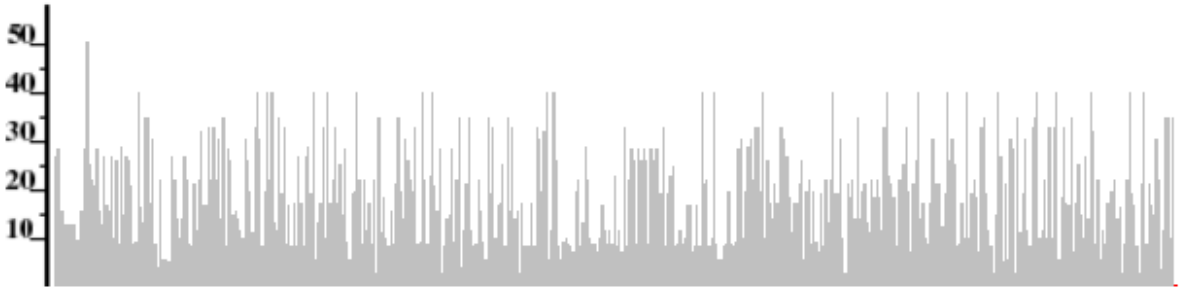
Zandleven, J., Sorensen, S. O., Harholt, J., Beldman, G., Schols, H. A., Scheller, H. V., & Voragen, A. J. (2007). Xylogalacturonan exists in the cell walls from various tissues of *Arabidopsis thaliana*. *Phytochemistry, 68*(8), 1219–1226. https://doi.org/10.1016/j.phytochem.2007.01.016.

Zhu, R., Wang, C., Zhang, L., Wang, Y., Chen, G., Fan, J., Jia, Y., Yan, F., & Ning, C. (2019). Pectin oligosaccharides from fruit of *Actinidia arguta*: Structure-activity relationship of prebiotic and antiglycation potentials. *Carbohydrate Polymers, 217*(3), 90–97. https://doi.org/10.1016/j.carbpol.2019.04.032.

Appendices

Appendix 1 Codon analysis of protein 0050.

Bar chart of codon frequences



ATG GGC AGC AGC CAT CAT CAT CAT CAT CAC AGC AGC GGC CTG GTG CCG CGC GGC AGC CAT ATG GCT AGC ATG ACT GGT GGA CAG CAA ATG GGT CGC GGA TCC GAA TTC TTG AAA AAA TAT ATT TCG CUA TTT AUA AUA TGT TGT ATG TTT TTA TCT TTA ATG CCG GGA ACA GCA GTA TTT GCG GCT GCT GAT TTT GAT TTT ATT TTA AAA TCA GGC GGT CAA AGC AGC GTA ACT ATT GGT GTT CGC GAT GAA ATT ACA GTT GAA TTT GAA TTG GTA AAA AAT GAT GGA GCT TCA TAT TCA ATG TAT TCA ATG CAG AAT GAA AUA TTG TAT GAT ACT GAA TAT TTT GAT TAT GTG CAA GGC AGT AUA AAT GTT GAA AAC GGA TTT AAG TAT GGA TTT AGA AAA CTT GAG TCT TCA AGC GGA GCA AAA GTT TTA ATT GGT TTT GTT GAT TCG AGT GAA AAC GGA ACC GAA CGC AGC GGC AGA ACA TTA GTC GGC AGT TTT AAA CUA AAG GCA AAA AAG ACA GGA AAC AGC AGT AUA AAA AAT GAT TCT GCT TAT GTG ACA AAA AGC GAT TTA AGC AGA TAT TCA TCA TCA TAT TCA GAT ATT GAC GCG GAA AUA GTA GAA GGT TCA AUA TCC ACT CCA ACA CCT GTT CCG TCA TTG CAG CCG ACT TCG GGA CCT ACT GCT AAG CCA AAG CCA ACC TCA AAG CCT GAT ACA AAC GGC GGT GGA GGC GGT GGC GGT GGA GGC GGT GGC GGC AAT GAT ACA AAT ACC GCC ACA CCA AAG CCA ACT GCT GCT CCT TCA GCT TCA GAA GCA CCG ACA TCT GAA CCA AUA AUA ACA CCA GTT CCA ACA AGT GGC ATT TCT GTT CAG ATT TTT GAT GAT GTT GAA ACT GGT TAT TGG GCA TAT GAT GAT ATT ATG GAG CTT TAT TAT GCA GGT AUA GTT AAC GGA GAC AGT CCT AAT ACA TTT TTG CCG GAA AAT AAT ATT ACT AGA GCA GAG TTT ACG AAA TTA GTT GCA TTG CTT TTT GAG TTT GAG GTA GAT GAA ACC GCA GAG ACA AAC TTT GTG GAT GTT CCT GCA GGT GAA TGG TAT ACT CCA TAT ATT GCA GCA GCA TAC AAT GAA GGT ATT GTG ACA GGA TAT TCT GAA ACT AAT TTT GAG CCT GAT AAA AAT GTA TCA AGA CAA GAA ATG TGT GCA AUA ATT GGC AGA AAA CTT AAT ATT GTA TCA GAT AAA GAA ACT GTA TTT ACT GAT TCT GAT GAA AUA GAG GAT TAT GCT AAA CCT TAT GTG CAA TCT ATG ACG GAA GCG GGA TTT AUA AAG GGA TAT GAC GAC AAC ACG TTC AGA CCA TTT GAA AAT GCT ACA AGA GCA GAA TCG GCA GCT GTC ATT AAC CGA GTA AAA AAA TCT AAA TAG

The Number of Bases in the above Sequence = 1386

The Number of Codons in the above Sequence = 462

Amino Acid	Rare Codon	Frequency of Occurrence
Arginine	CGA	1
	CGG	0
	AGG	0
	AGA	8
Glycine	GGA	16
	GGG	1
Isoleucine	AUA	13
Leucine	CUA	2
Proline	CCC	0
Threonine	ACG	4

Appendix 2 Mass spectrometry results of protein 0050 and wild types.

Protein 0050

Checked	Protein FD	Master	Accession	Description	Exp. q-val	Sum PEP S	Coverage	# Peptides	# PSMs	# Unique I	# AAs	MW [kDa]	calc. pI	Score Seq	# Peptides	# Protein Gr
FALSE	High	Master Pr	CAQ30532	chaperone with DnaK; heat shoc	0	12.361	12	2	5	2	376	41	7.66	9.15	2	1
FALSE	High	Master Pr	QJZ11694	porin OmpF [Escherichia coli BL2	0	7.393	12	2	4	2	362	39.3	4.96	0	2	1
FALSE	High	Master Pr	QJZ12582	bifunctional histidinol-phosphat	0	7.34	8	2	5	2	355	40.2	6.18	3.66	2	1
FALSE	High	Master Pr	QZI62512	catabolic alanine racemase Dad	0	7.061	9	2	5	2	356	38.8	7.05	0	2	1
FALSE	High	Master Pr	QZI63784	D-alanyl-D-alanine carboxypepti	0	6.855	9	2	5	2	403	44.4	8.28	0	2	1
FALSE	High	Master Pr	CAQ30666	glutamate-1-semialdehyde amin	0	6.807	8	2	3	2	426	45.3	4.86	3.98	2	1
FALSE	High	Master Pr	CAQ31525	glucose-1-phosphatase / 3-phyt	0	6.497	10	2	5	2	413	45.7	5.9	1.73	2	1
FALSE	High	Master Pr	ACT43847	D-tagatose 1,6-bisphosphate ald	0	6.177	7	2	5	2	420	47	5.78	0	2	1
FALSE	High	Master Pr	QZI64482	preprotein translocase subunit S	0	6.1	3	2	72	2	901	102	5.6	0	2	1
FALSE	High	Master Pr	CAQ32518	putative ATPase [Escherichia col	0	5.558	9	2	4	2	369	39.9	6.3	0	2	1
FALSE	High	Master Pr	ACT43984	ribonucleoside-diphosphate red	0	5.414	8	2	3	2	376	43.5	4.81	0	2	1
FALSE	High	Master Pr	QJZ11676	30S ribosomal protein S1 [Esche	0	5.248	5	2	4	2	557	61.1	4.98	0	2	1
FALSE	High	Master Pr	ACT44067	erythronate-4-phosphate dehyd	0	5.247	10	2	5	2	378	41.4	6.7	2.13	2	1
FALSE	High	Master Pr	QJZ13252	N-acetylmuramoyl-L-alanine am	0	4.732	7	2	2	2	417	45.6	9.58	0	2	1
FALSE	High	Master Pr	ACT42575	periplasmic protein [Escherichia	0	4.626	11	2	4	2	430	45.9	7.59	2.14	2	1
FALSE	High	Master Pr	ACT43572	succinylornithine transaminase,	0	4.557	10	2	5	2	406	43.6	6.35	0	2	1
FALSE	High	Master Pr	P15252	Rubber elongation factor protei	0	4.279	28	2	3	2	138	14.7	5.19	1.71	2	1
FALSE	High	Master Pr	CAQ32434	6-phosphogluconate dehydroge	0	4.211	7	2	4	2	468	51.5	5.07	0	2	1
FALSE	High	Master Pr	CAQ32445	ybl97 [Escherichia coli BL21(DE3	0	3.976	9	2	3	2	370	41.3	6.95	0	2	1
FALSE	High	Master Pr	CAQ33160	putative NAD(P)-linked reductas	0	3.904	13	2	2	2	346	38.5	6.76	0	2	1
FALSE	High	Master Pr	CAQ31614	3-oxo-acyl-[acyl-carrier-protein]	0	3.768	14	2	2	2	244	25.5	6.6	0	2	1
FALSE	High	Master Pr	QZI65833	maltoporin LamB [Escherichia co	0	3.657	8	2	2	2	446	49.9	4.98	0	2	1
FALSE	High	Master Pr	ACT44135	cysteine synthase A, O-acetylser	0	3.603	14	2	3	2	323	34.5	6.06	0	2	1
FALSE	High	Master Pr	WP_1529	aromatic amino acid transamina	0	3.291	9	2	2	2	402	44.1	5.59	0	2	1
FALSE	High	Master Pr	QZI65635	sn-glycerol-3-phosphate ABC tra	0	2.758	8	2	2	2	438	48.4	6.73	0	2	1
FALSE	High	Master Pr	QZI63606	30S ribosomal protein S4 [Esche	0	2.632	16	2	2	2	206	23.5	10.05	0	2	1
FALSE	High	Master Pr	WP_1532	dTDP-glucose 4,6-dehydratase [	0	2.58	10	2	3	2	368	41.5	5.67	0	2	1
FALSE	High	Master Pr	QJZ13744	elongation factor G [Escherichia	0	2.539	5	2	2	2	704	77.5	5.38	0	2	1
FALSE	High	Master Pr	QZI63997	Cu(+)/Ag(+) efflux RND transport	0	2.16	6	2	2	2	457	50.3	6.4	0	2	1
FALSE	High	Master Pr	P02662	Alpha-S1-casein OS=Bos taurus (	0	1.827	15	2	3	2	214	24.5	5.02	0	2	1

Wild type construct P1

Checked	Protein FD	Master	Accession	Description	Exp. q-val	Sum PEP S	Coverage	# Peptides	# PSMs	# Unique P	# AAs	MW [kDa]	calc. pI	Score Seq	# Peptides	# Protein Group
FALSE	High	Master Prc	B9O19_00050	SLH-domain containing protein	0	170.771	32	8	382	8	425	45.9	4.6	1346	8	1
FALSE	High	Master Prc	CAQ34326.1	elongation factor Tu [Escherichia	0	103.983	37	8	122	8	394	43.3	5.45	246.57	8	1
FALSE	High	Master Prc	P00761	Trypsin OS=Sus scrofa PE=1 SV=1	0	93.575	23	4	162	4	231	24.4	7.18	414.55	4	1
FALSE	High	Master Prc	QJZ13339.1	phosphoglycerate kinase [Escher	0	49.144	36	7	28	7	387	41.1	5.22	48.36	7	1
FALSE	High	Master Prc	QJZ11694.1	porin OmpF [Escherichia coli BL21	0	40.793	41	7	25	7	362	39.3	4.96	39.42	7	1
FALSE	High	Master Prc	ACT42555.1	citrate synthase [Escherichia coli	0	38.42	23	7	21	7	427	48	6.68	27.59	7	1
FALSE	High	Master Prc	QJZ13216.1	phosphopyruvate hydratase [Esch	0	36.762	30	7	15	7	432	45.6	5.48	28.66	7	1
FALSE	High	Master Prc	QJZ12827.1	acetate kinase [Escherichia coli B	0	29.259	26	5	13	5	400	43.3	6.28	18.66	5	1
FALSE	High	Master Prc	QZI66031.1	thioredoxin TrxA [Escherichia col	0	27.991	37	3	29	3	109	11.8	4.88	37.92	3	1
FALSE	High	Master Prc	CAQ31080.1	AhpC component, subunit of alky	0	23.971	21	3	7	3	187	20.7	5.17	20.4	3	1
FALSE	High	Master Prc	ACT42562.1	succinyl-CoA synthetase, beta su	0	23.952	16	4	13	4	388	41.4	5.52	26.6	4	1
FALSE	High	Master Prc	ACT42827.1	aspartate aminotransferase, PLP-	0	23.517	20	4	14	4	396	43.5	5.77	21.81	4	1
FALSE	High	Master Prc	4BLB	A Chain A, MALTOSE-BINDING PE	0	21.158	12	6	11	6	753	83.5	5.34	8.45	6	1
FALSE	High	Master Prc	QJZ11847.1	beta-ketoacyl-ACP synthase II [Es	0	20.401	21	4	10	4	413	43	6.09	15.29	4	1
FALSE	High	Master Prc	QJZ11672.1	3-phosphoserine/phosphohydro	0	20.052	24	6	13	6	362	39.8	5.74	9.52	6	1
FALSE	High	Master Prc	QJZ13700.1	DNA-directed RNA polymerase su	0	19.853	16	4	12	4	329	36.5	5.06	9.53	4	1
FALSE	High	Master Prc	ACT44070.1	3-oxoacyl-[acyl-carrier-protein] s	0	19.307	25	4	9	4	406	42.6	5.54	20.13	4	1
FALSE	High	Master Prc	ACT45387.1	tryptophanase/L-cysteine desulf	0	19.238	22	6	12	6	471	52.7	6.23	9.38	6	1
FALSE	High	Master Prc	P02769	Serum albumin OS=Bos taurus GN	0	18.674	6	3	19	3	607	69.2	6.18	30.35	3	1
FALSE	High	Master Prc	QJZ13023.1	cysteine desulfurase [Escherichia	0	17.942	18	5	12	5	404	45.1	6.37	10.06	5	1
FALSE	High	Master Prc	ACT45454.1	transcription termination factor [	0	16.921	13	4	11	4	419	47	7.25	2.21	4	1
FALSE	High	Master Prc	QZI63522.1	threonine synthase [Escherichia c	0	13.56	22	5	7	5	428	47.1	5.4	0	5	1
FALSE	High	Master Prc	ACT44002.1	uridine 5'-(beta-1-threo-pentapy	0	13.411	14	3	4	3	385	42.2	6.89	5.85	3	1
FALSE	High	Master Prc	ACT42585.1	Lactose operon repressor [Escher	0	12.948	11	2	8	2	360	38.6	6.89	16.38	2	1

Wild type construct P2

Checked	Protein FD	Master	Accession	Description	Exp. q-val	Sum PEP S	Coverage	# Peptides	# PSMs	# Unique I	# AAs	MW [kDa]	calc. pI	Score Seq	# Peptides	# Protein Gr
FALSE	High	Master Pri	B9O19_00050	SLH-domain containing protein	0	205.282	35	10	753	10	425	45.9	4.6	3313.06	10	1
FALSE	High	Master Pri	CAQ34326.1	elongation factor Tu [Escherichia	0	97.036	44	10	130	10	394	43.3	5.45	239.19	10	1
FALSE	High	Master Pri	QJZ13339.1	phosphoglycerate kinase [Esche	0	63.963	47	9	31	9	387	41.1	5.22	73.49	9	1
FALSE	High	Master Pri	P00761	Trypsin OS=Sus scrofa PE=1 SV=1	0	55.508	29	5	126	5	231	24.4	7.18	271.97	5	1
FALSE	High	Master Pri	QJZ11221.1	multidrug efflux RND transporte	0	50.072	39	8	29	8	397	42.2	7.99	59.91	8	1
FALSE	High	Master Pri	CAQ30612.1	essential cell division protein Fts	0	48.34	50	12	25	12	383	40.3	4.78	48.33	12	1
FALSE	High	Master Pri	ACT42555.1	citrate synthase [Escherichia col	0	44.315	23	8	26	8	427	48	6.68	59.35	8	1
FALSE	High	Master Pri	QJZ12827.1	acetate kinase [Escherichia coli I	0	41.265	36	7	23	7	400	43.3	6.28	46.01	7	1
FALSE	High	Master Pri	ACT42504.1	heat shock protein, putative NTF	0	40.096	33	7	24	7	359	40.6	6.67	35.69	7	1
FALSE	High	Master Pri	ACT42562.1	succinyl-CoA synthetase, beta su	0	38.168	28	7	31	7	388	41.4	5.52	64.61	7	1
FALSE	High	Master Pri	QJZ13023.1	cysteine desulfurase [Escherichia	0	36.05	27	9	28	9	404	45.1	6.37	46.06	9	1
FALSE	High	Master Pri	ACT43028.1	isocitrate dehydrogenase; e14 pr	0	32.851	30	7	23	7	416	45.7	5.4	32.19	7	1
FALSE	High	Master Pri	ACT45454.1	transcription termination factor	0	32.681	25	7	21	7	419	47	7.25	26.77	7	1
FALSE	High	Master Pri	CAQ32886.1	1-hydroxy-2-methyl-2-(E)-butenyl	0	31.355	26	6	17	6	372	40.7	6.21	23.33	6	1
FALSE	High	Master Pri	QJZ11847.1	beta-ketoacyl-ACP synthase II [E	0	26.838	21	4	13	4	413	43	6.09	29.52	4	1
FALSE	High	Master Pri	CAQ32433.1	UDP-glucose 6-dehydrogenase [	0	25.809	22	7	15	7	388	43.6	6.46	11.31	7	1
FALSE	High	Master Pri	QJZ11196.1	ATP-dependent protease ATP-bi	0	25.272	29	8	15	8	424	46.3	5.35	11.45	8	1
FALSE	High	Master Pri	P02769	Serum albumin OS=Bos taurus GI	0	24.196	13	6	24	5	607	69.2	6.18	32.82	6	1
FALSE	High	Master Pri	QZI63784.1	D-alanyl-D-alanine carboxypepti	0	23.22	25	6	12	6	403	44.4	8.28	13.92	6	1
FALSE	High	Master Pri	CAQ31080.1	AhpC component, subunit of alky	0	23.2	29	3	14	3	187	20.7	5.17	27.61	3	1
FALSE	High	Master Pri	ACT45387.1	tryptophanase/L-cysteine desulf	0	21.379	22	6	16	6	471	52.7	6.23	10.98	6	1
FALSE	High	Master Pri	QJZ13216.1	phosphopyruvate hydratase [Esc	0	21.027	23	6	10	6	432	45.6	5.48	10.54	6	1
FALSE	High	Master Pri	QJZ13700.1	DNA-directed RNA polymerase s	0	20.981	20	4	11	4	329	36.5	5.06	5.42	4	1
FALSE	High	Master Pri	CAQ30532.1	chaperone with DnaK; heat shoc	0	20.863	18	4	15	4	376	41	7.66	55.89	4	1

Appendix 3 Initial yes or no binding of wild type proteins and pectin substrates.

Substrate	P1 – Max. binding (nm)	P2 – Max. binding (nm)
RG-I	No binding	No binding
D-(+)-xylose	~0.27 nm	~0.08 nm
Poly galacturonic Acid	~0.045 nm	No binding
Potato Galactan (PG)	~0.15 nm	~0.05 nm
Hortgem Tahi	~0.24 nm	N/A
Takaka Green	~0.25 nm	N/A
Marju Red	No binding	N/A
Citrus Pectin (CP)	~0.3 nm	~0.06 nm
Linear-arabinan	~0.4 nm	~0.1 nm
Arabinan	~0.125 nm	No binding
Di-galacturonic Acid	~0.225 nm	~0.35 nm
Tetra-galacturonic Acid	~0.1 nm	~0.07 nm
Cellulose	No binding	~0.04 nm

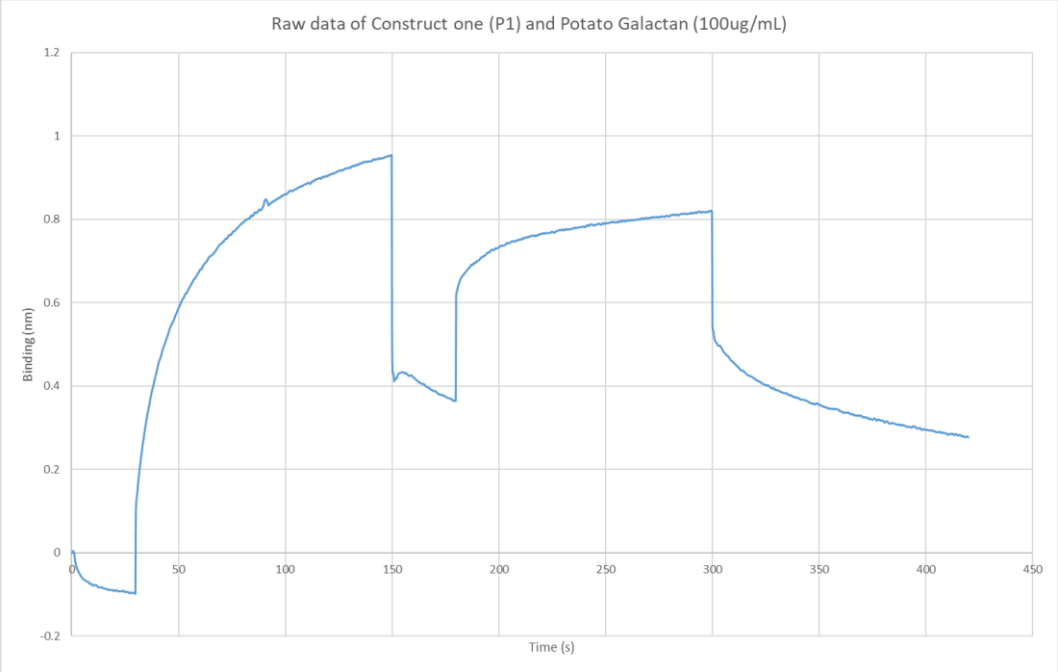
Note: RG-I: Rhamnogalacturonan I; N/A: Not applicable.

Appendix 4 A representation of the BLItz Binding Curves created.

Graphs here are a single representation of how the binding data was processed for Table 5.6.

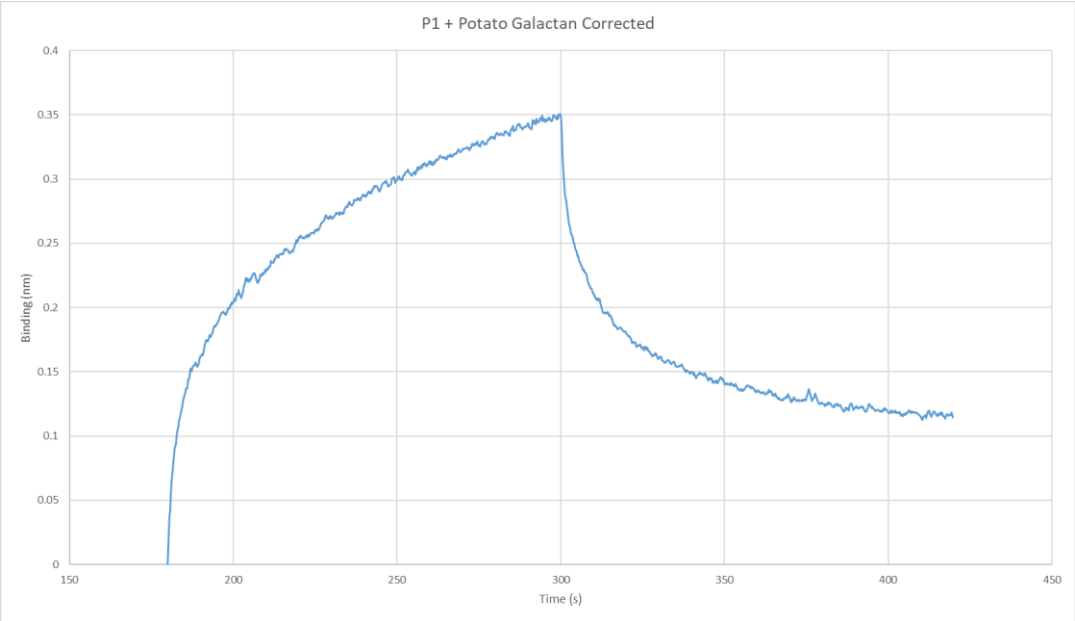
Raw data

The graph here is the raw data with no corrections following methods from 2.5.6



Data corrected for BLItz probe.

The sharp spike in binding is the protein 0050 construct one (P1) associating with potato galactan. The decline at 300 seconds is the disassociation of the potato galactan to construct one (P1).



Data corrected for BLItz probe and non-specific binding of substrate to the His-Tagged probe

The binding of substrate to the His-Tagged probe is removed from the corrected BLItz probe data.