



**Aalto University
School of Chemical
Engineering**

Vilma Jäämuru

**EXPRESSION SCREENING, PRODUCTION AND
FUNCTIONAL CHARACTERIZATIONS OF
BACTERIAL EXPANSIN-RELATED PROTEINS**

Master's Programme in Chemical, Biochemical and Materials Engineering
Major in Biotechnology

Master's thesis for the degree of Master of Science in Technology submitted for
inspection, Espoo, 29 July, 2022.

Supervisor

Professor Emma Master

Instructor

Majid Haddad Momeni, PhD

Copyright ©2022 Vilma Jäämuru

Author Vilma Jäämuru

Title of thesis Expression screening, production and functional characterizations of bacterial expansin-related proteins

Degree Programme Chemical, Biochemical and Materials Engineering

Major Biotechnology

Thesis supervisor Professor Emma Master

Thesis advisor(s) / Thesis examiner(s) Majid Haddad Momeni, PhD

Date 29.07.2022**Number of pages** 49+12**Language** English

Abstract

Expansins are small cell wall loosening proteins first discovered in the '90s as mediators of "acid growth" in plants. With the discovery of plant expansins, sequence alignment searches revealed expansin-like proteins from non-plant sources like bacteria, fungi and other prokaryotes. Without bona fide catalytic activity, expansins and expansin-like proteins are hypothesized to act on certain "biomechanical hotspots" to induce cell wall creep, slippage and loosening. As a result, expansins have gathered a lot of attention as possible surface-acting proteins that can modify abundant biopolymers like cellulose into different functional materials in various industries.

This thesis aims to (1) complete further expression screening and purification of bacterial expansin-like proteins, (2) functionally characterize target bacterial expansin-like proteins, and (3) examine the synergistic activities of target bacterial expansin-like proteins. Small scale production of target proteins was completed with MagicMedia dual temperature protocol or lysogeny broth (LB) media and isopropyl β -d-1-thiogalactopyranoside (IPTG) induction. Large scale production was completed only with LB media and IPTG induction but at various production conditions. Different lysis methods were compared, with sonication being used for large scale production. Characterization of the proteins was done with insoluble polysaccharide pulldown (IPP) assays on commercial substrates (Avicel, chitin, and oat-spelt xylan) as well as hardwood and softwood pulps. Synergistic activities was determined in various reaction conditions with cellulase from *Trichoderma reesei* and xylanase from a fungal source.

From expression screening trials three new bacterial expansin-like proteins were identified. However, large scale production of these proteins were produced with no yield. Production of LT367 and CPO928 was completed successfully with yields of 6 mg/L and 1 mg/L, respectively. All three bacterial expansin-like proteins tested for binding (HE673, LT367 and AJC165) showed significant binding to oat-spelt xylan and both pulps. HE673 also showed significant binding to Avicel and chitin due to its additional carbohydrate binding domain (CBM2). Possible synergistic activity was observed in LT367 and AJC165. The further characterization of LT367 and AJC165 with more binding and synergy studies along with bioreactor production of the protein is recommended. Studies examining the cell wall loosening activity or filter paper weakening are also suggested for further characterization steps.

Keywords expansins, microbial expansin-like proteins, protein expression, protein purification, characterization, enzyme synergy

Acknowledgements

This thesis could not be possible without the wonderful support of the Protein Technology group. I would like to thank my supervisor, Professor Emma Master, for giving me the opportunity to work on this project and team. A big thanks is also needed for my advisor, Dr. Majid Haddad Momeni, who guided me through every step of the way. I am grateful for the support and advice Emma, Majid and everybody in the Protein Technology group gave me through my experimental work.

My gratitude would not be complete without mentioning my family. I am forever grateful for their continuous support and belief in me. They are always encouraging me to expand my boundaries and challenge myself. Finally, I want to thank my boyfriend, Rob, who supported me through my writing process and always knows how to make me smile. I would not be in my position today without them.

Thank you!

Abbreviations

Asp	Aspartic Acid
B-PER	Bacterial Protein Extraction Reagent
BSA	Bovine Serum Albumin
CBM	Carbohydrate Binding Domain
CV	Column Volume
D1	Domain 1
D2	Domain 2
EXLA	expansin-like A
EXLB	expansin-like B
EXLX	expansin-like from non-plant sources
EXPA	α -expansins
EXPB	β -expansins
FT	Flow Through
GH45	Glycosyl Hydrolases 45
Glu	Glutamic Acid
HC	Heated Cells
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
IPP	Insoluble Pulldown Polysaccharide
IPTG	Isopropyl β -d-1-thiogalactopyranoside
LB	Lysogeny Broth
NaOAc	Sodium Acetate
OD	Optical Density
PAHBAH	4-hydroxybenzhydrazide
SDS-PAGE	Sodium Dodecyl Sulfate–Polyacrylamide Gel Electrophoresis
Thr	Threonine

Contents

Abstract	3
Acknowledgements	4
Abbreviations	5
Contents	6
1.0 Introduction	8
1.1 Thesis Structure and Aim	9
2.0 Background on Expansins	10
2.1 Discovery of Expansins	10
2.2 Expansin Classification.....	11
2.3 Expansin Structure	11
2.4 Expansin Activity.....	13
2.5 Production of Microbial Expansins	16
3.0 Materials & Methods	18
3.1 Microbial Expansin-Related Sequences	18
3.2 Small-scale Production.....	20
3.2.1 MagicMedia Production	20
3.2.2 LB Media	21
3.3 Large-scale Production with LB Media.....	22
3.4 Cell Lysis.....	23
3.4.1 B-PER	23
3.4.2 MgCl ₂ Buffer.....	23
3.4.3 Sonication	24
3.5 Purification	24
3.5.1 Small-scale Purification	24
3.5.2 Large-scale Purification	25
3.6 Characterization.....	26
3.6.1 IPP Assay.....	26
3.6.2 Synergism Reactions.....	27
4.0 Results and Discussion	29
4.1 Small-scale Production.....	29

4.1.1	MagicMedia.....	29
4.1.2	LB Trial 4.....	30
4.2	Lysis Comparison	31
4.3	Large-scale Production and Purification	33
4.3.1	AP099, JAB176 and JAB041.....	34
4.3.2	LT367, CPO928 and SFM691	35
4.4	Characterizations.....	37
4.4.1	Binding Assays	38
4.4.2	Synergisms	39
5.0	Conclusion	43
	References.....	45
A.	Equipment & Chemicals	50
B.	DNA and Gene Accession Numbers.....	52
C.	Supplementary Figures.....	58

1.0 Introduction

Biopolymers are a sustainable and renewable alternative material to conventional petrochemical plastics with numerous applications in biofuel production, textiles, and healthcare. As some of the most abundant biopolymers on Earth, cellulose, hemicellulose, and chitin are getting attention for their possible modifications and transformations into different functional materials and industrial applications (Ling et al., 2018; Raud et al., 2019). However, efficient treatment of these natural biopolymers remains expensive and difficult due to the complexity of the feedstock macrostructure (Lakhundi et al., 2015; Lynd et al., 2002; Schwarz, 2001).

Cellulose, a major component of cell walls, is a linear polymer of β -1-4 linked glucose monomers. Chains of these glucose polymers pack together to form microfibrils, which in turn pack together to form macrofibrils. Inter- and intra-chain hydrogen bonding between parallel chains and layers in cellulose provides rigidity and strength (Lakhundi et al., 2015). Despite the simple chemistry of the glucose monomer in nature cellulose rarely exists in a purely packed or crystalline structure. The partial ordered crystalline and disordered amorphous structure of natural cellulose makes treatment via enzymatic hydrolysis difficult (Schwarz, 2001). Expansins, a cell-wall-loosening protein first discovered in plants, may help play a role in the treatment of abundant natural biopolymers into functional materials by disrupting the macrostructural integrity of well-packed and chemically unavailable monomer units (Ding et al., 2022).

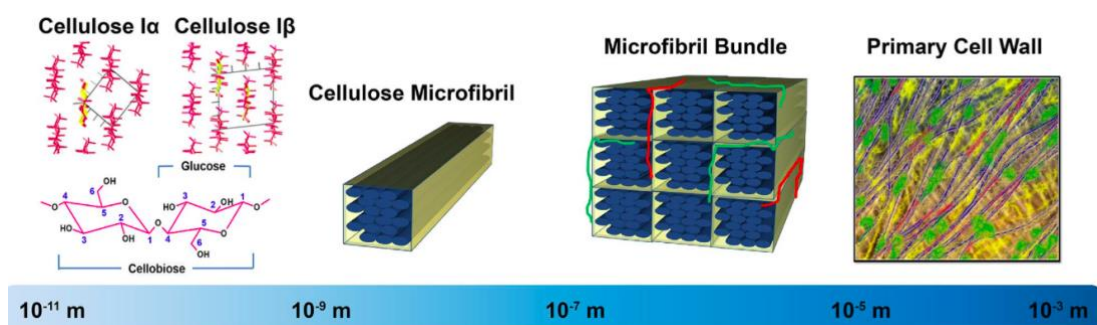


Figure 1.1 | Cellulose structure. The figure describes how cellulose is structured and its size in cell walls. Image from Rongpipi et al. (2017).

First discovered in the early 1990's from cucumber seedlings, expansins are small proteins that mediate a form of acid induced cell wall extension or "acid growth" without any detectable catalytic activity (McQueen-Mason et al., 1992). Now, 30 years later, four large families of expansins and expansin-like genes and their respective proteins have been discovered in plants, as well as bacteria, fungi, and other prokaryotes. Expansins are hypothesized to act on certain mechanical hotspots that loosen microfibrils from the cellulosic matrix, relaxing and opening the cellulosic material to be more accessible for cellulases (Cosgrove, 2000). Despite no catalytic activity being detected and the mechanism of action remaining unclear, expansins have garnered a lot of attention for their possible applications in biofuels, textiles and many other areas that rely on cellulose breakdown.

1.1 Thesis Structure and Aim

This thesis is part of the larger work called BioUPGRADE that aims to transform abundant biopolymers into high-value resource-efficient bio-based materials (BioUPGRADE, 2020). Specifically, this thesis focuses on the BioUPGRADE key deliverable of developing a toolbox of surface-acting proteins, such as expansins, for bio-based materials engineering. The thesis had three main aims that are outlined below:

1. Further expression screening and purification of target bacterial expansin-like proteins with modified protocols
2. Functional characterization of target bacterial expansin-like proteins via binding assays using commercial and industrial biomasses
3. Examination of synergistic activities of target bacterial expansin-like proteins with canonical hydrolytic enzymes

The thesis is broken down into five different sections. After the introduction, a background on expansins and their properties is given. Third, all the materials and methods of the different experiments completed are covered. Then, the results of said experiments are reviewed and discussed. Finally, the thesis is summarized in the conclusion. Supplementary information can be found in the appendices at the end of the report.

2.0 Background on Expansins

In this chapter, background information on expansins is provided. Specifically, expansin classification, structure, activity, and production are discussed.

2.1 Discovery of Expansins

In plants, the cell wall is responsible for protection, structure, and strength, however as the plant grows the rigid structure of the cell wall needs to loosen and expand to allow for volumetric growth. Before the discovery of expansins, *in vivo* hydrolytic enzymes were believed to be the sole catalysis for plant cell enlargement. However, exogenously added enzymes failed to induce isolated cell wall extension (Ruesink, 1969). Finally, in the early 1990's, three proteins, later classified as α -expansins, were identified to have the characteristic ability to induce cell wall extension missing from hydrolytic enzymes (McQueen-Mason et al., 1992; Li et al., 1993).

Soon after the initial discovery and the sequencing of the first expansins, the number of expansins discovered and identified quickly skyrocketed. Cosgrove (1997) identified group 1 allergens from grass pollen that had the ability to induce wall creep similar to that of expansins. However, the group 1 allergens had notable differences to the original expansins. Despite predicted similar secondary structures, the proteins shared only 20-25% amino acid identity (Cosgrove et al., 1997). As a result, these expansins would be categorized as β -expansins.

In the late 90s and early 2000s, sequence similarity of expansins was being noted in proteins of non-plant sources like bacteria and fungi (Cosgrove, 1996; Laine et al., 2000; Li et al., 2002). Many expansin-like proteins from non-plant origins are in fact from plant pathogens or species closely associated with plants. Despite rare occasions of genetic transfer between eukaryotes and prokaryotes, two instances of horizontal gene transfer from plants to bacteria and fungi are identified to be the origin for microbial expansins (Nikolaidis et al., 2014). The crystal structure determination of YoaJ protein (EXLX1) from *Bacillus subtilis* confirmed unequivocally the existence of expansin-like proteins in bacterial sources (Kerff et al., 2008). *B. subtilis* is a gram-positive bacterium frequently found in soil with close association with plant roots. The protein, later renamed BsEXLX1, was determined to have structural similarity to the

grass pollen allergens (β -expansins) and the expansin characteristic of inducing plant cell wall extension, but ten times weaker than for plant β -expansins.

2.2 Expansin Classification

With the presence and identification of expansins from varying organisms, a classification system for all expansins and expansins-like proteins was developed (Kende et al., 2004). In order to be classified as an “expansin” or “expansin-like”, the protein must contain the characteristic two domain structure of expansins. Other putative proteins that contain only the D1 domain or retain an additional domain are called “Loosenins”, or “Swollenins”, respectively.

In plants, four large families of expansins are recognized based on sequence-based phylogenetic analysis: α -expansin (EXPA), β -expansin (EXPB), expansin-like A (EXLA) and expansin-like B (EXLB). The first expansins discovered in cucumber seedlings and phylogenetically related expansins are designated α -expansins. The proteins similar to the group 1 grass pollen allergens are β -expansins. EXPA and EXPB have both experimentally been shown to exhibit cell wall loosening, whereas expansin-like proteins A and expansin-like proteins B are only known from their genetic sequence (Sampedro & Cosgrove, 2005).

For non-plant organisms, expansin-like family X (EXLX) nomenclature is used (Kende et al., 2004). The terminology EXLX is a “catch-all” term reserved for genes and their respective proteins that fulfill the following criteria:

1. They occur in non-plant organisms.
2. Their protein products have homologies to both domains.
3. They do not fall in within the established expansin gene families.

2.3 Expansin Structure

With the crystallization of plant expansin EXB1 and bacterial expansin *BsEXLX1* the secondary structure of expansins was solved (Kerff et al., 2008; Yennawar et al., 2006). Despite low sequence similarity between different expansin families, the superimposed crystallographic structures show congruent proteins with low divergence (Ding et al., 2022; Georgelis et al., 2015). Both expansins are small,

elongated ellipsoidal proteins ~6 nm long and ~4 nm wide (Cosgrove, 2017). Both bacterial and plant expansins are roughly 25 kDa (~225 amino acids) in size with two compact domains. Domain one (D1) is a six-stranded double- ψ β -barrel domain and domain two (D2) is a β -sandwich domain with a distinct topological fold. A short linker connects the two domains.

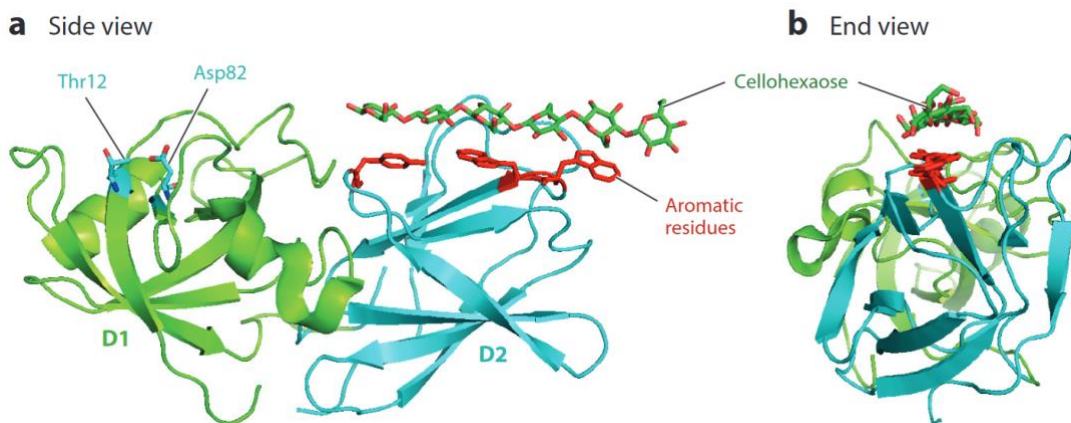


Figure 2.1 | Expansin Structure. Crystal structure of *BsEXLX1* in complex with cellohexaose (Cosgrove, 2017). The conserved catalytic GH45 residues are Thr12 and Asp82. Domain 2 is responsible for cellulose binding via three aromatic residues depicted in red in the figure.

The D1 structure resembles a family of endoglucanases, glycosyl hydrolases 45 (GH45), that also have the double- ψ β -barrel fold. Endoglucanases exhibit a deep cleft surrounding the catalytic site that enable binding with glucans. Expansins, however, only contain a shallow groove in D1, which exhibits negligible binding to polysaccharides (Georgelis et al., 2011; Georgelis et al., 2012). The catalytic machinery of GH45 is partially conserved in the D1 domain including the two residues Thr12 and Asp82. Although, a second critical Asp residue is missing, resulting in no bona fide catalytic activity being detected in expansins (Kerff et al., 2008; McQueen-Mason & Cosgrove, 1994, 1995; Yennawar et al., 2006). Despite the conservation, Thr12 was not found to be essential for expansin wall creep activity (Georgelis et al., 2011). Mutations of the Asp82 residue indicate that the carboxyl group on aspartate is essential for wall extension activity. Another recent study on fungal expansins also found D1 residues Asp237, Glu238 and Asp248 to be essential in extension activity due to their hydrogen bonding to glucoses (Ding et al., 2022).

The D2 domain is the founding member of family 63 carbohydrate binding module (CBM) and is primarily responsible for mediating binding of expansins to

polysaccharides (Georgelis et al., 2015). CBMs are divided into two categories: type-A and type B. Type-A bind to crystalline cellulose, while type-B bind to less organized areas of cellulose (Georgelis et al., 2015). The D2 domain was found to exhibit characteristics of type-A CBM, supporting that binding occurs with the hydrophobic face of crystalline cellulose (Georgelis et al., 2012). The binding to cellulose is mediated by three conserved aromatic residues (W125, W126, and Y157) that are linearly arranged to form an open planar surface. This binding occurs primarily through the hydrophobic CH- π interactions. Conversely, binding to whole cell walls was not mediated by the three aromatic residues, but via electrostatic forces to positively charged nonconserved residues on the opposite surface of the D2 domain (Georgelis et al., 2011).

2.4 Expansin Activity

Expansins were first discovered for their ability to induce “acid growth” or plant cell enlargement. They have since been associated with numerous other plant functions where wall loosening occurs such as fruit ripening, seed germination, and leaf shedding (Belfield et al., 2005; Chen & Bradford, 2000; Cosgrove, 2000). Expansins were first categorized as inducers of “acid growth” due to the observation that cell wall extension and plant growth occurred in acidic conditions. Numerous studies have confirmed the pH optimum for expansins around 4-5 with rate of cell wall relaxation decreasing when carried out at higher pH (Li et al., 1993; Li et al., 2003; McQueen-Mason & Cosgrove, 1995, McQueen-Mason et al., 1992). However, bacterial expansin-like protein *BsEXLX1* induced wall creep from pH 5.5 – 9.5 (Georgelis et al., 2011). At lower pH the protein precipitated.

In order to induce cell wall loosening, expansins need to be able to correctly bind to plants' cell walls as a substrate, which typically consists of complex polysaccharides such as cellulose or xylan (a type of hemicellulose). As mentioned previously, binding is mediated by the D2 domain either through three aromatic residues on a planar surface or through electrostatic interactions with charged residues on the opposite side. Binding via the aromatic residues has been associated with wall loosening activity whereas binding via electrostatic interactions is non-specific and have no correlation to expansins activity. A bacterial expansin *HcEXLX2* showed binding to Avicel (microcrystalline cellulose), and oat-spelt xylan at optimum temperature and

pH of 4°C and pH 6.0 (Lee et al., 2010). Bacterial expansin *BsEXLX1* bound to cellulose with an apparent $K_D \sim 1 \mu\text{M}$ and had high binding capacity to wheat arabinoxylan and wheat coleoptile cell walls (Georgelis et al., 2011; Kim et al., 2013). Other microbial expansin-like proteins have similar binding affinities (Bunterngsook et al., 2015; Olarte-Lozano et al., 2014). However, a native plant expansin bound much stronger to cellulose with $K_D \sim 0.2 \mu\text{M}$ (Georgelis & Cosgrove, 2015; McQueen-Mason & Cosgrove, 1995). The weaker binding to substrate may explain why microbial expansins have weaker cell wall creep activity than plant expansins (Georgelis et al., 2013; Kerff et al., 2008). Mutagenesis of the aligned three aromatic residues in D2 responsible for hydrophobic binding of substrate, resulted in no binding to pure cellulose and eliminated cell wall creep activity, further suggesting hydrophobic interactions with cellulose appear essential for cell wall loosening activity (Georgelis et al., 2011). Hepler and Cosgrove (2019) tested this theory and found mutants with stronger binding had greater wall creep activity. However, the relationship was nonlinear and plateaued at about 40% higher than wild-type.

With no catalytic activity being detected in expansins, their mechanism of action still remains unclear (Kerff et al., 2008; McQueen-Mason & Cosgrove, 1994, 1995; Yennawar et al., 2006). Expansins are known to act quickly and not affect the plasticity of the cell wall (Yuan et al., 2001). This differs from endoglucanases that take longer to cause cell wall creep and induce structural changes to the cell wall that increase wall plasticity. A nonenzymatic approach for cell wall extension has been hypothesized, where expansins disrupt the hydrogen bonds between cellulose microfibrils and matrix polysaccharides (Cosgrove, 2000; Ding et al., 2022; McQueen-Mason & Cosgrove, 1994, 1995). Expansins are predicted to act on certain limited “biomechanical hotspots” where cellulose-xyloglucan-cellulose contacts are made that are key sites for wall mechanics and cell loosening (Cosgrove, 2015, 2017; Park & Cosgrove, 2012). Using NMR technology, Wang et al. (2013) showed *BsEXLX1* targets binding to a distinct specific region of cellulose with proximity to xyloglucan, similar to the mentioned “biomechanical hotspots”. In this model, the small expansin proteins are predicted to inch along cellulose microfibrils and disconnect them from each other. Turgor driven pressure and stress cause slippage between glucose chains and loosening of the cell wall inducing cell wall extension.

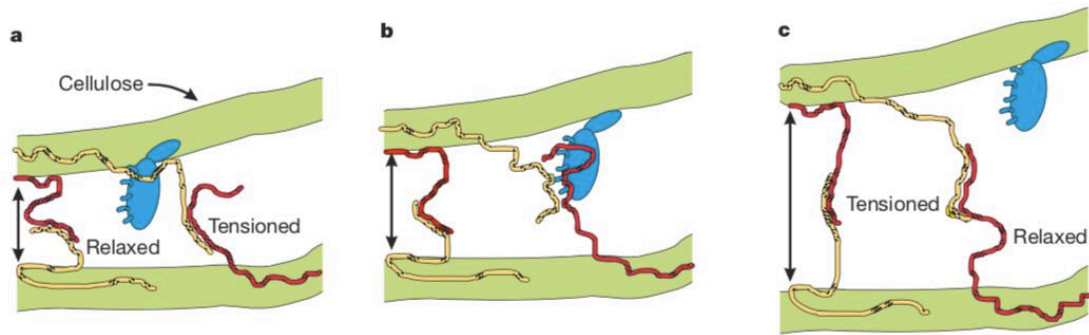


Figure 2.2 | Expansin Mode of Action. Cellulose microfibrils (green) are connected to each other by glycans (red and yellow) (Cosgrove, 2000). The expansin (blue) disrupts the non-covalent bonds between cellulose and glycans (a) or between glycans (b). Under mechanical stress, displacement of wall polymers occurs, loosening the cell wall (c).

By disrupting the hydrogen bonds and loosening the cell wall, expansins are hypothesized to open up the cellulose macrostructure and make surface glucans more available for enzymatic attack by various cellulases such as exoglucanases, endoglucanases and/or glucosidases. Numerous studies have examined this possibility for synergistic effects of microbial expansins on enzymes with varying results. Kim et al. (2009) found 5.9-fold synergistic effects of *BsEXLX1* with cellulase on hydrolysis of filter paper compared to cellulase alone, but only at very low concentrations of cellulase. As the concentration of cellulase was increased the synergistic effects diminished. Another study with the same group showed that bacterial expansin *HcEXLX2* released 4.6-fold more reducing sugars than with cellulase alone and 3.1-fold more than with xylanase alone (Lee et al., 2010, 2013). *HcEXLX2* with cellulase was even about two times greater than the positive control *CtCBD* with cellulase (a cellulose binding domain from *C. thermocellum* that is known to enhance cellulase activity). However, the negative control bovine serum albumin (BSA) also showed enhancement of cellulase activity that was not seen in the previous study by Kim et al. (2009). Two other studies also noted the enzyme enhancement seen from BSA, that is believed to be from non-specific effects (Georgelis et al., 2014; Olarte-Lozano et al., 2014). The same two studies also noted enhancement of enzymatic activity with microbial expansins, but not beyond what was seen with BSA. Bunternngsook et al, (2015) found three microbial expansins that depicted varying levels of synergistic effect and no enhancement was seen with BSA. In a more recent study, quartz crystal microbalance with dissipation (QCM-D) was used to examine enzymatic hydrolysis of cellulose with an exoglucanase (Cel7A) or endoglucanases (Cel7B) and *BxEXLX1* (Zhang et al., 2021). They found that with a

1:1 ratio of enzyme and expansins, the rate of hydrolysis was 5 times the rate for enzyme alone. Another recent study, found a fungal expansin exhibiting synergistic activity with a commercial cellulase while BSA at 0.4 $\mu\text{g}/\text{mL}$ showed no obvious signs of synergy (Ding et al., 2022).

2.5 Production of Microbial Expansins

The discovery of bacterial expansin *BsEXLX1* proved to be a big milestone as bacterial expansins can be easily expressed in *Escherichia coli* (*E. coli*) where plant expansins have proved to be challenging (Georgelis et al., 2015). *E. coli* is an ideal host for recombinant protein production from prokaryotic sources due to its simplicity, fast growth, and greater economic feasibility. Additionally, in *E. coli* up to 50% of total protein content can be the desired recombinant protein (Francis & Page, 2010). Typically, production of microbial expansins occurred in *E. coli* strain BL21 using lysogeny broth (LB) and isopropyl β -d-1-thiogalactopyranoside (IPTG) for culture media and induction (Bunterngsook et al., 2015; Georgelis et al., 2014, 2015; Kim et al., 2009; Lee et al., 2010; Olarte-Lozano et al., 2014). Other expression parameters varied greatly between the studies. IPTG induction concentration varied between 0.1-1.0 μM . Too low concentration could result in insufficient induction, while high IPTG concentration can be toxic (Dvorak et al., 2015). Induction temperatures and length varied from 16°C to 37°C and 3 to 16 hours, respectively. Lower temperatures improve the solubility of recombinant proteins and decrease protease activity (Francis & Page, 2010). However, the cell processes slow down thus longer incubation times are needed. As a result, in this thesis various expression parameters were used for examining optimal production conditions of microbial expansin-like proteins.

Despite, the many benefits of recombinant production in *E. coli*, secretion of the target protein to extracellular space is more complex. As a result, once production is complete, the cell membranes need to be disrupted to extract the target protein from the cells and then the target protein purified from the rest of the cell debris. Due to the complexity, several methods have been developed for cell disruption, such as physical, or chemical lysis (ThermoFisher Scientific). One physical method is to use a sonicator that emits high frequency sound waves that agitate and break up the cells. This method is highly effective for large volumes but requires a lot of energy and

produces excess heat that can damage and denature the target proteins. Chemical detergent-based methods provide a softer and easier alternative for cell lysis. They break down the cell membrane by solubilizing proteins and disrupting lipid and protein interactions within the membrane. However, if using a commercially developed extraction solution for large-scale productions, the method can become expensive and inefficient. If the protein resides in the periplasm, the space between the two cellular membranes in gram negative cells, alternative protein extraction methods that do not disrupt both cell membranes can be utilized. By disrupting only the outer membrane, downstream processing of the protein can be eased as only 4-8% of native *E. coli* protein reside in the periplasm and other cellular components remain intact (Wurm et al., 2017). Kerff et al. (2008) reported that the *BsEXLX1* protein was expressed in the periplasm, indicating this may be a possibility for other expansin-like proteins as well. As a result, in addition, to the main aims of this thesis discussed in section 1.1, a small parallel experiment was completed comparing different lysis methods for protein extraction from *E. coli*. A lysis buffer that was theorized to only disrupt the outer membrane was compared to traditional sonication and chemical extraction solutions.

3.0 Materials & Methods

The following section discusses the materials and experimental procedures used to achieve the aims of this thesis. First, the expansin-like sequences used will be presented, followed by the different small-scale and large-scale production trials. Finally, the binding and synergism assays used to characterize the bacterial expansin-like proteins are detailed. Tables of the common chemicals and equipment used with manufacture details can be found in the appendix.

3.1 Microbial Expansin-Related Sequences

The ten different microbial expansin-related protein sequences analyzed in this thesis along with their properties are shown in Table 3.1. Accession numbers and sequences can be found in the appendix. Molecular weight, amino acids, pI and extinction coefficient (ξ) were calculated for the expressed protein using an online server ([ProtParam](#)). The first eight sequences were part of production trials, while the last three were used for characterization experiments. Five sequences (AP099, JAB176, JAB041, VFW002, and SMS108) were used in the small-scale studies. Three promising sequences from the small-scale production along with SFM691, CPO928 and LT367 were assessed in large-scale production trials. SFM691, CPO928 and LT367 had already been expressed in a previous students work and were chosen here in order to produce more protein for possible characterization studies.

The microbial expansin-like sequences were already designed and transformed to BL21 *E. coli* strains by previous members of the Protein Technology group. A pET-22b(+) expression vector with an ampicillin resistance and a T7 promoter that is activated by isopropyl β -D-1-thiogalactopyranoside (IPTG) were used.

Table 3.1 A | Microbial Expansin-Like Sequences. The label, source organism, accession numbers are shown in the table.

Label	DNA Accession Number	Gene Accession Number	Source Bacteria
AP099	AP024099.1	BCL92919.1	<i>Ralstonia solanacearum</i>
JAB176	JABDRL010000176.1	NNJ61190.1	<i>Dactylosporangium sp.</i>
JAB041	JABDRL010000041.1	NNJ59718.1	<i>Dactylosporangium sp.</i>
VFW002	VFWZ01000002.1	TPN87812.1	<i>Flavobacteriaceae bacterium</i>
SMS108	SMS01108	WP_087481129.1	<i>Vibrio mangrove</i>
CPO928	CP016928.1	AUC40652.1	<i>Dickeya aquatica</i>
SFM691	FOUV01000008.1	SFM96691.1	<i>Streptomyces sp.</i>
LT367	LT615367	SLM63089.1	<i>Dickeya aquatica</i>
AJC165	AJC47165	WP_043094747	<i>Xanthomonas sacchari</i>
HE673	HE614873.1	CCE74122.1	<i>Clavibacter nebraskensis</i>

Table 3.1 B | Microbial Expansin-Like Sequences. The label, size in kDa, number of amino acids, pI, extinction coefficient and modularity are shown in the table. The properties shown are based on the expressed protein sequence.

Label	Size (kDa)	AA	pI	ξ $\left(\frac{L}{g\ cm}\right)$	Modularity
AP099	24.8	230	9.24	1.595	Exp
JAB176	36.9	360	4.54	1.753	Exp-CBM2
JAB041	26.7	259	8.64	1.279	Exp
VFW002	50.9	461	5.92	1.282	Exp-CBM6
SMS108	34.9	324	4.39	1.987	Exp-CBM2
CPO928	23.6	216	6.16	1.755	Exp
SFM691	23.0	211	7.84	1.554	Exp
LT367	23.6	217	8.42	1.819	Exp
AJC165	23.4	237	6.82	1.710	Exp
HE673	35.6	341	5.78	1.975	CBM2-Exp

3.2 Small-scale Production

Small-scale production was completed with two different media: MagicMedia™ from Invitrogen and typical LB media with IPTG induction. Three different trials were completed utilizing MagicMedia and the dual temperature protocol to examine the expression of five difficult to produce proteins: APO99, JAB176, JAB041, VFW002, and SMS108. One trial was completed with the LB media and IPTG induction on the three best proteins selected from the MagicMedia trials. The different production conditions and protocols are elaborated on in the following sections.

3.2.1 MagicMedia Production

In general, the BL21 *E. coli* strains were streaked onto LB plates supplemented with ampicillin, incubated overnight at 30°C and then stored in the fridge (8°C). Colonies from the plates were used to grow precultures in 15 mL sterile culture tubes and covered with sterile breathable tape. Precultures were incubated at 300 to 350 RPM and 30 to 35°C using Thermomixer C (Eppendorf). Optical density (OD₆₀₀) measurements were taken at 600 nm using BioPhotometer plus (Eppendorf). LB media was used as a blank and for sample dilutions. Once OD₆₀₀ was at least > 0.7, main cultures were inoculated and covered with breathable tape. Duplicates of each protein was used in main cultures. Cultures were incubated in a tabletop incubator, Certomat R (B. Braun). After an initial 6 hours of growth, the temperature was decreased to the induction temperature indicated in the table below. Cells were harvested by centrifugation until supernatant was clear and not cloudy. Cell dry weights were measured and recorded.

Table 3.2 | MagicMedia Production. The following table shows the different expression conditions tested in the three different MagicMedia production trials.

Trial	Size (mL)	RPM	Temperature (°C)	Time hours
1	15	240	20	18
2	15	265	18	18
3	25	265	18	16

In the first trial, 2.5 mL of MagicMedia was used in the preculture. Once OD₆₀₀ was at least 1.0, the main culture (15 mL of MagicMedia in 150 mL Erlenmeyer flasks) was

inoculated with 35 to 40 μL of preculture. The preculture was grown for another 3 hours and then harvested by centrifugation at 4°C and 3800 RPM for 30 min. The main cultures were incubated at 37°C and 230 RPM for 6 hours after which the temperature was decreased to 20°C . After 18 hours of incubation at 20°C , the cells were centrifuged for 30 min at 4°C and 3900 RPM.

In the second trial, 2.5 mL of LB media and 100 $\mu\text{g}/\text{mL}$ of ampicillin as a selective antibiotic was used for the preculture. The main cultures (15 mL of MagicMedia in 150 mL Erlenmeyer flasks) were inoculated with 350 to 400 μL of preculture depending on OD_{600} values. Main cultures were incubated at 265 RPM and 30°C for 6 hours. Before decreasing the temperature to 18°C , 2 mL samples were collected for analysis. After 18 hours of incubation at the lower temperature, the cells were harvested.

In the third trial the volumes were slightly increased. The preculture was 3 mL of LB media (100 $\mu\text{g}/\text{mL}$ ampicillin) and the main culture was 25 mL of LB (100 $\mu\text{g}/\text{mL}$ ampicillin) in 250 mL Erlenmeyer flasks. The main cultures were inoculated with 850 to 900 μL of preculture, and incubated at 265 RPM and 30°C . Similar to trial 2, a 2 mL sample was taken after 6 hours of growth before the temperature was decreased to 18°C . The cells were harvested after 16 hours at the lower temperature.

3.2.2 LB Media

After identifying successful expressions in the MagicMedia trials, the same needed to be completed with LB media, a cheaper and more practical media for large-scale production. Trial 4 was completed with only the promising proteins from the three MagicMedia trials: AP099, JAB176 and JAB041. Using the same plates from the MagicMedia trials, colonies were inoculated in 3 mL of LB (100 $\mu\text{g}/\text{mL}$ ampicillin) in 15 mL sterile culture tubes and covered with breathable tape. Precultures were incubated at 300 to 350 RPM and 35 to 37°C until $\text{OD}_{600} > 1$. Main cultures (30 mL of LB + ampicillin in 250 mL Erlenmeyer flasks) were inoculated with 240 μL of preculture. Cultures were incubated at 30°C and 255 RPM until OD_{600} was around 0.6 to 0.8. Samples were induced with 1 mM IPTG and incubated at 20°C and 265 RPM for 18 hours. Cells were harvested by centrifugation at 3800 RPM and 4°C for 30 min.

3.3 Large-scale Production with LB Media

Large-scale production of multiple different proteins was completed in LB media with IPTG induction. Various culture sizes were done with multiple 2 or 3 L Erlenmeyer flasks covered with breathable tape. The LB volume was one fifth of the flask size with ampicillin supplemented (100 μ g/mL) as an antibiotic. Colonies from plates were used to inoculate precultures that were covered with breathable tape. Once OD₆₀₀ had reached around 1.0, the precultures were used to inoculate the main cultures. The main cultures were grown at 30 to 37°C and 170 to 230 RPM in Innova 44 incubator shaker series (New Brunswick Scientific). Once OD₆₀₀ was around the target value, the cultures were induced with either 0.5 mM or 1.0 mM IPTG. Induction was carried out at set temperature and duration that are depicted in the table below. After induction cells were harvested by centrifugation (SORVALL LYNX 4000, ThermoFisher Scientific). Supernatant was discarded and cell pellets were resuspended with LB media and transferred to pre-weighed 50 mL falcon tubes. After another round of centrifugation, the supernatant was removed, and cell dry weights were measured from pellets.

Table 3.3 | Large-scale Production Conditions. The production conditions (size, RPM, temperature, time, OD before induction, and IPTG induction concentration) for each large-scale production is summarized in the table below.

Protein	Size (L)	RPM	Temp (°C)	Time (hour)	OD₆₀₀	IPTG (mM)
AP099	3.2	230	18	15	0.8-1.0	1
JAB176	2.8	215	16.5	14	0.7-0.8	1
JAB041	3.2	215	16.5	14	0.8	1
LT367	3.6	230	16	12	0.8-1.0	1
CPO928	3.2	230	18	15	0.4-0.8	0.5
CPO928	3.2	230	16	12	0.8	1
SFM691	2.8	230	16	12	1.0	1

3.4 Cell Lysis

For extracting the desired protein from the cells, three different cell lysis methods were used: Bacterial Protein Extraction Reagent (B-PER) (ThermoFisher Scientific), a MgCl_2 lysis buffer (50 mM TRIS, 5 mM MgCl_2 , 1 mM EDTA, pH 7.8), and sonication with Q500 Sonicator (QSONICA). The procedures for each are outlined below. Cells from LB production Trial 4 were used to compare B-PER and sonication lysis methods. Another lysis comparison experiment was done with AP099, JAB176 and JAB041 from the large-scale production to compare B-PER vs. MgCl_2 and sonication vs. MgCl_2 .

3.4.1 B-PER

B-PER was used for lysis of small volumes, such as in the MagicMedia productions. Harvested cell pellets were resuspended in 4 mL per gram of cells of B-PER. Resuspended cell pellets were vortexed until samples were homogenous. Next, samples were rotated for 20 min at 22 RPM and room temperature. Then, samples were centrifuged for 30 to 40 minutes at 15000 g (12650 RPM). Supernatant and cell pellets were collected and stored separately in -20°C for further analysis. An unpurified sample was collected from the supernatant, and the protein concentration measured via NanoDrop Lite Spectrophotometer (ThermoFisher Scientific).

3.4.2 MgCl_2 Buffer

A MgCl_2 lysis buffer (50 mM TRIS, 5 mM MgCl_2 , 1 mM EDTA, pH 7.8) was a new lysis method analyzed for extracting proteins from within cells. In the initial experiment comparing B-PER vs. MgCl_2 , cells (AP099, JAB176 and JAB041) were resuspended in 4 mL of lysis buffer (50 mM TRIS, 1 mM EDTA, pH 7.8) per gram of cells. Samples were incubated at room temperature and 300 RPM for 60 min. Next, 5 mM MgCl_2 was added and samples incubated for 5 minutes. Samples were centrifuged for 20 minutes using 8000 RPM at 4°C . Supernatant and cell pellets were collected and stored separately. An unpurified sample was collected from the supernatant, and the protein concentration measured via NanoDrop.

In the analysis comparing MgCl₂ vs. SONICA, cell pellets were resuspended in 2 mL/g of lysis buffer already containing the MgCl₂ (50 mM TRIS, 5 mM MgCl₂, 1 mM EDTA, pH 7.8). Samples were incubated at 300 RPM and room temperature for 60 minutes. Samples were centrifuged at 15000 g for 20 minutes. Supernatant and cell pellets were collected and stored separately in -20°C until further analysis. An unpurified sample was collected from the supernatant, and the protein concentration measured via NanoDrop Lite Spectrophotometer (ThermoFisher Scientific).

3.4.3 Sonication

The Q500 Sonicator (QSONICA) was used for lysis of large volumes in the large-scale expression trials. The cell pellets were resuspended in 2 mL/g of buffer A (150 mM NaCl, 50 mM TRIS, 10 mM imidazole, pH 7.8). The cells were vortexed until homogenous. The samples were placed on ice during sonication with QSONICA. Three different settings were used on the sonicator and are depicted in Table 3.3.

Table 3.4 | Sonication settings. The following table shows the different settings used for cell lysis with the sonicator.

Setting	Pulse on (seconds)	Pulse off (seconds)	Total Pulse (min)	Amplitude (%)
1	2	5	2	30
2	2	5	1.5	25
3	2	5	1	25

3.5 Purification

After cell lysis samples were purified either with manual His GraviTrap™ TALON® (GE healthcare) columns or ÄKTA Purifier 10 FPLC System (Amersham Biosciences). The procedure for each is described in the next two sections.

3.5.1 Small-scale Purification

His GraviTrap™ TALON® (GE healthcare) columns were used for small-scale purification, volumes less than 2 mL. The columns were prepared by removing the ethanol storage solution and washing with at least 15 mL of MQ water. The column

was then equilibrated with 10 mL buffer A (150 mM TRIS, 50 mM NaCl, 10 mM imidazole, pH 7.8). The supernatant after cell lysis was loaded onto the column, and the column washed with 15 to 20 mL of buffer A. The flow through was collected and the target protein was eluted with 2.5 to 3 mL of elution buffer B (500 mM imidazole, 150 mM TRIS, 50 mM NaCl, pH 7.8). The protein concentration of the purified sample was measured in triplicates with NanoDrop. The columns were washed with buffer A, MQ water, and then stored in 20% ethanol.

3.5.2 Large-scale Purification

ÄKTA Purifier 10 FPLC System (Amersham Biosciences) was used for purification of sample volumes greater than 15 mL. All large-scale productions were purified using ÄKTA and the His Trap™ ff crude 5 mL column by Cytiva. All buffers and MQ water were filtered with 0.45 µm Millipore Express® PLUS filter paper by Merck and degassed before installation in ÄKTA. Supernatant from cell lysis was also filtered with Whatman Puradisc™ 25 mm. Main pumps and sample pumps were washed with five times column volume (CV) of MQ water and buffer A (150 mM TRIS, 50 mM NaCl, 10 mM imidazole, pH 7.8). After the column was equilibrated with sufficient buffer A, the filtered supernatant was loaded onto the column and the flow through collected. Next, the column was washed with A buffer until UV-280 values had returned to preloading values. Then, the protein was eluted into 2 mL fractions with a gradual increase of buffer B (500 mM imidazole, 150 mM TRIS, 50 mM NaCl, pH 7.8) concentration to 45% over 3 minutes. Once the protein was eluted, buffer B was increased to 75 to 80% to wash the column. Finally, the column was washed with buffer A to prepare for next purification. Fractions were analyzed on SDS-PAGE for purity. The top 3 to 4 fractions with high purity and concentration of desired protein were combined. The protein concentration of the pooled-down fractions was measured via NanoDrop. The elution buffer B was exchanged to a sodium acetate (NaOAc) buffer pH 5.5 with Vivaspin® 20, 10000MWCO (Sartorius). After buffer exchange the protein concentration was measured via NanoDrop and stored for 3 days at -20°C. After 3 days the protein concentration was remeasured, and flash frozen with liquid nitrogen prior to storage at -80°C.

Mini-PROTEAN® TGX™ (Biorad) SDS-PAGE gels were used for protein identification during production and protein purity analysis during purification. Samples were diluted

with MQ water or concentrated depending on NanoDrop readings in order to achieve a visible band on the gel. Heated cell samples were prepared by resuspending lysed cells in buffer A, heating at 90°C for 30 minutes, and then centrifuging to separate the supernatant and cell pellet. Samples were loaded onto gel along with Precision Plus Protein™ Dual Color (Biorad). The gel chamber was filled with SDS buffer and powered with PowerPac 3000 (Biorad) until the samples reached the end of the gel. The gel was stained with PageBlue protein staining solution (ThermoFisher Scientific) for 1 hour with gentle shaking. The gel was de-stained with distilled water overnight and gentle shaking. The gels were imaged with ChemiDox XRS+ System (Biorad).

3.6 Characterization

HE673, LT367 and AJC165 were examined for expansin-like characteristics with Insoluble Polysaccharide Pulldown (IPP) assay to various commercial substrates and synergism studies with cellulase from *Trichoderma reesei* (Celluclast, Sigma) and xylanase from a non-pathogenic fungal source (Ecopulp TX-800 A, AB Enzymes GmbH).

3.6.1 IPP Assay

IPP assays were completed in 25 mM HEPES, pH 7.5 buffer and 25 mM NaOAc pH 5.5 buffer. Commercial products (Avicel PH-101, oat spelt xylan, and chitin) and industrial biomasses (bleached, oven dried at 60°C hardwood and softwood pulps) were used as substrates. Stock solutions of 1% w/v of commercial substrates were prepared by weighing 1 g of substrate and adding 100 mL of MQ water. Stock solutions were well mixed using magnetic stirrer. From well-mixed homogenous solution 250 µL was pipetted to 2 mL Eppendorf tubes. Tubes were centrifuged for 20 minutes at 15000 RPM and 4°C. The supernatant was removed and 250 µL of water was added. Samples were vortexed and centrifuged again for 20 minutes at 15000 RPM and 4°C. Supernatant was removed, and samples were washed with 250 µL of ethanol, and then twice more with water. After final washing step, supernatant was completely removed, and substrate pellet was stored at 8 °C. The pulps were prepared by weighing 2.5 mg into 2 mL Eppendorf tubes.

Reactions were completed with a triplicate of a negative control (substrate but no protein) and each sample (HE673, LT367, AJC165 and BSA) having triplicates and two negative samples (protein with no substrate). Proteins were prepared from frozen stocks by diluting with sodium acetate buffer pH 5.5 until protein concentration was around 2 mg/mL as measured with Bradford assay. The reaction was completed with a reaction protein concentration of 0.2 mg/mL and a reaction volume of 250 μ L. The reaction tubes were briefly vortexed to start the reaction and then incubated for 90 minutes at 1100 RPM and 25°C. The reaction was stopped by centrifuging at 15000 RPM and 4°C for 15 minutes. The supernatant from samples was used to determine unbound protein concentration with a Bradford assay. Protein samples with no substrate were used for determining initial protein concentration.

Bradford was used to determine protein concentrations. The Bradford reagent was prepared by filtering a 5X dilution of Bio-Rad Protein Assay Dye Reagent Concentrate with 0.45 mM Whatman filter paper. A standard curve was prepared with serial dilutions (0.5, 0.25, 0.125, 0.0625, 0.031, and 0.015 mg/mL) of bovine serum albumin (BSA). In a 96 well plate, 10 μ L of standards and samples were pipetted in triplicates. With a multichannel pipette 200 μ L of prepared Bradford reagent was pipetted into the well plate. The plate was incubated at room temperature for at least 10 minutes. The absorbances at 595 nm was measured with EON plate reader by BioTek.

3.6.2 Synergism Reactions

An initial synergism test was completed with used sample tubes from hardwood pulp IPP assay pH 5.5. To the reaction tubes already containing proteins from the IPP assay, cellulase from *Trichoderma reesei* (Celluclast, Sigma) was added to reach a total reaction volume of 250 μ L and a cellulase concentration of 0.1 mg/mL. A zero hour sample was collected before vortexing the reaction tubes. Samples were incubated at 40°C and 1100 RPM for 72 hours. Another 20 μ L sample was collected after 8 hours. The 0, 8 and 72 hour samples were analyzed with 4-hydroxybenzhydrazide (PAHBAH) assay to determine the total amount of reducing sugars that were released into the solution.

After the initial synergism test, four other experiments were run, all on hardwood pulp at 25 mM NaOAc pH 5.5. Hardwood pulp (2.5 mg) was weighed into 1.5 mL Protein

LoBind Eppendorf tubes. All experiments were completed with 1% w/w_{substrate} of protein at 35°C and 1100 RPM with samples collected at 0, 2, 6, 8, 24 and 48 hours. The zero hour samples were collected after the addition of each component to reaction tube and before a brief vortex of the tubes to begin the reaction. Final sample was collected by centrifugation of the reaction tubes for 15 min at 15000 RPM and 4°C and collecting the entirety of the supernatant. Other samples were collected by taking a small sample from the clear supernatant after a brief centrifugation. The different conditions for each test are highlighted in the table below.

Table 3.5 | Synergism Reactions. The table summarizes the conditions that varied in each synergy reaction, such as hardwood substrate amount, reaction volume, enzyme concentration in relation to substrate amount and sample volume.

Experiment	Hardwood (mg)	Reaction Volume (μL)	Enzyme (w/w_{sub} %)	Sample Volume (μL)
1	2.5	250	1% cellulase Celluclast (Sigma)	17
2	2.5	250	1% xylanase Ecopulp TX-800 A (AB Enzymes)	17
3	2.5	500	1% xylanase Ecopulp TX-800 A (AB Enzymes)	34
4	2.5	500	0.5 % xylanase Ecopulp TX-800 A (AB Enzymes)	34

The samples were then analyzed with a PAHBAH assay to determine the quantity of reducing sugars released during the reaction. A standard curve was prepared with serial dilutions of glucose (0.015 – 0.75 mg/mL). Into a 96 well plate, 10 μ L of standard and sample were pipetted. The PAHBAH reagent was prepared by adding 5 mL of reagent A (5% w/w 4-hydroxybenzhydrazide in 0.5 M HCl) into 20 mL of reagent B (0.5 M NaOH). With a multichannel pipette 200 μ L of the PAHBAH reagent was added to the 96 well plate. The plate was covered and then incubated at 70°C for 30 minutes and then cooled down at 8°C for 10 minutes. The plate was centrifuged for a minute at 3900 RPM. The absorbances were read at 405 nm using EON plate reader by BioTek.

4.0 Results and Discussion

The following chapter is divided into four main sections: small-scale production, lysis comparison, large-scale production, and characterization. The main results from each experiment are discussed and analyzed. Additional figures are provided in the appendix.

4.1 Small-scale Production

In the following sections, the results from the three magic media trials and the first LB trial will be discussed. The different conditions tested in the three MagicMedia trials and the first LB trial are briefly summarized in Table 4.1. The outcome of analyzing trials 1 – 4, was that lower temperatures and shorter incubation times were more promising for successful expression of AP099, JAB176 and JAB041.

Table 4.1 | Small-scale Production Conditions. The following table shows the different expression conditions tested in the three different MagicMedia production trials and the small-scale LB trial.

Trial	Size (mL)	RPM	Temperature (°C)	Time hours	IPTG
1	15	240	20	18	-
2	15	265	18	18	-
3	25	265	18	16	-
4	30	265	20	18	1 mM

4.1.1 MagicMedia

The expression of 5 proteins (AP099, JAB176, JAB041, VFW002, and SMS108) that were difficult to express, were examined with a MagicMedia *E. coli* expression system (Invitrogen). MagicMedia promotes high yield growth of *E. coli* and high expression of T7 heterologous proteins (Life Technologies, 2014). In MagicMedia trial 1, the samples were not purified, and loading was too heavy on the gel, leaving it difficult to analyze the success of desired protein expression (Figure C.1 in appendix). Samples from trial 2 were more successfully loaded on the gel and can be seen in Figure 4.2 A (left panel). The left panel illustrates an SDS-gel with weak bands around 25 kDa

for purified AP099 and JAB041, and around 37 kDa for purified protein JAB176. This corresponds to the theoretical size of AP099 (25 kDa), JAB041 (27 kDa) and JAB176 (37 kDa). Samples from trial 3 shown in Figure 4.2 B (right panel) reconfirm the bands seen in trial 2. The three proteins with the strongest and thickest bands at the correct theoretical size (AP099, JAB176 and JAB041) were chosen for further optimization.

A thick band can be seen around 50 kDa in purified SMS108 in both gels. However, this does not correspond to the theoretical size of 35 kDa for SMS108. No post translational modifications are expected since the production host and source organism are prokaryotes. The acidity of the protein ($pI = 4.39$) may cause slightly higher apparent molecular weights on SDS-PAGE (Alves et al., 2004; Graceffa et al., 1992). The error could have also occurred at the transcriptional or translational level with amino acid mismatching (Wong et al., 2018). Further analysis of the protein band is needed for identification of the protein.

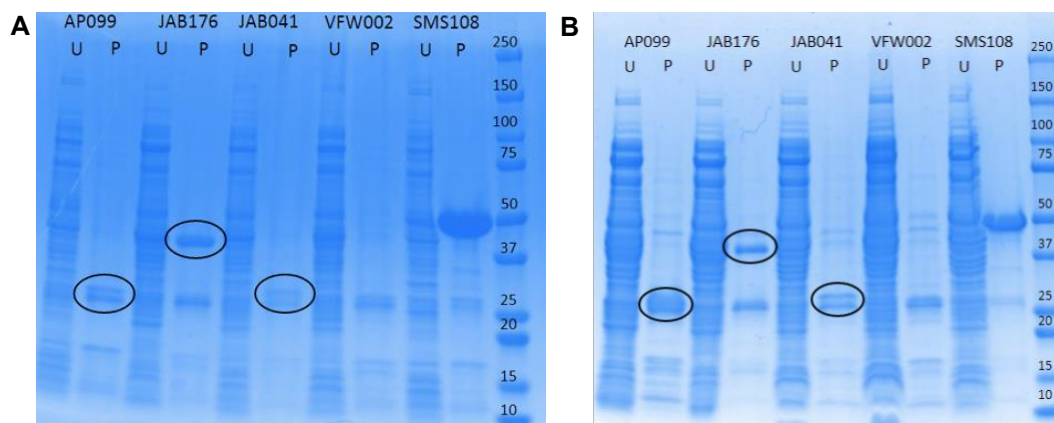


Figure 4.2 | Magic Media Production. The SDS protein gels from small-scale production trials 2 and 3. Purified (P) and unpurified (U) samples from cell cultures produced in trial 2 (A) and trial 3 (B).

4.1.2 LB Trial 4

MagicMedia is an expensive *E. coli* expression media and is not practical for large-scale productions, especially when the expression has not been confirmed nor the protein characterized for its desired functions. As a result, another small-scale trial was conducted with LB media, a cheaper and more practical media for large-scale production. A faint band around 37 kDa is detected in purified JAB176, corresponding to theoretical size of the protein. No distinguishable bands were observed in purified

AP099 nor JAB041. This can be due to many reasons other than a successful expression, such as insufficient quantity of protein loading onto the gel for the bands to be visible. Regardless, in the unpurified batch of AP099 and JAB041, there are thick dense bands around 25 kDa, suggesting possible desired protein.

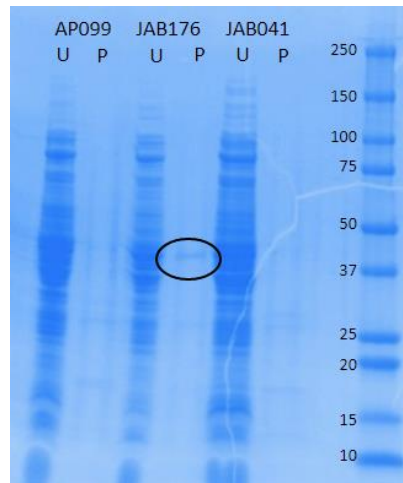


Figure 4.3 | Small-scale LB Production. The protein gel for purified (P) and unpurified (U) samples from LB trial 4 completed at 20°C and 265 RPM for 18 hours with 1 mM IPTG induction.

4.2 Lysis Comparison

Three different trials were completed to analyze different lysis methods for extracting the desired proteins AP099, JAB176 and JAB041 from cell debris. In the first trial, cells from production trial 4 were used to compare B-PER and sonication methods (Figure C.3). No clear bands are visible in either the purified samples from B-PER or sonication. The second and third trial, comparing B-PER with $MgCl_2$ lysis buffer, and $MgCl_2$ lysis buffer with sonication were completed with cells produced in large-scale production trials (Figure 4.4).

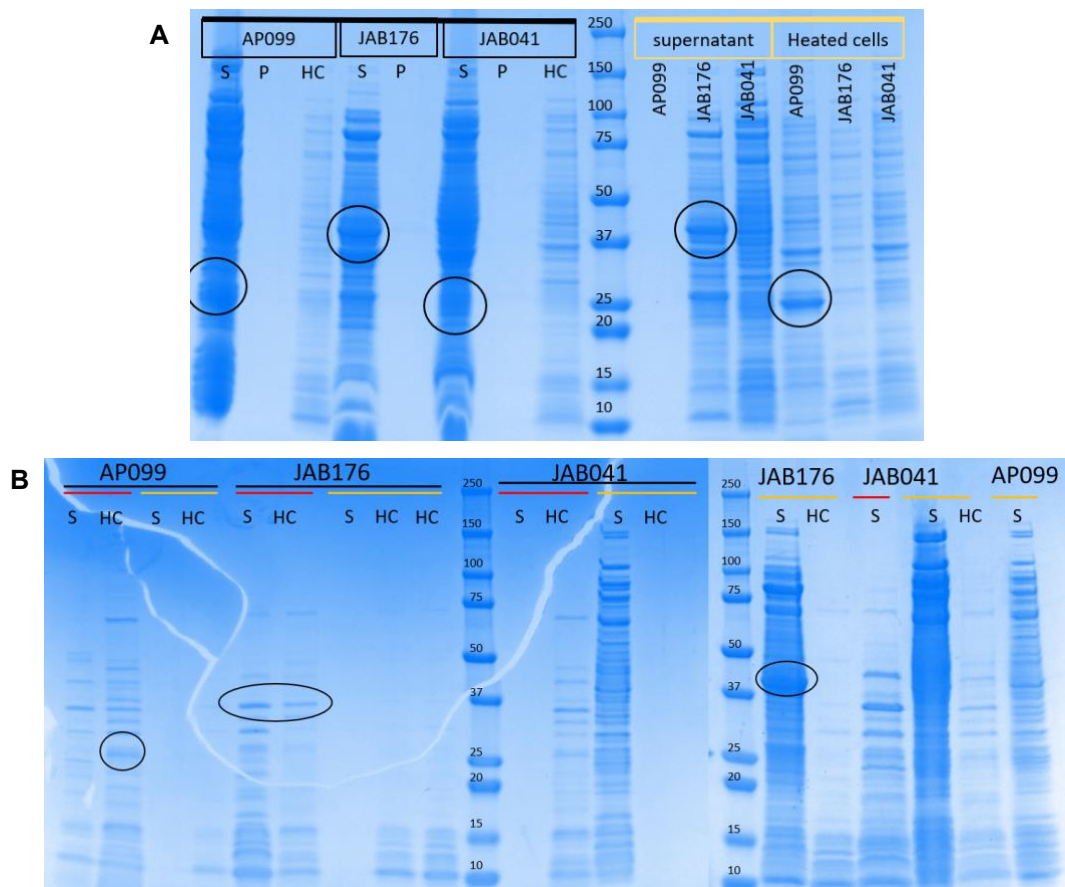


Figure 4.4 | Lysis Comparison. Protein gels with purified (P) protein, unpurified supernatant (S) and supernatant from heated cells after lysis (HC). Gel A compares B-PER results (black) with MgCl₂ lysis buffer results (yellow). Gel B shows comparison between sonication (red) and MgCl₂ lysis buffer (yellow).

In the upper panel (A), with the purified samples no bands were seen at all, indicating that the concentration of the purified samples was not sufficient to be illuminated on the gel. However, the unpurified supernatant samples from B-PER have thick bands around 25 kDa for AP099 and 37 kDa for JAB176. In the heated-lysed cells from B-PER no thick bands are observed at the theoretical size for AP099 or JAB041, indicating that the B-PER lysis was successful in removing all the desired proteins and very little target protein if any is left over in the lysed cells. This is not the case with the AP099 sample lysed with MgCl₂ lysis buffer. In the heated cells from the MgCl₂ lysis a thick band is seen around 25 kDa for AP099. This means that MgCl₂ lysis was not successful in removing AP099 from the cell and some target protein was still left in the cells. In contrast, the unpurified JAB176 supernatant from MgCl₂ lysis had a thick band around 37 kDa, implying that the JAB176 protein was successfully extracted with MgCl₂ lysis buffer and thus resides in the periplasm. This result is

reconfirmed during the sonication vs. MgCl₂ analysis. However, the sonication JAB176 samples also show bands around 37 kDa for both the protein supernatant and the heated cells, suggesting the protein also resides within the cell and the rounds of sonication were insufficient for extracting all the protein. Initial testing indicates that the MgCl₂ lysis buffer could be an alternative extraction method for JAB176. Further testing, however, is still needed to confirm whether the method could improve overall production yields. As a result, for large-scale production lysis, the standard three rounds of sonication were used to guarantee extraction of the target proteins.

4.3 Large-scale Production and Purification

The three proteins AP099, JAB176, and JAB041 that were shown to be promising in the MagicMedia trials were used for large-scale production and purification. An additional 3 other proteins that had previously been successfully produced were also produced at large-scale in order to produce more protein for further characterization analysis of bacterial expansin-like proteins. The production conditions for each protein are summarized in the table below.

Table 4.5 | Large-scale Production. The production conditions (size, RPM, temperature, induction length, IPTG induction concentrations) along with cell dry weights and amount of expressed protein are summarized.

Protein	Size (L)	RPM	Temp (°C)	Time (hour)	IPTG (mM)	CDW (g)	Amount (mg)	Yield (mg/L)
AP099	3.2	230	18	15	1	7.9	-	-
JAB176	2.8	215	16.5	14	1	9.8	-	-
JAB041	3.2	215	16.5	14	1	15	-	-
LT367	3.6	230	16	12	1	11	~22	6
CPO928	3.2	230	18	15	0.5	7.6	~3.5	1
CPO928	3.2	230	16	12	1	8.3	~1	0.3
SFM691	2.8	230	16	12	1	7.8	-	-

4.3.1 AP099, JAB176 and JAB041

The large-scale production and purification of AP099 yielded negligible amounts of protein. In total 3.2 L was produced and lysed with sonication and the protein supernatant was purified using ÄKTA. The desired protein was identified with SDS-PAGE in samples A₇₋₁₁. The samples early in the elution had high impurities and were primarily of a protein around 70 kDa in size. Fractions A₇₋₁₀ were pooled-down together with an A₂₈₀/mL concentration of 0.637 as measured with NanoDrop. Based on volume and final concentration of pooled-down eluted fractions, only 3 mg of protein was produced. Therefore, no buffer exchange was done and no AP099 protein was collected.

During resuspension of cells to prepare for lysis with sonication, it was noted that the AP099 cells took longer and were more difficult to resuspend. Based on previous results with lysis analysis indicating that AP099 was still found in lysed cells, it was hypothesized that perhaps some protein was still left in the lysed cells causing the low yield. Samples of the lysed cells were resuspended in Buffer A and heated to 95°C to extract all proteins and run on SDS-PAGE to see if any protein was still left in the cell. The gel showed, no strong bands around 25 kDa.

Both JAB176 and JAB041 also had negligible amounts of protein produced with large-scale production. In the chromatogram, unlike AP099 with broad peaks and an initial shoulder, JAB176 and JAB041 had narrow straight peaks. SDS-PAGE of the pooled-down fractions confirmed bands around the theoretical sized of 25 kDa and 37 kDa for JAB041 and JAB176 respectively. However, the SDS-PAGE showed numerous other bands of various sizes, including a similar band at 70 kDa that was seen in AP099 as well. By increasing the time gradient in which 35% buffer B concentration was reached and broadening the narrow elution peaks of JAB176 and JAB041, an elution peak similar to AP099 can be reached and the 70 kDa protein may elute earlier creating higher purity elution fractions of the desired protein. Regardless of the impurities fractions A₄₋₈ for JAB041 and A₄₋₉ for JAB176 were pooled-down together. The A₂₈₀/mL of pooled-down JAB176 and JAB041 were 1.445 and 1.417, respectively, keeping in mind the high imidazole concentration of the elution buffer and impurities. Considering total volume and extinction coefficient, this was about 10 mg of total proteins. JAB176 also showed a band at 37 kDa in the flow through,

suggesting even more protein. During buffer exchange to remove the imidazole content, a white precipitate was seen in both JAB176 and JAB041, causing a loss of all potential protein and resulting in a negligible yield of protein.

With all three proteins, no quantifiable yield was achieved. However, with the larger volumes during production and purification the confirmation of protein expression could be seen more readily in the SDS-PAGE gel at corresponding theoretical sizes, indicating that the desired protein is expressed in *E. coli* despite low yields. Further optimization of production conditions and purification methods is recommended to achieve quantitative yields. For example, JAB176 and JAB041 successfully produced protein but was isolated with high impurities and subsequently lost during buffer exchange to precipitation. Modification of the purification protocol and methods that improve the purity of the proteins and prevent precipitation can help achieve quantitative protein yields. Initial assessments of different lysis methods indicated that JAB176 may reside in the periplasm and can be extracted with a $MgCl_2$ lysis buffer. By using a different lysis method that breaks down only the outer membrane of the cells and releases a smaller fraction of cellular protein, the purity of JAB176 may be improved and help prevent loss of protein that occurred due to precipitation.

4.3.2 LT367, CPO928 and SFM691

Large-scale production of LT367 was successful with a final yield of around 6 mg/L. Running the fractions on an SDS-PAGE gel revealed strong bands around the theoretical size of LT367 at 23.6 kDa (Figure C.8 in appendix). The fractions early in the peak showed multiple impurities, but further into the peak the impurities decreased. Fractions A₅₋₁₀ containing the desired protein were pooled-down together, and buffer exchanged to remove the high concentration of imidazole. In total 22 mg of protein (2 mL of 11 mg/mL) was collected and stored at -80°C for further characterization.

Two large-scale productions of CPO were done at different conditions each containing 3.2 L of LB media. In the first production, illustrated on the SDS-PAGE (appendix), thick bands around the theoretical size of 23.6 kDa confirmed the production of CPO928 protein. Similar to that of LT367, earlier elution fractions had higher number of impurities. Fractions A₅₋₉ were pooled-down together, and buffer exchanged to a

50 mM sodium acetate buffer (pH 5.5). In total 4.5 mg/mL of CPO928 protein was collected. This equates to a yield of 1 mg/L, 6-fold smaller than LT367 despite similar chromatogram results from purification for both proteins. It is possible that protein precipitated during buffer exchange. It was noted that CPO928 looked milky or cloudy couple of days after buffer exchange, indicating possible protein precipitation.

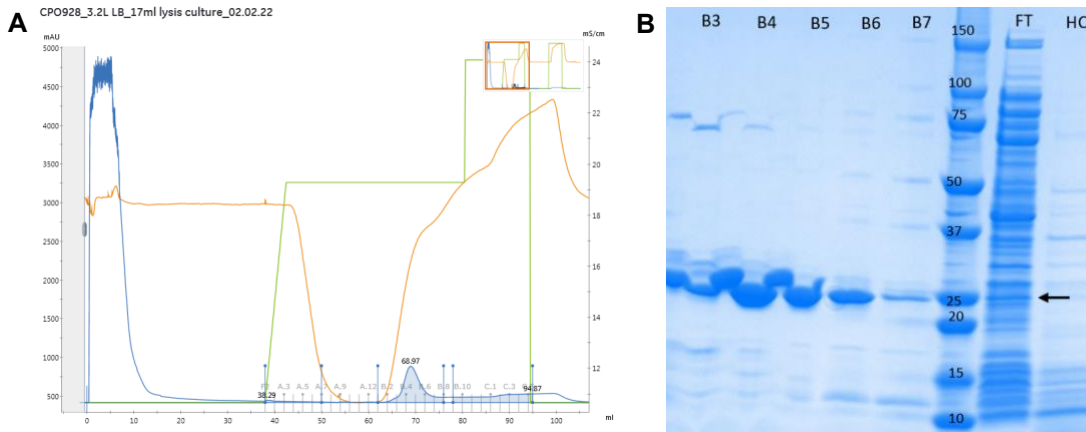


Figure 4.6 | Large-scale Production of CPO928. The chromatogram (A) and SDS-PAGE gel (B) from the second large-scale production and purification trial of CPO928. The chromatogram shows the UV280 (blue), concentration gradient of buffer B (green), and conductivity (orange). The gel was run with elution peak samples B3-B7, flow through (FT) and the supernatant from heated leftover cells (HC).

Since CPO928 is from the same organisms as LT367, *Dickeya aquatica*, it was hypothesized that same production conditions that were used to obtain a 6 mg/L yield could be utilized to boost the production yield of CPO928. Thus, the second CPO928 production was completed at 16°C, 12 hours, and 1 mM IPTG induction. Based on the chromatogram (Figure 4.6 A) the purification was not successful. The elution peak occurred very late in fractions B₃₋₇ and the conductivity (orange line in figure) was observed to drop after the concentration of buffer B was increased. The SDS-PAGE (Figure 4.6 B) confirms the production of CPO928 at theoretical size of 23.6 kDa. After pooling together fractions B₃₋₇ and buffer exchange, 0.4 mL of 3.37 mg/mL protein was collected. Just as in the previous CPO928 production, after a couple of days the protein solution appeared cloudy, and the concentration was measured at 2.53 mg/mL using NanoDrop. The final yield was as low as 0.3 mg/L. However, analysis of the SDS-PAGE, revealed a thick band around 23 kDa in the flow through sample, suggesting a large amount of protein could still reside in the flow through that was not successfully purified. Sadly, due to the breakdown of the ÄKTA purifier, the quantity of possible protein left in the flow through could not be collected and quantified. At

least another 2.3 mg of protein would have been needed to be collected from the ~15 mL of flow through to match the 1 mg/L yield observed in the first CPO production. The flow through had a A280/mL of 54 as measured with the NanoDrop. Estimating that CPO928 is 1% of the total protein in the flow through and a 50% recovery from purification and buffer exchange, the necessary 2.3 mg of CPO would be collected from the flow through. As a result, it is hard to determine which production conditions provided a larger yield without retrieving and quantifying the possible protein still left in the flow through.

The final protein produced in large-scale was SFM691. Like proteins AP099, JAB176 and JAB041 no quantifiable yield was produced. SDS-PAGE confirmed the protein around the theoretical size of 23 kDa. Similar to the other proteins initial fractions early in the peak contained more impurities. Fractions A₄₋₇ were pooled-down together, and concentration was measured (0.433 A280/mL on NanoDrop). Due to very small amount of protein (< 3 mg) no buffer exchange was performed and protein was not stored.

Out of the three protein LT367, CPO928 and SFM691, two were successfully produced at quantifiable yields. LT367 was the only protein produced at a large enough quantity for characterization studies with a yield of about 6 mg/L. However, this is still only about half the 10 mg/L yield reported for *BsEXLX1* (Kim et al., 2009). CPO928 was also successfully produced but had an even smaller yield of around 1 mg/L. SFM691 produced no protein. Since previous expression trials have already been completed with LT367, CPO928 and SFM691 and no improvement was seen in these trials, it is now recommended that other methods besides changing culture parameters are utilized to improve production yields.

4.4 Characterizations

Bacterial expansin proteins that were in available large enough quantity (HE673, LT367, and AJC165) were utilized for characterization studies to examine whether the produced proteins indeed exhibited expansin-like characteristics. As mentioned in the introduction, expansins will bind to different carbohydrate substrates via the D2 domain and will improve the hydrolysis yield of hydrolytic enzymes. As a result,

binding was examined with Insoluble Pulldown Polysaccharide assays (IPP) and synergisms were examined with cellulase and xylanase.

4.4.1 Binding Assays

Binding assays were completed on commercial products Avicel, chitin, and oat-spelt xylan in two pH conditions, 25 mM HEPES pH 7.8 and 25 mM NaOAc pH 5.5. AJC165 was not tested with every substrate due to time and limited quantity of the protein. The lower pH was tested first as many studies show expansins being most effective in acidic pH (Li et al., 1993; Li et al, 2003; McQueen-Mason & Cosgrove, 1995, McQueen-Mason et al., 1992). With the reference protein (BSA) binding to the substrates at low pH and some expansins showing wall creep activity at more neutral pH levels, binding was also tested in pH 7.8 HEPES buffer.

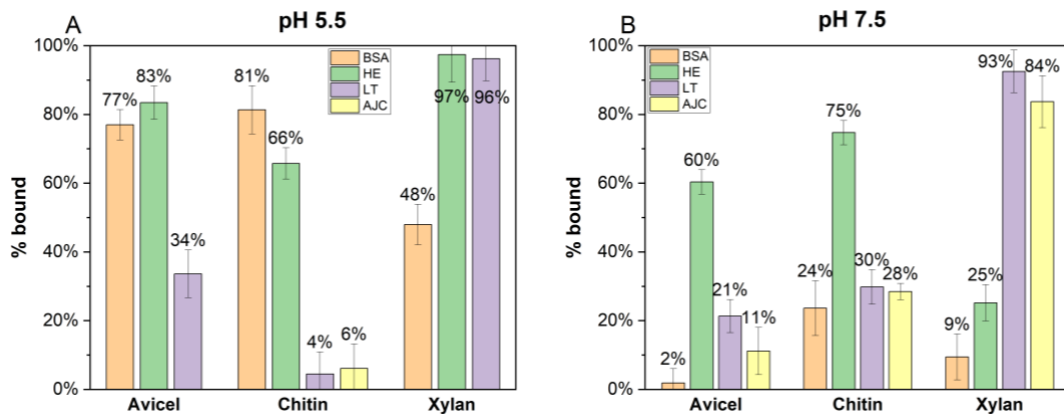


Figure 4.7 | Binding to Commercial Substrates. The graphs depict the percentage of bound protein to commercial substrates calculated from the initial and final unbound protein concentrations. Binding is shown for Avicel, chitin, and oat spelt xylan at pH 5.5 (A) and pH 7.5 (B). AJC165 was not tested for every substrate due to time constraints and protein quantity.

The best binding from all proteins was seen with xylan. LT367 and AJC165 did not show much binding to Avicel and chitin whereas HE673 had significant binding. This is most likely explained by the additional carbohydrate binding domain (CBM2) appended to the expansin from the N terminal side. Interesting to note is that binding to Avicel and xylan decreased with higher pH, but binding to chitin increased with all three proteins HE673, LT367 and AJC165. This could indicate binding to chitin, a positively charged biopolymer, is mediated by electrostatic forces. As pH increases,

the proteins generate more negative charge across the molecule, thus strengthening the electrostatic attraction to positively charged chitin (Georgelis et al., 2015).

After observing binding to commercial substrates binding to more complex carbohydrates were tested. Hardwood and softwood pulp contain varying compositions of celluloses, hemicelluloses, and negligible amounts of lignin. Similar to the commercial substrates, binding to hardwood and softwood were tested at pH 5.5 and pH 7.8. BSA indicated very little binding to hardwood or softwood, as expected. All three proteins showed significant binding to hardwood and softwood at pH 5.5, which trends similarly to commercial substrates previously tested. Additionally, as noted in Georgelis et al. (2011), binding to whole cell walls is mainly mediated by electrostatic forces. This may also explain the better binding at lower pH when the net charge of the proteins are positive. Conversely, LT367 deviates from this trend with higher binding affinity to hardwood at higher pH. This might indicate that LT367 binding to hardwood was not mediated by electrostatic forces but with the conserved aromatic residues in D2 that have been associated with wall loosening activity. However, the exact mechanism for binding affinity requires more advance binding studies and other means of characterizations.

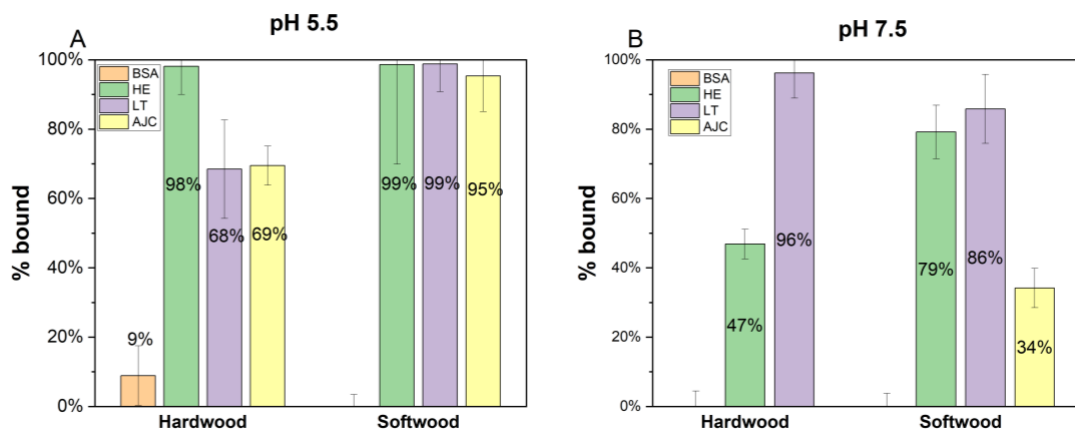


Figure 4.8 | Binding to Pulps. The graphs depict the percentage of bound protein to pulps calculated from the initial and final unbound protein concentrations. Binding is shown for hardwood, and softwood at pH 5.5 (A) and pH 7.8 (B).

4.4.2 Synergisms

The ability of the expansin-like proteins to enhance enzymatic hydrolysis was tested with synergism studies on hardwood substrate. An initial synergy experiment was run

on the samples from the IPP hardwood assay at pH 5.5, effectively adding cellulase to a pretreated sample. This differed from the subsequent experimental setups where protein and enzyme were added simultaneously. This is an important note, as in some cases synergy of bacterial expansins was only seen with pretreated samples (Cosgrove, 2017). After observing a boosting impact of protein compared to cellulase alone, different reaction conditions were examined with cellulase and xylanase.

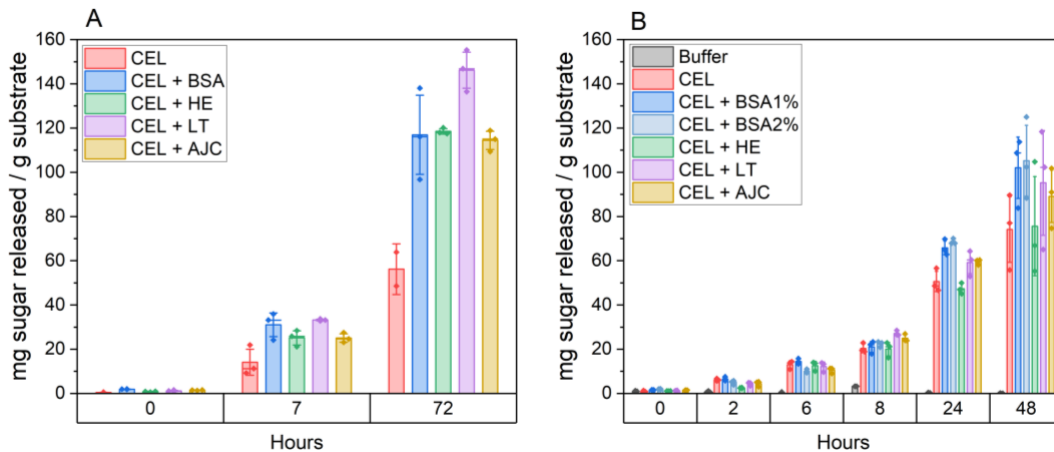


Figure 4.9 | Impact of Bacterial Expansins on Celluclast Hydrolysis of Hardwood Pulp. The graphs depict the total amounts of reduced sugars released per gram of substrate at several different time points. The initial synergy (A) was done with 10 mg/mL hardwood substrate, 2% w/w protein and 1% w/w cellulase (Celluclast, Sigma) at 40°C and 1100 RPM. The first synergy experiment (B) was completed with 10 mg/mL hardwood substrate, 1% w/w protein, 1% w/w cellulase (Celluclast, Sigma) at 35°C and 1100 RPM.

In the first synergy experiment with cellulase, the boosting effects of LT367 and AJC165 compared to cellulase alone can already be observed at the 8 hour mark as indicated in Figure 4.9 B. This impact is still observed at 48 hours but is no longer better than the non-specific boosting observed with BSA. Additionally, large deviations in the triplicates at the 48-hour mark lowers robustness of the analysis. It was noted that the BioTek instrumentation used to measure the absorbance from the PAHBAH assay would yield greater deviations on samples with high absorbance that were placed on the outside edges of the 96-well plate. For subsequent analyses, the 48-hour samples were not placed on the outer wells, reducing the large error seen in triplicates.

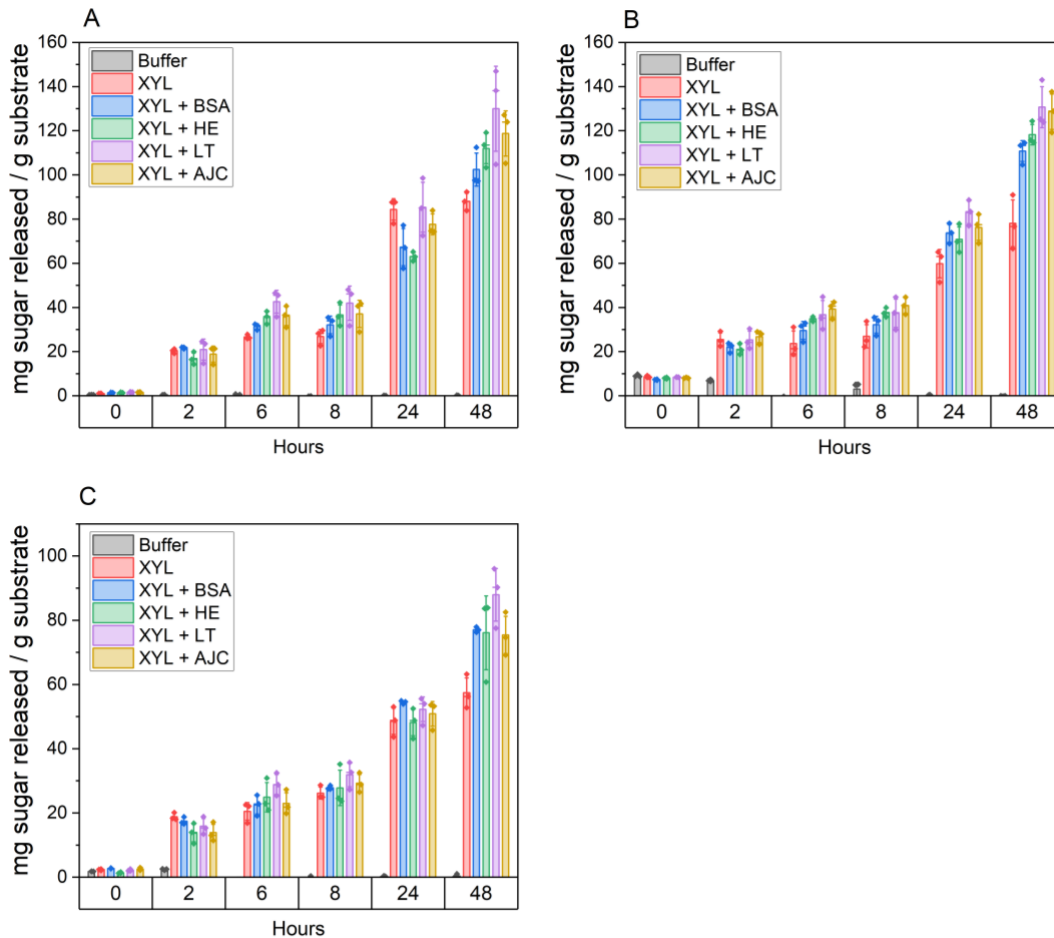


Figure 4.10 | Impact of Bacterial Expansins on Xylanase Ecopulp TX-800 A Hydrolysis of Hardwood Pulp. Three different synergism experiments were run with xylanase Ecopulp TX-800 A (AB enzymes): (A) 10 mg/mL hardwood substrate, 1% w/w protein, 1% w/w xylanase at 35°C and 1100 RPM, (B) 5 mg/mL substrate, 1% w/w protein, 1% w/w xylanase at 35°C and 1100 RPM, and (C) 5 mg/mL substrate, 1% w/w protein, 0.5% w/w xylanase at 35°C and 1100 RPM.

In addition to cellulase, synergism of the expansin-like proteins was also determined with xylanase. In the first xylanase synergy experiment (Figure 4.10 A), boosting was detected clearly for LT367 at the 6 and 8-hour mark. This boosting was not observed at the 24-hour sampling time, but was apparent again at 48 hours. The synergistic impact of LT367 with xylanase released the most sugars at the end of the reaction. A one tailed t-test revealed LT367 was significantly greater than xylanase alone but not significantly greater than the non-specific effects of BSA. In the second experiment with xylanase, the substrate concentration was lowered to 5 mg/mL (Figure 4.10 B). The results were similar to that of the first experiment with smaller deviations between biological triplicates. A one tailed t-test showed LT367 was significantly greater than

xylanase alone or BSA with xylanase. This suggests that the synergism effect seen with LT367 is more than just the non-specific boosting seen with BSA. As synergisms were noted mostly with low enzymatic concentrations, the xylanase concentration was halved for the third trial (Cosgrove, 2017). Again, similar results were seen with boosting effects, most notably with LT367 (Figure 4.10 C). However, unlike the second xylanase trial, a one tailed t-test showed LT367 was not significantly greater than BSA.

These initial synergism studies show possible synergist effects of LT367 that go beyond the non-specific effects of BSA. However, the statistical significance of the results varied between trials, due to the large variation between triplicate samples. It is suggested that sample variation is reduced with more experiments that utilize larger sample sizes to confirm the possible synergistic effects seen with LT367 and possibly AJC165.

5.0 Conclusion

This thesis had three main objectives that were achieved with varying levels of success. The main aims are reiterated below:

1. Further expression optimization, screening and purification of target bacterial expansin-like proteins
2. Functional characterization of target bacterial expansin-like proteins via binding assays using commercial and industrial biomasses
3. Determination of synergistic activities of target bacterial expansin-like proteins on industrial pulps with canonical hydrolytic enzymes

Through expression screening of small-scale production trials with MagicMedia three new potential bacterial-expansins were identified AP099, JAB176 and JAB041. However, large-scale production and purification of said samples yielded zero protein. Large-scale production of the previously screened proteins LT367, CPO928 and SFM691 yielded 6, 1, and 0 mg/L, respectively. Despite two different production conditions tested, the yield for CPO928 was considerably low.

Three proteins HE673, LT367 and AJC165 were successfully characterized with binding assays using commercial substrates and industrial biomasses. All three proteins showed the expansin-like characteristic of binding to xylan, hardwood and softwood pulp substrates. HE673 also showed binding to Avicel and chitin. As mentioned in the Background of Expansins, expansins exhibit two modes of binding: hydrophobic interactions via aromatic residues that are associated with expansin cell wall loosening activity, or alternatively through electrostatic interactions via positively and negatively charged residues that are not correlated to expansin activity. Additional binding and structural studies that elucidate the mechanism of binding can provide more insight into possible expansin-like activity of the tested proteins.

The outcome of the synergistic activities of HE673, LT367 and AJC165 to cellulase and xylanase revealed possible synergism between LT367 and xylanase. Despite low binding to pure cellulose (Avicel). LT367 and to some extent AJC165 also revealed possible synergistic activity with cellulase. Since only one reaction condition was examined with cellulase, further studies with different reaction conditions (different temperatures, pH, concentrations) should be implemented to find the optimal ratio for

synergy between cellulase and expansin concentration. An optimal condition was found in the xylanase trials with low xylanase concentration that revealed a statistically significant difference between LT367 and BSA or xylanase alone, despite large variations in the data.

As this thesis was part of the larger work BioUPGRADE, it is important to keep in mind their objective of developing surface acting proteins for bio-based materials engineering as well. This thesis successfully contributed to achieving this aim. With the observed synergistic effects, LT367 and/or AJC165 show promise as expansin-like proteins to be utilized for materials engineering. The next crucial steps are up-scaling the protein production of the target bacterial expansins using bioreactors, and further characterization of LT367 and AJC165 with additional characterization approaches to elucidate more novel information about the mechanism of binding and confirm the reproducibility of the synergy results observed in this thesis.

References

- [1] Alves, V. S., Pimenta, D. C., Sattlegger, E., & Castilho, B. A. (2004). Biophysical characterization of Gir2, a highly acidic protein of *Saccharomyces cerevisiae* with anomalous electrophoretic behavior. *Biochemical and biophysical research communications*, 314(1), 229–234. <https://doi.org/10.1016/j.bbrc.2003.12.08>
- [2] Belfield, E. J., Ruperti, B., Roberts, J. A., & McQueen-Mason, S. (2005). Changes in expansin activity and gene expression during ethylene-promoted leaflet abscission in *Sambucus nigra*. *Journal of experimental botany*, 56(413), 817–823. <https://doi.org/10.1093/jxb/eri076>
- [3] BioUPGRADE. (2020). *Scope and Targets*. <https://biouprgrade.eu/project-overview/scope-and-targets/>
- [4] Bunternngsook, B., Eurwilaichitr, L., Thamchaipenet, A., & Champreda, V. (2015). Binding characteristics and synergistic effects of bacterial expansins on cellulosic and hemicellulosic substrates. *Bioresource technology*, 176, 129–135. <https://doi.org/10.1016/j.biortech.2014.11.042>
- [5] Chen, F., & Bradford, K. J. (2000). Expression of an expansin is associated with endosperm weakening during tomato seed germination. *Plant physiology*, 124(3), 1265–1274. <https://doi.org/10.1104/pp.124.3.1265>
- [6] Cosgrove D. J. (1996). Plant cell enlargement and the action of expansins. *BioEssays: news and reviews in molecular, cellular and developmental biology*, 18(7), 533–540. <https://doi.org/10.1002/bies.950180704>
- [7] Cosgrove, D. J., Bedinger, P., & Durachko, D. M. (1997). Group I allergens of grass pollen as cell wall-loosening agents. *Proceedings of the National Academy of Sciences of the United States of America*, 94(12), 6559–6564. <https://doi.org/10.1073/pnas.94.12.6559>
- [8] Ding, S., Liu, X., Hakulinen, N., Taherzadeh, M. J., Wang, Y., Wang Y., Qin, X., Wang, X., Yao, B., Lao, H., Tu, T. (2022). Boosting enzymatic degradation of cellulose using a fungal expansin: Structural insight into the pretreatment mechanism. *Bioresource technology*, 358. <https://doi.org/10.1016/j.biortech.2022.127434>
- [9] Dvorak, P., Chrast, L., Nikel, P. I., Fedr, R., Soucek, K., Sedlackova, M., Chaloupkova, R., de Lorenzo, V., Prokop, Z., & Damborsky, J. (2015). Exacerbation of substrate toxicity by IPTG in *Escherichia coli* BL21(DE3) carrying a synthetic metabolic pathway. *Microbial cell factories*, 14, 201. <https://doi.org/10.1186/s12934-015-0393-3>

- [10] Francis, D. M., & Page, R. (2010). Strategies to optimize protein expression in *E. coli*. *Current protocols in protein science, Chapter 5*(1), 5241–5.24.29. <https://doi.org/10.1002/0471140864.ps0524s61>
- [11] Georgelis, N., Nikolaidis, N., & Cosgrove, D. J. (2014). Biochemical analysis of expansin-like proteins from microbes. *Carbohydrate polymers, 100*, 17–23. <https://doi.org/10.1016/j.carbpol.2013.04.094>
- [12] Georgelis, N., Nikolaidis, N., & Cosgrove, D. J. (2015). Bacterial expansins and related proteins from the world of microbes. *Applied microbiology and biotechnology, 99*(9), 3807–3823. <https://doi.org/10.1007/s00253-015-6534-0>
- [13] Georgelis, N., Tabuchi, A., Nikolaidis, N., & Cosgrove, D. J. (2011). Structure-function analysis of the bacterial expansin EXLX1. *The Journal of biological chemistry, 286*(19), 16814–16823. <https://doi.org/10.1074/jbc.M111.225037>
- [14] Georgelis, N., Yennawar, N. H., & Cosgrove, D. J. (2012). Structural basis for entropy-driven cellulose binding by a type-A cellulose-binding module (CBM) and bacterial expansin. *Proceedings of the National Academy of Sciences of the United States of America, 109*(37), 14830–14835. <https://doi.org/10.1073/pnas.1213200109>
- [15] Graceffa, P., Jancsó, A., & Mabuchi, K. (1992). Modification of acidic residues normalizes sodium dodecyl sulfate-polyacrylamide gel electrophoresis of caldesmon and other proteins that migrate anomalously. *Archives of biochemistry and biophysics, 297*(1), 46–51. [https://doi.org/10.1016/0003-9861\(92\)90639-e](https://doi.org/10.1016/0003-9861(92)90639-e)
- [16] Hepler, N. K., & Cosgrove, D. J. (2019). Directed in vitro evolution of bacterial expansin BsEXLX1 for higher cellulose binding and its consequences for plant cell wall-loosening activities. *FEBS letters, 593*(18), 2545–2555. <https://doi.org/10.1002/1873-3468.13528>
- [17] Invitrogen by Life Technologies (2014). Magic Media *E. coli* Expression Medium. Life Technologies Corporation.
- [18] Kende, H., Bradford, K., Brummell, D., Cho, H. T., Cosgrove, D., Fleming, A., Gehring, C., Lee, Y., McQueen-Mason, S., Rose, J., & Voesenek, L. A. (2004). Nomenclature for members of the expansin superfamily of genes and proteins. *Plant molecular biology, 55*(3), 311–314. <https://doi.org/10.1007/s11103-004-0158-6>
- [19] Kerff, F., Amoroso, A., Herman, R., Sauvage, E., Petrella, S., Filée, P., Charlier, P., Joris, B., Tabuchi, A., Nikolaidis, N., & Cosgrove, D. J. (2008). Crystal structure and activity of *Bacillus subtilis* YoaJ (EXLX1), a bacterial expansin that promotes root colonization. *Proceedings of the National*

Academy of Sciences of the United States of America, 105(44), 16876–16881. <https://doi.org/10.1073/pnas.0809382105>

- [20] Kim, E. S., Lee, H. J., Bang, W. G., Choi, I. G., & Kim, K. H. (2009). Functional characterization of a bacterial expansin from *Bacillus subtilis* for enhanced enzymatic hydrolysis of cellulose. *Biotechnology and bioengineering*, 102(5), 1342–1353. <https://doi.org/10.1002/bit.2219>
- [21] Kim, I. J., Ko, H. J., Kim, T. W., Choi, I. G., & Kim, K. H. (2013). Characteristics of the binding of a bacterial expansin (BsEXLX1) to microcrystalline cellulose. *Biotechnology and bioengineering*, 110(2), 401–407. <https://doi.org/10.1002/bit.24719>
- [22] Laine, M. J., Haapalainen, M. L., Wahlroos, T., Kankare, K., Nissinen, R., Kassuwi, S., & Metzler, M. C. (2000). The cellulase encoded by the native plasmid of *Clavibacter michiganensis* ssp *sepedonicus* plays a role in virulence and contains an expansin-like domain. *Physiological and Molecular Plant Pathology*, 57, 221–233. <https://doi.org/10.1006/pmpp.2000.0301>
- [23] Lakhundi, S., Siddiqui, R. & Khan, N.A. Cellulose degradation: a therapeutic strategy in the improved treatment of *Acanthamoeba* infections. *Parasites & Vectors* 8, 23 (2015). <https://doi.org/10.1186/s13071-015-0642-7>
- [24] Lee, H. J., Kim, I. J., Kim, J. F., Choi, I. G., & Kim, K. H. (2013). An expansin from the marine bacterium *Hahella chejuensis* acts synergistically with xylanase and enhances xylan hydrolysis. *Bioresource technology*, 149, 516–519. <https://doi.org/10.1016/j.biortech.2013.09.086>
- [25] Lee, H. J., Lee, S., Ko, H. J., Kim, K. H., & Choi, I. G. (2010). An expansin-like protein from *Hahella chejuensis* binds cellulose and enhances cellulase activity. *Molecules and cells*, 29(4), 379–385. <https://doi.org/10.1007/s10059-010-0033-z>
- [26] Li, L. C., Bedinger, P. A., Volk, C., Jones, A. D., & Cosgrove, D. J. (2003). Purification and characterization of four beta-expansins (*Zea m 1* isoforms) from maize pollen. *Plant physiology*, 132(4), 2073–2085. <https://doi.org/10.1104/pp.103.020024>
- [27] Li, Y., Darley, C. P., Ongaro, V., Fleming, A., Schipper, O., Baldauf, S. L., & McQueen-Mason, S. J. (2002). Plant expansins are a complex multigene family with an ancient evolutionary origin. *Plant physiology*, 128(3), 854–864. <https://doi.org/10.1104/pp.010658>
- [28] Li, Z. C., Durachko, D. M., & Cosgrove, D. J. (1993). An oat coleoptile wall protein that induces wall extension *in vitro* and that is antigenically related to a similar protein from cucumber hypocotyls. *Planta*, 191(3), 349–356. <https://doi.org/10.1007/BF00195692>

- [29] Ling, S., Chen, W., Fan, Y., Zheng, K., Jin, K., Yu, H., Buehler, M. J., & Kaplan, D. L. (2018). Biopolymer nanofibrils: structure, modeling, preparation, and applications. *Progress in polymer science*, 85, 1–56. <https://doi.org/10.1016/j.progpolymsci.2018.06.004>
- [30] Lynd, L. R., Weimer, P. J., van Zyl, W. H., & Pretorius, I. S. (2002). Microbial cellulose utilization: fundamentals and biotechnology. *Microbiology and molecular biology reviews: MMBR*, 66(3), 506–577. <https://doi.org/10.1128/MMBR.66.3.506-577.2002>
- [31] McQueen-Mason, S. J., & Cosgrove, D. J. (1995). Expansin mode of action on cell walls. Analysis of wall hydrolysis, stress relaxation, and binding. *Plant physiology*, 107(1), 87–100. <https://doi.org/10.1104/pp.107.1.87>
- [32] McQueen-Mason, S., & Cosgrove, D. J. (1994). Disruption of hydrogen bonding between plant cell wall polymers by proteins that induce wall extension. *Proceedings of the National Academy of Sciences of the United States of America*, 91(14), 6574–6578. <https://doi.org/10.1073/pnas.91.14.6574>
- [33] McQueen-Mason, S., Durachko, D. M., & Cosgrove, D. J. (1992). Two endogenous proteins that induce cell wall extension in plants. *The Plant cell*, 4(11), 1425–1433. <https://doi.org/10.1105/tpc.4.11.1425>
- [34] Nikolaidis, N., Doran, N., & Cosgrove, D. J. (2014). Plant expansins in bacteria and fungi: evolution by horizontal gene transfer and independent domain fusion. *Molecular biology and evolution*, 31(2), 376–386. <https://doi.org/10.1093/molbev/mst206>
- [35] Olarte-Lozano, M., Mendoza-Núñez, M. A., Pastor, N., Segovia, L., Folch-Mallol, J., & Martínez-Anaya, C. (2014). PcEx1 a novel acid expansin-like protein from the plant pathogen *Pectobacterium carotovorum*, binds cell walls differently to BsEXLX1. *PloS one*, 9(4), e95638. <https://doi.org/10.1371/journal.pone.0095>
- [36] Park, Y. B., & Cosgrove, D. J. (2012). A revised architecture of primary cell walls based on biomechanical changes induced by substrate-specific endoglucanases. *Plant physiology*, 158(4), 1933–1943. <https://doi.org/10.1104/pp.111.192880>
- [37] Raud, M., Kikas, T., Sippula, O., & Shurpali, N. J. (2019). Potentials and challenges in lignocellulosic biofuel production technology. *Renewable and Sustainable Energy*, 111, 44-56. <https://doi.org/10.1016/j.rser.2019.05.020>
- [38] Rongpipi, S., Ye, D., Gomez, E. D., & Gomez, E. W. (2019). Progress and Opportunities in the Characterization of Cellulose - An Important Regulator

- of Cell Wall Growth and Mechanics. *Frontiers in plant science*, 9, 1894. <https://doi.org/10.3389/fpls.2018.01894>
- [39] Ruesink, A. W. (1969). Polysaccharidases and the Control of Cell Wall Elongation. *Planta*, 89(2), 95–107. <http://www.jstor.org/stable/23367888>
- [40] Sampedro, J., & Cosgrove, D. J. (2005). The expansin superfamily. *Genome biology*, 6(12), 242. <https://doi.org/10.1186/gb-2005-6-12-242>
- [41] Schwarz W. H. (2001). The cellulosome and cellulose degradation by anaerobic bacteria. *Applied microbiology and biotechnology*, 56(5-6), 634–649. <https://doi.org/10.1007/s002530100710>
- [42] ThermoFisher Scientific. (n.d.). *Traditional Methods of Cell Lysis for Protein Extraction*. <https://www.thermofisher.com/fi/en/home/life-science/protein-biology/protein-biology-learning-center/protein-biology-resource-library/pierce-protein-methods/traditional-methods-cell-lysis.html>
- [43] Wang, T., Park, Y. B., Caporini, M. A., Rosay, M., Zhong, L., Cosgrove, D. J., & Hong, M. (2013). Sensitivity-enhanced solid-state NMR detection of expansin's target in plant cell walls. *Proceedings of the National Academy of Sciences of the United States of America*, 110(41), 16444–16449. <https://doi.org/10.1073/pnas.1316290110>
- [44] Wong, H. E., Huang, C. J., & Zhang, Z. (2018). Amino acid misincorporation in recombinant proteins. *Biotechnology advances*, 36(1), 168–181. <https://doi.org/10.1016/j.biotechadv.2017.10.006>
- [45] Wurm, D. J., Slouka, C., Bosilj, T., Herwig, C., & Spadiut, O. (2016). How to trigger periplasmic release in recombinant Escherichia coli: A comparative analysis. *Engineering in life sciences*, 17(2), 215–222. <https://doi.org/10.1002/elsc.201600168>
- [46] Yennawar, N. H., Li, L. C., Dudzinski, D. M., Tabuchi, A., & Cosgrove, D. J. (2006). Crystal structure and activities of EXPB1 (Zea m 1), a beta-expansin and group-1 pollen allergen from maize. *Proceedings of the National Academy of Sciences of the United States of America*, 103(40), 14664–14671. <https://doi.org/10.1073/pnas.06059791>
- [47] Yuan, S., Wu, Y., & Cosgrove, D. J. (2001). A fungal endoglucanase with plant cell wall extension activity. *Plant physiology*, 127(1), 324–333. <https://doi.org/10.1104/pp.127.1.324>
- [48] Zhang, P., Su, R., Duan, Y., Cui, M., Huang, R., Qi, W., He, Z., & Thielemans, W. (2021). Synergy between endo/exo-glucanases and expansin enhances enzyme adsorption and cellulose conversion. *Carbohydrate polymers*, 253, 117287. <https://doi.org/10.1016/j.carbpol.2020.117287>

A. Equipment & Chemicals

Item	Use	Product Name	Manufacturer
ÄKTA	Purification	ÄKTA Purifier 10 FPLC System	Amersham Biosciences
Bradford reagent	IPP Assay	Bio-Rad Protein Assay Dye Reagent Concentrate	Biorad
Breathable Tape	Production	Nunc Sealing Tape, White Rayon, Breathable, Sterile	ThermoFisher Scientific
BSA	Reference protein	Bovine serum Albumin	Sigma-Aldrich
Cellulase	PAHBAH assay	Cellulase from <i>Trichoderma reesei</i> , Celluclast	Sigma
Centrifuge (Large)	Production	SORVALL LYNX 4000	ThermoFisher Scientific
Column (ÄKTA)	Purification	HisTrap™ FF crude (5 mL)	Cytiva
Column (manual)	Purification	His GraviTrap TALON column	Cytiva
EDTA	Lysis buffer	Ethylenediaminetetraacetic acid disodium salt dihydrate	Sigma-Aldrich
Gel imager	SDS-PAGE	ChemiDox XRS+ System	Biorad
Gel power source	SDS-PAGE	PowerPac 300	Biorad
Incubator (Large)	Production	innova 44 incubator shaker series	New brunswick scientific
Incubator (Tabletop)	Production	Certomat R	B.Braun
MagicMedia	Production	MagicMedia E. coli Expression Medium	ThermoFisher Scientific
MgCl ₂	Lysis buffer	Magnesium chloride hexahydrate	Supelco

NanoDrop	Purification	NanoDrop Lite spectrophotometer	ThermoFisher Scientific
OD cuvettes	Production	Semi-Micro Cuvette, PMMA	BRAND
OD machine	Production	Biophotometer plus	Eppendorf
Plate reader	IPP Assay, Synergism Assay	EON	BioTek
Protein ladder	SDS-PAGE	Precision Plus Protein™ Dual Color Standards	Biorad
SDS gel	SDS-PAGE	Mini-PROTEAN TGX	Biorad
SDS-PAGE stain	SDS-PAGE	PageBlue Protein Staining Solution	ThermoFisher Scientific
sonicator	Lysis Production, IPP	Sonicator Q500	QSONICA
Thermoshaker	and Synergism Assays	Thermomixer C	Eppendorf
TRIS	Lysis buffer	Trizma base	Sigma-Aldrich
VivaSpin	Purification	Vivaspin® 20, 10000MWCO	Sartorius
Xylanase	Synergism Assay	Ecopulp TX-800 A (from non-pathogen fungi)	AB Enzymes

B. DNA and Gene Accession Numbers

Sequences are for the expressed protein.

1. AP099

DNA Accession Number: AP024099.1

GCGCCGGCCAACACCTGGGGCAGCACGTTACCGGCATTGCGACCGCAACCGGC
TCGGGCTATTCGGGCGGTGCTTTCTGCTCGATCCGATACCAAGGACCGGGAAA
TCACCGCGCTCAACCCGGCGCAGGCAAACCTCGGCGGCATCCCCGCGGCAATGG
CCGGGGCCTATCTGCGCGTGCAGGGCCCCAAGGGCTGTACCACGGTGTACGTGA
CCGATCTTACCCCGAAGCAGCATCGGGCGGTCTGGATCTTTCATACAACGCCTTC
GCCAAGATCGGGCAGCTGCAGCAGGGACGGATTCCGGTCCAGTGGAGGCTGATC
CCGGGCCCGGTCACCGGCAACGTCGTCTACCGCATCAAAGAGGGCAGCACGATGT
GGTGGGCCGCGATCCAGGTGCGCAATCACACCTTCCCGTGGTGAAGCTGGAAGT
CTTCCAGGGCAAGGCCTGGGTGAGCCTGCCGAAAGCGGACTACAACCACTTTGTC
GGCACGCAGCTCGGCGACAAGCCCCTGGTTATCCGGATCACCGACATCCGAGGGC
GGATTCTCGTCGACAAGCTGCCCCGCTGCTCAAGGATTGCACACCCAAGCAGGC
GGGCGAGGCATCGCCGTGCAGCAAGCCCTATTTTGTCCAGGGGAAGGTGCAGTTT
TCCGAGGCCCGCCGCCATCATCATCATCATTGA

Gene Accession Number: BCL92919.1

APANTWGSTFTGIATATGSGYSGGAFLLDPITKDREITALNPAQANLGGIPAAM
AGAYLRVQGPKGCTTVYVTDLYPEAASGGLDLSYNAFAKIGDLQQGRIPVQWR
LIPGPVTGNVVYRIKEGSTMWWAAIQVRNHTFPVKLEVFQGWVSLPKADY
NHFVGTQLGDKPLVIRITDIRGRILVDKLPPLLKDCTPKQAGEASPCSKPYFVQG
KVQFSEAAHHHHHH

2. JAB176

DNA Accession Number: JABDRL010000176.1

GCCGTCGCGGACACCGACTCGCACTCCGGCCGGGCGACCTTCTACTCCCT
GGGCGACAGCGGGCTGGGCAACTGCTCGTTCCCGAGGTTCCCGCCGAT
CGCTTCTACGCGGCGCTCGGGCCGGAGGACTACGACCATGCCGGGGGCT
GCGGTGGCTATCTCGAGGTGACCGGTCCGCTGGGTACCGTTCGCGTGCTG
GCCGCCGACAAGTGCCCGGAATGCGAGTCCGGGCATCTCGACATGAGCC
CAGAGGCCTTCGCGATGATCGGACCGCTGGCCGACGGCGTTATCCCGATC
ACATACCGGGGCGTGCGGGATCCGGACGGAGTCCGCCCCGTCGGGGTCC
GGGTCAAGGAGGGCTCCTCGGCATGGTGGATCGGGTTTTTGATCATGGAT
CACGGCAACCCGCTGGCCTCCGTGGAGTACCTCGCGGACGCTGACGCGG
ACGAGTGGCGCGGCCTGGCGCGTGCCGAGTACAACACTGGCTCCAGGA
GGATGGCGCCGGCCCGGGGCCGTTCACTGCGCCTGACCGACATCTAC
GGTCAACAGGCCGTGTCGAGGGCATCGAACTGTCGCCCATGGTCGATCA
GGTGACAAATGTGCGGATGTACGAGACCGACGCATCCTCGCCGACTGGAT
CCGCGTCTCCGCGTCTCCGCGTCTCCGCGTCTCTCCGGCCCCGACG
GCGAGCGTCCCGGGCCCCGCCGAATCGGGGGCCTGTGCCGCGACCGCC
GCCGTCAACACGATGTGGCCCACCGGGTTCCAGGCGACCGTCACCGTGC
GGAACACCGCCGCGGCGGGCCGGGTCGTGGGTGTCACCTGGATCGT
GCCGGCGGACGTACCATCGGCACGGCATGGAACGCCACCGTCGGGCAG
AGCGGCACCGCCGTGACCGCCAGCGTCCGGTCTGGAGTCCACCCCTCG
CGGTAGGAGAGACCGTCTCGATCGGATTCAACGCCGACCGACCTATCGGC
CCGGGAACGGTCGAGGCCGCGGTGGCTGGCTCGATCACCTCAACGGGG
AGTCATGCGTCGGTTTTGCGCGCCGCCATCATCATCATCATTGA

Gene Accession Number: NNJ61190.1

AVADTDSHSGRATFYSLGDSGLGNCSFPEVPADRFYAALGPEDYDHAGGCG
GYLEVTGPLGTVRVLAAADKCPECESGHLDMSPFAFAMIGPLADGVIPITYRGVR
DPDGVGPVVRVKEGSSAWWIGFLIMDHGNPLASVEYLADADADEWRGLARA
EYNYWLQEDGAGPGPFTLRLTDIYGHQAVVEGIELSPMVDQVTNVRMYETDA
SSPTGSASSASSASSASSPAPTASVPGPAESGACAATAAVNTMWPTGFQATV
TVRNTAAAAAGSWVVTWIVPADVTIGTAWNATVGQSGTAVTASAPVWSPTLA
VGETVSIGFNADRPVGPVVEAAVAGSITLNGESCVGFAAAHHHHHH

3. JAB041

DNA Accession Number: JABDRL01000041.1

GCCGGGCCGCTCGCCGCGGGAGAGTCCCGCAGCGGCAAGGCGACCCATT
ACGACGCCACCGGCGGGCAACTGTTCCCTTCGAGGCCTCGGGGCCTGG
GCTCTACGTCGCGCTGTCGCCAGCGAGTACGCGGCCGGTGCGGCCTGC
GGCGGTACCTTGACGTCACGAATGCCCGGGGCACGGTGCGGGTCAAGG
TCACCGATCAGTGCCCGGGGTGCCAGCTGGCCATATCGACCTGAGCAAG
GAAGCGTTCGCCAGGCTGGGAGCGCTGGCCACCGGCGAACTCGCGGTCA
CCTACCGGACGGTGGTGAATCCGGCCCTGCCCGTCCGATACCGTACGG
GTCAAGGATGGGGCATCCCGCTGGTGGTTTGCCGCGCGGATCGATAATCA
CGGCAATCCGCTGTCCCGGTAGAGGTCAAGACCGCCTCCGGATGGGCC
GCGCTCGATCGGACCGACTACAATACTGGCTCAAGCCGGACGGCGCCG
GGCCCGGGCCGTTACGATTCGCATCGTTGACTCCCGGGGCCACACCGC
GGCGGCAACGGGGGTTTCCCTGGCCCCGAACAGGTTACGCGCACCTCC
GTGCGGATGTACGCGACATCAATTCCGTCGAGCGATGTCCCCAGGAGTCC
GTCCCCCGCGAGATCAGATCTCCCGCCAGCGAGGCCTCGTCGCCGGAG
CCGGGAGGGGGCGGTGCGTTGCCGTCTTCGGCGATGCCGTCCGCTACCC
GGCCTCGTTGCGCCGCCGCCATCATCATCATCATTGA

Gene Accession Number: NNJ59718.1

AGPLAAGESRSGKATHYDATGGGNCSFEASGPGLYVALSPSEYAAGAACGG
YLDVTNARGTVRVKVDQCPGCPAGHIDL SKEAFARLGALATGELAVTYRTVV
NPALPGPITVRVKDGASRWWFAARIDNHGNPLSRVEVKTASGWAALDRTDYN
YWLKPDGAGPGPFTIRIVDSRGHTAAATGVSLAPEQVQRTSVRMYATSIPSSD
VPRSPSPARSDLPPSEASSPEPGGGGALPSSAMPSATRPRCAAHHHHHH

4. VFW002

DNA Accession Number: VFWZ01000002.1

CAATGTAATGACAGTATAAAATCCGGAGAAGCTACTTTTTATGGTGGTGTGC
CAGGAAGCGAAGGAGGTAATTGTTTATTACCAGTAGCAGTAAATGATTTTAT
GCATTGCGCATTGAATAATACTGACTACGATGAAAGTAATGCCTGTGGCGC
TTGCCTTGAGGTAACGGGTTCTCGAGGCAAACTATCGTACAGGTAGTGGA
TCGTTGTCCTGAATGTAAACCTGGTGATGTAGATCTAACACAACAAGCATT
GCTAAGGTTGCGAATCCGATAGATGGGAGAGTACCCATTACTTGGAATTT
GTACCATGTCCATTGGCAGCGAATAAAACAATTAAGGTCAATTTTAAATCAG
GTTCTTCAAATATTGGACTGCAATACAATTTTCGAGATCTCAAACATGCAGT
ATCAAAAATGGAATATCGAAAAATAATACATGGATAAATGTAGAACGAAAA
TTATTTAATTTTTTATTGAACCAACGGGTATCGAATCTCCAATGCAGCTACG
AATTACTTCTGTAGTAGGAGAAGAATTAATTCTTAATGATATTGTCATAAATA
CAAGCGCAGATTTTACTCCAAAGTTCAATTTAGTACACCACCCAATTGCGA
AGAAGAAGGATCGATAACCGAAAATCAATCTTTTAAAAATCATATAATCCCT

GGTCGTGTAGAAGCAGAAGATTTTGATAAAGGAGGGCAAGGTATAGCCTAC
AAGGATAGTGATACAAATAATAGAGGAGGTCAATATCGAAACGAAGGAGTT
GATATTCAAACACACAAGATACTAACGGGCTTTTTAATGTGGGGTGGACC
AATGCAGGTGAATGGTTAACATATACTGTAGAAGTTAATAAGTCTGGAATAT
ATGATTTTATATTGAGAACAGCTTCTATTTTTAATACATCATCTCTACAGTTA
CTAGTGGATGGAGTAACTGTCTCAGAAGAGATTATGCTGCCGAATACTGGG
AACTGGCAACAATACTGATGTTGTAATCCCTAATGTTTCTTTATCTCAAGG
AAAACACAAAATTAGGTTGAATGTGATAAAACCAGGAATGAATTTGAATTAC
TGGTCAGCTTCCTTAGCTAACTCAAACAGTAGTAAAGTTAATTCTAATCAA
CAAGCACTGTAACACAGGAAACGAATTCAAAGTATATCCAAATCCAACCTC
TTCTATAGTTACTTTACAAGGAAATACAAGTCATTGGAGACTACAAAATATAA
AGGGAGAAATCTTTCCGAAGGAAAAGGAAATACAATAGACCTATCACTTTA
CCCTAAAGGATATTATATTATAAAGATAGATGATAGAACAGTCAAAGTAGTA
AAAGAAGCCGCCGCCCATCATCATCATCATTGA

Gene Accession Number: TPN87812.1

QCNSIKSGEATFYGGVAGSEGGNCSLPVAVNDFMHCALNNTDYDESNACGA
CLEVTGSRGKTIVQVDRCPCKPGDVLTLQQAFKVANPIDGRVPITWKFP
CPLAANKTIKVNFKSGSSKYWTAIQFRDLKHAVSKMEYRKNNTWINVERKLFNF
FIEPTGIESPMQLRITSVVGEELINDIVINTSADFDSKVQFSTPPNCEEEGSITEN
QSFKNHIIPGRVEAEDFDKGGQGIAYKSDTNNRGGQYRNEGVDIQTQDTN
GLFNVGWTNAGEWLTYTVEVNKSGIYDFILRTASIFNTSSLQLLVDGVTVSEEIM
LPNTGNWQQYTDVVIPNVLSQGKHKIRLNVKPGMNLNYWSASLANSNSSKV
NSNQTSTVTTGNEFKVYPNPTSSIVTLQGNTSHWRLQNIKGEILSEGKNTIDL
SLYPKGYIYIIDDRTVKVVKEAAAHHHHHH

5. SMS108

DNA Accession Number: SMS01108

GCTGACGACTCGATTACACACAGGAGAAGGTACATTCTACGGTTATGGCGGT
GGTGGAAACTGTAGCTTTCCGCTCCCGGATGACAGTATTTACACCGCAGCA
ATGAATGCGACAGACTATAATAATTCGGCTGCTTGTGGTGCTGTTATTGAG
GTAACAAATACTAAAACCAATCAGACCGTCACCGTGCGCATCGATGATCAA
TGCCCCGAATGTGCCAAAGGCGATGTCGATCTGGATCAAGATGCCTTCGC
GGAAATTTCACTACTCGAAGCCGGGCGGATTCCAATTTCTTGAAATATGT
AGCCAATGAGCAGGCCGGTAATATGAAGCTCTTTTTCAAAGAAGGTTCCAG
CCAGTGGTGGACTGCAGTTCAGGCTCGGGATCACCGCTATCCAATCACCG
CCATTGAATACCGTGTTTCCGGCTCCGGTAACAGTTACGAACTTAGAGC
GTAAGCCATATAACTATTTTGAAAAGCTGACGGATTCGGGGTTGGTCCTTA
TGATTTCCGTATCACCGATTTTTGGGGACAACTGTAGAAGTGCTGTCAGT
GCCTCTTATTCTGACAAATGAAATTGATACCGCGACTCAGTTCCCGGTTTCAT
CAGGGTGACGGATCTTCTGGTGATACCGGAGGCTCAGGCGATGATGGCTC
GACACCACCTGATGATAACAATACCCCAACCTCTGACAACATTGCGGTTTCT
GAAGCAACGACAGTTCATGGAACGATGGCTACTGTGAAACCGTTACTGTA
CCAATAACAATGACCACGGAGTTGTCTGGGCAGTCACACTGGATATTACTG
GTAGCGTATACAATCTGTGGGACGGACAATGGTCTCAGTCTGGTAAACTT
TATCTGTCACCGGAGCTTCATGGAATGACTCGCTGTCAGCCGGAGCAAGTA
CAACGTTCCGGCTTCTGTGCTAATCGTGCCGCCGCCCATCATCATCATC
ATTGA

Gene Accession Number: WP_087481129.1

ADDSIHTGEGTFYGYGGGGNCSFPLPDDSIYTAAMNATDYNNSAACGAVIEVT
NTKTNQTVTVRIDDQCPECAKGDVLDQDAFAEISPLEAGRIPISWKYVANEQA

GNMKLFFKEGSSQWWTAVQARDHRYPIAIEYRVSGSGNSYANLERKPYNYF
VKADGFGVGPYDFRITDFWQTVSVPLILNEIDTATQFPVHQDGGSSGD
TGGSGDDGSTPPDDNNTPTSDNIAVSEATDSSWNDGYCETVTVTNNDHGVV
WAVTLDITGSVYNLWDGQWSQSGKLSVTGASWNSLSAGASTTFGFCANR
AAAHHHHHH

6. CPO928

DNA Accession Number: CP016928.1

CCTGGGAACCTTGGCGATATCTGCTACGGCTATGCAACAGCCACAGGTTCCGG
GTTATTCCGGCGGGCGCTTTATTGCTGGATCCGATCCCTGCCAATATGGAGA
TAACCGCGTTAAACAGAACGCAACTGGACTATAAAGGCGTTAAAGCGGCGC
TCGCCGCGCTTATCTCAAAGTCACCGGGCCTAAGGGAAGCACAATTGTTT
ATGTTACTGATTTATATCCTGAAGGTGGCGATTGTGCGCTGGACTTGTCATT
CAATGCGTTCGAGAAAATTGGCAATTTACAGGATGGAAAAATTAATATCCAA
TGGGAGCTGGTAAAAGCGCCGGTAAACGGTAATGTGGTATACAGAGTTAAA
GAGGGGTCTAATCCTTATTGGGCGGCAGTGCAATTCAGAAACGTCAAATAT
CCGATTATTGAGATGAAGTATCTGCGAGGTAGTCAATGGGTCAGCGCGCCA
AAAACAGATTACAATCACTTTATTCTGGAATTCGTTGGAAAAAATGATATCCC
GATCGAGTTTACCGATATCAAAGGCAATATATTAAGCGATACGCTCCCGCCA
ATGTCAGACAGCACGTCATCCGCCTATTTGATCACCGGAAAAGTACAACCTG
CCTGCCGCCGCCATCATCATCATCA TCATTGA

Gene Accession Number: AUC40652.1

AWELGDICYGYATATGSGYSGGALLLDPIPANMEITALNRTQLDYKGVKAALAG
AYLKVTGPKGSTIVYVTDLYPEGGDCALDLSFNAFEKIGNLQDGKINIQWELVKA
PVNGNVVYRVKEGSPYWAAVQFRNVKYPIIEMKYLRGSQWVSAPKTDYNHFI
LEFVGKNDIPIEFTDIKGNILSDTLPMSDSTSSAYLITGKVQLPAAAHHHHHH

7. SFM691

DNA Accession Number: FOUV0100008.1

GCGTACGCCAGTGTCCAGCCCGGCGTCACCTACTCCGGCGAAGGCACCTT
CTACGGCGCGACCGGCGTCCGGAACTGCCTCTACGACGCCACGAGCGACA
TCGCCATAGCGGCCCTCAACCACACCGACTACGACAACGCCCGTATGTGC
GGTGCCTTCATCCGGGTCAAGGGCCCGCGCGGCGAGCTGACGGTGAGGA
TCGTCGACCGCTGTCCGGAGTGCCGCCCGGCGACGTCGACCTCGGACAG
CAGGCCTTCGCCCGGATCGCCGACCCCGTGGCCGGCCGGGTGCCATCA
CCTGGACGCTGGTGAGCCCGGACCTCGCCGGACCCGTCTCCTACCGTTAC
AAGGAAGGCTCCACCCAGTGGTGGTGCGGCATCCAGGTCCGCAACCACCG
CAACCCCGTCGCCACCCTGGAGGTCCGCACCGGCACCACCTGGCGGCAAC
TCCCGCGCCAGGAGTACAACACTTTCGTCTCGGCGGACGGCGCGGGCTGC
GGCTCCGACATCCGCGTCAAGGACATCCACGGCCAGACCCTCACCGACAC
CGGCATCGCCCTCACCCCGAACGTGACCCAGCCGGGCCGCGCCAGTTCA
CCAAGCGCGCCGCCGCCATCATCATCATCATTGA

Gene Accession Number: SFM96691.1

AYASVQPGVTYSGEGTFYGATGVGNCLYDATSDIAIALNHTDYDNARMCGAFI
RVKGPARGELTVRIVDRCPICRPDGDVLDGQQAFARIADPVAGRVPITWTLVSPD
LAGPVSRYRYKEGSTQWWCGIQVRNHRNPVATLEVRTGTTWRQLPRQEYNYFV
SADGAGCGSDIRVKDIHGQTLTDTGIALTPNVDQPGRQFTKRAAAHHHHHHH

8. **LT367**

DNA Accession Number: LT615367

GCAGCCTGGCAACTCGGTGATATTTGCTACGGCTATGCAACAGCAACAGGC
TCAGGATATTCAGGAGGTGCTGTATTGCTGGACCCGATTCCAGCCGATATG
GAAATCACAGCACTGAACCCCGCACAGCTCAACTATAAAGGCGTTAAAGCT
GCATTGGCCGGTGCTTATCTGAAAGTCACCGGCCCTAAAGGCAGCACGATT
GTTTATGTCACGGACCTTTATCCTGAAGGAGGCAATTGCGCACTGGACTTGT
CATTTAATGCATTCAAGAAAATTGGCAATCTGCAAGATGGCAAATCAATATT
CAGTGGGAATTGGTTAAGGGGCCTGTTACCGGGAATGTTGTCTACCGCGTT
AAAGAGGGTTCTAACCATACTGGGCCGCGAGTACAGATTAGAAATGTTAAGT
ACCCTATTCTCTCACTGAGATATTTACGTGATGGGAAATGGATAAGCCCGCC
GAAAACAGATTACAACCACCTTTATTCTGGAGAATTTAGGTAAAACTATACAC
CTATCGAATTCACCGACGTGAAAGGCAATATATTAAGCGATACGTTGCCGCC
AATGCCTGATAACACATCCTCTGCCTATCTGATTACTGGCAAGGTGCAGTTA
CCGGCCGCCGCCATCATCATCATCATTGA

Gene Accession Number: SLM63089.1

AAWQLGDICYGYATATGSGYSGGAVLLDPIPADMEITALNPAQLNYKGVKAALA
GAYLKVTPKPGSTIVYVTDLYPEGGNCALDLSFNAFKKIGNLQDGKINIQWELVK
GPVTGNVVYRVKEGSNPYWAQVQIRNVKYPILSLRYLRDQKWKISPPKTDYNHFI
LENLGNKYTPIEFTDVKGNILSDTLPPMPD NTSSAYLITGKVQLPAAAHHHHHH

9. **AJC165**

DNA Accession Number: AJC47165

GCCGCCTGGAACGATGTCTGCAGCGGCACCGCCACCTATAACCAGCTCCGG
CTATTCCGGCGGCGCGCTGCTGCTGGATCCCATCGATCCCAATGCGACGAT
CACCGCGTTGAACCCCAAGCAGCTTAATTACGGCGGAATCCAGGCGGCGC
TGGCCGCGCCTACCTGCAGGTGCAGGGGCCGCGCGGCATGGTCACCGT
GTATGTCACCGACCTGTATCCGGAAGGCGCCGACTGCGGGCTGGACCTGT
CGCCCAATGCCTTCGCCGCGATCGGCGACACCAGCGCCGGGCGCATCCC
GATCCGCTGGCAGGTGGTGGCCGCGCCGGTCACCGGCAACGTGGTGTAC
CGGATCAAGGAGGGCAGCTCGCAATACTGGGCGGCGATCCAGGTGCGCAA
CCACCGCTATCCGGTGGTGAAGTTCGAATACAAGAAGAACGGCGACTGGGT
CAGCCTGCCCAAGACCGCGTACAACCATTTCTCGGCGAGCAGATGGGCG
CGCAGCTGCTGGAGATCCGCCTGACCGACATTCTGGCCAGGTGGTGACC
GACACGCTCGGCGCGCTGCCAAGCCAGGGCGACAAGGGCGTGTACTTCGC
CGATGGGCATGTGCAGTTCCCGCGCGCCGCCGCCATCATCATCATCA
TTGA

Gene Accession Number: WP_043094747.1

AAWNDVCSGTATYTSSGYSGGALLLDPIDPNATITALNPKQLNYGGIQAALAGA
YLQVQGPRGMVTVYVTDLYPEGADCGLDLSPNAFAAIGDTSAGRIPRWQVVA
APVTGNVVYRIKEGSSQYWAAIQVRNHRYPVVKFEYKKNGDWVSLPKTAYNHF
VGEQMGAQLLEIRLDIRGQVVDLTLGALPSQGDQKGVYFADGHVQFPRAAAHH
HHHH

10. **HE673**

DNA Accession Number: HE614873.1

TCGCAGCCAGGTCCGGCGAAGCCAGGCGCGATTTACGCGAAGTGGCAGC
CGGGCGGGTCTGGGCATCCGGCTACGTGGCGAACCTCGACGTCACCGC

GACAGACGCTGTCACGGGATGGACCGTCTCATGGGCCAGCCCGGAAACCA
CCGGCGTCGTCAACAGTTGGGGCATGCGCTGTAGCGTCGCCTCCAACACC
GTGACCTGCACCGGCACGGACTGGGCAAGCGAGCTCGCCGCCGGACAGA
CCGTTAGCGTCGGCGTGCAGTTGGCGGGTGGACCCGCTCCCTCTTCGCT
CAAATCAGTGCCACAGCGGCCGGAACGCCGCCGTCCCAGCCCACGCCGC
CGTCCCAGCCCACGCCGCCGTCCCAGCCCACGCCGCCGTCCCAGTCGGC
GACGCACGGCCGTGCCACGCACTACTCGCTCGGTACCGGCAACACGATCG
CGAATGGCAACTGCTCCATGCCGGCTGTCCCCGCCGACCGCATGTACGTT
GCGGTCAGCAGCCCCGAGTACAGCGGAGCCGCCGCGTGGCGTTTCGTACC
TCCTCGTCACGGGCCCAAGGGCACCGTCCGCGTCCAAATCGTCGACCAA
TGCCACGAGTGCGAGATCGGACATCTCGACTTGAGCGAGGAAGCGTTCCG
TGCCGTCGGCGACTTCGATGCCGGCGTCATCCCGATCAGCTACACCACCG
TCCGGGATCCGGACGTGCCCGATGTCCCGTCCGGGTCAAGGAGGGCTC
ATCCCGATGGTGGGCCGGTCTGCAGATTCTGAACGCCGGCAACCCCATCG
ACCATGTTGAAGTCCAAGCCGATGGACAGTGGCTGGGCCTGAGCCGCACG
ACTACGGGTACTGGGTGACACCTTACCTATCCCGGACGGCCCGATGACC
GTTAGAGTGACTGACCAATACGGTCGCTCGGTCTTCTCCCCGGAATCCGA
ATGGCACCAGGGGAGATCCAGAGTACGGCGCGCCGGTCTATCCGGTGCA
CGCCGCCGCCCATCATCATCATCATTGA

Gene Accession Number: CCE74122.1

SQP G P A K P G A I S A K W Q P G G S W A S G Y V A N L D V T A T D A V T G W T V S W A S P E T T G
V V N S W G M R C S V A S N T V T C T G T D W A S E L A A G Q T V S V G V Q L A G G P A P S S P Q I S
A T A A G T P P S Q P T P P S Q P T P P S Q S A T H G R A T H Y S L G T G N T I A N G N C S
M P A V P A D R M Y V A V S S P E Y S G A A A C G S Y L L V T G P K G T V R V Q I V D Q C H E C E I G H
L D L S E E A F R A V G D F D A G V I P I S Y T T V R D P D V P D V A V R V K E G S S R W W A G L Q I L N
A G N P I D H V E V Q A D G Q W L G L S R T D Y G Y W V T P S P I P D G P M T V R V T D Q Y G R S V V L
P G I R M A P G E I Q S T A R R F Y P V H A A A H H H H H H

C. Supplementary Figures

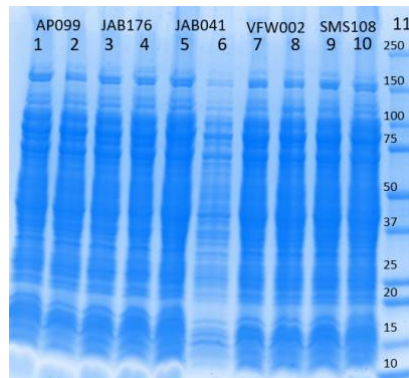


Figure C.1 | MagicMedia Trial 1. The SDS-PAGE gel from the first MagicMedia trial. Duplicate unpurified samples are shown. First trial was completed at 20°C and 240 RPM for 18 hours.

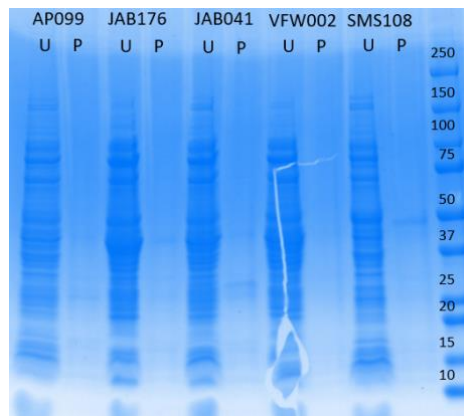


Figure C.2 | First Gel from MagicMedia Trial 2. The SDS-PAGE gel from the first duplicates in the second MagicMedia trial, completed at 18°C and 265 RPM for 18 hours. Unpurified (U) and purified (P) protein samples are shown.

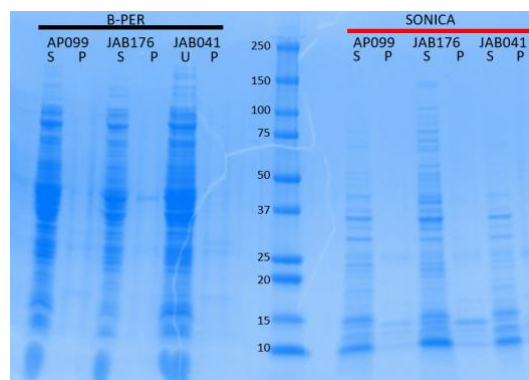


Figure C.3 | Lysis Analysis with B-PER and Sonication. SDS-PAGE from lysis analysis comparison done with samples from small-scale production trial 4 with LB media. Purified (P) and unpurified supernatant (SN) samples are shown for cells lysed with B-PER (black) and sonication (red).

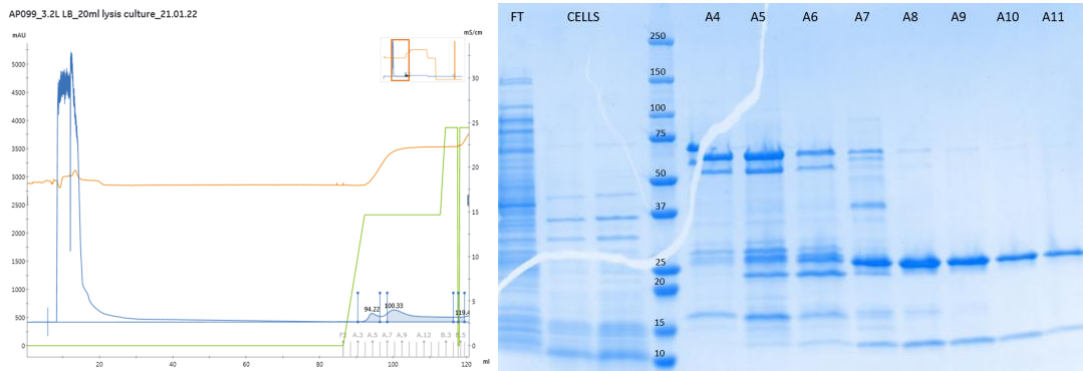


Figure C.4 | Large-scale Production of AP099. The chromatogram (left) and SDS-PAGE gels (right) from the large-scale production and purification trial of AP099. The chromatogram shows the UV280 (blue), concentration gradient of buffer B (green), and conductivity (orange). The gel was run with elution peak samples A4-A11, flow through (FT) and the supernatant from heated leftover cells (cells).

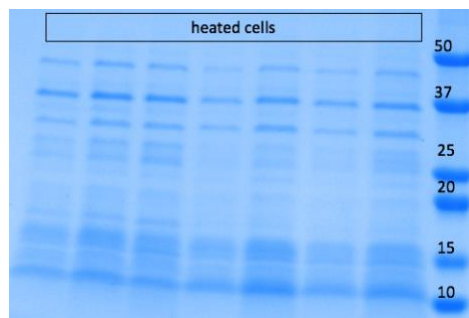


Figure C.5 | Additional SDS-PAGE Analysis for AP099. Multiple heat cells were tested to examine if any protein still resided in the cells.

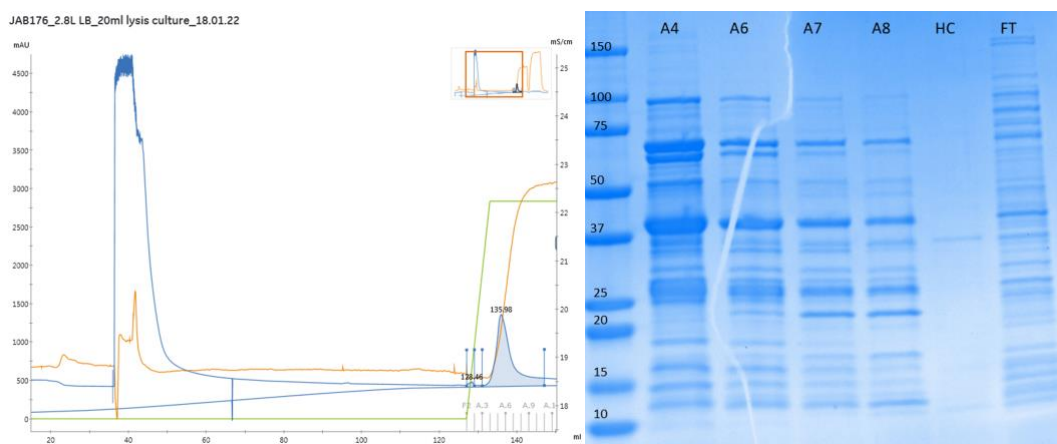


Figure C.6 | Large-scale Production of JAB176. The chromatogram (left) and SDS-PAGE gel (right) from the large-scale production and purification trial of JAB176. The chromatogram shows the UV280 (blue), concentration gradient of buffer B (green), and conductivity (orange). The gel was run with elution peak samples A4, A6-A8, flow through (FT) and the supernatant from heated leftover cells (HC).

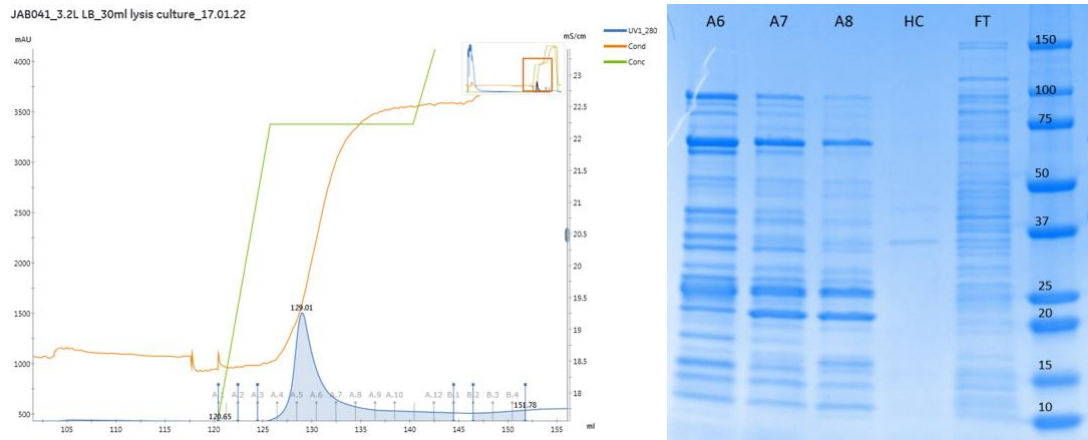


Figure C.7 | Large-scale Production of JAB041. The chromatogram (left) and SDS-PAGE gel (right) from the large-scale production and purification trial of JAB041. The chromatogram shows the UV280 (blue), concentration gradient of buffer B (green), and conductivity (orange). The gel was run with elution peak samples A6-A8, flow through (FT) and the supernatant from heated leftover cells (HC).

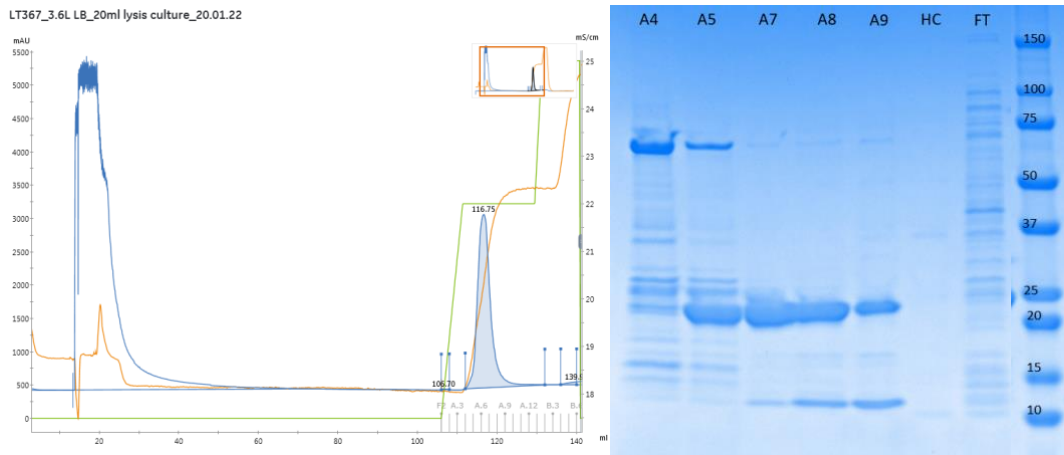


Figure C.8 | Large-scale Production of LT367. The chromatogram (left) and SDS-PAGE gel (right) from the large-scale production and purification trial of LT367. The chromatogram shows the UV280 (blue), concentration gradient of buffer B (green), and conductivity (orange). The gel was run with elution peak samples A6-A8, flow through (FT) and the supernatant from heated leftover cells (HC).

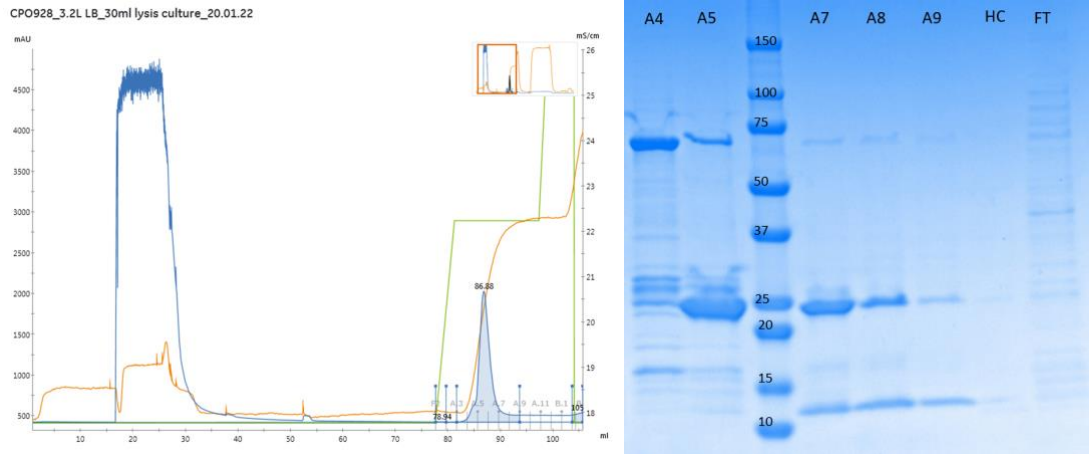


Figure C.9 | Large-scale Production of CPO928 Trial 1. The chromatogram (left) and SDS-PAGE gel (right) from the first large-scale production and purification trial of CPO928. The chromatogram shows the UV280 (blue), concentration gradient of buffer B (green), and conductivity (orange). The gel was run with elution peak samples A4-A5, A7-A9, flow through (FT) and the supernatant from heated leftover cells (HC).

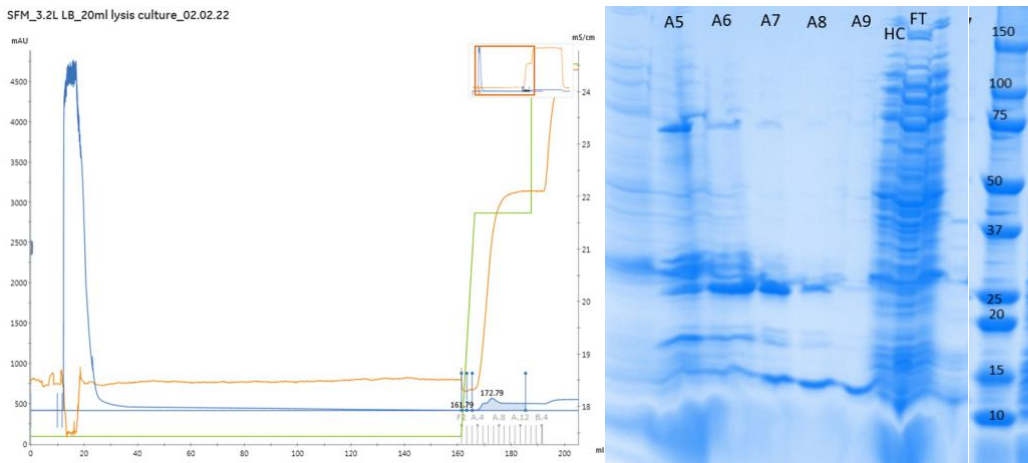


Figure C.10 | Large-scale Production of SFM691. The chromatogram (left) and SDS-PAGE gel (right) from the first large-scale production and purification trial of SFM691. The chromatogram shows the UV280 (blue), concentration gradient of buffer B (green), and conductivity (orange). The gel was run with elution peak samples A5-A9, A7-A9, flow through (FT) and the supernatant from heated leftover cells (HC).