2019

Characterization of WssF; a putative acetyltransferase from Achromobacter insuavis and Pseudomonas fluorescens

Cody Reese
rees2190@mylaurier.ca

Characterization of WssF; a putative acetyltransferase from *Achromobacter insuavis* and *Pseudomonas fluorescens*

By

Cody William Reese

Thesis

Submitted to the Department of Biology

Faculty of Science

In Partial Fulfillment of the

Master of Science in Integrative Biology

Wilfrid Laurier University

2019

Table of Contents

List of Abbreviations:

4-MUB Ac: 4-methylumbelliferyl acetate

A_{280}: Absorbance at 280nm wavelength

AChE: Acetylcholinesterase

Blast: Basic Local Alignment Search Tool

c-di-GMP: Bis-(3´-5´)-cyclic dimeric guanosine monophosphate

CF: Cystic Fibrosis

C-terminal: Carboxyl terminal

ddH2O: Double Distilled Water

DMSO: Dimethyl sulfoxide

EDTA: Ethylenediaminetetraacetic acid

EPS: Exopolysaccharide

FPLC: Fast protein liquid chromatography

His_6: Hexahistidine affinity tag

IMAC: Immobilized metal ion affinity chromatography

IPTG: Isopropyl β-D-1-thiogalactopyranoside

kDa: Kilodalton

LB: Lauria Broth

MES: 2-(N-morpholino)ethanesulfonic acid

Ni-NTA: Nickel-nitrilotriacetic acid chelate

N-terminal: Amino-terminal

OD_{600}: Optical density at 600nm wavelength

Phyre2: Protein homology/analogy recognition engine, version 2

SB: Super broth media

SDM: Site-Directed Mutagenesis

SDS-PAGE: Sodium dodecyl sulfate-polyacrylamide gel electrophoresis

TDLO: The lowest dose causing a toxic effect.

Tris: Tris(hydroxymethyl)aminomethane

List of Figures:

List of Tables:

Abstract

Biofilms are a survival mechanism commonly employed by communities of bacteria for adherence and protection. Bacteria produce a matrix of polymers (e.g. exopolysaccharides, such as cellulose) that allow them to exert control on their local environment. In the case of cellulose biofilms, acetylation (addition of acetate on carbohydrates) is paramount for polymer integrity and in some cases virulence. For this research, the wrinkly spreader (WS) genotype of the emergent human pathogen *Achromobacter insuavis* facilitates infections of the eyes of contact lens wearers and the lungs of Cystic Fibrosis patients (CF). Chronic infections have created a growing concern for the protective role biofilms play in antimicrobial resistance. Cellulose biofilms are an important mechanism for other pathogenic bacteria such as *Pseudomonas fluorescens*, which affect produce and water sources by forming a robust cellulose-containing biofilm that enhances the ability of the bacteria to spread across surfaces. Both *A. insuavis* and *P. fluorescens* contain the *wss*F gene that encodes a putative O-acetyltransferase protein, in that site-specific mutagenesis of *wss*F resulted in non-acetylated cellulose. The purpose of this research was to characterize WssF from both species, to understand its role in the production of acetylated cellulose and investigate its role as a putative O-acetyltransferase protein. This research progressed successfully through three primary objectives; expression and purification of WssF, functional characterization of acetyltransferase activity and structural characterization of the WssF protein. Esterase activity, the first step in acetyl transfer, was successfully characterized through a kinetic workup, which found K_M (2.88 – 7.18 mM), k_{cat} (0.24 – 0.34 s^{-1}) and, V_{max} (47 – 83 $s^{-1}*M^{-1}$) values comparable to homologous acetyltransferases. Following this, acetyl transfer was confirmed through the use of Thin Layer Chromatography (TLC) and Mass Spectrometry, both mono and diacetylated cellopentaose was produced in the presence of

unmutated WssF. Site Directed Mutagenesis (SDM) of the predicted catalytic triad residues (Ser, Asp and, His) resulted in up to 85% inhibition of esterase activity and at least 95% of acetyl transfer thereby confirming their role in catalysis. Successful functional characterization of WssF has allowed for screening of potential enzyme inhibitors by High Throughput Screening (HTS) assays. A set of 18 compounds has been selected for further study as potential WssF acetyltransferase inhibitors. Finally, structural characterization of WssF has been advanced, with optimal conditions for crystallization having been determined and (~2Å) data sets collected but solving a structural model of WssF is ongoing. The findings of this research successfully outline the functional and structural characteristics of WssF, which are the first for a cellulose acetyltransferase involved in biofilm formation. This thesis further develops our understanding of biofilm modification, a vital step in combating antimicrobial resistance.

1.Introduction

1.1 Biofilms

A common survival strategy for many bacteria is the biosynthesis and export of polysaccharides to immobilize the cells within a protective biofilm together with fragments of DNA and peptides. Biofilms are a complex matrix of synthesized exopolymers utilized for adherence to various surfaces (1). A growing body of literature underscores the importance of further research into biofilms. For example, the Center for Disease Control (CDC) estimates 48 million Americans suffer annually from foodborne illnesses (many caused by biofilm-forming bacteria) costing between $50-70 billion per annum (2). Unofficial estimated costs of biofilm infections, not including foodborne illness, was $94 billion in the USA alone, with an additional $105 billion spent on dental care in 2007 (3–5). In the healthcare setting nosocomial, or hospital acquired, infections frequently arise due to patients being exposed to hard-to-clean equipment ranging from catheters to ventilators that harbour biofilms (3, 6). The costs associated with treating these biofilm-related infections have created a large need for further research into biofilm forming pathogens.

Even outside of the clinical healthcare setting personal hygiene and comfort are affected by biofilm forming bacteria. Cultures grown from the contact lenses of Corneal Infiltrative Event (CIE) patients showed 61 % of the cases had *Acromobacter* spp. present (7). All the contact len cases examined used multipurpose disinfecting solutions (MPDS), which are tested for a reduction in *P. aeruginosa*, *S. marcescens* and *S. aureus* growth (8). These MPDS are not tested against a multitude of other bacteria, such as, the various *Achromobacter* spp. As a result of the inability of many MPDSs to fully sterilize the contact lens cases CIEs, such as, in **Figure 1.1**, occur after prolonged usage (7). In response to increased bacterial load, the typically avascular

cornea will have white blood cells migrate into the area to form white plaques (9). In severe cases, infectious events may threaten to destroy eyesight completely. The incidence for adverse effects, such as, CIE or keratitis are estimated to range from approximately 40 in 100,000 users up to 25 in 10,000 for individuals who wear the contact lens for longer than 10-12 hours and/or greater than 5 days a week (7, 10, 11). Due to the lack of efficacious treatments infections by *Achromobacter* spp. remain a prominent risk for all contact lens wearers. Therefore, treating the contact lenses more effectively will provide not only health benefits but also increase convenience for the average contact lens wearer.

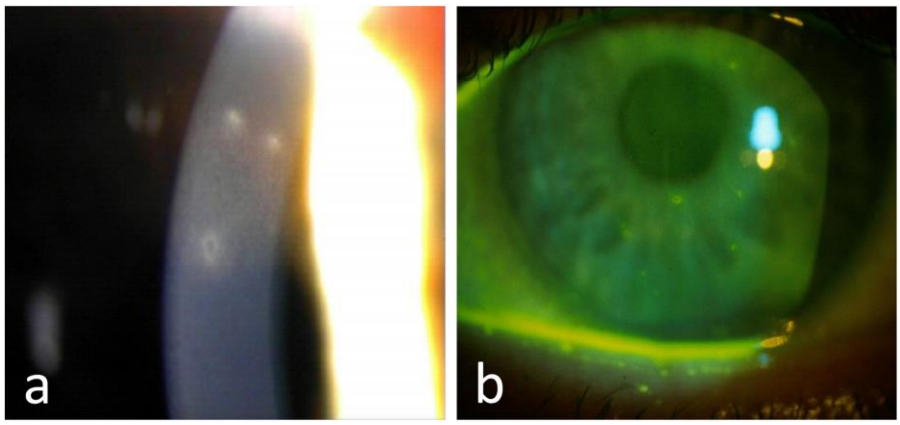


Figure 1.1 Presence of CIEs after extended contact lens case use with MPDS as only cleaning agent. Multipurpose disinfecting solutions have been found to be inadequate in preventing CIEs without also replacing contact lens cases at regular intervals. (Adapted from Kilvington *et al.* (7))

Preventing negative healthcare outcomes is one of the benefits of maintaining the quality of any society's food and water supplies. Unfortunately, our food and water supply chains are susceptible to contamination biofilm forming bacteria. The adverse effects of biofilm forming bacteria are not only related to health, but also strongly connected to economic outcomes. For example, the production of biofilms by *Pseudomonas* within raw milk allow for bacteria communities to survive the pasteurization process (12). Even biofilms of nonpathogenic bacteria

present within food processing facilities have recently been identified as overlooked sources of contamination that might potentially harbour pathogenic bacteria (13). Allowing this bacterial presence within treated milk reduces the shelf life of the final products and reduces the quality of products that companies sell. The damaging effects of biofilm production on agriculture are not limited to the dairy industry. *Pseudomonas* spp. have also caused an epidemic of the kiwi canker, which has threatened kiwi crops worldwide (14–16). Repercussions from the loss of the kiwi will have international economic effects, as kiwis are a main export of many Pacific countries. This loss will not just be economic as kiwi fruits are also a major cultural export of some countries, like New Zealand. The impact of not being able to stop these infections ranges from agricultural sustainability, and economic loss to the diminishing of cultural/national identity.

While most of the focus on biofilms has revealed the negative impacts of formation, there are emerging cases where biofilm formation could be productive. In North America the presence of some *Pseudomonas* spp. have been found to protect crops of tomatoes and radishes from wilting (17, 18). As cash crops, tomatoes and radishes not only provide food security but also economic stability for vast communities (19, 20). The cause of wilting is due to a fungus called Fusarium; fortunately multi-species communities, including *Achromobacter* and *Psuedomonas*, outcompete the fungus for iron within the soil; thereby making it inhospitable for further infestation (17, 18). Biofilms allow the bacteria to resist the physical forces, such as, rainwater, which in turn allows the bacteria to survive longer on the crops as they combat the fungal infestation. In these cases, the ability of the biofilm to protect the symbiotic bacteria is an asset for economic gain instead of an adverse health effect. Thus, there are potential benefits to learning not just how to disrupt biofilms, but how to manipulate certain community members to thrive within a desired environment.

1.2 Benefits of Biofilm Production

Adaptation for increased survival, through the production of a protective extracellular polymeric substance (EPS) barrier, is common in environments ill-suited for “free living” (planktonic) lifestyles (21, 22). Consequently, biofilms (composed of EPS) are produced by many different species of bacteria living as a community for protection (23, 24). Multispecies biofilms allow each species to perform specialized tasks a homogenous microbial community may not be able to, allowing the multispecies biofilm improved utilization all the potential metabolites within their environment (25). In the case of Inflammatory Bowel Disease (IBD) the intestinal microbiota changes significantly, producing larger biofilms, with increased affinity for iron abundance (25). While these changes negatively affect the hosts overall health, the bacteria have an increase in viability. More specifically in the case of Crohns disease, inflammation due to an abnormal immune response decreases the functionality of the intestines without eradicating the underlying infection (25, 26). Reducing intestinal efficiency increases the nutrients available for the pathogenic bacteria, while simultaneously reducing the viability of symbiotic gut microflora. Bacteria within the biofilm form controlled extracellular environments, which allow them to share DNA and metabolites. These systems become even more complicated as gradients of nutrients, waste, chemical signals, toxins, antibiotic chemicals, etc. form. Shared metabolite utilization, however, is not the only benefit as biofilms also increase bacterial adhesion to solid surfaces (27, 28). Enhanced adhesion is extremely valuable to bacteria that occupy irrigation systems or the digestive tract where supplied nutrients flow through the system from a single source. The ability to remain attached to a location with a constant nutrient source avoids excretion, while increasing exposure to metabolites, increasing fitness and viability. Enhanced adhesion avoids excretion as it allows the bacteria to remain on crops and storage containers

when experiencing shearing forces (e.g. rain or cleaning cycle) (29). Thus, the benefit of EPS production (inherent in biofilms) is not only protection, but increased control over the external environment of the bacteria.

Biofilms are able to protect microbial communities from both host immune functions and antibiotic treatments. By producing a barrier, the multitude of antibodies and neutrophils produced by a host are unable to interact with, identify or remove the bacterial presence (30). In some cases, the biofilm communities are mostly non-pathogenic bacteria, like those that form the gut microbiota. A problem arises when pathogenic bacteria join these non-pathogenic communities, like in the case of IBD as previously discussed (25). The once commensal community of bacteria breaks the tenuous balance with the host and its immune system relying on the biofilm community for survival. Biofilms offer embedded bacteria more than just protection from the immune system. Embedded bacteria also benefit from antibiotic resistance, which is increasingly being recognized as an urgent problem that threatens to make current medications obsolete (31, 32). When antimicrobials are introduced to a biofilm community a concentration gradient is formed as they slowly diffuse through the biofilm matrix, which generates lower antibiotic dosages among the less metabolically active bacteria, therefore allowing potentially sensitive cells to survive (33). These lower antibiotic dosages are less effective allowing resistant bacteria the opportunity to transfer plasmids, containing resistance genes, to neighboring bacteria that would otherwise be susceptible (3). New approaches to increase antibiotic efficacy and/or generate new antibiotics are crucial to our ability to control these bacterial communities (34). In summary, once the biofilm matrix surrounds the cell it offers protection from antibiotics, competing organisms, harsh environmental factors, and host

immune systems (3, 33). This complex defense mechanism has created an urgent need for research into remediating the biofilm barrier (3).

1.3 Biofilm Composition and Development

A biofilm is a complex matrix comprised primarily by EPS that can be produced by either a homogenous or heterogenous community of bacteria (1, 30). Microcolonies characterized by their EPS make up neighborhoods, which build the foundations that determine the structural formation of the maturing biofilm later in development (30, 35). Each microcolony may form a different EPS or proliferate in a way that changes their neighborhoods structural foundation (35, 36). While the EPS provides the majority of the structure and strength found within a biofilm, the composition is far from homogenous as seen in **Figure 1.2**. Biofilms are diverse mixture comprised of lipids, nucleic acids, proteins and, carbohydrates, which are the most abundant substance within the matrix (36, 37).

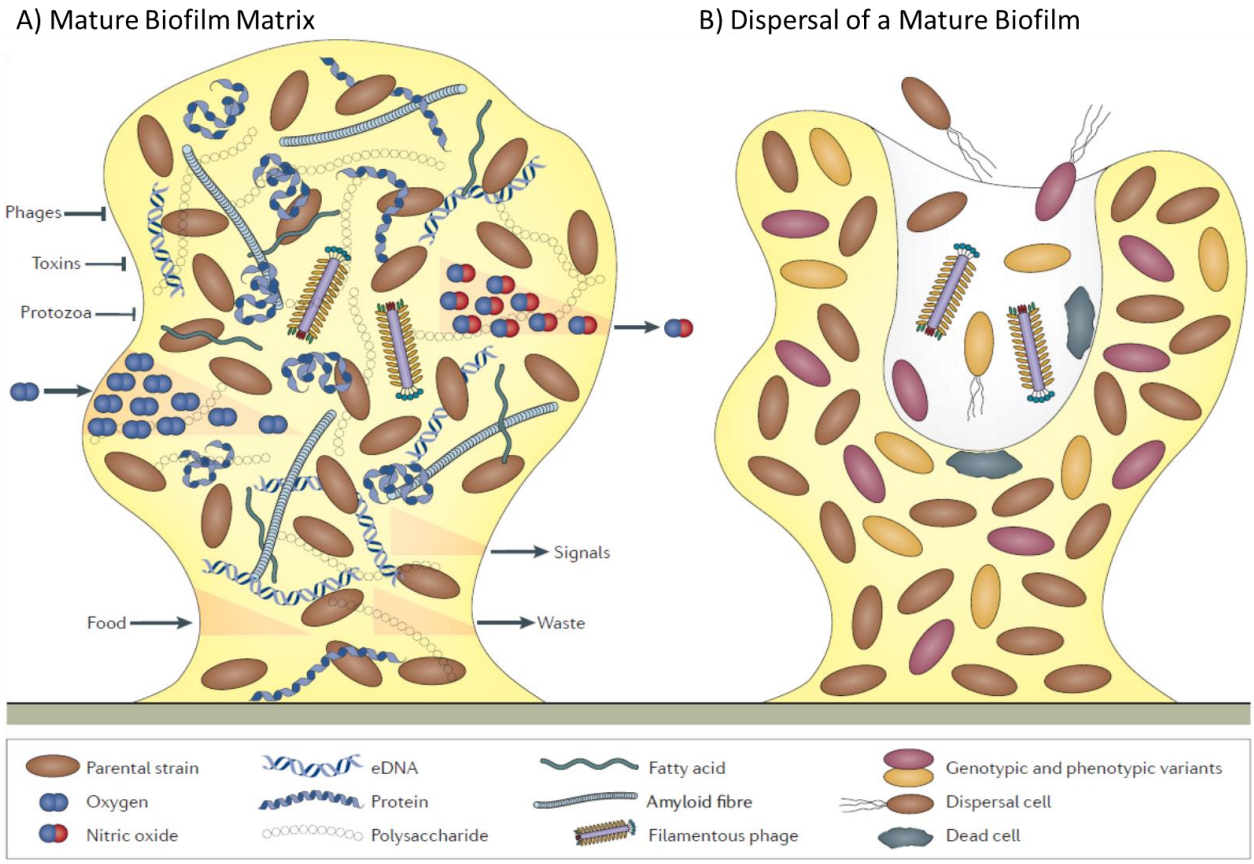


Figure 1.2 The composition of a developing biofilm (**Adapted from** (38)**).** On the left (A) is a representation of the complex environment produced by biofilm forming bacteria. The EPS produced forms a protective barrier reducing the effects of toxins, lytic phages and, protozoa. An environment of various proteins, eDNA, fatty acids and nutrients forms allowing the cells to create a shared extracellular metabolism. The production of nutrient, waste and chemical signal gradients also form as a result of the thick EPS, which force the cells to adapt. On the right (B) is a figure representing dispersion of a mature biofilm. After/during maturation portions of the biofilm begin to hollow out where cells revert to planktonic lifestyles. Dispersing cells leave behind the dead cells trapped within the matrix.

Production of a biofilm begins when the bacteria detect sufficient bacterial density within an environment suitable for formation. This detection is mediated through Quorum sensing, where chemical signaling molecules allow bacteria to communicate information like population density (39, 40). Different species use different pathways to detect environmental conditions and use these signals to help regulate inter and intra species cooperation (40, 41). Adhesion, population density, and neighboring biofilm production all signal a change in the cyclic

diguanylate (c-di-GMP) pathway (42–45). Bacteria deploy their survival mechanisms through control of cytosolic c-di-GMP levels. An example of this control is when phosphodiesterase enzymes hydrolyze c-di-GMP deactivating the secondary messenger and upregulating motility (*ie.* a mechanism of dispersal as depicted in Figure 1.1) (44, 45). Increased motility is achieved through the production and activation of cellular structures, such as, flagella, which are used to propel the cell. In the case that external signals indicate sufficient population density at an interface, diguanylate cyclases increase the levels of c-di-GMP, which in turn up regulate proteins contributing to adhesion and biofilm formation (42, 46). Maintaining a balance of c-di-GMP production and hydrolysis allows the cell to respond with planktonic (motile) or sessile (biofilm production) modes of growth **(Figure 1.3)**, depending on what environmental pressures act upon the cell (43).

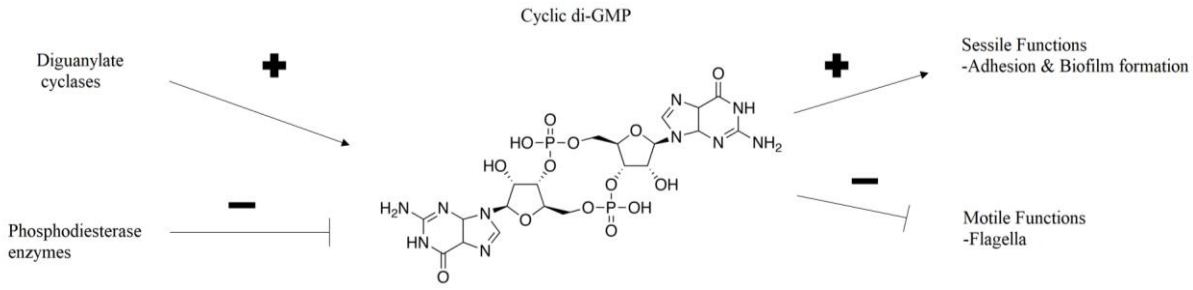


Figure 1.3 Diagram illustrating Cyclic di-GMP pathway. (Generated in Microsoft Paint)
Diguanylate cyclase increases c-di-GMP levels while phosphodiesterase hydrolyzes c-di-GMP. The presence of c-di-GMP increases sessile functions and decreases motile functions.

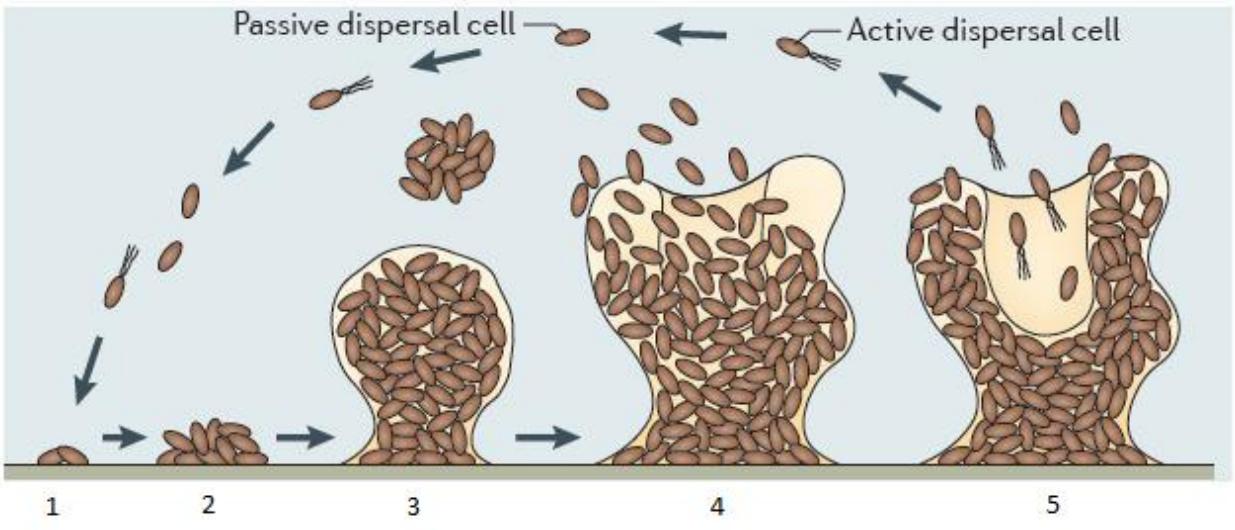


Figure 1.4 Stages of Biofilm Development (Modified(38))**.** Biofilm development proceeds in 5 stages. 1) **Planktonic** cells within the local environment determine whether conditions are favourable for biofilm formation. 2) **Initial attachment** of cells is facilitated through utilization of membrane bound constructs, such as, flagella or pili. 3) **Irreversible attachment** through the production of exopolysaccharides, which provide strength and organization. 4) **Maturation** of biofilm as the organization of the biofilm allows for the formation of water channels and stratification of the biofilm. 5) **Dispersal** of the mature biofilm through both active and passive dispersal allows for cells to either reattach elsewhere or reenter a planktonic lifestyle.

Once the bacteria receive a signal that a suitable environment is present, the development of the biofilm can progress through five subsequent stages **(Figure 1.4)**; **Planktonic**, **Initial attachment**, **irreversible attachment, maturation** and, **dispersal**. The first stage in biofilm development is set in motion by the environmental signals discussed previously. Chemical signals (e.g. N-acyl homoserine lactones, ion concentrations, etc.) cause upregulation in transcription of the genes encoding diguanylate cyclases, which increase levels of c-di-GMP, causing a switch to sessile functions (42, 46). With increased transcription of genes encoding for adhesion proteins, such as, curli fimbriae, there is a parallel down (-) regulation for motility proteins, such as, flagellar components (46, 47). This change allows for **initial attachment** where anchoring of bacteria to various surfaces (e.g. host cell membrane proteins or nutrient rich soil/sediment) begins. As more cells attach, potentially from diverse species, the cells divide and

grow forming microcolonies (48). This density change causes upregulation in the genes responsible for EPS synthesis, modification and, export (as will be discussed in detail later in section 1.4), which begins stage three, **irreversible attachment**. As a major component of the biofilm matrix, EPS forms around and throughout the microcolonies, beginning to create the extracellular environment previously discussed within the biofilm (49, 50). This causes the microcolonies to grow and connect forming the neighborhoods that will later dictate the overall structure of the greater community (30). The formation and development of the microcolonies in stage three are drastic as seen in **Figure 1.5.** Early in biofilm development the cells are sparsely arranged within the biofilm matrix. As the biofilm grows there are significant increases in cell and matrix density. Cells become densely packed as the biofilm develops from the stage of **irreversible attachment** into stage four **maturation**.

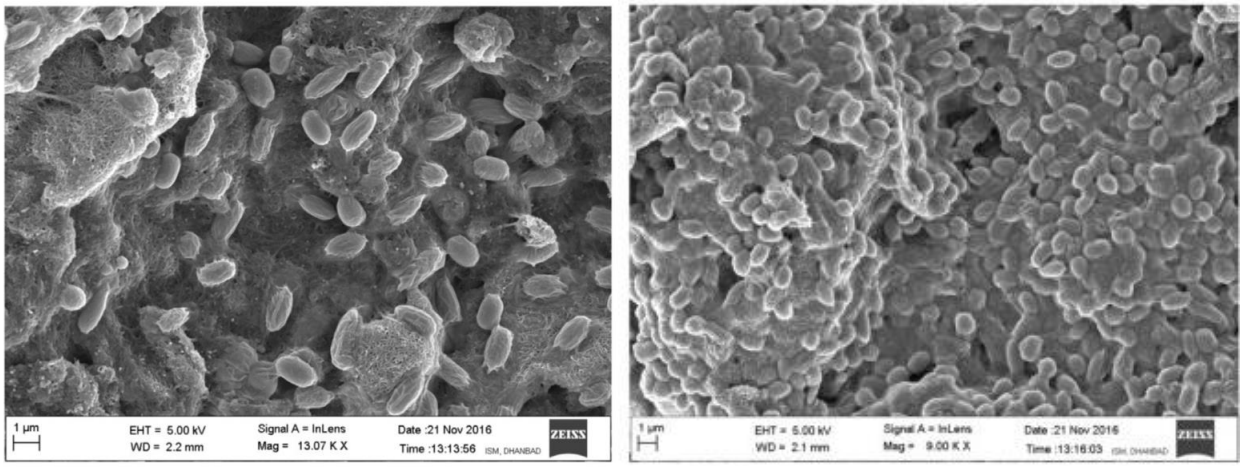


Figure 1.5 Fluorescence Emission Scanning Electron Microscopy images of *Bacillus thuringiensis*. (Modified from (51)**)** Early biofilm development displaying sparsely packed cells Left. Mature biofilm development displaying densely packed cells (15 days) Right.

Stage four, biofilm **maturation**, is marked by copious amounts of EPS and pili but also the development of community structure. The structure of the growing community changes and expands based on multiple factors. As previously mentioned, the initial attachment and the early

formation of microcolonies guides how and where the mature biofilm has formed. A driving factor of this formation is of course access to nutrients, which not only signals a place for initial attachment but also the formation of channels. Channels for external resources, such as, water and nutrients form, allowing the network of neighborhoods access (48, 50). However, within those neighborhoods access to the external environment is not uniform as deeper cells have limited access to water or nutrients. This creates cells that are less metabolically active as previously discussed in section 1.2.

As the biofilm matures some areas within the biofilm begin to enter stage 5, dispersion. **Dispersion** occurs when external portions are removed from the community by passive or active dispersal (52). Passive dispersal is when cells are sheared away from the biofilm, by physical forces, such as, currents or predator grazing (52, 53). Alternatively, active dispersion is an individual or coordinated community signaling that cause cells on the periphery of the biofilm to change back to a planktonic lifestyle leaving the biofilm community. These chemical signals regulate gene transcription to produce phosphodiesterase enzymes that decrease the levels of c-di-GMP in the cytoplasm, which in turn result in the promotion of transcription/translation of proteins involved in motile functions (42, 46). Whether actively or passively dispersed, cells may move back into a planktonic lifestyle or find another location to reattach and undergo the stages of biofilm formation once more.

1.4 Cellulose

Biofilm production is a complex process requiring various proteins to perform a multitude of functions. Once bacteria sense a need for biofilm production, one or more operons begin transcription of the required protein complex. Bacterial cellulose synthase (*bcs*) produces cellulose, a β1,4 linked glucose polymer, to aid and protect bacteria, such as, *E. coli* or

Salmonella spp. Cellulose synthesis and export is achieved primarily by four proteins, BcsA,B,Z and,C (54, 55). BcsA polymerizes UDP-glucose into the β 1,4 linked polymer concomitant with transport into the periplasm through its transmembrane pore region (56–58). Interestingly, the role of BcsA in polymerization is in some biofilm synthase complexes completed by multiple proteins instead of one. Within the alginate system, for example, the bacteria utilize two well described proteins, alg44 and alg8 (59, 60). Having multiple solutions for the task of polymerization displays the complex nature of EPS production. Moving toward the outer membrane, the next protein is BcsB, which complexes with BcsA and is believed to guide the growing cellulose polymer through the periplasm (56–58). BcsZ then acts as an endoglucanase; cleaving the polymerized cellulose chain for release from the cell (61, 62). The function of BcsC is yet unknown, however, it is believed to function similar to the analogous AlgE/AlgK complex that exports alginate across the outer membrane in *Pseudomonas* spp. (63). This would mean BcsC contains both the β-barrel porin of AlgE and the tetratricopeptide repeat of AlgK allowing it to guide cellulose through the periplasm and outer membrane. While the Bcs complex does not directly acetylate cellulose, recent research has indicated the presence and importance of EPS modifications in the Wss (cellulose) and Alg (alginate) systems. Although the Bcs system does not acetylate cellulose it is useful for comparing against other cellulose synthase systems like Wss for determining function of the various proteins.

In the case of *A. insuaivs* and *P. fluorescens*, the wss operon is responsible for the acetylated cellulose synthase apparatus. Potential functions of the proteins within the *wss* system were determined using mutant knockouts (22, 28). Mutant knockouts determined *wss*B,C,D and,E (analagous to the *bcs*A,B,Z and,C genes from other systems) are required for cellulose synthesis. Once removed, the absence of any one of the genes resulted in a mutant that did not

produce cellulose when stimulated (28). Although important for the synthesis of cellulose, the function of WssC is currently uncharacterized but believed to be a hypothetical binding domain (analogous to BcsB). After polymerization, cellulose is translocated through the periplasm where modification can occur before exportation. Further knockouts found WssF,G,H and,I are responsible for acetylation of the cellulose polymer (22). Mutants produced a non-acetylated cellulose polymer that had a decreased capacity for protecting the bacteria from anti-microbial forces, such as shearing forces (64). Further importance of polymer modification will be discussed later in section 1.5. In order to visualize the Bcs and Wss systems **Figure 1.6** has been generated, which includes the well described Alginate system for comparison.

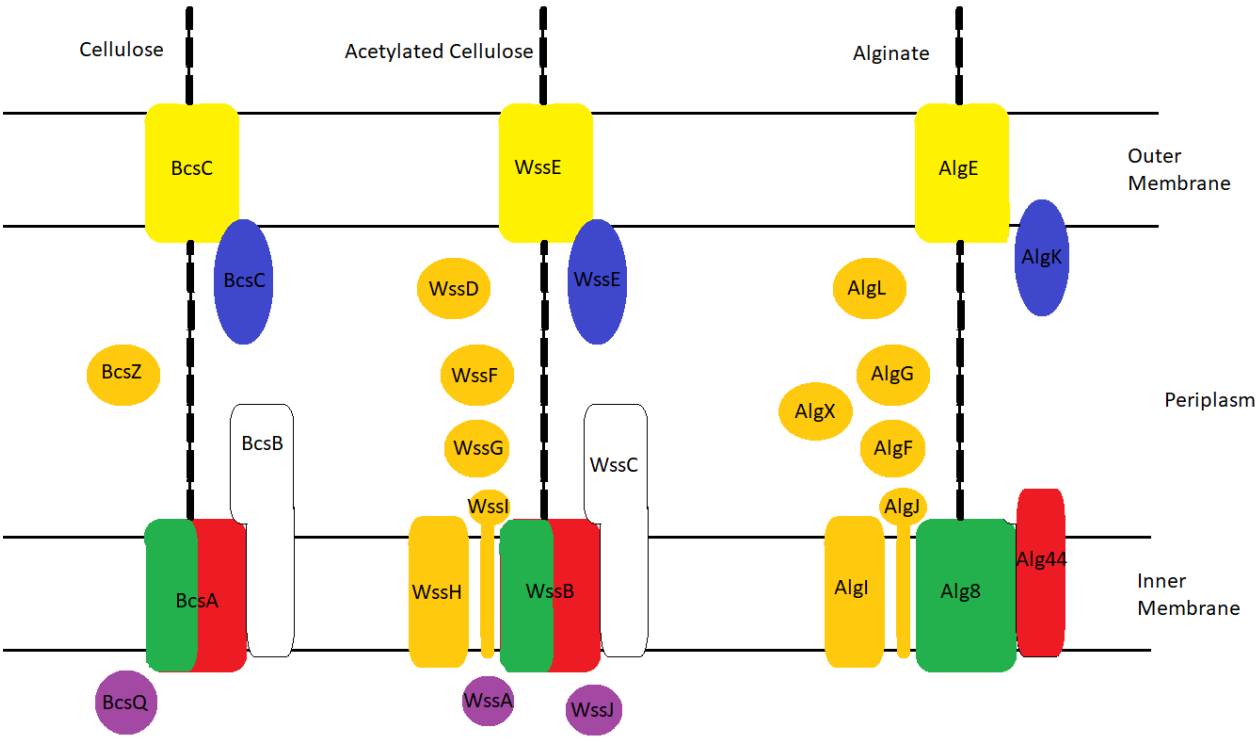


Figure 1.6 Graphic representation of alginate and cellulose biosynthetic apparatus The protein components are color-coded according to predicted function as follows: purple, MinD homolog; green, synthase; red, c-di-GMP, bis-(3′-5′)- cyclic dimeric guanosine monophosphate binding domain/protein; orange, exopolysaccharide-modifying enzyme; blue, tetratricopeptide repeat; yellow, β-barrel porin; and white, hypothetical binding domain. **(Generated in Microsoft Paint based on** (65)**)**

Another example that has been extensively researched is the synthesis and modification of alginate biofilms. Even though the present research has focused on WssF, which is part of the *wss* operon for cellulose synthesis, the alginate system provides a useful comparison to the less characterized cellulose system **(Figure 1.6)** (65). The *alg*D operon contains genes for the 10 proteins that form an alginate biosynthetic apparatus (66–68). Unlike the alginate system, the Wss system has two MinD homologues WssA and WssJ, which are believed to be involved in the localization of the cellulose synthase and O-acetylation proteins respectively (28). Additional proteins are then needed to coordinate c-di-GMP signal detection and polymer synthesis. In the case of alginate, the inner membrane located Alg44 protein detects the presence of c-di-GMP and acts as a positive activator of Alg8 to convert molecules of guanosine 5'-diphosphate-mannuronic acid into poly-β-D-mannuronic acid (59, 60). In the cellulose system, the activities of alg44 and alg8 are hypothesized to be contained in one multi-domain protein, denoted WssB or BcsA depending on the bacteria it was identified in (28). Following synthesis, certain proteins perform polymer modifications, such as, O-acetylation, epimerization, cleavage, and more. For example, AlgF,I,J,X act to O-acetylate the polymer, which has been shown to increase the strength and organization of the final product (69–71). In the cellulose system WssG,H,I,F bear sequence homology to the alginate proteins, respectively, and are hypothesized to be responsible for the acetylation of cellulose biofilms (28). The locations for AlgF,I,J were determined, using alkaline phosphatase (AP) gene fusions and this work supported bioinformatics sorting programs (72, 73). Amoung all the randomly inserted AP gene fusions the in-frame fusions to the proteins algF and algJ had Alkaline Phosphatase activity, indicating that both AlgF and AlgJ were exported to the periplasm. Further, immunoblot analysis demonstrated that AlgF is a soluble periplasmic protein, while AlgJ is anchored to the inner membrane (73). Alg J was found to have

seven transmembrane domains, which anchor the protein, followed by a large periplasmic domain. AlgF,I,J are believed to transfer an acetyl group across the inner membrane where AlgX and AlgJ acetylate the alginate polymer (69). Analogous proteins in the teichoic acid alanylation system flip wall teichoic acids (WTA), produced on the undecaprenyl carrier lipid, through the inner membrane of Gram-positive bacteria (74). Within this system DltD has recently been described as an O-acyltransferase in the DLT pathway (74). DltD is believed to utilize a Ser-His-Asp catalytic triad. Both AlgX and AlgJ possess a SGNH hydrolase-like domain that contains a catalytic triad of Ser-His-Asp to acetylate the alginate polymer. SGNH hydrolases are a vast and diverse super family of esterases and lipases, which have an alpha/beta/alpha structure, where the middle beta sheet is normally found to be five parallel strands (75). Based on similarities between the alginate and cellulose systems, WssI and WssF may also serve as acetyltransferases due to their SGNH hydrolase-like domain.

Following acetylation, the alginate and cellulose systems have additional similarities and differences. For example, AlgG from the alginate system has been shown to act as a C5-mannuronan epimerase (76). That converts alginate from a solely β-D-mannuronic acid to a polymer that is infrequently spaced with α-L-guluronic acid; thereby altering the viscosity and polarity of the final product. In the case of the Wss system, cellulose is produced as a homopolymer and no equivalent protein to AlgG is present. The similarities between the two systems then converge again with respect to the other proteins in the periplasm and outer membrane. The alginate polymer is cleaved by AlgL, similar to WssD/BcsZ, allowing for a controlled release of the polymer (62). Finally exportation of the polymer is achieved by AlgK (a tetratricopeptide repeat (TPR) containing protein) that binds the polymer guiding it through the β-barrel porin AlgE (77–79). In the cellulose system, there are also TPR and β-barrel regions, but

they are contained within a single WssE/BcsC protein (28). Strong similarities between the alginate and cellulose synthetic pathways allow for some confidence when predicting the functions of various Wss proteins.

1.5 Polysaccharide Modification

Comprehending the biological relevance of biofilms requires understanding the importance of polysaccharide modification. Modified polysaccharides are a common mechanism for protecting the cell. Acetylation is emerging as one important way of modifying the EPS used by bacteria for protection. The importance of acetylation, however, is far more wide spread than just modifying biofilms. Briefly, acetylation is known to play an integral role ranging from mammalian development to providing increased protection from heat and oxidative stress for many *E.coli* proteins (80–82). In the case of acetylating exopolysaccharides, O-acetylation of the cellulose is used to improve biofilm strength and efficacy (28, 83). For bacterial mutants that form cellulose without O-acetylation the biofilms where observed to be stringy and thinner than wildtype biofilms (28). Spiers and colleagues (2003) successfully visualized the effects of non-acetylated cellulose as seen in **Figure 1.7**. When each of the proteins required for acetylation were mutated, *wss*FGHIJ, the resulting biofilm phenotypes were noticeably different. One of these mutants was designated WS-18 where a mini insertion was located on the *wss*F gene (64). Analysis of the polar mutants, mutations that affect expression of downstream genes, showed that when any of the proteins were removed from the system it caused a disruption in O-acetylation of the polymer. Without the proper cohesion, provided by polymer modification, the bacteria are unable to form a thick biofilm, as seen with the wildtype. Non-acetylated cellulose is more disorganized and less cohesive, resulting in the cellulose polymer displaying a decrease in mechanical protection for the community (84). Those bacteria unable to successfully form/join

biofilms are more susceptible to antibacterial compounds or sudden changes in environmental forces (e.g. rain or rivers current) (36, 37). As such the presence of a strong biofilm is important for the success of many bacterial communities as they explore nutrient rich, yet dangerous locations. (28, 64, 85). As such, it is important for the biofilm to have an organized and cohesive structure, which in the case of cellulose is often dependent upon acetylation.

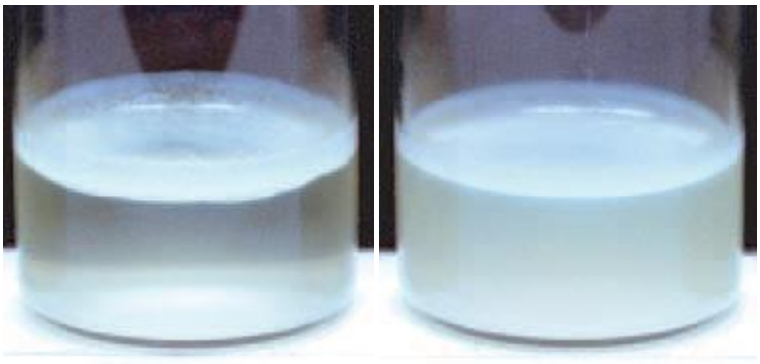


Figure 1.7 Comparison of biofilm formation. Wildtype air-liquid interface *P. fluorescens* biofilm (left) compared against WS-18 (WS mini-Tn5::*wsp*F – *wss*FGHIJ polar mutant) biofilm (right). (28)**)**

While beneficial, O-acetylation of a polymer is a complex process. As seen previously **(Figure 1.6)**, multiple proteins can be required for polymer modification. This is because O-acetylation of materials, like cellulose or alginate, requires more than one reaction (86). First, an enzyme must remove the acetate from a donor and transport it through the inner membrane into the periplasmic space. To facilitate this one or more membrane bound O-acetyltransferases (MBOATs) are required (70, 86, 87). Next the acetate must be transferred onto the growing polymer. Once more this is achieved one of several ways, in the case of the wss system it is believed to rely on multiple periplasmic proteins that function as O-acetyltransferases WssFGI (28, 69). To predict how the Wss acetyltransferase system works it helps to look at the Alginate system where a “bi bi ping pong” or double displacement system has been described for AlgX

and AlgJ (88). This reaction proceeds in two steps; AlgJ utilizes an esterase reaction to remove the acetate from a donor, which allows AlgX to catalyze the O-acetylation of the growing alginate polymer (70). Which Wss enzymes contribute to each part of the O-acetylation of cellulose is still unknown.

2. Research Need

Tomatoes are a significant economic crop for North American Countries, within the US tomatoes account for $2 billion in annual farm cash receipts (20). In Canada vegetable crops in 2016 were valued at $1 billion with carrots and tomatoes representing over $100 million in crops each (89). Microbial biofilm forming species, such as, *P. fluorescens,* are capable of improving crops yields through increased nutrient utilization (90, 91). With increased nutrients accessible to crops these bacteria expand the growing regions of crops into otherwise arid or uncultivatable lands. Additionally these bacteria are protecting against fungal infestation (17, 18, 92). Protection against tomato and radish wilt, as mentioned before, will greatly increase yields and therefore profits. *Pseudomonas spp.* have been found to restore seedling emergence in tomato plants infected with *Ralstonia solanacearum* (92). Treated plants found 75 % seedling emergence compared to untreated plants with 40 % emergence, which was almost a full restoration with uninfected controls at 76 %. Beyond the economic benefits, concern for *P. fluorescens* is warranted from a healthcare perspective. *P. fluorescens* has largely been overlooked as its universal presence in places like plant soils and waste drainage is well known (93). However, recent research has discovered the ability of *P. fluorescens* to cause infections within immunocompromised hosts (94–97).

The most common population for chronic infection globally are cystic fibrosis (CF) patients (98). Infection rates of cellulose biofilm-forming *Achromobacter* species vary among CF patients: *A. xylosoxidans* rates vary from 17 % in Belgium (99) to as high as 20 % in Brazil, (100) while *A. insuavis* infections in the UK has been reported at 12 % (101). The global impact of CF is significant, as 1 in 3500 children are born with CF in the United States, which corresponds to more than 85,000 sufferers (98). However, immunocompetent populations are also at risk of infection notably observed in the case of contact lenses as mentioned previously. With the ubiquity of acute and chronic infections by bacterial species like *Pseudomonas* and *Achromobacter it* is important to understand what mechanisms they utilized to survive.

As one of the major survival mechanisms employed by bacteria, biofilms and specifically modified EPS are an important area of study. In particular the acetylation of cellulose, in *P. fluorescens* and *A. insuavis*, can be studied through elucidation of the WssGHIF cassette. Mutants of *wssF* have shown the encoded protein is required for the production of acetylated cellulose (28). However, due to a lack of structural and functional data, the role of WssF in acetylating cellulose remains biochemically uncharacterized. ***Our hypothesis is that characterizing the structure and function of WssF will confirm its function as an acetyltransferase that infers its role as a cellulose modifying enzyme in the cellulose biosynthetic apparatus.***

To address our hypothesis, we will perform the following objectives:

Objective #1: Protein expression and purification of the WssF *A. insuavis* and *P. fluorescens* constructs.

Objective #2: Perform a functional analysis of both WssF constructs for acetyltransferase activity. Functional characterization of esterase activity will be achieved through colour and fluorescence spectrometry. Transferase activity will be characterized through Thin-layer chromatography (TLC) and Mass Spectrometry of treated cellulose samples. Finally, profiles will be created against known inhibitor compounds as well as High Throughput Screening (HTS) on vast libraries of potential inhibitors.

Objective #3: Perform a structural analysis of both WssF constructs through a combination of X-ray crystallography and small-angle scattering experiments.

3. Material and Methods

3.1 Reagents, Chemicals, and Media

The chemicals purchased from Fisher Scientific were sodium chloride, tris(hydroxymethyl)aminomethane (Tris), Coomassie R250, acetic acid, methanol and glacial acetic acid. Fisher Bioreagents was used to purchase both Tetramethylethylenediamine Dithiothreitol (DTT) and Sodium Dodecyl Sulfate. From Fisher Chemical sodium phosphate mono- and dibasic were purchased. The products purchased from BioBasic were Tryptone powder, isopropyl β-D-1-thiogalactopyranoside (IPTG), yeast extract, kanamycin sulfate, agar, glycerol and, ammonium persulfate. From BioShop Imidazole, RNaseA and, DNaseI, Protein Ladder and, Acrylamide/Bis-Acrylamide (29:1) 30 % (v/v) solution were purchased. A Precision Plus Protein™ Dual Color Standard was purchased from Biorad. The His-pur Ni-NTA Resin used for purification was purchased from Thermo Scientific. For plasmid preparation a QIAprep Spin Miniprep Kit was purchased from Qiagen. Vivaspin centrifugal ultrafiltration devices were obtained from GE healthcare. The crystal screens Top 96, MCSG-1, 2, 3, and 4 were purchased

from Microlytic. The 96-well INTELLI-PLATEs were purchased from Art Robbins Instruments while Crystal Clear sealing tape was obtained from Hampton Research. Both cellulose and chito oligomers were purchased from Megazyme. Any reagents or chemicals not listed above were purchased from Sigma-Aldrich. When culturing *E. coli* there were two medias used primarily. When culturing stocks overnight LB or Luria Bertani broth was used. which contained 10 g/L tryptone, 5 g/L yeast extract and 10 g/L sodium chloride. LB agar plates were made by adding 1% (w/v) agar. When growing cultures for expression a Super Broth or SB media was used. which contained 32 g/L tryptone, 20 g/L Yeast Extract and, 10 g/L sodium chloride. Any media containing kanamycin sulfate or chloramphenicol were made to a concentration of 50 µg/mL.

3.2 Bioinformatics analysis of WssF

In order to begin elucidating WssF multiple programs were used to analyze the sequence of WssF for a predictive model. One such program was Protparam, which utilized the amino acid sequences found in **Appendix A** to predict various chemical and physical features of WssF (102). This process of analysis was also utilized for each of the constructs ordered or generated in lab (**Appendix A**). After Protparam, the sequences were further analyzed with the help of the BLAST (Basic Local Alignment Search Tool) servers (103). With the use of BlastP, a program used to search protein databases, the FASTA sequence of WssF was compared for homology against other proteins, including acetyltransferases, in order to predict possible active residues and regions containing similar structures. Through the use of Phyre2 (Protein Homology/AnalogY Recognition Engine Version 2) the FASTA sequence of WssF was used to create a predictive three dimensional model (PDB or Protein Data Bank file) of the proteins structure based on homology detection (104). The Consurf program was then used to determine amino acid conservation against potential homologues (105, 106). Consurf utilized PSI-Blast

(Position-Specific Iterative Basic Local Alignment Search Tool) to collect homologues from the UniProt database and the program Muscle for multiple sequence alignments (107–109). As a result, a conservation scoring system was used to colour code the Phyre2 predictive model (.PDB file) based on specific amino acid conservation. In order to visualize the predictive model, with conservation colouring, Pymol was used to display 3D models, which would be used to direct research decisions (110).

3.3 Expression and Purification of WssF

3.3.1 Isolation and replication of *wssF* plasmids

Two *wssF* genes were ordered from Genscript, with a pET24 C-Terminal His_6 tag, one from *Acromobacter insuavis* (pJPJTW5) and the second from *Pseudomonas fluorescens* (pJPJTW3). Plasmid stocks were replicated by inoculating *E. coli* Top 10 cells, grown in a Super Broth (supplemented with 50 µg/mL kanamycin) at 37 ˚C for 16 h. A Plasmid Mini prep kit (Axygen, Union City California or Qiagen: QIAprep Spin Miniprep Kit) was used to prepare 50 µL plasmid aliquots with roughly 180 ng/µL of DNA. These samples were stored at -20 ˚C until needed.

A 2 µl aliquot of *wssF* plasmid (pJPJTW5 or pJPJTW3) was transformed into $CaCl_2$ *E. coli* BL-21 competent cells (Novagen) using the standard heat shock method (111). Plasmid was added to a standard 1.5 mL microcentrifuge tube containing 100 µL of BL-21 competent cells, which was then incubated at 4 °C for 30 min. The sample was heat shocked (32 ˚C) for 1.5 min, before being left at 4 °C for 5 min. After mixing for 5 min the sample was added aseptically to 500 µL of pre-warmed (37°C) LB broth and incubated at 37 °C for 1 h shaking at 220 rpm. After transformation the cells were plated on LB agar (50 µg/mL kanamycin and chloramphenicol) before being incubated for 16 h at 37 °C. A single colony was taken from the plated cultures to

inoculate new cultures, of LB with a kanamycin and chloramphenicol concentrations of 50 µg/L. Cultures were used to create stocks of the transformed BL21 cells (cell cultures containing 15 % (v/v) glycerol) and stored at -80 ˚C until needed for future inoculations.

3.3.2 Expression of WssF

Protein expression proceeded with the goal of optimal production of carbon terminal tagged His_6 WssF proteins for *A. insuavis* (*Ai*WssF) and *P. fluorescens* (*Pf*WssF). In order to achieve this goal, overproduction of each protein was accomplished through the use of transformed *E. coli* BL21 (pET24) cells. Cultures were grown for 16 h at 37 ˚C shaking at 220 rpm and then used to inoculate larger SB cultures (1/50) for expression. Cultures grew at 37 ˚C shaking at 220 rpm until an OD_{600nm} of 0.6 was reached and then Isopropyl β-D-1-thiogalactopyranoside (IPTG; Biobasic) was added to a final concentration of 1 mM to induce expression from the pET plasmid constructs through activation of the T7 RNA polymerase within the *lac* operon. (112, 113) Induced cultures were grown for 16 h at 220 rpm and 30 ˚C. Post induction, the bacterial cultures were subjected to centrifugation for 15 min at 5000 x *g* and 4 ˚C. The supernatant was decanted, and cell pellets were stored at -20 ˚C until required for purification.

3.3.3 Purification of WssF

Lysis buffer (50mM Tris-HCl pH 8.0, 150mM NaCl, 2 % (v/v) glycerol (Biobasic)) was used to suspend thawed cell pellets. The cell suspension was then lysed through mechanical pressure using a cell disruptor (Constant Cell Disruption Systems, Deventry United Kingdom) at 17.4 kPsi with one passage through the system. The cell lysate was subjugated to centrifugation for 45 min, at 28 000 x *g* and 4 ˚C, to remove the insoluble fraction that included inclusion bodies. The clarified supernatant was subjected to Immobilized Metal Affinity Chromatography

(IMAC) for the purification of His-tagged proteins (114). Briefly, the supernatant was mixed with 1 mL of His-pur Ni-NTA Resin (Thermo Fisher Scientific) for 1 h. The sample was then applied to an open column and the resin was washed first with 25 mL of Lysis buffer followed by 25 mL of wash buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 2 % (v/v) glycerol, 50 mM imidazole (Bioshop)). Elution of bound protein from the column was accomplished by applying 5 mL of Elution buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 2 % (v/v) glycerol, 500 mM imidazole) to the column and allowing it to incubate with the column for 10 min prior to elution. The elution fractions containing WssF (24.6 kDa) were concentrated using 10 kDa molecular mass cut off filters (Vivaspin) to a volume of 1 mL by subjecting the samples to centrifugation at 4000 x *g* at 4 ˚C. Due to the size of the WssF proteins they would not pass through the filter allowing for the sample to be concentrated. Filtered samples were additionally subjected to centrifugation for 10 - 15 min at 13 400 x *g* at 4 ˚C to remove any insoluble particulate or precipitate. Buffer exchange was performed to remove the imidazole and exchange the 50 mM pH 8.0 Tris-HCl for 50 mM pH 7.0 Sodium Phosphate buffer. Other concentrations, such as, NaCl and glycerol, remained constant. Purified protein following these steps was assessed from Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) analysis, and routine functional assays were used with protein purified to this level. Protein concentrations were determined through the use of a Cytation 5 Cell Imaging Multi-Mode Reader (Biotek), which utilized the Beer lambert law of $\varepsilon = cl\lambda$ where the extinction coefficient (ε) 47690 M^{-1} cm^{-1} of WssF against absorbance measurements (λ) at 280 nm with a constant pathlength (l) of 1 cm.

For structural analyses, protein was further purified by fast protein liquid chromatography (FPLC) with a SuperDex 75 Size Exclusion column (SEC) to separate out contaminants based on size. Following equilibration of the SEC column in Gel filtration buffer (50 mM Tris-HCl pH

7.4, 150 mM NaCl, 2 % Glycerol (v/v)), 1 mL of protein sample was applied to the column with a flow rate of 1 ml/min (115). Fractions were collected in 2 mL aliquots and WssF-containing fractions eluted reliably from the column in the 11-14 mL fractions. SDS-PAGE was used to confirm the presence and purity of protein in all fractions.

3.3.4 Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE was used to determine the presence and purity of WssF throughout purification. SDS-PAGE gels were performed using standard procedures to separate molecules based on molecular mass. (116, 117) Routine gels consisted of a 12 % (v/v) separating gel (2 mL Acrylamide/Bis-Acrylamide (29:1) 30 % (v/v) solution (Bioshop), 1.675 mL of ddH_2O (double distilled water), 1.25 mL of 1.5 M pH 8.8 Tris-HCl, 50 µL of 10 % (w/v) SDS, 25 µL of Ammonium persulfate (Biobasic), 5 µL of Tetramethylethylenediamine or TEMED (Fisher Bioreagents)) and a 4 % (v/v) stacking gel (0.65 mL Acrylamide/Bis-Acrylamide (29:1) 30 % (v/v) solution, 3.05 mL of ddH_2O water, 1.25mL of 0.5M pH 6.8 Tris-HCl, 50 µL of 10% (w/v) SDS, 25 µL of Ammonium persulfate, 10 µL of TEMED). Purification fractions were mixed with 5 times concentrated SDS sample buffer, containing 1.54 µg/ml Dithiothreitol (DTT), in a 1:1 ratio. Mixed samples were heated to 90 ˚C for 10 min to denature the protein, cooled and then loaded onto the gel along with Precision Plus Protein™ Dual Color Standards, that was used as a mass comparison. After running, at 200 V for 45 min, the proteins in the gel were stained with Coomassie brilliant blue (0.1% (w/v) Coomassie R250, 10% (v/v) acetic acid and, 40 % (v/v) methanol) and destained in destain solution (10% (v/v) glacial acetic acid, 20 % (v/v) methanol and, 70 % (v/v) water) until the desired contrast was achieved (116).

3.4 Functional Characterization of WssF

3.4.1 Esterase activity assays

Routine monitoring and functional characterization of WssF activity was achieved using the colorimetric substrates 4-nitrophenyl acetate (pNP Ac) and 4-methylumbelliferyl acetate (4-MUB Ac). Assays were done in quadruplicate unless otherwise stated. Standard assays (unless noted otherwise elsewhere) consisted of using a pNP Ac substrate and were conducted in 96 well microtiter plates (Sarstedt) read at an absorbance of 405 nm (Cytation5) for 30 min in 1 min intervals following a 5 s initial linear shake to mix contents thoroughly. For 4-MUB Ac assays were tested in fluorescence compatible 96-well plates (Grenier BioOne) with excitation and emission wavelengths of 365 and 450 nm, respectively, used to monitor the assay. Esterase rates were corrected for the background rate of spontaneous hydrolysis of the substrate used. All assays were designed to minimize differences in well conditions between assays and run in quadruplicate unless otherwise specified. Well conditions had a maximum volume of 300 µL. Common Assay Conditions were 50 % (v/v) 100 mM pH 7 Sodium Phosphate (Fisher Chemical), 5 % (v/v) 4 µM protein (2mg/mL stock), 5 % (v/v) pNP Ac (EtOH) or 4-MUB Ac (DMSO) 80 mM stock, 40 % (v/v) ddH_2O water. pH 7 Sodium Phosphate buffers were made with 9.75 mL of 200 mM monobasic Sodium Phosphate, 15.25 mL of 200 mM dibasic Sodium Phosphate and, 25mL of H_2O. Assays were initially done with a maximum volume of 300 µL before being scaled down to 150 µL and 94 µL. Scaling of the assays allowed for less protein consumption, which allowed for more assays and/or replicates per mL of purified protein. Specific activity was reported as 1 unit of esterase activity, which was defined as the amount of enzyme required to release a µmol of *p*-nitrophenol min^{-1} (mg of protein)$^{-1}$.

3.4.1.1 Storage profiles

Short and long-term storage profiles were developed for both WssF constructs. Assays were conducted with 4-MUB Ac as substrate. When samples were stored at -20 ˚C or -80 ˚C the final well conditions contained 10 % (v/v) glycerol in place of an equal volume of ddH_2O water.

3.4.1.2 pH Profile

A pH profile of both WssF constructs was investigated to determine the optimal pH and buffer for activity assays. A 2-(*N*-morpholino) ethanesulfonic acid (MES) buffer was chosen between pH ranges of 5.5 - 6.5, sodium phosphate buffer for pH of 6.5 - 7.5 and Tris buffer for the final range of 7.5 - 9.25. The assay was performed in quadruplicate with 50 mM buffer of each of the various pH values in the presence of 4 μM purified WssF, and 6 mM *p*NP-Ac. The data are presented as percent activity of WssF compared to the highest value at pH 8.5.

3.4.1.3 Inhibitor Profile

An inhibitor profile was performed to determine if the compounds phenylmethanesulfonyl fluroide (PMSF) (a serine protease inhibitor), methanesulfonyl fluoride (MSF) (an acetylcholinesterase and serine protease inhibitor) and EDTA (metal chelator), would elicit any inhibition of WssF. Assays were completed using 5 mM pNP Ac as substrate after a 1 h incubation at 22 ˚C. All compounds were tested against both constructs with inhibitor concentrations ranging from 0 mM to 5 mM while maintaining 5 % (v/v) final well volume of inhibitor in place of an equal volume of ddH_2O.

3.4.1.4 Acceptor Profile

An acceptor profile was done on both WssF constructs using 4 mM pNP Ac as substrate. The acceptors utilized were cellobiose, cellotriose, cellotetraose, cellopentaose and cellohexose. Following this, varying lengths of a Chito-oligomer, from chito-biose to chitohexaose, was used

to test for the specificity of WssF. In both cases acceptor concentrations of 1 mM were attained while maintaining final well volume of 5% (v/v) acceptor replacing an equivalent volume of ddH_2O.

3.4.2 Site Directed Mutants

Ser17, Asp188 and, His191, were replaced with Ala by site-directed mutagenesis of the *A. insuavis wssF* gene. Primers used for this mutagenesis are listed in **Table 3.1** mutagenesis was carried out using JPJTW5 as a template for pfu Turbo DNA polymerase according to the QuickChange site-directed mutagenesis kit. Following PCR, 1 mL of *Dpn*1 was added to a 50 µL PCR mixture and incubated for at least 1 h at 37 °C to remove the original, methylated template. The resulting plasmids were used to transform TOP10 Chemically Competent *E. coli*, and clones were screened for the correct mutations by DNA sequencing. Each of the protein products was purified to apparent homogeneity, as described above for wild-type *Ai*WssF. New purification buffers were used for each mutant purification to avoid potential contamination of the various mutant forms of the enzymes with each other.

Table 3.1 SDM Primers sequences. Primer sequences for mutants *Ai*WssF S17A, D188A and, H191A.

Mutation	Primer sequence (5'-3')
*Ai*WssF_S17A Fwd	CTGCTGATCGGTGATGCCCATCTGGCGGCCCC
*Ai*WssF_S17A Rev	GGGGCCGCCAGATGGGCATCACCGATCAGCAG
*Ai*WssF_D188A Fwd	GGCCGACCACCGAGGTCAGCACTATACCAATACGGGC
*Ai*WssF_D188A Rev	GCCCGTATTGGTATAGTGCTGACCTGCGGTGGTCGGCC
*Ai*WssF_H191A Fwd	GGCCGACCACCGATGGTCAGGCTTATACCAATACGGGC
*Ai*WssF_H191A Rev	GCCCGTATTGGTATAAGCCTGACCATCGGTGGTCGGCC

*Highlighted codon represents the alanine point-mutation

3.4.3 Thin Layer Chromatography (TLC)

In order to test whether both constructs were capable of cellulose modification Thin Layer Chromatography (TLC) was utilized. After running a specific activity assay on either construct in the presence of cellulose as an acceptor for 24 h, samples were filtered using CarboGraph Columns. The activated carbon columns required a preparation protocol before samples could be filtered to remove everything except the cellulose (**Appendix B**). The Preparation protocol utilized 1 column volume (CV) of 100 % (v/v) acetonitrile and 2 CVs of ddH_2O. At this point the column was ready to filter the samples using a filtration protocol (**Appendix B**). Briefly, reaction samples were diluted by adding 150 µL of sample to ddH_2O to create a 1 mL final volume. The diluted mixture was added to the column for absorption before a 2 CV wash with water and 1 CV elution with 50 and then 100 % (v/v) acetonitrile. After collecting the elution of cellulose, the activated carbon columns were cleaned with a CarboGraph cleaning protocol (Appendix B). Briefly, columns were cleaned using 2 CVs of 50 % (v/v) Tetrahydrofuran (THF) and washed with 2 CVs of ddH_2O. Finally, TLC was performed on the now filtered samples of cellulose (Appendix B). A 10 x 10 cm aluminum-backed silica TLC plate was spotted with samples using precision capillary tubes (0.4 mm x 75 mm). An 8:3:4 Mobile phase (53.33 mL ethyl acetate, 20 mL ddH_2O, 26.67 mL Methanol) was used to separate samples in a rectangular developing chamber (General GlassBlowing Company Inc.). Following separation, the plate was air-dried and sprayed using a 25:3:2 Naphthol Stain (208.33 mL Methanol, 25 mL H_2SO_4, 16.67 mL ddH_2O) before being heated to 100 °C for visualization of the stain. Images were taken in top light mode in a Versadoc molecular imager (Biorad).

3.4.4 Mass Spectrometry

Following TLC, 75 µL (approximately 300 µM) of the cellulose samples, filtered through the active carbon columns, were sent for mass spectrometry. All samples sent for mass spectrometry were filtered through activated carbon columns in order to retain only the modified or unmodified cellulose sample. The samples were analyzed at the University of Guelph by Dyanne Brewer on a Bruker AmaZon SL LC-MSn instrument (118). Negative controls and 24 h time point samples were sent for both *Ai*WssF and *Pf*WssF constructs. Following this each of the *Ai*WssF site direct mutants (S17, D188, H191) and the previous *Ai*WssF wild type as a control control sample were sent for mass spectrometry. The size range scanned began at 400 *m/z* and ended at 1000-1200 *m/z*, which allowed for identification of both modified (850 Da) and unmodified cellulose (828 Da) adducts.

3.4.5 High Throughput Screening

While attempting to structurally and functionally elucidate WssF, stocks of *Ai*WssF protein (2 mg/mL) were sent to the SPARK facility at the University of Toronto. High Throughput Screening (HTS), done by Leanne Wybenga-groot, was utilized in determining potential inhibitory compounds. Fluorescence assays were run on 384-well microplates using 4-MUB Ac as an acetate donor analogue. Before any inhibitor compounds were added to a screen Z' scores were determined for positive control plates. Next a set of control plates were left with protein at room temperature to be read periodically throughout the day. Well conditions contained 50 mM pH 7 sodium phosphate with 0.3-1.5 µM AiWssF protein with 0.5 mM 4-MUB Ac at 5 % (v/v) DMSO of the final well volume of 47 µL. These plates were to control for any loss in activity seen throughout the day and determine the effect on Z' values. In addition, the control plates were used to determine if wells closest to the edge of the plate would see non-

uniform readings. Finally, when measuring the inhibitory effect of various compounds on esterase activity, Z scores were taken after 10 minutes a sufficient amount of time for the Z′ of the reaction to reach a reliable threshold of 0.7. In other words, the reaction was allowed to procced far enough for a sufficient difference between the positive and negative wells to produce a large statistical difference or effect. A 2 µL aliquot of compound was added to each well for a final concentration of 10 µM. Positive control wells were 1A - 1P and 2A - 2P, which lined one end of the plate, while wells 23A - 23P and 24A - 24P were used as negative controls lining the opposite end of the plate. Primary screening of approximately 64 000 compounds was performed in single trials and was followed by secondary screening of 103 hits (in triplicate) along with the addition of potential inhibitors, for similar putative acetyltransferases (WssI, PatB and, OatA), resulting in a total secondary screen of 146 compounds. Next, dose response screens were conducted on 83 compounds from the secondary screen. Dose response assays tested each of the selected potential inhibitory compounds from 0-100 µM. Finally, 18 compounds were selected for in-house testing, where dose response assays were run with 3 µM protein concentration.

3.5 Structural Characterization of WssF

3.5.1 Crystallization and Diffraction

Efforts to structurally characterize WssF focused initially on crystallography. Multiple screens were used in an attempt to find successful crystallization conditions. Crystal reservoir solutions were created from Top 96, MCSG-1, 2, 3, and 4 (Microlytic) premade formulations with the addition of *Ai*WssF (pJPJTW5) protein. Protein stocks of *Ai*WssF were comprised primarily of the buffer used during SEC (50 mM Tris-HCl pH 7.4, 150 mM NaCl, 2 % (v/v) Glycerol). The screens were created using a Gryphon 96 head dispenser (Art Robbins Instruments) to dispense the 96-well plate screens in an efficient and organized manner. 1 - 2 µL

of protein (10 - 29 mg/mL) was added to create a sitting drop with 1 - 4 µL from the well condition. The reservoir of each well contained 60 µL of a condition obtained from the preformulated screens. Crystal screens were observed, using an Olympus SZX16 Stereomicroscope, for successful crystal hits.

Successful crystal hits were then frozen, so they could be sent to the Canadian Light Source (CLS) for X-ray diffraction analysis. In order to cryo-protect the crystals, 50% (v/v) glycerol was added to the drop were the crystal was located. Next the crystals were picked or scooped out of their drops using a loop from Mitegen. Next, the crystals were submerged into a second premixed drop (1:1, reservoir solution: 50 % (v/v) glycerol). Finally, the loop and crystal would be placed into a container of liquid nitrogen for flash freezing and storage until analysis at the CLS. Diffraction data sets were collected using synchrotron radiation on the 08B1-1 beamline at the Canadian Macromolecular Crystallography Facility (Canadian Light Source (CLS), Saskatoon) using a CCEL MD2 microdiffractometer and MarMosaic mx300 CCD X-ray detector. Typical datasets at the CLS consisted of 360 images at 1˚ oscillations with a 250 nm crystal-to-detector distance and an exposure time of 0.2 s per image. The data were integrated, reduced and scaled using AutoProcess (119). Additional datasets were also collected at the University of Waterloo using a Rigaku rotating copper anode X-ray generator and an R-axis IV++ detector. These datasets consisted of 360 images at 1˚ oscillations with a 275 nm crystal-to-detector distance and an exposure time of 2 min per image. The data were integrated and scaled using HKL-2000 (120).

Expansion plates were made manually to recreate and expand on the previously successful conditions. Examples of expansion plate layouts can be found in **Appendix B**. Gradients of PEG concentration and protein concentration were created directly around those

used in the successful crystal conditions. For the Top 96 formulations, protein concentrations of 17, 15 and, 13 mg/ml were used with PEG 3350 concentrations of 20, 23, 25, 27 and, 30 % (w/v) to create an expansion of the successful condition. To expand the conditions of well H10 from MCSG1 protein concentrations 29.5 and 25 mg/ml were used with PEG concentrations of 20, 23, 25, 27 and, 30 % (w/v). Next, an expansion of well condition F1 from screen MCSG1 was created from mixing 29.5 and 25 mg/ml with PEG 3350 concentrations of 15, 18, 20, 22, 25 % (w/v). In addition to changing protein concentration a set of expansion plates were made with variations of pH and PEG concentrations. PEG concentrations ranged from 19 to 31 % (w/v) with pH ranging from 5.6 to 8.4.

3.5.2 Molecular Replacement

Molecular replacement was attempted to solve phases for structure determination. The primary molecular replacement programs used were MRage and/or Phaser modules of the Phenix suite (121). Alternatively, automated molecular replacement was also attempted through the CCP4 online interface with the BALBES, MrBump, and MoRDa programs (122–125). Either way molecular replacement was conducted using PDB files from structures with possibly significant levels of homology to *Ai*WssF found in Phrye[2] and BLASTP searches, the top 5 hits were recorded and analysed for possible usage (see **Tables 4.2 and 4.3** in the Results section).

3.5.3 Small Angle X-Ray Scattering

While working on solving the phase problem, efforts were made to determine generic structural characteristics through the use of Small Angle X-Ray Scattering or SAXS. This was achieved by sending protein samples (1, 5, and 10 mg/mL) to the SIBYLS beamline (12.3.1) at the Advanced Light Source part of the Lawrence Berkeley National Laboratory (126). The X-ray wavelength was 1.0 Å, with a flux of 1013 photons per second, and a sample-to detector distance

of 1.5 m. Scattering images were collected using a Pilatus 2 M detector every 0.3 s, with a total of 33 samples images. All sampling was performed at 10 °C on multiple samples at various protein concentrations (1, 5, and 10 mg/mL) to remove concentration-dependent patterns (127). Negative control samples were also collected for background subtraction (128, 129). Every collected image was circularly integrated then normalized for beam intensity to generate a 1-dimensional scattering profile. Background subtraction of each profile was done against the negative controls to give two background subtracted sample profiles. Averaging of the subtracted profiles was achieved through the use of Frameslice (with the assistance of William Scott, Weadge Lab).

With the SAXS profiles corrected for concentration dependence and background intensities, further analysis was required for interpretation of the data. SCÅTTER was used to generate Rg (radius of gyration), Rc (cross-sectional radius), *Q* (Porod volume), and D_{max} (Dimension Maximum) values that best matched the averaged and background subtracted 1-dimensional scattering profiles (130, 131). These values were determined through linearization of the scattering profiles to create scatter curves and pattern models (132). Multiple automated transformations and calculations were applied to the scattering profiles (Guinier, Zimm, Holtzer, Kratky, Porod). From here the values obtained from SCÅTTER were exported to GNOM. GNOM was used to adjust where along the scatter curve the 3D model would begin and end processing of the SCÅTTER profile. GNOM provided a total estimate value, which varied from 0 to 1, with 1 representing ideal P(r) (Pair distance distribution function), R_g and, I_0 (Molecular mass) values (133). The total estimate value was used to determine the optimal positions to start and end modelling of the scatter curve.

Generation of 3D models required multiple programs for imaging and conversion. First the program GASBOR was used to image the 3D model parameters refined in GNOM (134). GASBOR operated through the online data analysis server ATSAS 2.8.4 (135). Finally, DAMAVER, run through the SIBYLS beamline website for SAXS applications, was utilized in conjunction with GASBOR to produce averaged PDB files that were visualized in Pymol (136).

Once exported to Pymol the legacy plugin SASpy was used to compare and orient the SAXS model to the Phyre2 model previously created (137). SASpy function crysol was used to fit the SAXS data with the predictive model while the functions sasref and supalm were used to refine and superimpose the two models respectively.

4. Results

4.1 Bioinformatics

Bioinformatic analysis for both *A. insuavis* and *P. fluorescens* full length WssF proteins and C terminal His_6 constructs was completed. Protparam was used to determine various properties for all four sequences. As can be seen in **Table 4.1,** molecular mass (23495 - 24560.1Da), theoretical pI (8.79 - 8.58), extinction coefficients (47690M^{-1} cm^{-1}) and absorbance (2.03 - 1.942 g/l) were all predicted based on the primary amino acid sequences (provided in **Appendix A**). Both full proteins and His_6 tagged constructs were predicted to be stable with instability scores ranging from 21.51 to 22.65. Of note, each construct analyzed had four cysteine residues and six methionine residues. Cysteine residues are important to consider as they can participate in disulfide bond formations that effect the folded structure of the protein (138). A disulfide bond has been found in AlgX a homolog of WssF (69). Finally, Methionine residues

would be instrumental in deciding how to proceed during structural characterization (Selenomethionine) (139).

Table 4.1 ProtParam Protein Output. Predicted chemical and physical features for Full *Ai/Pf* WssF Proteins and C terminal His_6-tagged constructs.

ProtParam Output	*Ai*WssF Full protein	*Ai*WssF pET24-CtermHis$_6$	*Pf*WssF Full protein	*Pf*WssF pET24-CtermHis$_6$
Number of amino acids	221	229	221	229
Molecular mass	23495	24560.1	23495	24560.1
Theoretical pI	8.79	8.58	8.79	8.58
Extinction coefficients (in units of M^{-1} cm^{-1}, at 280 nm measured in water)	47690	47690	47690	47690
Absorbance (g/l)	2.03	1.942	2.03	1.942
Instability index	22.4	21.26	22.65	21.51
Cystine residues	4	4	4	4
Methionine residues	6	6	6	6

Primary amino acid sequences were also utilized to predict the functionally of the WssF proteins. Comparing the amino acid sequence of WssF against the National Center for Biotechnology Information's (NCBI) Conserved Domain Database (CDD) predicted an oxyanion hole formed by residues 17 (Serine), 69 (Glycine) and, 103 (Threonine) (140). The CDD also predicted a catalytic triad formed by residues 17 (Serine), 188 (Aspartic acid) and, 191 (Histidine). These predictions are made based on the conservation of specific amino acids compared to a position-specific score matrices or PSSM (140, 141). Proteins are then placed into families, within the NCBI database, that share functionality and sequence conservation. As a result of the CDD predictions, WssF is predicted to share functionality with the SGNH hydrolase family, which also contains members of an acetyltransferase subgroup (AlgX in particular) (69). Active site prediction based on sequence alignment supports the possibility that WssF proteins are capable of esterase activity. Ester hydrolysis would be required as the first step in the

functionality of WssF as an acetyltransferase. In order to further support this hypothesis, tertiary structure models and amino acid conservation scores were determined. Phyre2 was used to model a hypothetical structure for both truncated WssF constructs based on the provided amino acid sequences. Low sequence identity was achieved with the highest matches reaching 19% for *Ai*WssF and 18% for *Pf*WssF (**Tables 4.2 and 4.3**). For those that had the highest identity, confidence was close to 99% with 90% coverage.

Table 4.2 Top 5 Phyre2 Predicted Homology Results for *Ai*WssF.

PDB	Coverage (%)	Confidence (%)	% i.d.	PDB molecule	PDB title
1.c2o14A	89	99.9	16	hypothetical protein yxim	X-ray crystal structure of protein yxim_bacsu from bacillus2 subtilis. northeast structural genomics consortium target3 sr595
2.c3p94A	92	99.9	13	gdsl-like lipase	Crystal structure of a gdsl-like lipase (bdi_0976) from2 parabacteroides distasonis atcc 8503 at 1.93 a resolution
3.c3milA	90	99.9	19	soamyl acetate-hydrolyzing esterase	Crystal structure of isoamyl acetate-hydrolyzing esterase from2 saccharomyces cerevisiae
4.c2vptA	88	99.9	16	lipolytic enzyme	Clostridium thermocellum family 3 carbohydrate esterase
5.c3hp4A	88	99.9	17	gdsl-esterase	Crystal structure of psychrotrophic esterase esta from2 pseudoalteromonas sp. 643a inhibited by3 monoethylphosphonate

Table 4.3 Top 5 Phyre2 Predicted Homology Results for *Pf*WssF

PDB	Coverage (%)	Confidence (%)	% i.d.	PDB molecule	PDB title
1.c4lhsA	87	99.9	18	uncharacterized protein	crystal structure of a hypothetical protein (bacova_00914) from2 bacteroides ovatus atcc 8483 at 1.40 a resolution
2.c2o14A	93	99.9	17	hypothetical protein yxim	X-ray crystal structure of protein yxim_bacsu from bacillus2 subtilis. northeast structural genomics consortium target3 sr595
3.c4iyjA	91	99.9	15	gdsl-like protein	Crystal structure of a putative acylhydrolase (bacuni_03406) from2 bacteroides uniformis atcc 8492 at 1.37 a resolution
4.c2vptA	88	99.9	18	lipolytic enzyme	Clostridium thermocellum family 3 carbohydrate esterase
5.c3p94A	89	99.9	16	gdsl-like lipase	Crystal structure of a gdsl-like lipase (bdi_0976) from2 parabacteroides distasonis atcc 8503 at 1.93 a resolution

Next Consurf was used to determine amino acid conservation with potential homologues. Consurf utilized PSI-Blast to collect homologues from the UniProt database and Muscle for multiple sequence alignments. Conservation scores were determined utilizing the empirical Bayesian paradigm, which is used to determine the rate of evolutionary conservation at each amino acid site (142). Consurf scoring is broken down into 9 categories where 1 is the least conserved and 9 is the highest conserved residues (105). Although a site may be rated the highest conserved score this does not mean it is 100 % conserved only that it is the most conserved in the sequence. Investigation of conservation among top homologues showed the highest conservation scores of 9 for 17 (Serine), 188 (Aspartic Acid), 191 (Histidine) and surrounding residues, which form the proposed active site of WssF. The hypothetical models were coloured from cyan (low conservation) to purple (high conservation) and are represented in the WssI model in Figure 6.2. All of the 50 closest homologues were predicted or are shown to be SGNH hydrolases, have SGNH hydrolase like domains, or are an esterase of some form. This supports the possibility that WssF has an SGNH hydrolase like domain, which facilitates the ability of the protein to undertake esterase. Both constructs are also predicted to have the prototypical alpha/beta/alpha fold of an SGNH esterase **(Figure 4.1)**. However, the beta sheet of *Ai*WssF is predicted to only contain 4 parallel strands instead of 5, which normally make up the beta sheet found within SGNH hydrolases. Surface models **(Figure 4.1)** show areas of high conservation within the predicted active sites, which form a pocket within the hypothetical structure. Limited conservation is seen throughout the surface models outside the predicted active sites representing the low sequence identity previously determined.

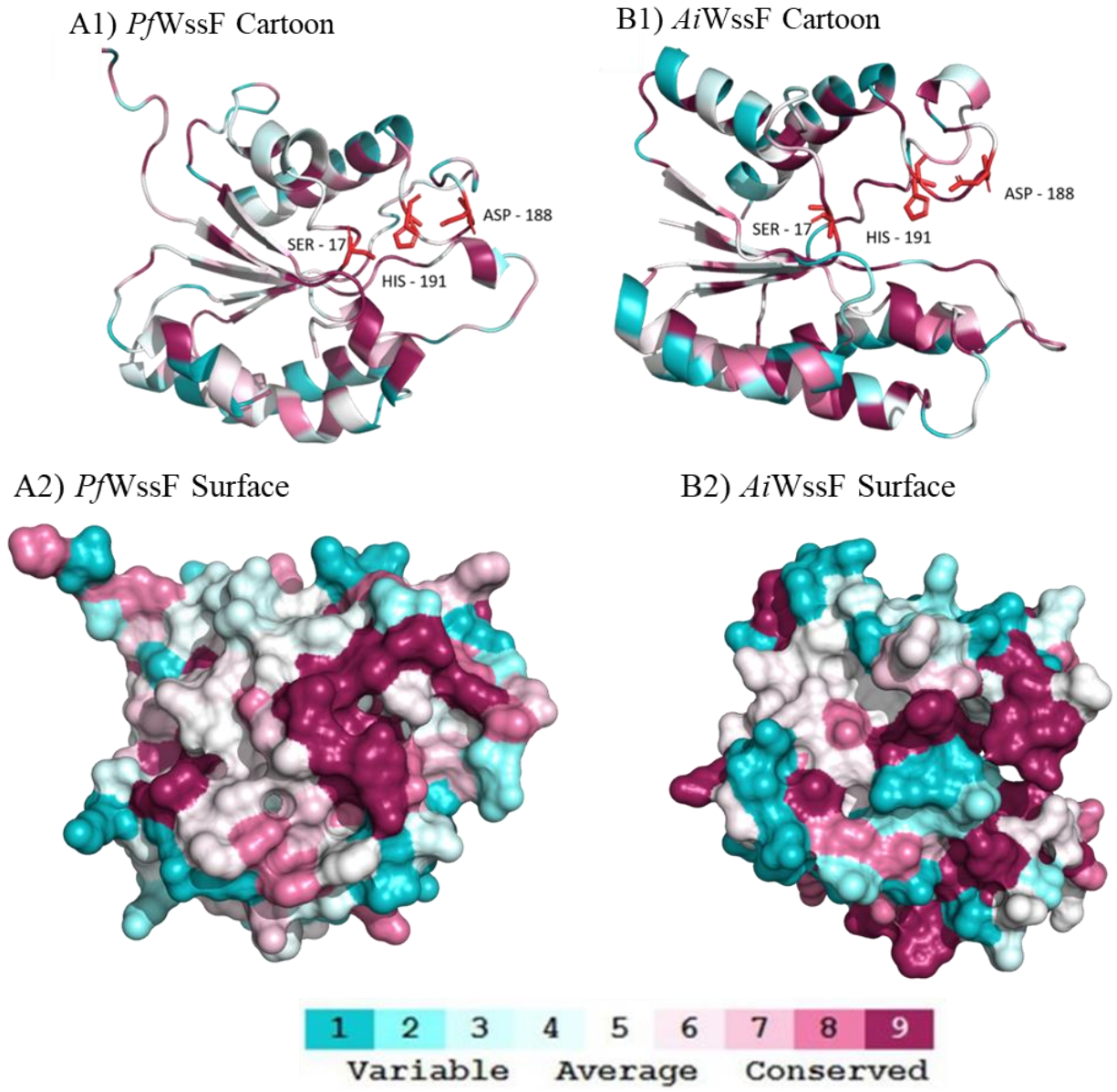


Figure 4.1 Phyre2 Predicted Structural Models of WssF. Phyre2 models of WssF (based on PDB: c2o14A, 16 % amino acid identity) in its native state adopt an alpha/beta/alpha motif, shown in 2 separate angles, for both *P. fluorescens* and *A. insuavis* constructs (A, & B). *Pf*WssF cartoon model with stick active site residues (A1), *Ai*WssF cartoon model with stick active site residues (B1), *Pf*WssF surface model (A2), *Ai*WssF surface model (B2). Each construct is colour coded using shades of purple to represent high conservation and green to represent low conservation. All images were rendered in PyMOL. (104, 110)

Analysis using Clustal Omega found a sequence alignment along the catalytic active site with one deviation of a single nucleic acid site H191 (**Figure 4.2**). Models for AlgX (PDB:4knc (yellow)) and AlgJ (PDB:4O8V(green)) were aligned with the predictive model of *Ai*WssF

(blue). The Carbohydrate binding module (CBM) domain of AlgX was removed to aid in visualization of the three aligned protein models. As was expected the CBM domain of AlgX did not align with either WssF or AlgJ. Normalized Spatial Discrepancy (NSD) values were determined by SASpy for the alignment of WssF/AlgX (2.36) and WssF/AlgJ (2.23). Pymol alignment of the models gave root-mean-square deviation (RMSD) scores of 13.86 for WssF/AlgX, 4.78 for WssF/AlgJ and, 1.80 for AlgX/AlgJ. SASpy was used for model alignment due to a better fit of the NSD values compared to RMSD values from Pymol alignment alone.

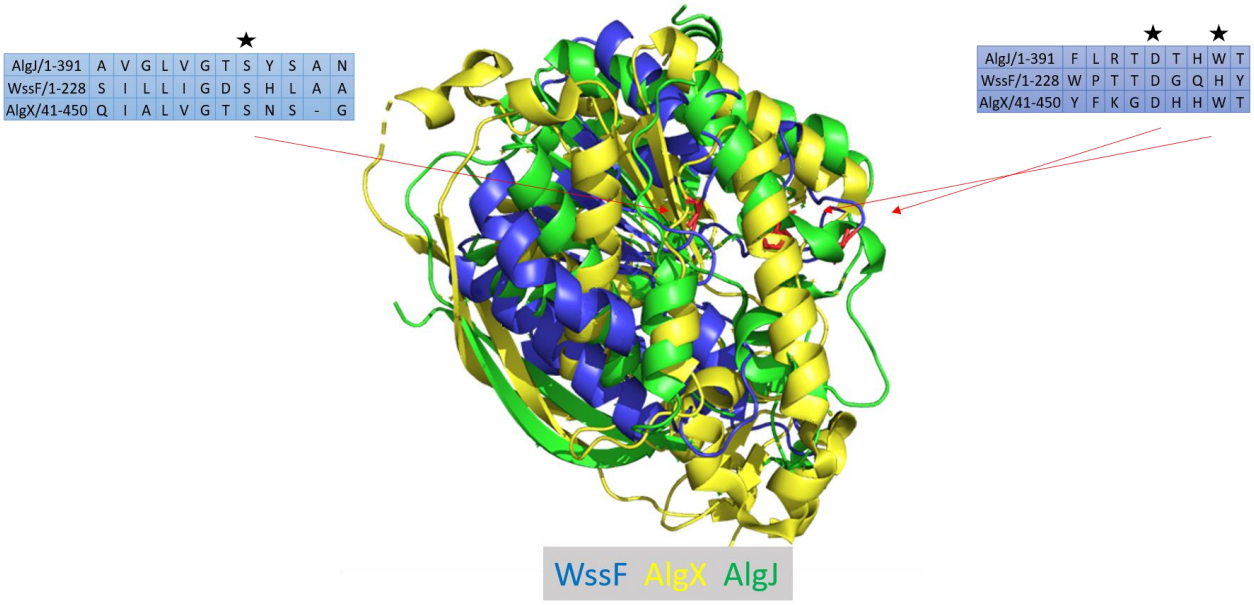


Figure 4.2 Three-Dimensional Alignment of WssF Homologs. Three-dimensional alignment of predicted structure of WssF(blue), AlgX (PDB:4knc (yellow) and, AlgJ (PDB:4O8V) (Green). CBM domain of AlgX has been removed for visualization. Conserved active site residues S17, D188 and, H191 are depicted in red to outline the predicted active site. Figure rendered in PyMol and superimposed by SASpy. Normalized Spatial Discrepancy (NSD) value for WssF and AlgX alignment was 2.36. NSD for WssF and AlgJ alignment was 2.23.

4.2 Profile of Optimal Protein Expression and Purification (Objective 1)

4.2.1 Protein Expression

Bioinformatic analyses to predict the structure and function of WssF, was done concurrently with optimization of protein expression and purification. Optimal protein expression was determined through various trials, where the duration and temperature of expression were found to be important. A wide range of temperatures for expression were attempted and 16 ˚C, 26 ˚C and, 37 ˚C expressions resulted in insufficient protein production. Expression at 30 ˚C for 4 h led to moderate protein production, whereas a 16 h expression at 30 ˚C gave a strong band as seen in **Figure 4.3**.

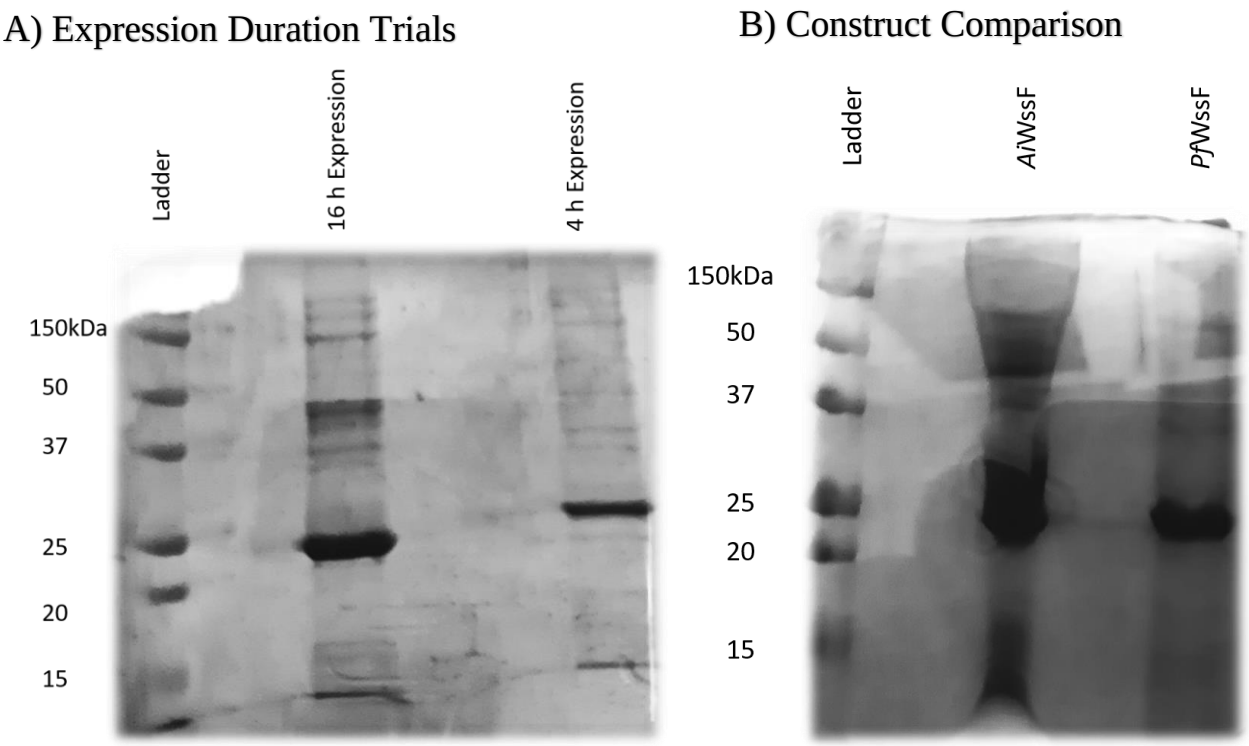


Figure 4.3 SDS-PAGE Protein Expression of *Ai*WssF and *Pf* WssF. SDS-PAGE analysis (12 % (v/v)) of *Ai*WssF expression of 1 L cultures incubated at 30°C and 200 rpm (A). Lane 1, molecular-weight markers (kDa); Lane 2, 16 h expression; Lane 3, 4 h expression. SDS-PAGE analysis (12 % (v/v)) of *Ai*WssF and *Pf*WssF expression of 1 L cultures incubated for 16 h at 30°C and 200 rpm (B). Lane 1, molecular-weight markers (kDa); Lane 2, *Ai*WssF; Lane 3, *Pf*WssF.

4.2.2 Protein Purification

Both versions of WssF were subjected to an initial purification round that involved immobilized metal affinity chromatography (IMAC) that took advantage of the His_6 tags on the C-terminus of the expressed constructs. Briefly, WssF was left to incubate with Ni-NTA resin. Bound protein was then washed with increasing imidazole concentrations to remove binding contaminates **Figure 4.4**. Purity of WssF during purification can be seen in **Figure 4.4A** and was sufficient for functional characterization. After primary purification samples were typically found to retain contaminate bands and a possible protein dimer, as seen in Lane 5 of **Figure 4.4B**. SEC through FPLC was used in an attempt to further remove contaminants and increase the purity of the protein samples (**Figure 4.4**; Panel B, lane 6). Protein production yielded a consistent 3 mg of purified protein per L of culture expressed when the optimal conditions were met. With the successful and reliable production of protein from both constructs optimized, there was enough protein for subsequent characterization assays (outlined in sections that follow).

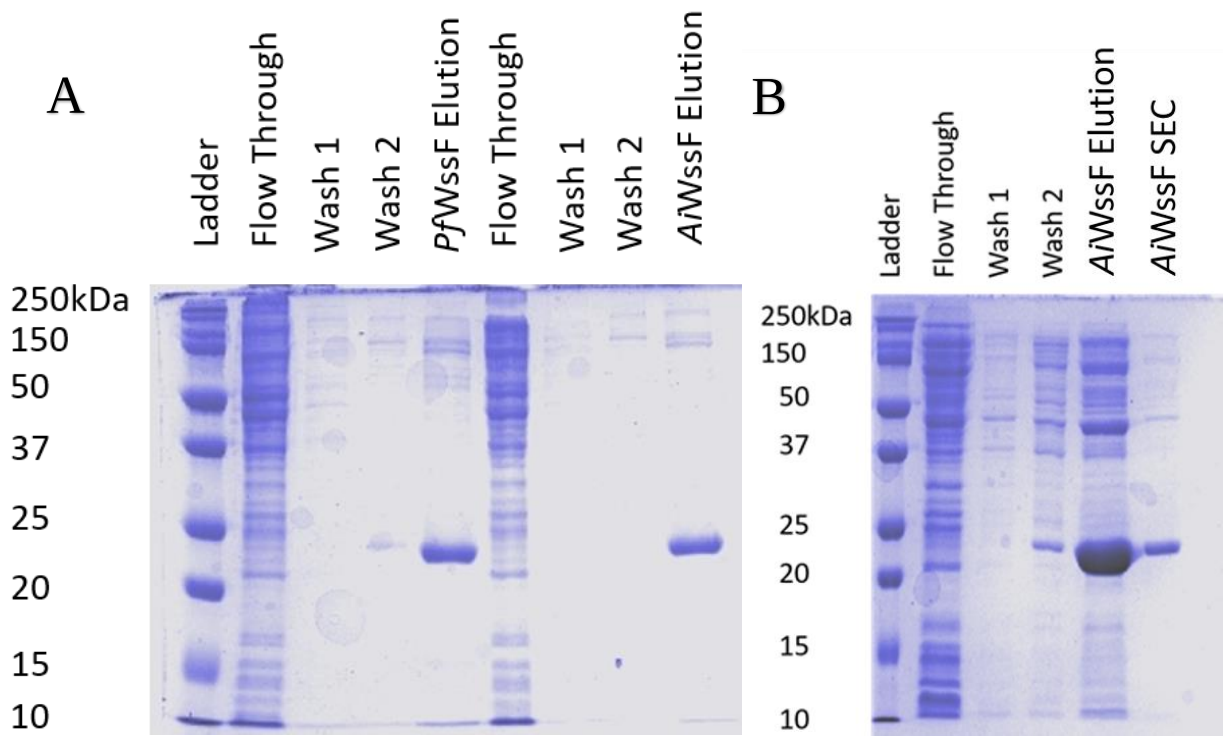


Figure 4.4 SDS-PAGE Analysis of the Purification of Expressed WssF protein samples. SDS-PAGE analysis (12 % (v/v)) of *Ai* and *Pf* WssF Ni-NTA IMAC purification fractions (A). Lane 1, molecular-weight markers (kDa); lane 2, Ni-NTA column unbound lysate; lane 3, wash with lysis buffer (no imidazole); lane 4, wash buffer (50 mM imidazole); lane 5, elution buffer (500 mM imidazole); lane 6, Ni-NTA column unbound lysate; lane 7, wash with lysis buffer (no imidazole); lane 8, wash buffer (50 mM imidazole); lane 9, elution buffer (500 mM imidazole) Typical SDS-PAGE analysis (12 % (v/v)) of *Ai*WssF Ni-NTA IMAC purification fractions and Size Exclusion chromatography (B). Lane 1, molecular-weight markers (kDa); lane 2, Ni-NTA column unbound lysate; lane 3, wash with lysis buffer (no imidazole); lane 4, wash 1 (50 mM imidazole); lane 5, elution buffer (500 mM imidazole); lane 6, combined SEC elution fractions.

4.2.3 Further Purification Optimization

Although a working optimum of protein yield had been achieved, time was also taken to slowly determine improvements for yielding pure protein. A fresh plasmid preparation from a Qiagen midiprep kit along with the use of Rosetta™ (DE3) Competent Cells found improvements to protein purity and yield as seen in **Figure 4.5**. An added benefit was that the potential dimerization band seen earlier was found to no longer exist.

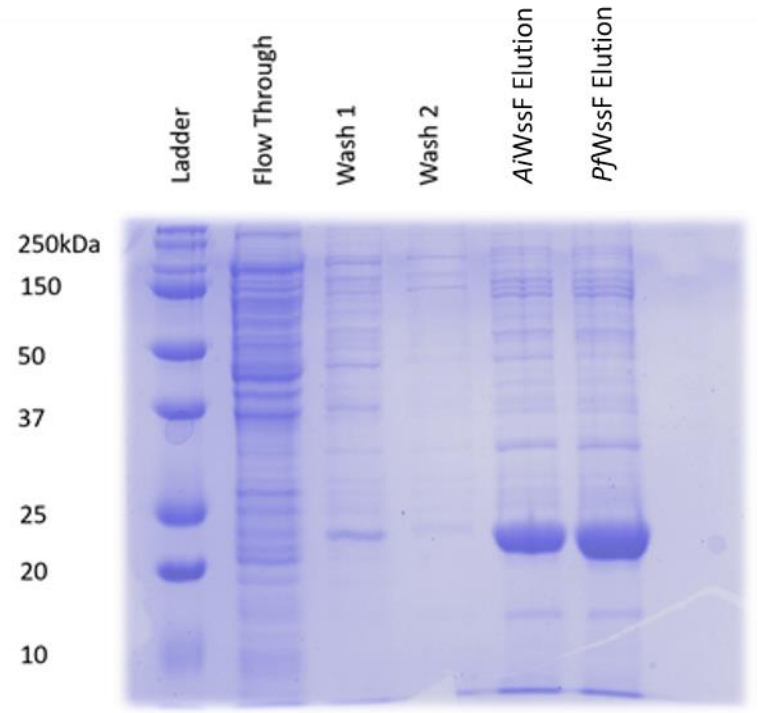


Figure 4.5 Purification of Expressed *Ai*/*Pf*WssF protein sample. SDS-PAGE analysis (12 % (v/v)) of *Ai* and *Pf* WssF Ni-NTA purification fractions. Lane 1, molecular-weight markers (kDa); lane 2, *Ai*WssF Ni-NTA column unbound lysate; lane 3, wash with lysis buffer (no imidazole); lane 4, wash buffer (50 mM imidazole); lane 5, *Ai*WssF elution buffer (500 mM imidazole); lane 6, *Pf*WssF elution buffer (500 mM imidazole).

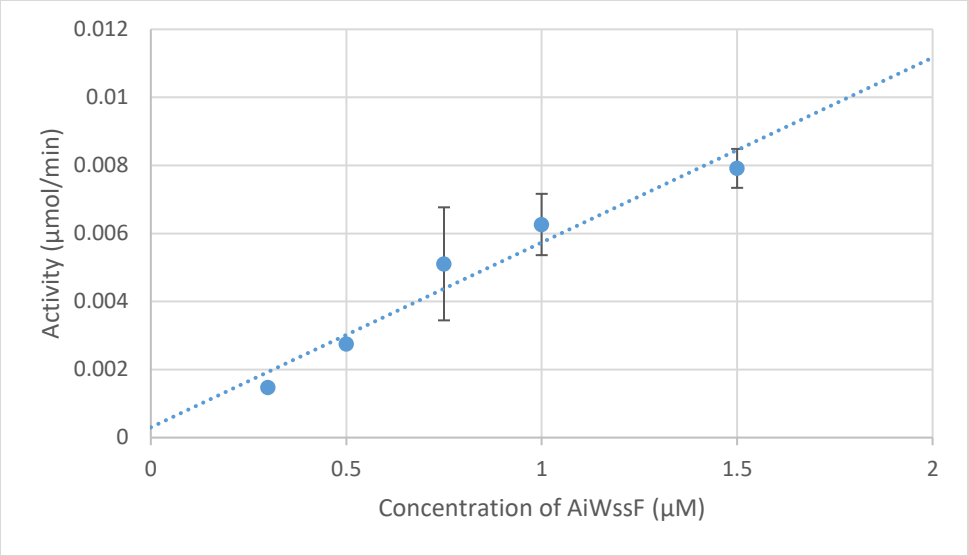
4.3 Functional Characterization of WssF (Objective 2)

4.3.1 Verification of WssF Esterase activity

The mechanism of action for WssF was explored through various *in vitro* assays. As a putative acetyltransferase the function of WssF was able to be tested in two separate reactions. First hydrolysis would be required for initial cleavage of the acetyl donor substrate followed by transferase activity or O-acetylation of cellulose. Predictions of potential active site residues, based on homology of WssF with other O-acetylation proteins, made testing for esterase activity the next logical step (69, 143). As such, purified protein was added to reaction mixtures with 4-MUB Ac or 4-Nitrophenyl acetate (*p*NP Ac) acting as a substrate. Testing for the potential esterase activity of WssF was done through adding increasing concentrations of protein to

reaction mixtures containing a constant concentration of 2 mM 4-MUB Ac or pNP Ac **(Figure 4.6)**. Changing the concentration of protein directly affected the concentration of 4-methylumbelliferyl or 4-nitrophenol produced based on the shifts in the florescence and colourimetric profiles. These results confirm that the esterase activity observed was due to the presence of WssF, since the appropriate controls lacked this correlation, and that the substrates could be measured consistently throughout the colourimetric assay. Assays were performed for 30 min, but only the linear portion of the curve was used for calculations in this and subsequent sections. Specific activity values for *Ai*WssF was determined for the hydrolysis of both 4-MUB Ac (1.46 ± 0.040 μmol*min^{-1}*mg^{-1}) and *p*NP Ac (0.039 ± 0.0057 μmol*min^{-1}*mg^{-1}). The measurable presence of substrate hydrolysis was suitable for further functional characterization of WssF.

A) 4-MUB Ac



B) *p*NP Ac

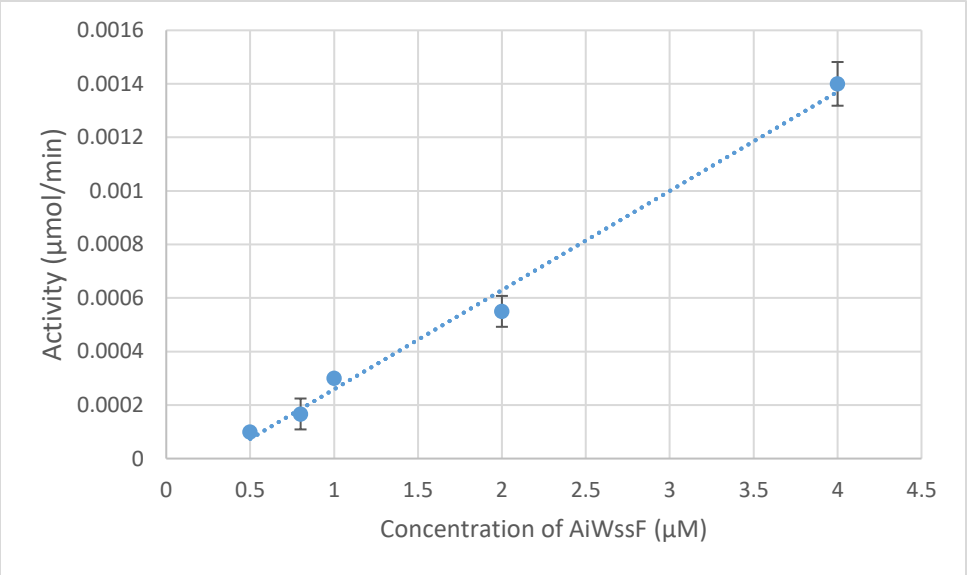


Figure 4.6 Rate of Reaction with Increasing Concentrations of WssF and specific activity with Acetyl Donors pNP-Ac and 4-MUB-Ac. A) Reaction conditions performed in 8 replicates, which included 50mM sodium phosphate buffer pH 7, 0.3 - 1.5 µM *Ai*WssF and 2 mM 4-MUB Ac. B) Reaction conditions performed in quadruplicate included 50mM sodium phosphate buffer pH 7, 0.5 - 4 µM *Ai*WssF and 2 mM pNP Ac.

4.3.2 Protein Storage

The use of 10% (v/v) glycerol was assayed as a potential suitable cryo-stabilizing agent for the long-term storage of purified WssF. Interestingly, the presence of glycerol in standard esterase assays, with 4-MUB-Ac as substrate, led to increased esterase activity. In fact, the activity approximately doubled in the presence of 10 % (v/v) glycerol **(Figure 4.7)** after accounting for the appropriate controls (*ie.* assay conditions in the absence of protein, but presence of glycerol). Protein stocks containing 10 % (v/v) glycerol were always diluted into reaction buffers so that the glycerol concentration was never higher than 0.5 % (v/v). This dilution was done so that the presence of glycerol was consistent across all of our assays (especially the HTS assays where minimization of components like glycerol was required to prevent unwanted interactions/complications with the compounds assayed). The only exemptions to this was the short and long-term storage assays, which contained 10 % (v/v) glycerol.

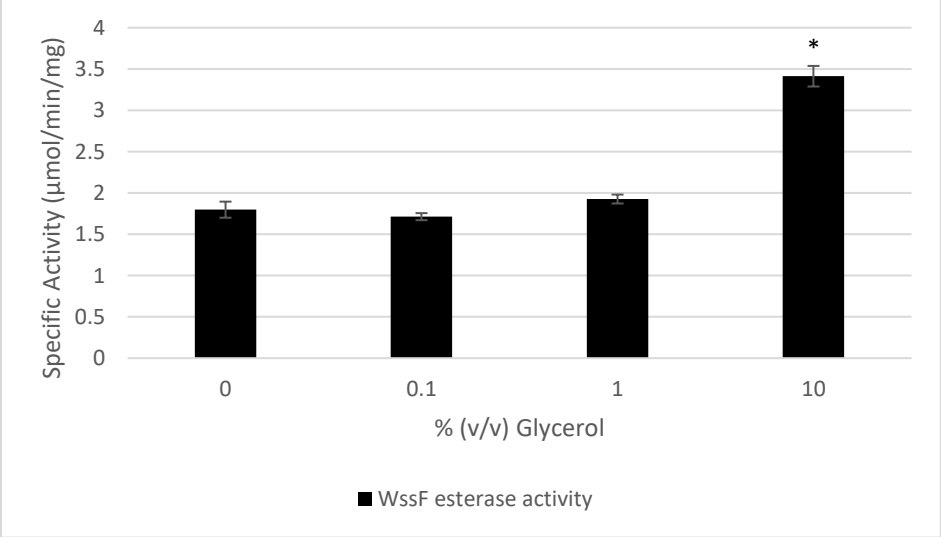


Figure 4.7 Specific activity of *Ai*WssF in the presence of glycerol. 0.5µM *Ai*WssF tested with 2mM 4-MUB-Ac standard esterase assays in increasing concentrations of glycerol. Glycerol increases from 0 % (v/v) to 10% (v/v) of final assay volume. Assays performed in quadruplicate. $p < 0.05$ for samples indicated by the asterisk, which were therefore statistically significant

Although protein production and purification were optimized, the amount of protein required to complete both structural and functional characterization was considerable. In order to improve workflow and minimize protein degradation, storage trials were undertaken to keep protein stocks viable long enough for increased rounds of analysis. Storage trials of 2mg/mL protein samples, in enzyme buffer (sonicated and filtered solution of 50mM Sodium Phosphate, 150mM sodium Chloride), were assessed at 4 ˚C, -20 ˚C and, -80 ˚C, with all samples stored below zero degrees containing 10 % (v/v) glycerol to act as a protectant and assayed over time for remaining esterase activity **(Figure 4.8)**. Protein samples stored at 4 ˚C saw a significant decrease in esterase activity after only 3 d relative to the sample assayed on the day of purification. However, protein stored below freezing retained full activity.

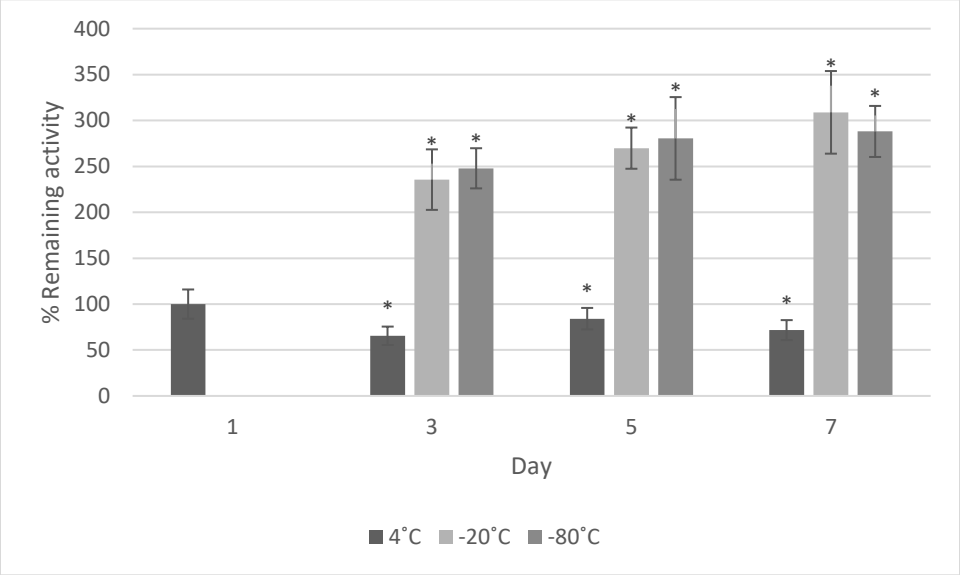


Figure 4.8 Storage trials of *Ai*WssF. The amount of retained esterase activity (with 4-MUB-Ac as substrate in standard esterase assays) after storage at the indicated temperatures was used to assess the storage condition. Samples stored at -20 ˚C and -80 ˚C contained 10 % (v/v) glycerol in the enzyme buffer, while 4 ˚C was stored in enzyme buffer without glycerol. All trials are calculated relative to an initial assay performed on the day of purification in the absence of glycerol (day zero = 100 % activity). Trials consisted of 4 replicates. $p < 0.05$ for samples indicated by the asterisk, which were therefore statistically significant.

Following the success of these short-term (7 day) trials, long-term storage assays were also undertaken to analyze the longevity of the activity of WssF under these conditions **(Figure 4.9)**. Fresh protein samples were frozen at -80 ˚C with 10 % (v/v) glycerol and storage trials monitoring over a 38 d period noted that this condition stabilized the activity over the course of the trial (relative to day zero pre-freezing trials as a reference point). Each day corresponds to a separate aliquot of enzyme to avoid multiple freeze/thaw cycles of one sample. With long-term storage of protein sorted, expression and purification of WssF was scaled up and protein stored for downstream analyses with confidence that the protein was active and stable.

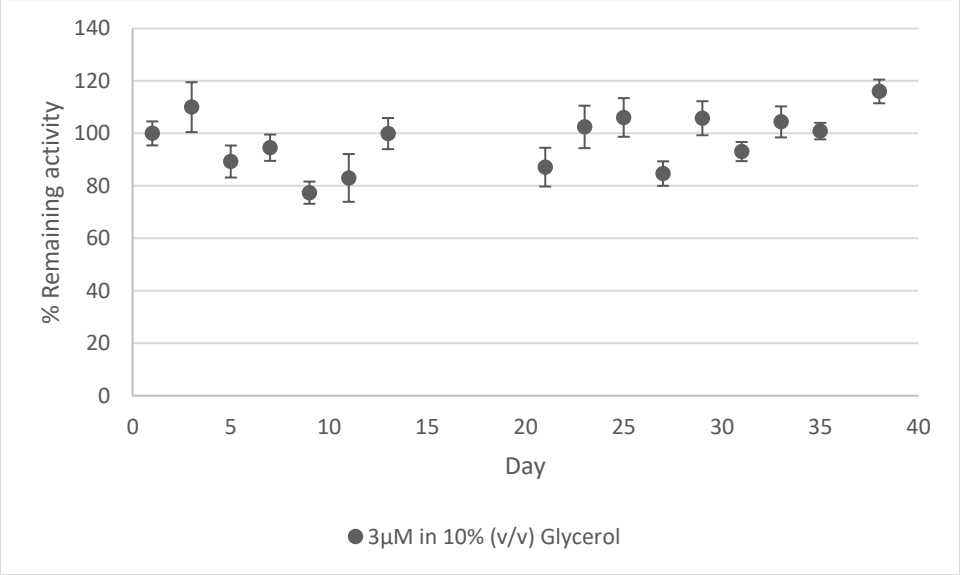


Figure 4.9 Long term storage trials of *Ai*WssF. The amount of retained esterase activity (using the standard esterase assay with 4-MUB-Ac as substrate) after storage at the indicated temperatures was used to assess the storage condition. Samples stored at -80 ˚C contained 10 % (v/v) glycerol in the enzyme. Trials consisted of 4 replicates.

4.3.3 Characterization of Esterase Activity

4.3.3.1 Kinetic Analysis

Once the esterase activity of WssF was reliably verified, efforts were made to determine the kinetic parameters of this reaction. Initial rates were determined using a reaction mixture of WssF protein with increasing concentrations of *p*NP-Ac acting as a substrate donor (**Figure 4.10**, **Table 4.4)**. The V_{max} was determined to be 0.0063µmol/min for *Ai*WssF and 0.0043 µmol/min for *Pf*WssF, indicating the maximum rate of reaction when all units of WssF are saturated with substrate. The binding affinity of WssF for *p*NP-Ac (K_M)was determined to be 7.18 ± 2.71 mM for *Ai*WssF and 2.88 ± 1.38 mM for *Pf*WssF. The turnover rate or k_{cat} for *Ai*WssF was 0.34 ± 0.06 s^{-1}, whereas *Pf*WssF had a k_{cat} of 0.24 ± 0.033 s^{-1}. Finally, the catalytic efficiencies or k_{cat}/K_m values of *Ai*WssF and *Pf*WssF were 47 $s^{-1}*M^{-1}$ and 83 $s^{-1}*M^{-1}$, respectively. Thus, under these conditions, *Pf*WssF is almost twice as catalytically efficient as *Ai*WssF.

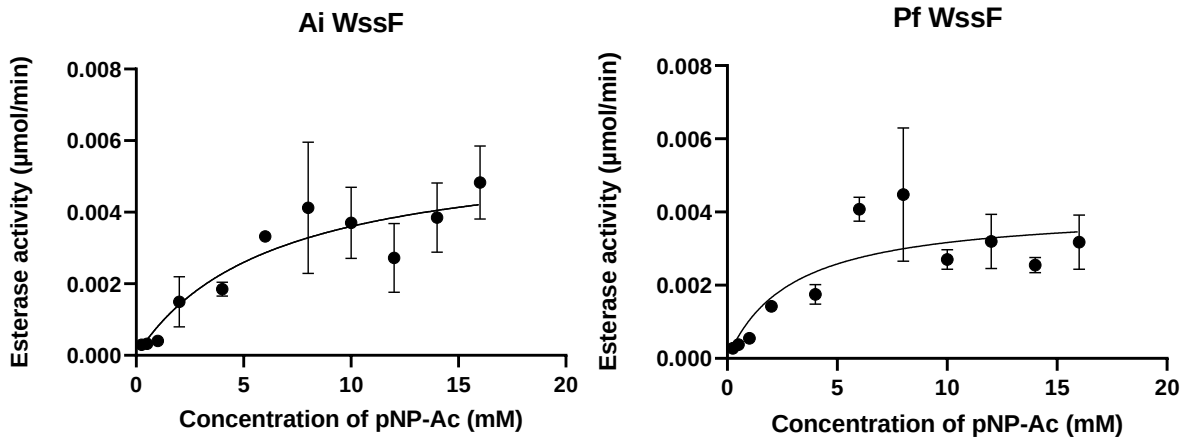


Figure 4.10 Michaelis-Menten Kinetic Curves of WssF Constructs. The assay was performed in quadruplicate and quantitatively assessed at 405 nm for 30 min with *p*NP-Ac concentrations ranging from 0-16mM.

Table 4.4 Kinetic analysis of WssF Enzymes

Construct	V_{max} (μmol/min)	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M ($s^{-1} \cdot M^{-1}$)
***Pf*WssF**	0.0043	2.88 ± 1.38	0.24 ± 0.033	83
***Ai*WssF**	0.0063	7.18 ± 2.71	0.34 ± 0.056	47

* Kinetic parameters were calculated using the Michaelis-Menten program from GraphPad non-linear regression one-binding site model.

4.3.3.2 pH profile of WssF activity

To further characterize the activity of both *Ai*WssF and *Pf*WssF constructs, a pH profile was created **(Figure 4.11)**. A buffer system, consisting of MES, phosphate and Tris, facilitated the assessment of the pH profile with overlapping pH values for a continuous assessment of activity. MES buffer was used from pH 5.5 - 6.5 and little esterase activity was noted below pH 6.0 and approximately 20 % activity at pH 6.5 for both constructs. From pH 6.5 to 7.5 phosphate buffers were used, where continued increasing activity was noted from 20 % at pH 6.5 up to approximately 60 % activity at pH 7.5 for both constructs. Finally, Tris buffer was used to test pH 7.5 to 9.5 and peak activity was at pH 8.5 (100 %), while above pH 8.5, activity gradually declined to 80 % by pH 9.5. The bell-shape of the pH curve is indicative of a catalytically important histidine, since it has two ionizable groups with pKa values of 6 and 9, respectively, that match this pH profile. While the optimal activity was at pH 8.5, routine assays were still conducted at pH 7, as this is more representative of the biological buffering *in vivo* (144).

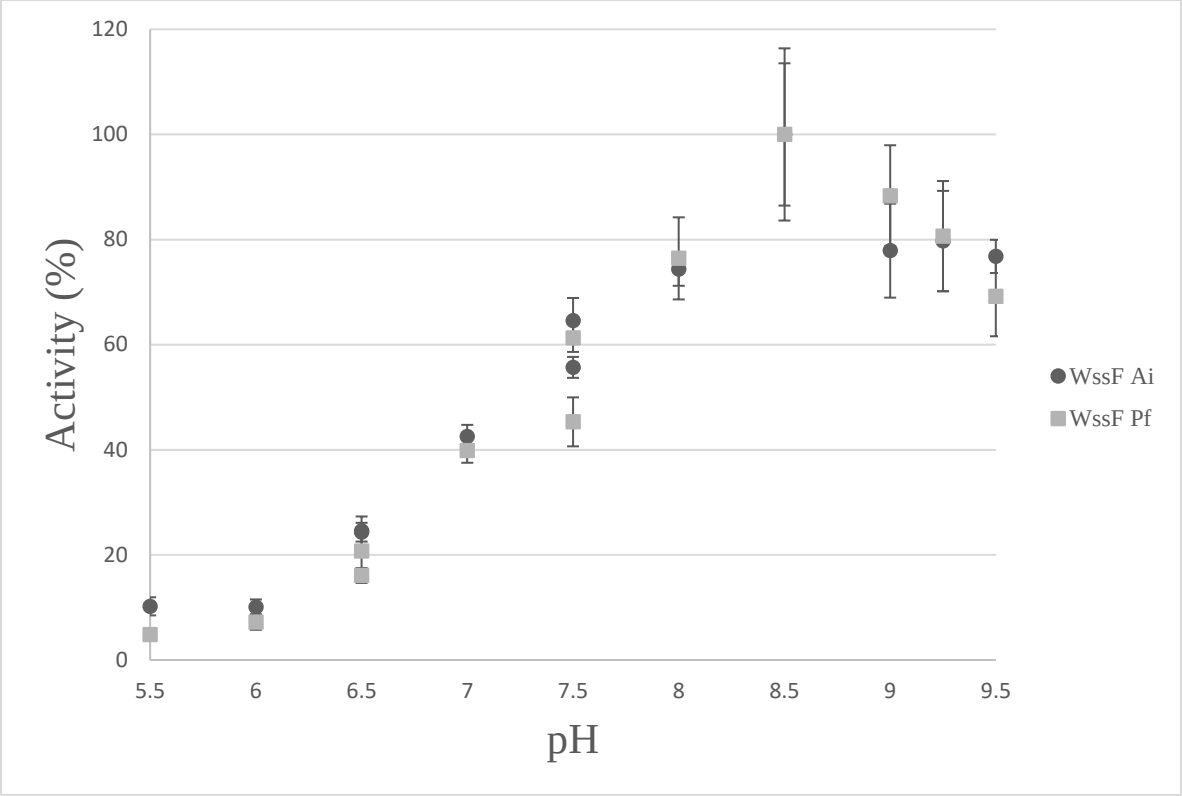


Figure 4.11 pH Profile of WssF Constructs. WssF constructs (4 µM) were assayed with 6 mM *p*NP-Ac substrate across various pH conditions using the following compliment of 50 mM buffers: MES from pH 5.5 - 6.5, phosphate from pH 6.5 - 7.5 and, Tris from pH 7.5 - 9.25. The normal distribution may indicate two ionizable groups on an active site residue. The assays were performed in quadruplicate; molar absorptivity's used the activity calculations can be found in appendix B.

4.3.3.3 Common SGNH Hydrolase Inhibitors

Based on the bioinformatic predictions (Section 4.1) the WssF proteins have homology to the SGNH hydrolase family. As such, common inhibitors for SGNH subfamilies were used to assess possible functional characterizations **(Table 4.5)**. EDTA, a metal chelator, was added to WssF reactions in increasing concentrations and at 5 mM (higher than biologically tolerant) EDTA was unsuccessful at knocking out esterase activity. *Ai*WssF retained 77 % activity while *Pf*WssF retained almost 60 % activity in the presence of 5 mM EDTA. As such WssF is unlikely to share a functional pathway with Metalloproteases to achieve esterase reactivity, which are inhibited by metal chelators, such as, EDTA (145, 146). Next, PMSF was used to determine if

WssF esterase activity shared a functional pathway with serine proteases (also includes serine esterases, lipases and others with a Ser-His-Asp active site). PMSF imitates the characteristics of a serine protease substrate and binds covalently to the nucleophilic serine within the active site, causing loss of activity. When biologically toxic levels of PMSF were added to WssF reaction mixtures, *Ai*WssF still retained 100 % activity and *Pf*WssF retained 80 % activity in the presence of 5 mM PMSF. As a follow up step, WssF was also tested for inhibition by MSF, which has the same inhibition mechanism as PMSF, but is a smaller chemical that lacks the bulky phenyl group. *Ai*WssF still retained 84 % activity and *Pf*WssF retained 60 % activity in the presence of 5 mM MSF. The common inhibitor profile was not conclusive as such other avenues to characterize the active site were then initiated with respect to mutagenesis of conserved active site residues (discussed in section 4.3.3) and high-throughput screening of unique inhibitors (section 4.3.4).

Table 4.5 WssF Common Inhibitor Profile.

Reagent	**Residual Activity (%)***	
	***Pf*WssF**	***Ai*WssF**
5 mM EDTA	57.9 ± 9.4	77.4 ± 1.5
5 mM PMSF	80.3 ± 5.2	100.0 ± 2.3
5 mM MSF	60.1 ± 1.2	84.2 ± 0.8

* WssF was incubated for 1 h with each listed inhibitor in standard assay conditions prior to the assay being initiated with the addition of 4 mM *p*NP-Ac. Assays performed in quadruplicate.

4.3.4 Characterization of O-acetylation activity

4.3.4.1 WssF specificity for acceptor substrates

With esterase activity confirmed, the ability of WssF to also perform acetyltransfer onto cellulose needed to be tested in order to support the role of WssF as an O-acetyltransferase. As an initial indication of cellulose oligomer specificity, standard esterase assays with WssF were conducted in the presence of cello-acceptors of defined length (*ie.*, cellobiose to cellohexaose) **(Figure 4.12)**. While *Pf*WssF esterase activity did not deviate beyond the standard error of the no-acceptor control (Figure 6.13; Donor only), *Ai*WssF esterase activity increased with the longer cellulose acceptors. *Ai*WssF had approximately 120 and 140 % activity, compared to the control reaction, for the cellopentaose and cellohexaose reactions, respectively. The cellohexaose reaction for *Ai*WssF was the only test with a significant value ($p < 0.05$). These results provide indirect evidence of substrate acceptor specificity for *Ai*WssF, which is further analyzed directly in section 4.3.5.

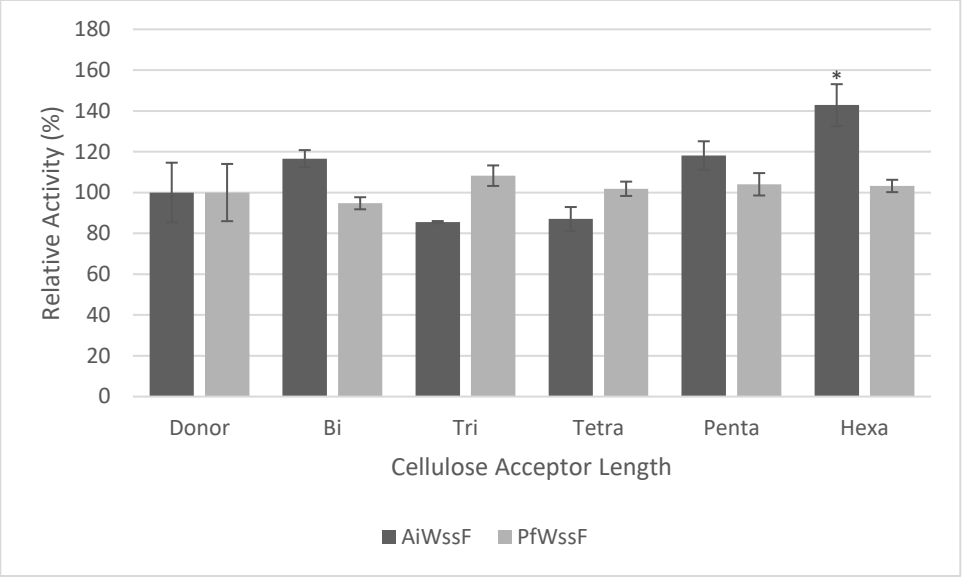


Figure 4.12 WssF Cellulose Acceptor Profile. Profile of acetate hydrolysis by WssF constructs using standard esterase assays in the presence of increasing lengths of cellulose-oligomers, 4 µM *Ai*WssF or *Pf*WssF and 5 mM pNP Ac. Assays performed in quadruplicate. All specific activities are relative to the no-acceptor (Donor Only) control. $p < 0.05$ for samples indicated by the asterisk, which were therefore statistically significant.

Chitin oligomers (analogous to cellulose but containing an *N*-acetyl group at position 2 of the sugar ring) were also assayed under the same conditions to determine if WssF acceptor substrate selectivity was specific for cellulose **(Figure 4.13)**. Increasing lengths of chito-oligomers, chitobiose to chitohexaose, were added to both *Ai*WssF and *Pf*WssF reaction mixtures and assayed under the same conditions as the cello-acceptor. In contrast to the cello-acceptor trials, none of the chito-acceptor substrates led to increases in activity for either version of WssF. In fact, there was a slight 10 % decrease in activity for most of the substrates, while chitotetraose reactions had the only significant ($p < 0.05$) decrease of approximately 65 % for *Ai*WssF and 75 % for *Pf*WssF relative activity. Thus, these results further corroborate the indirect esterase activity evidence that *Ai*WssF has specificity for cello-acceptor binding.

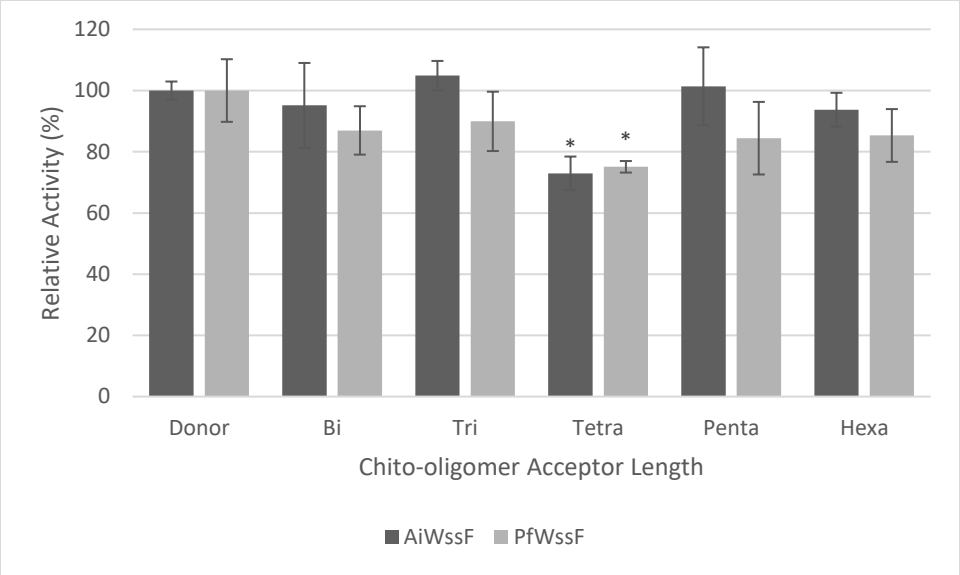


Figure 4.13 ***WssF Chito-oligomer Acceptor Profile.*** Profile of acetate hydrolysis by WssF constructs using standard esterase assays in the presence of increasing lengths of chito-oligomers, 4 µM *Ai*WssF or *Pf*WssF and 5 mM pNP Ac. All specific activities are relative to the no-acceptor (Donor Only) control. Assays performed in quadruplicate. $p < 0.05$ for samples indicated by the asterisk, which were therefore statistically significant.

4.3.4.2 Thin Layer Chromotography and Mass Spectrometry Assessment of WssF Acceptor Products

Thin layer chromatography (TLC) was initially used to assess acetyltransfer by WssF onto cellulose oligomers as a direct verification of acetyltransfer, which had previously been inferred from the esterase assays. After leaving reaction mixtures containing various lengths of cellulose for 24 h, the cellulose acceptor product was purified with SPE carbograph columns and analyzed by TLC, with assistance from Dave Sychantha in Anthony Clarke's lab (University of Guelph). Investigation through TLC revealed the cellulose oligomers were modified in the presence of *Ai*WssF **(Figure 4.14)**. As the length of the cellulose acceptor increased, the strength of a secondary band increased, and the presence of a faint third spot was observed with the cellopentaose and cellohexaose acceptors. These results indicate successful modification and that a possible second acetyltransfer event onto these substrates may have also occurred. To confirm

these TLC results, the cellopentaose sample was analyzed by mass spectrometry (**Figure 4.15).** Masses consistent with mono- and di- acetylated cellopentaose as sodium and potassium adducts (MS in positive ion mode) were detected in the spectra of *Ai*WssF but absent in the control (not shown).

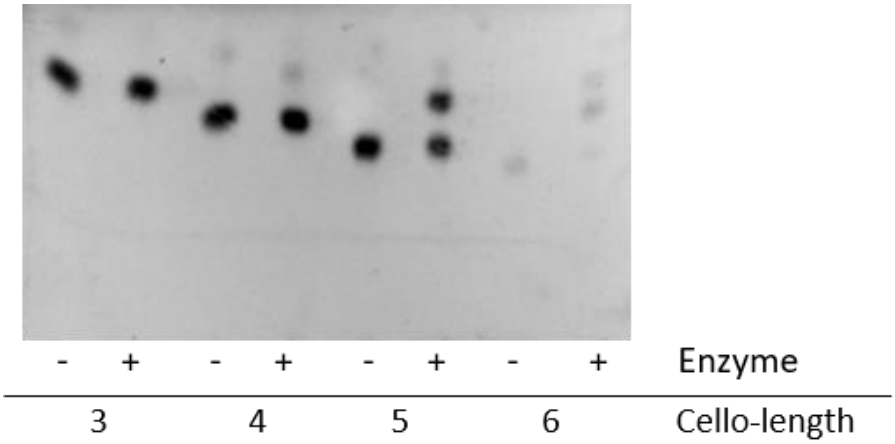


Figure 4.14 WssF Cellulose Acceptor Analysis by TLC. Assays were performed in 50mM sodium phosphate buffer (pH 7), 1mM cello-acceptor, 4µM *Ai*WssF, and 6mM pNP-Ac as the acetyl donor for 24 h. Lanes are in pairs for each cello-acceptor length (triose – hexaose), where the (-) denotes the negative control reaction with no enzyme and (+) denotes the sample of post-reaction purified cello-acceptor that was treated with *Ai*WssF. The addition of an acetate adduct onto the substrates alters the retention time so that there is the observable appearance of acetylated product above the control band. These reactions were purified utilizing SPE carbograph columns and the samples were separated with a mobile phase of ethyl acetate, water and methanol (6:3:6 v/v) and stained with α-naphthol staining solution (3 % (v/v) w/v α–naphthol in 25:3:2 methanol, sulfuric acid and water).

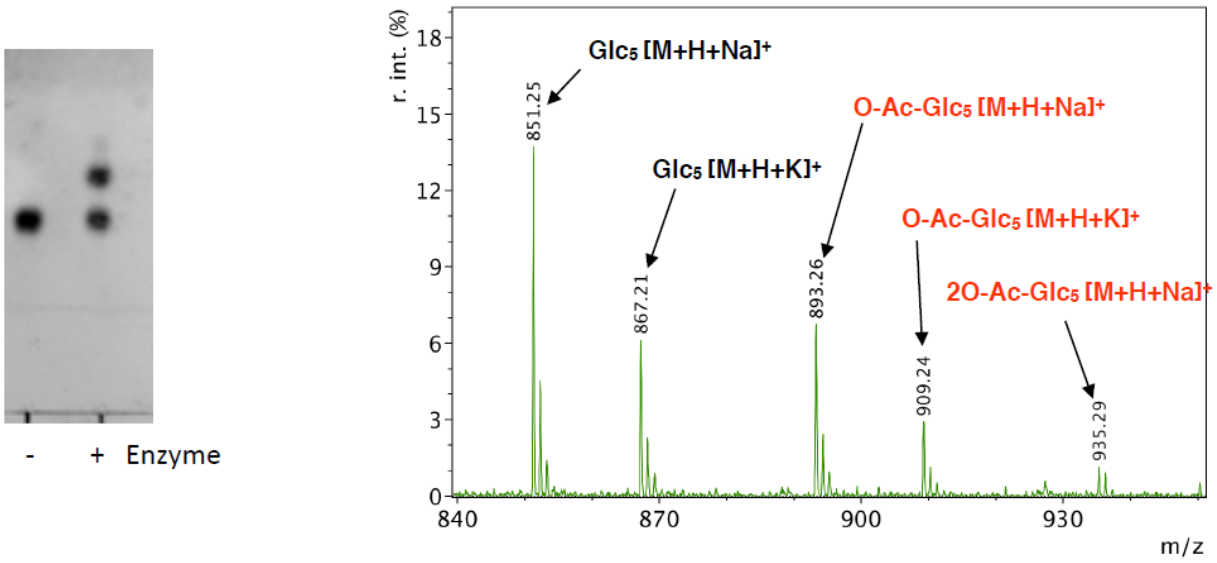
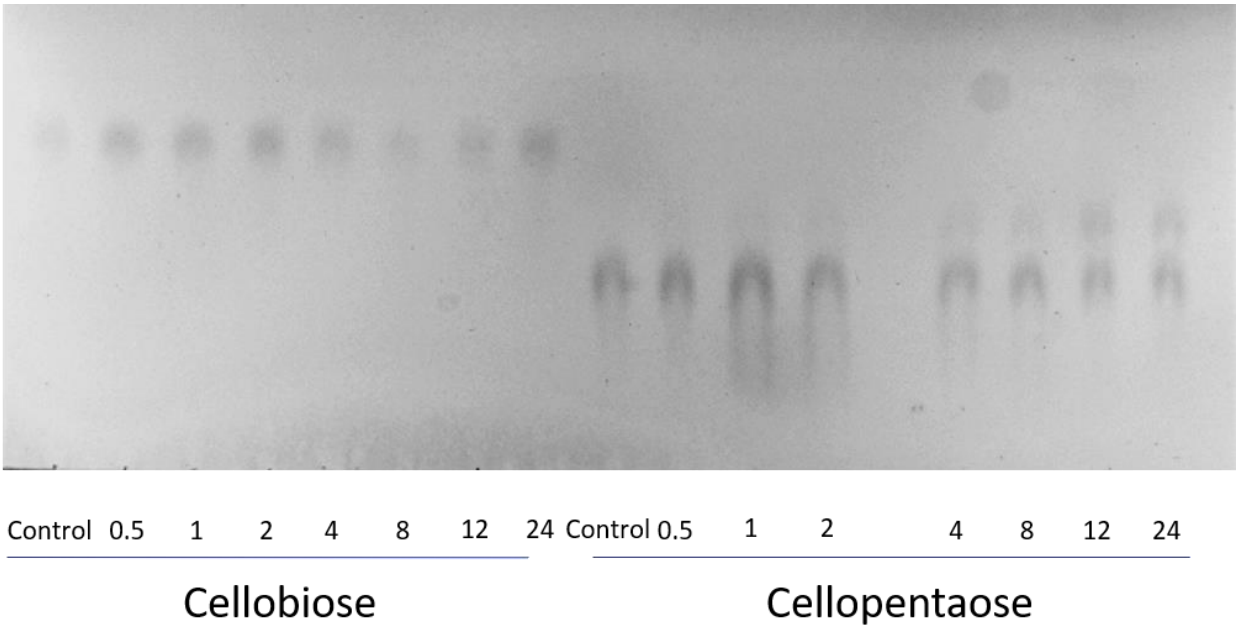


Figure 4.15 TLC and Mass Spectroscopy of Acetylated Cellopentaose. TLC of cellopentaose with and without *Ai*WssF present (Left panel). ESI Mass Spectrometry of the positive TLC sample to determine presence of O-acetylation (Right panel). M denotes the mass of cellopentaose (828 g/mol), while H^+, Na^+, K^+ and Ac denote the addition of a hydrogen (1g/mol), sodium (23 g/mol), potassium (39 g/mol) and/or acetate (42 g/mol) to the cellopentaose, respectively.

Once cellulose acetylation was observed, samples were prepared to create multiple time courses with cello-acceptors to obtain information regarding the rate of acetyltransfer **(Figure 4.16)**. Reaction mixtures containing either cellobiose or cellopentaose were left to react with the *Ai*WssF and *Pf*WssF enzymes and were sampled over a 24 h time period. None of the cellobiose samples analyzed had a secondary band indicating modified cellobiose, despite the fact that esterase activity (*ie.* an increase in *p*NP product) was observed. In contrast, the *Ai*WssF and cellopentaose containing samples displayed the appearance of a secondary band (Ac-modified cellopentaose product) after 1 h that increased in intensity over the course of the reaction, while the unmodified cellopentaose substrate (lower band) decreased in intensity (**Figure 4.16**; Panel A). While the results from TLC with *Pf*WssF were not as clear (despite multiple attempts), the results are consistent with *Ai*WssF, where an increasing intensity of the acetyl-product (noted as a smeared band above the unmodified reactant) was observed over the course of the experiment.

A) *Ai*WssF 24 h Time Course



B) PfWssF 24 h Time Course

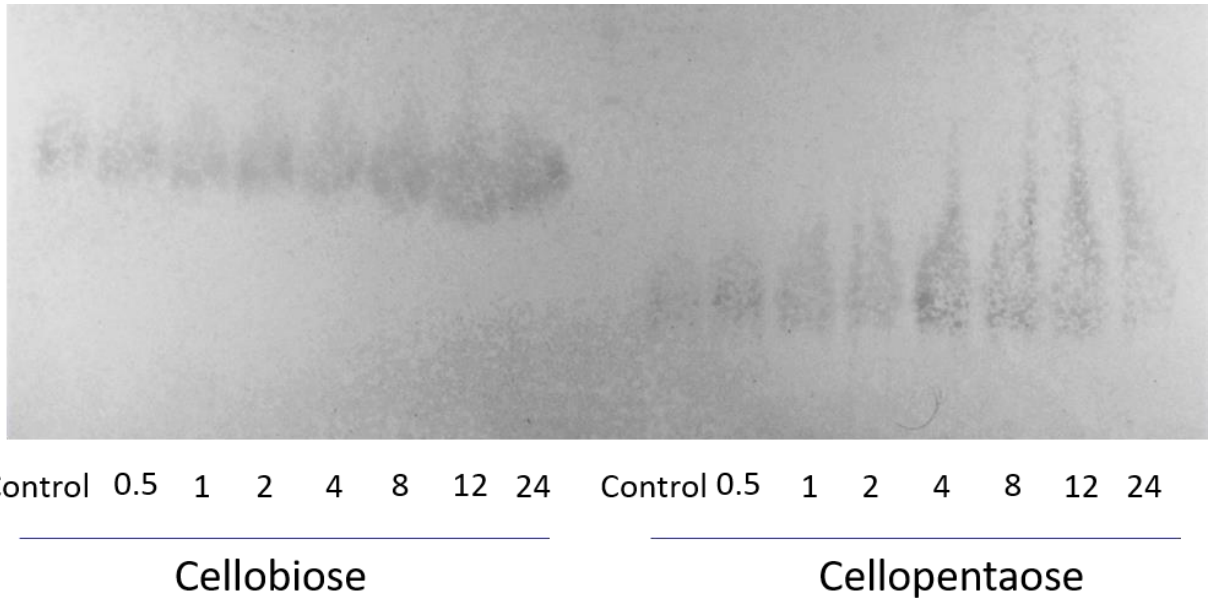


Figure 4.16 TLC Analysis of 24 h Time Course Reactions of WssF and Cello-acceptors. Assays were performed in 50mM sodium phosphate buffer (pH 7), 1mM cello-acceptor, 6 mM *p*NP-Ac as the acetyl donor and 4µM *Ai*WssF (Panel A) or *Pf*WssF (Panel B). For both Panel A and B, Lanes 1-8 display reactions with cellobiose and lanes 9-16 display reactions with cellopentaose as acceptor over the sampled 24 h time period (as indicated below the image).

Following TLC analysis, samples from the reaction mixtures were sent for MS verification of the acceptor products. For the control reaction mixture, a signal was detected at a mass of 848.52 *m/z*, which corresponds with an ammonia adduct of cellopentaose (**Figure 4.17**; Panel A). While the *Ai*WssF and *Pf*WssF treated samples (**Figure 4.17**; Panels B and C respectively) also had masses consistent with cellopentaose sodium and/or potassium adducts, they also contained peaks consistent with acetylated polymer. For *Ai*WssF, the spectral profile reconfirmed the previously analyzed sample (**Figure 4.15**), where a mass of 893.53 *m/z* was detected corresponding to a sodium adduct of acetylated cellopentaose. Weaker mass signals for the 909.52 *m/z* potassium adduct of acetylated cellopentaose were also noted (**Figure 4.17**; Panel B). For *Pf*WssF treated reactions, the cellopentaose sodium adduct (851.52 *m/z*) was observed, but the only other notable peak was the diacetylated cellopentaose potassium adduct (950.58 *m/z*). Comparison of the *Ai*WssF and *Pf*WssF results reveals profile differences that are characteristic to each enzyme, where after 24 h all of the detectable product for *Pf*WssF had undergone two acetyltransfer events, while *Ai*WssF under these conditions had both monoacetylated products.

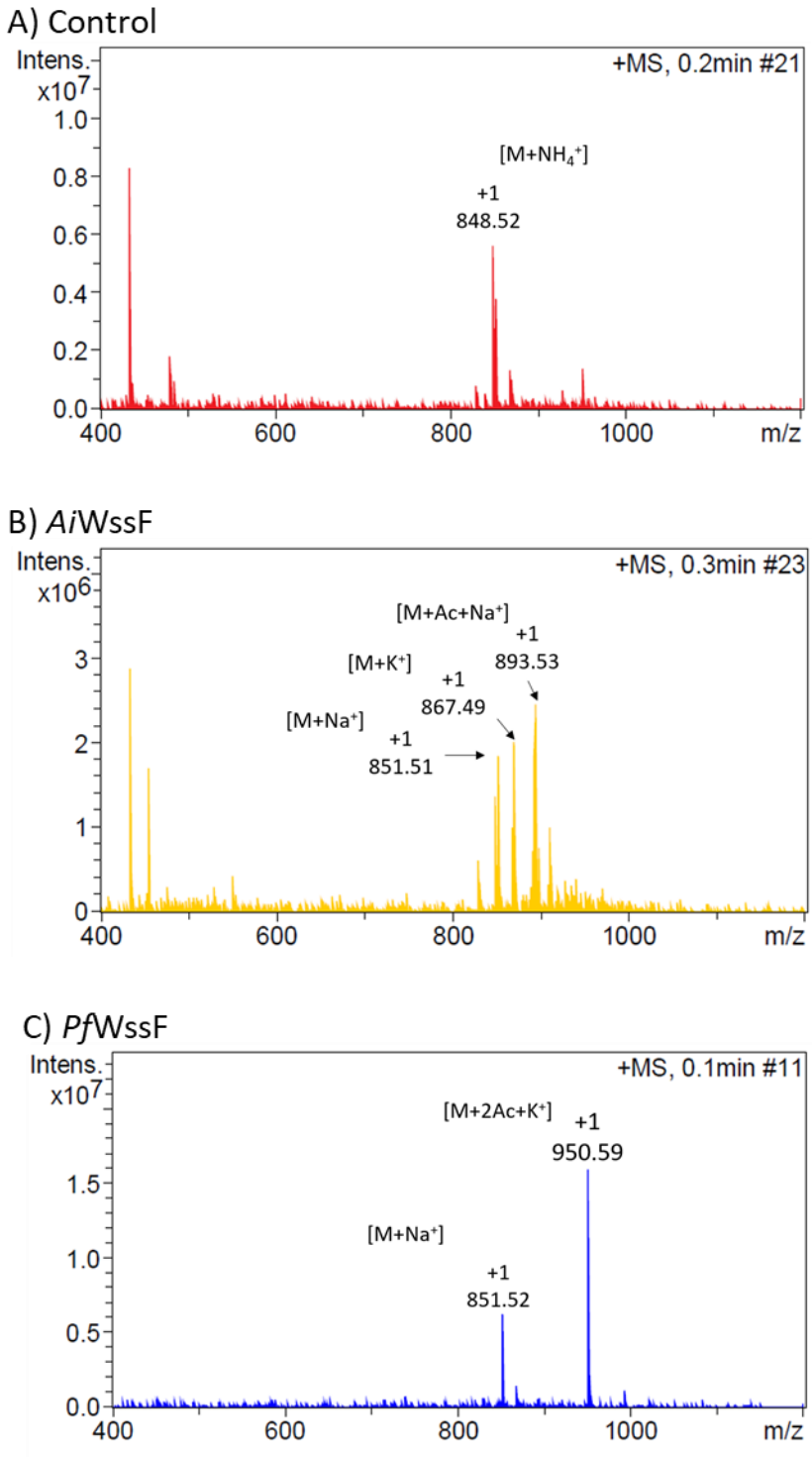


Figure 4.17 MS Analysis of WssF Time Course Assay. Mass spectra of the control sample (Panel A), *Ai*WssF treated (Panel B) and the *Pf*WssF treated (Panel C) 24 h samples from the time course assays are depicted. M denotes the mass of cellopentaose (828 g/mol), while H^+, Na^+, K^+, NH_4^+ and Ac denote the addition of a hydrogen (1g/mol), sodium (23 g/mol), potassium (39 g/mol), ammonium (18 g/mol) and/or acetate (42 g/mol) to the cellopentaose, respectively.

4.3.5 Site Directed Mutants

4.3.5.1 Expression

Based on bioinformatics results (section 4.1), alanine mutants of conserved potential active site residues (Ser17-His188-Asp191; *Ai*WssF amino acid numbering) were created in order to verify the role of these residues in catalysis. The mutated proteins were successfully expressed and purified using the same protocols as the unmutated original *Ai*WssF construct, albeit at lower yields. For example, yields of 2.7, 2.5 and 1.5 mg of protein/L of culture were obtained following IMAC purification for the S17A, H188A and D191A *Ai*WssF mutants, respectively, compared to 3 mg of protein/L of culture for the wild type (WT) enzyme **(Figure 4.18)**. Regardless, there was enough protein to concentrate and use in subsequent activity assays for each of the constructs.

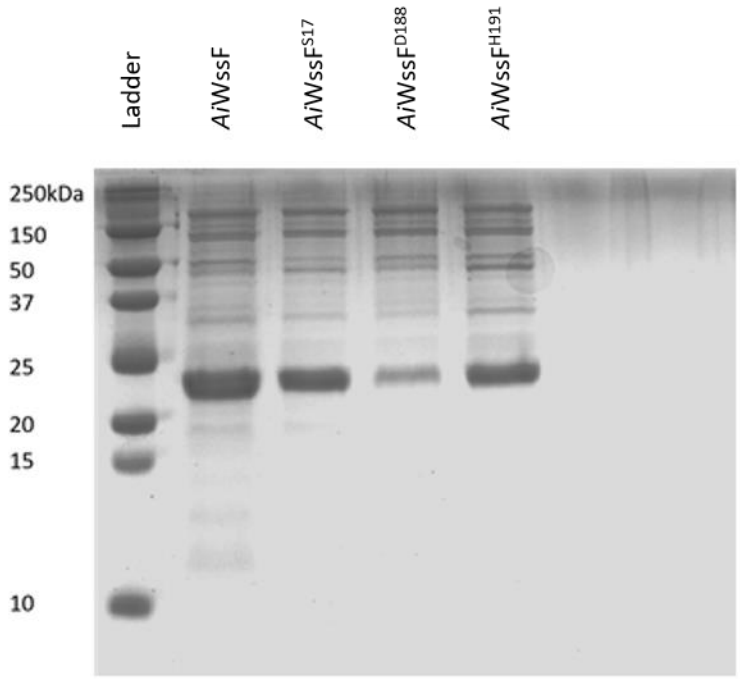


Figure 4.18 SDS-PAGE Purification of *Ai*WssF and the Active Site Mutants. SDS-PAGE analysis (12 % (v/v)) of *Ai*WssF and active site mutants concentrated eluate following IMAC purification. Lane 1, molecular-mass markers (kDa); lane 2, Wild Type *Ai*WssF; lane 3, *Ai*WssF^{S17A}; lane 4, *Ai*WssF^{D188A}; lane 5, *Ai*WssF^{H191A}. The expected molecular mass of each of the purified enzymes is 24.6 kDa.

4.3.5.2 Esterase Activity of Active Site Mutants of AiWssF

Each of the purified mutant proteins were assayed for esterase activity utilizing *p*NP Ac or 4-MUB Ac substrates in standard assays and compared to the WT enzyme **(Figure 4.19)**. Using the unmutated sample as the standard all mutants saw a statistically significant (t test, $p < 0.05$) reduction in activity for both substrates. When testing with *p*NP Ac residual activity for the *Ai*WssF^{S17A} mutant was 57.8 ± 4.7 % whereas the *Ai*WssF^{D188A} and *Ai*WssF^{H191A} mutants each retained 32.7 ± 1.6 % and 25.5 ± 3.9 % activity respectively. When tested in 4-MUB Ac residual activity for the *Ai*WssF^{S17A} mutant was 23.3 ± 1.9 % whereas the *Ai*WssF^{D188A} and *Ai*WssF^{H191A} mutants each retained 13.1 ± 1.8 % and 25.5 ± 4.1 % activity respectively. Residual activity for the *Ai*WssF^{H191A} mutant was consistent between both substrates whereas *Ai*WssF^{S17A} and *Ai*WssF^{D188A} had less than half the residual activity with 4-MUB Ac than they did with *p*NP Ac.

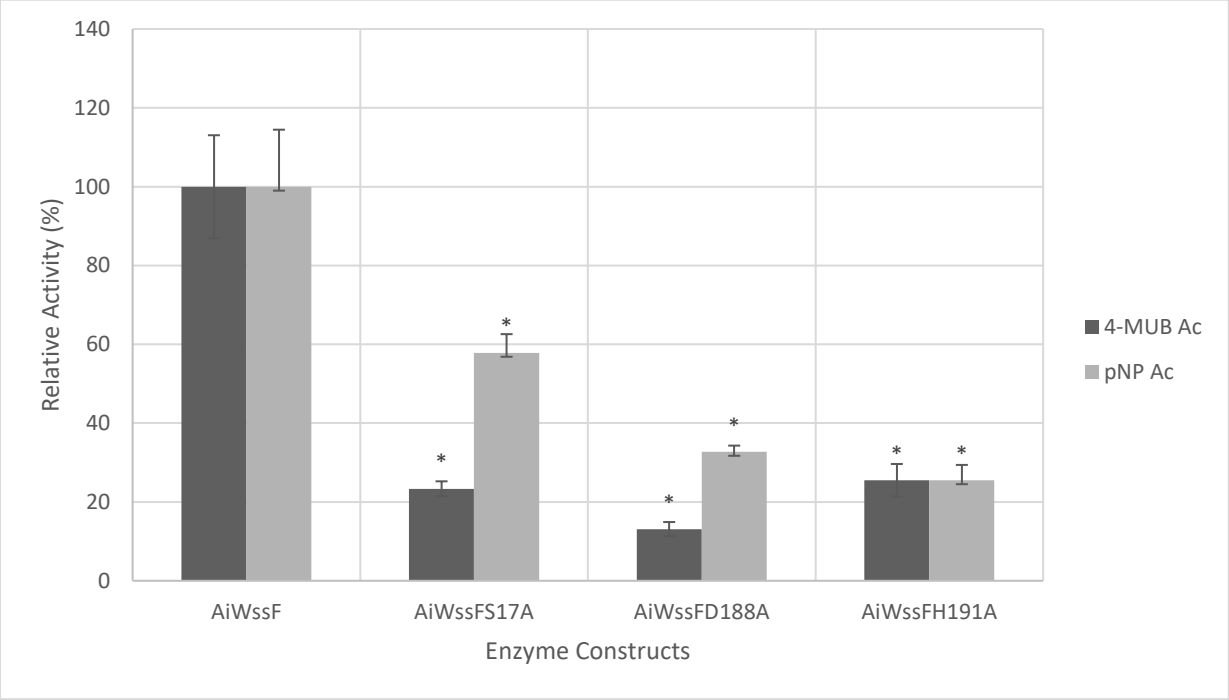


Figure 4.19 Relative Esterase Activity of the *Ai*WssF Catalytic Triad Mutants. The amount of retained esterase activity with either 6mM *p*NP-Ac or 4-MUB-Ac as substrate in standard esterase assays for the WT enzyme and the S17A, D188A, H191A mutants thereof. Trials consisted of at least 4 replicates. *$p < 0.05$ for all mutant samples was determined, which were therefore statistically significant.

4.3.5.3 Thin Layer Chromatography Analysis of Active Site Mutant Reactions

TLC analysis for actetyltransfer of each of the active site mutants was performed using standard reactions in the presence of cellopentaose **(Figure 4.20)**. As noted before, WT *Ai*WssF reactions had two distinct bands of similar intensity following 24 h reactions. However, no acetylated product (*i.e.* band above the control) was noted for either the S17A or the D181A mutant, despite the residual esterase activity (*p*NP product) observed for these mutants **(Figure 4.19)**. For the H191A mutant, a slightly less intense unmodified product band was noted along with a very faint potential second band; thereby suggesting that some residual acetyltransfer (despite the low residual esterase activity) may be occurring with this mutant.

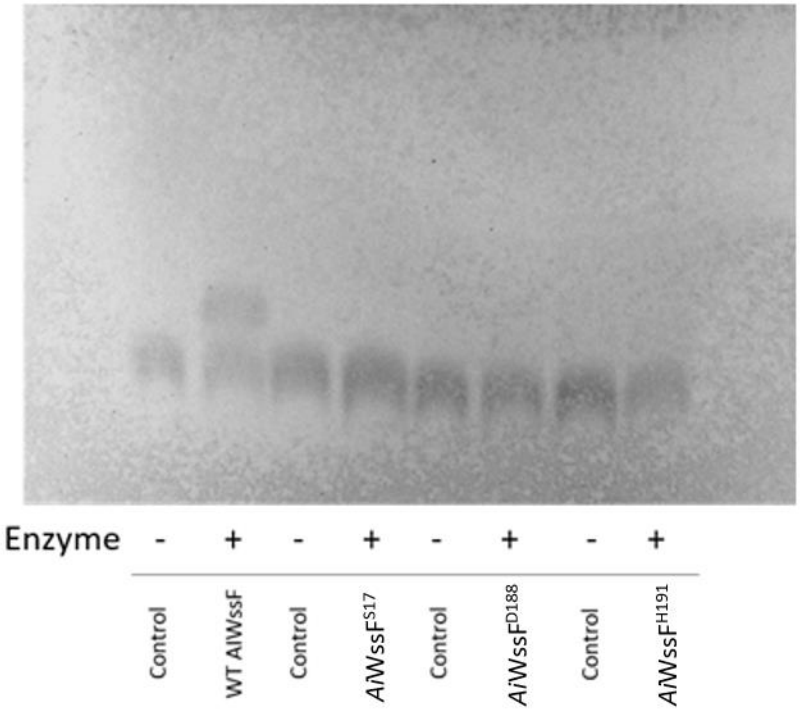


Figure 4.20 SDM Cellopentaose TLC. Assays were performed in 50mM sodium phosphate buffer (pH 7), 1mM cellopentaose-acceptor, 6mM *p*NP-Ac as the acetyl donor and 4 µM WT AiWssF (Lane 1), *Ai*WssF^{S17} (Lane 4), *Ai*WssF^{D188} (Lane 6) and, *Ai*WssF^{H191} (Lane 8) for 24 h. Lanes 1,3,5,7 represent negative control reactions, which contained no protein. (as indicated below the image).

4.3.5.4 Mass Spectrometry Analysis of Active Site Mutant Reactions

To further verify the TLC results with the active site mutants, samples from the reaction mixtures were sent for MS analysis. As before, the WT *Ai*WssF treated sample contained masses consistent with unmodified (851.51 and 867.49 *m/z*) and acetylated (893.53 *m/z*) cellopentaose (**Figure 4.21**; Panel A). Analysis of the each of the active site mutant spectra (**Figure 4.21**; Panels B-C) had a mass consistent with the unmodified cellopentaose sodium adduct (around 851 *m/z*). However, although other masses were noted in the spectra of the S17A and D188A mutants, none of these corresponded to masses consistent with an acetylated product (or derivative thereof); thereby indicating that acetyltransfer had not taken place to a degree that was detectable by MS analysis. For the H191A mutant, a weak signal near background levels corresponding to a sodium adduct of modified product was noted similar to what was observed during TLC. These results (together with the esterase reaction results), suggest that the Ser-His-Asp conserved residues of WssF do affect the activity of the enzyme and are part of the catalytic mechanism for acetyltransfer onto cellulose substrates.

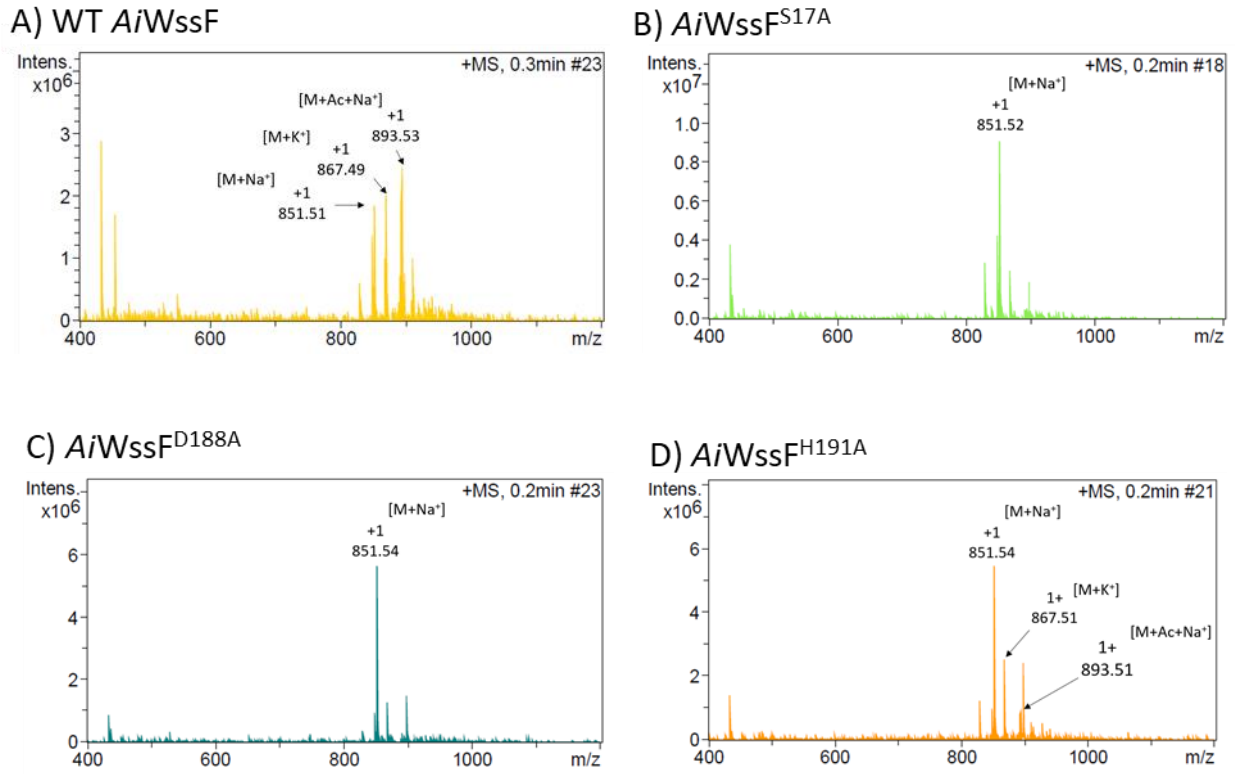


Figure 4.21 Site Directed Mutant Mass Spectrums. Mass spectra of the *Ai*WssF Wild Type positive control sample (Panel A), *Ai*WssF^{S17} treated (Panel B) *Ai*WssF^{D188} treated (Panel C), and the *Ai*WssF^{H191} Treated (Panel D) 24 h samples from the time course assays are depicted. M denotes the mass of cellopentaose (828 g/mol), while H^+, Na^+, K^+, NH_4^+ and Ac denote the addition of a hydrogen (1g/mol), sodium (23 g/mol), potassium (39 g/mol), ammonium (18 g/mol) and/or acetate (42 g/mol) to the cellopentaose, respectively.

4.3.6 High Throughput Screening for WssF Inhibitors

For high throughput screening (HTS), the standard esterase assay was first assessed and optimized prior to screening. As an initial step, a Z′ value was determined for the standard esterase reaction with *Ai*WssF **(Figure 4.22)**. Z′ is a unitless value that can be used to determine whether a reaction is amenable to HTS because it provides an indication that there is a great enough window between the detection of a positive reaction and the background. The assessment of inhibitors in the HTS screen can then be reliably detected in this window as the decrease in activity. The following formula was used to calculate Z′:

Formula1:

$Z' = 1 - (3 * (\text{Standard Deviation of Positive Reaction}) + (3 * (\text{Standard Deviation of the Negative Control}) \div |(\text{Mean of the Maximum Signal Control} - \text{Mean of the Minimum Signal Control})|)$.

Values for Z′ range from 0-1, with higher quality data the closer the reaction approaches 1. In the case of WssF, Z′ values consistently reached or exceeded 0.8 across separate purification samples, which was higher than the 0.7 that was deemed sufficient for reliable data collection. The assay was then miniaturized in volume (45 µL) and enzyme concentration (0.75 µM *Ai*WssF) to conserve substrates and enzyme, while maintaining a Z′ value of 0.8 (Figure 4.22).

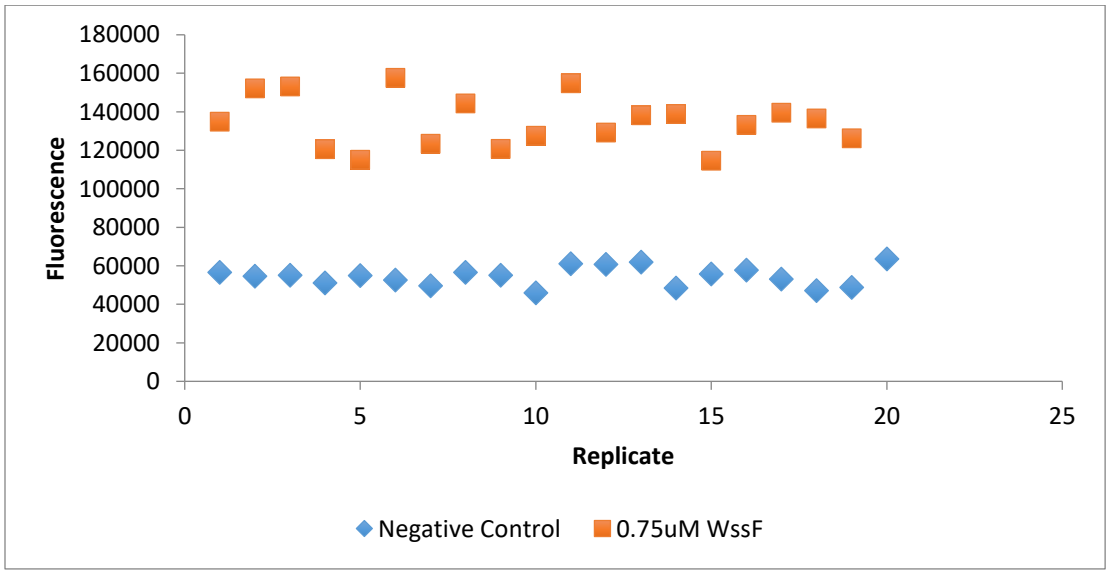


Figure 4.22 Representative Z′ 4-MUB-Ac Fluorescence Assay. WssF esterase activity measured through the fluorescence of 4-MUB as a reaction byproduct. Reaction was performed in a 45 µL reaction mixture with a final concentration of 0.5 mM 4-MUB-Ac and 5 % (v/v) DMSO. A trial with 20 replicates used for the Z′ calculation are depicted here.

4.3.6.1 HTS at SPARC Biocentre

Once the optimization of the reaction was accomplished, samples of *Ai*WssF were sent to the SPARC facility at The Hospital for Sick Children (Toronto) for HTS screening. Primary screening of approximately 64 000 compounds revealed 103 hits with Z scores ≤ -3 **(Figure 4.23)**. To clarify, a Z score of -3 indicated that the activity associated with 103 hits were at least 3 standard deviations reduced compared to the mean activity of the positive control (**Figure 4.23**; Orange to Green spheres). The 103 hits represent a hit rate of 0.16 %, which is expected following primary screening.

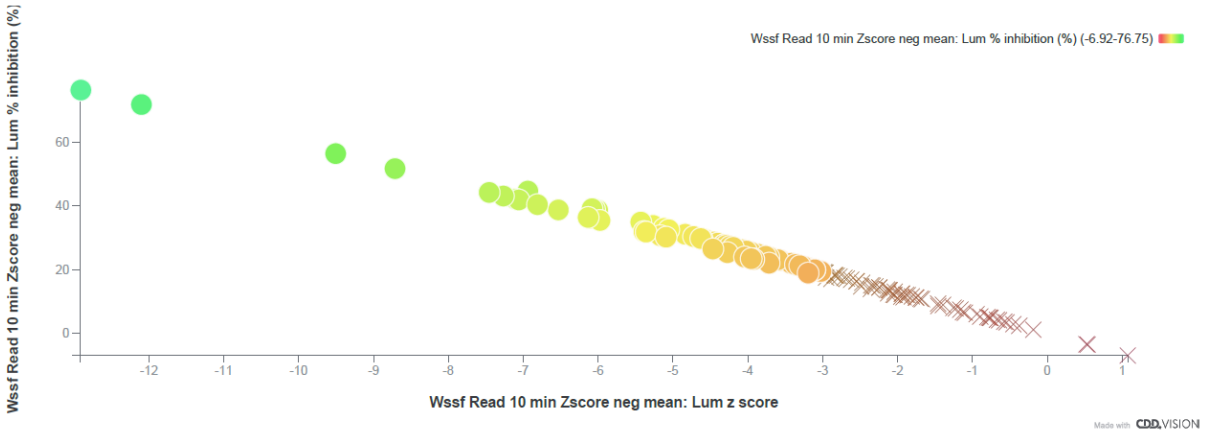


Figure 4.23 Comparison of % Inhibition vs Z score for Potential Inhibitory Compounds of *Ai*WssF. Compounds with a Z score at least 3 standard deviations below the positive control (Z score = 0) are depicted in orange to green circles and were selected for further study. Data from this primary screen represent single trials with compound at 20 µM.

Secondary screening of the 103 hits was done to repeat the results of the primary screen for these compounds in triplicate. Potential inhibitors for similar putative acetyltransferases (WssI, PatB and, OatA) uncovered in parallel screens were also included in the secondary screen of WssF for a total of 146 compounds. Following secondary screening, 83 compounds were further selected for dose response testing based on the removal of Pan-Assay Interference Compounds (PAINs) identified with consultation with Dr. Mark Nitz (University of Toronto). PAINS compounds

consist of those that have non-specificity in binding, disruptive functional groups and often give false positives. The remaining 83 promising compounds were found to reach 100 % inhibition **(Figure 4.24)** and IC_{50} values calculated from the curves were 9.49, 11.3, 19.5, 31.6 and 37.4 μM for compounds 1-5, respectively.

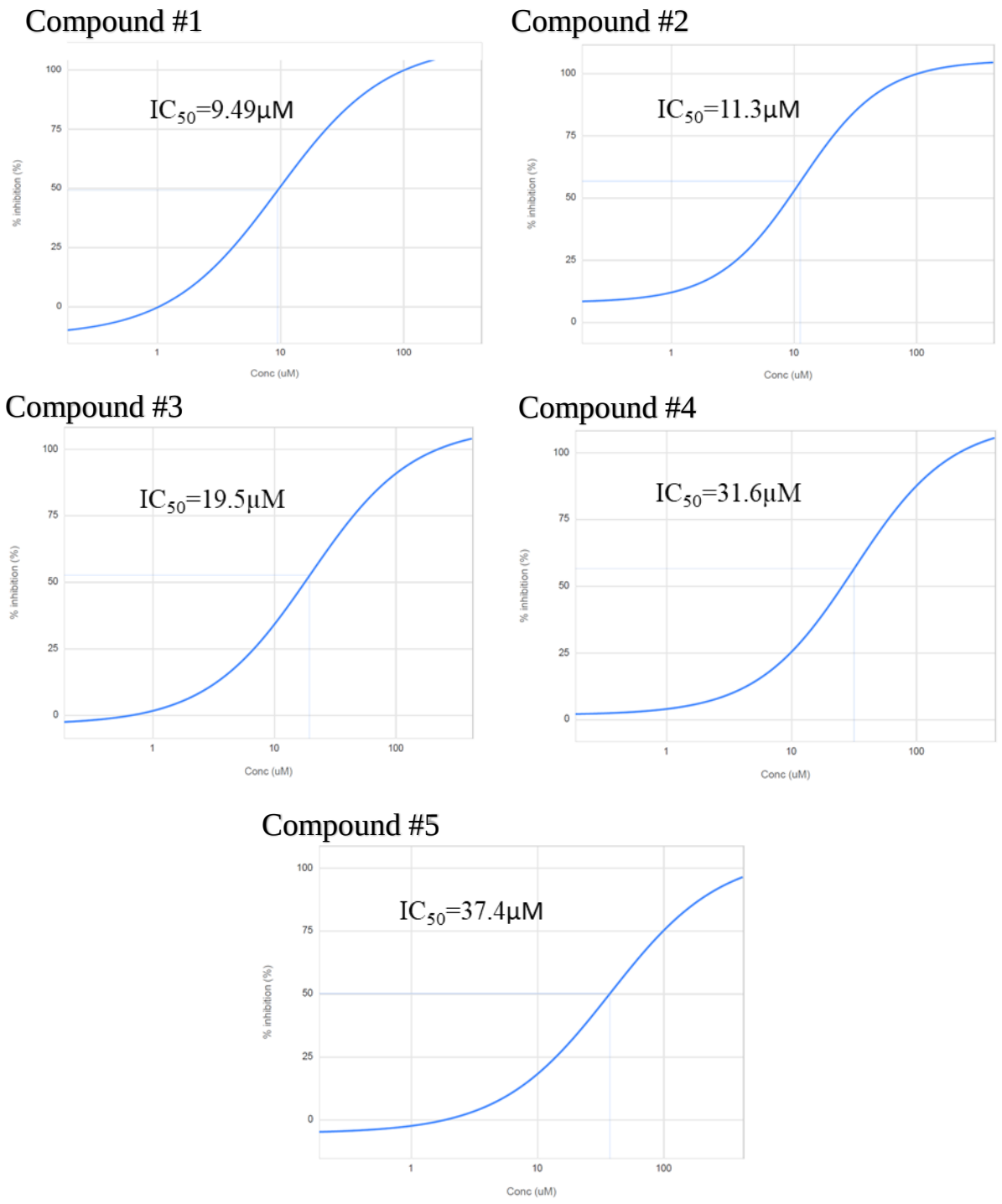


Figure 4.24 Dose Response Curves of the Top 5 Hit Compounds From HTS Screening of *Ai*WssF. Compounds were chosen based on IC_{50} values below 100µM. All assays were conducted in triplicate in 47 µL final well volumes with 0.5 µM protein concentration and 0.5 mM 4-MUB-Ac substrate.

4.3.4.2 Verification of Inhibition at WLU

From the dose response curves, identified compounds with at least 80 % inhibition and an IC_{50} value of less than 100 µM, a list of 47 compounds was created to undergo testing for false positives. All of the tested compounds were found to have quenched the fluorescence activity within the well condition and were false positive hits. As a result, 18 promising compounds were instead chosen from the primary and secondary screens for further analysis. All 18 compounds were subjected to dose response assays and Figure 4.25 depicts the curves from the top 3 hits. Compound A was found to have an IC_{50} value of 48.09 µM achieving 70 % inhibition by 100 µM, while compound B had an IC_{50} value of 43.21 µM and 70 % inhibition by 100 µM. Finally, compound C was found to have an IC_{50} value of 71.55 µM achieving 60 % inhibition by 100 µM.

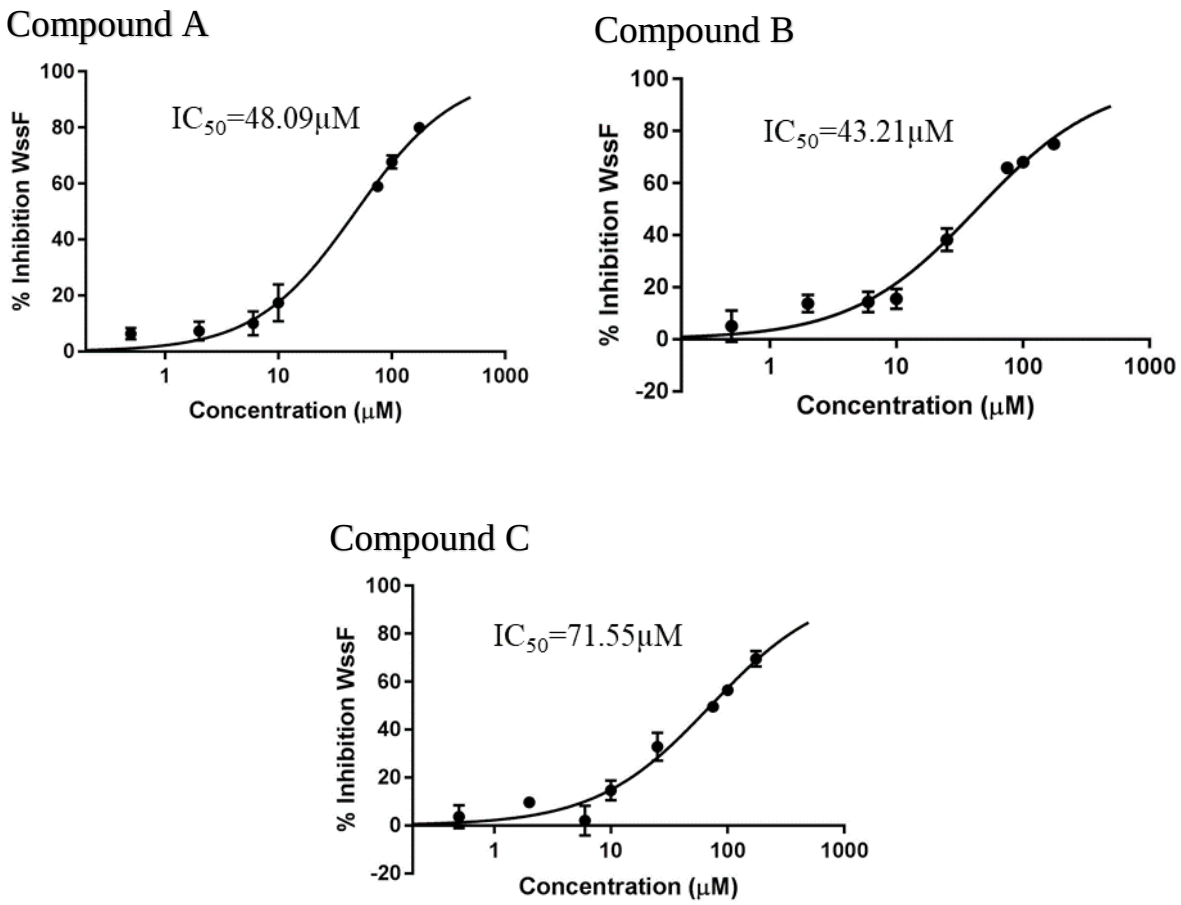


Figure 4.25 The Top 3 Hit Compounds from the Second Dose Response Assay of *Ai*WssF Inhibitors. Compounds were chosen based on IC_{50} values below 100µM. All assays were conducted in triplicate in 47 µL final well volumes with 3 µM protein concentration and 0.5 mM 4-MUB-Ac substrate.

4.4 Structural Characterization of WssF (Objective 3)

4.4.1 Crystallization Trials

In order to fully characterize WssF as an acetyltransferase, purified protein was used to create protein crystals for structural analysis. Multiple sparse matrix screens, including MCSG 1 through MCSG 4 and a top 96 formulation of conditions, were used in initial crystal trials to produce protein crystals (a conservative estimate of approximately 480 total conditions were assayed). After approximately 4 months, four crystal conditions led to the production of protein

crystals (Summarized in **Table 4.6**). Two crystal hits were observed from the MCSG-1 screen in these initial trials (conditions F1 and H10) when 29.11 mg/ml protein in Tris-HCl buffer (pH 7.4 was used and mixed in a 1:1 ratio with the crystal condition (see Table 6.4 for condition components). Two other hits; one from the MCSG 1 screen (condition B8) and another from the Top96 screen (condition E2) were created with 15 mg/ml of protein mixed 1:1 with well conditions (Table 6.4).

Table 4.6 Successful Crystal Conditions. Well conditions F1, G10 and, H10 from the MCSG1 screen were used with 29 mg/ml protein in Tris-HCl buffer. Well condition B8 from the MCSG1 screen used 15 mg/ml protein in Tris-HCl buffer. Well condition from Top 96 formulations was used with 15 mg/ml protein in Tris-HCl buffer.

Well Position	Condition	Concentration
F1	0.1 M Bis-Tris:HCl pH 6.5 and, 20 % (w/v) PEG MME 5000	29 mg/mL
H10	0.1 M HEPES:NaOH pH 7.5 and, 25 % (w/v) PEG 3350	29 mg/mL
B8	0.17 M Sodium Acetate, 0.085 M Tris:HCl pH 8.5 and, 25.5 % (w/v) PEG 4000, 15 %	15 mg/mL
G10	0.1 M Magnesium Formate and, 15 % (w/v)PEG 3350	29 mg/mL
E2	25 % (w/v) PEG 3350	15 mg/mL

From the wells that produced crystal hits, images were taken before the crystals were frozen for data collection. Example images of the successful crystal hits and diffraction patterns are displayed in **Figure 4.26**. The crystal formed from the Top 96 formulations had a yellowish appearance and formed in a wedge like shape on one end, whereas the crystal from well F1 formed in a wide hexagon like shape. Both crystals were subjected to X-ray analysis at the Canadian Light Source (CLS) and found to have diffraction patterns with a resolution of 2.1Å for the crystal formed in well E2 and 2.2Å for the crystal formed in well F1 (Data statistics are summarized in **Table 4.7** for the E2 crystal)

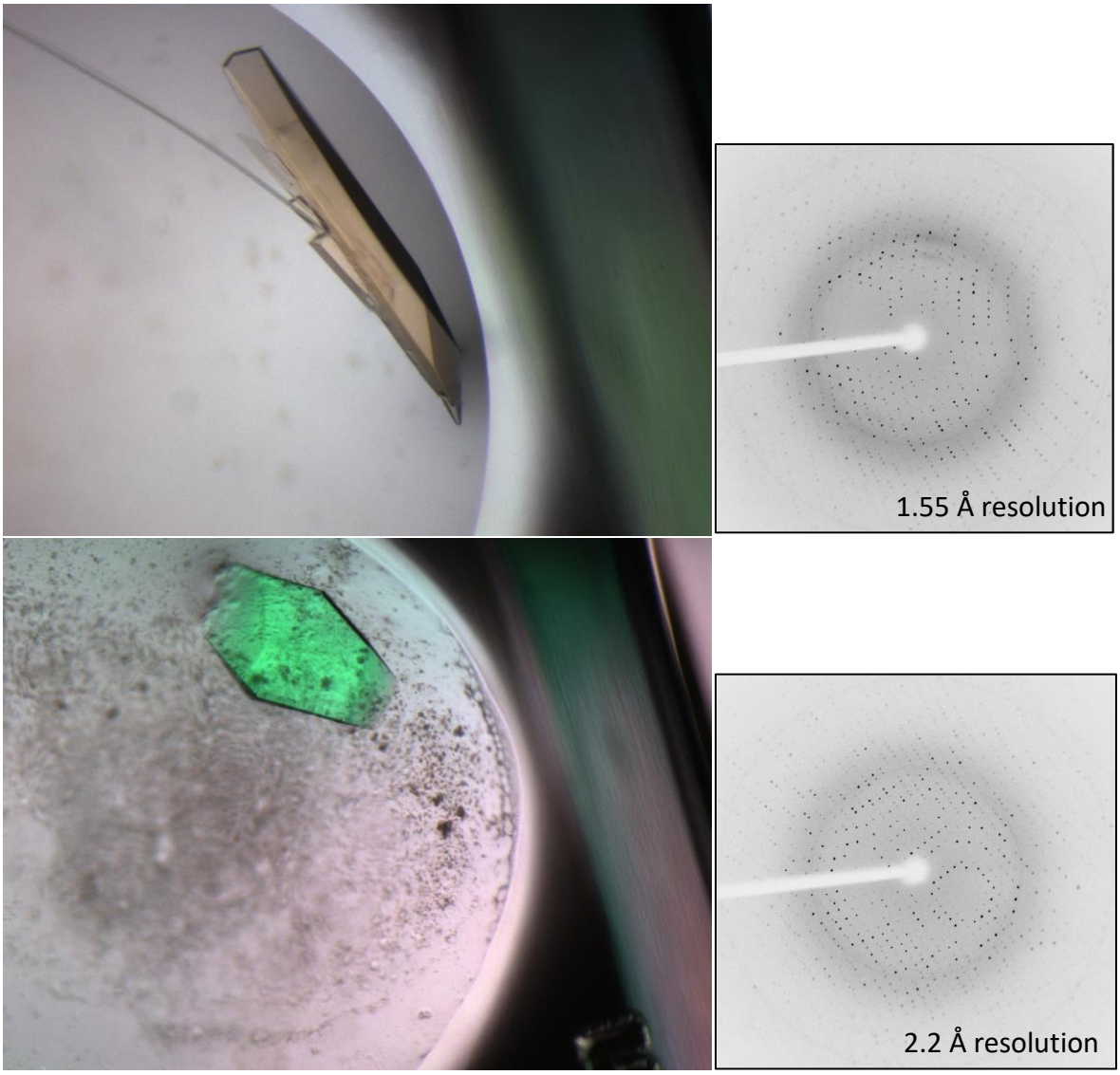


Figure 4.26 Crystal Hits and X-Ray Diffraction Patterns from Top Crystal Hits. 15 mg/mL protein (Top) TOP96-E2: 25 % (w/v) PEG 3350. 29.11 mg/mL protein (Bottom) MCSG1-F1: 0.1M Bis-Tris (pH6.5); 20 % (w/v) PEG 5000 MME.

4.4.2 Diffraction Data

Single crystals isolated from these drops were assessed to have resolutions of ~ 1.5 - 2 Å (**Table 4.7**) following X-ray diffraction analysis at the Canadian Light Source (CLS) and University of Waterloo. From these crystals, high resolution data sets were collected for processing. From the two examples shown in **Table 4.7** analysis of both crystals (~ 1.5 - 2 Å)

displayed strong quality indicators for their high-resolution shells. Data processed from the crystal from the Top 96 condition E2 was found to have strong $CC_{1/2}$ values close to 1 as the resolution shells increased (1.57 - 1.54 Å). $CC_{1/2}$ is a quality measure, which measures the correlation between random half data sets and has been found to decrease as the resolution of the data sets increase (147). A signal to noise ratio or I/σI for E2 was found to be 9.99 overall but decreased slightly to 8.7 at the highest resolution shells. Traditionally, this ratio is used as a cutoff measure where bins with values ≥ 2 being selected (147). Completeness for the higher resolution shells was found to be 98.8 %, which reflects how well the high-resolution data covers the whole structure. A strong redundancy value (number of reflections ÷ number of unique reflections) was found for E2 at 13.6 for the highest resolution shells. Based on these quality indicators the E2 data set is a good choice for further processing. In comparison data sets for the crystal from the MCSG-1 condition G10 (2.05 - 2.00 Å) was found to have a weak I/σI (1.3), and a lower completeness (62.6 %). G10 was found to have a strong $CC_{1/2}$ (0.836) value for the highest resolution shells. Overall E2 displayed stronger quality indicators and as such was the best candidate for further analysis.

Table 4.7 Statistics for X-ray Data Collection and Processing of WssF for Two Typical Datasets*.

	Top96_E2 Condition	**MCSG1_G10 Condition**
Diffraction Source	CMCF beamline 08B1-1	University of Waterloo
Wavelength	0.979	1.540
Temperature	100 K	100 K
Exposure time	10 s	2 min
Number of images collected	360	360
Oscillation range	1 degree per image	1 degree per image
Space group	Orthorhombic Primitive	Orthorhombic Primitive
Cell dimensions		
***a, b, c,* (Å)**	43.26, 60.08, 77.13	42.38, 59.90, 77.39
α, β, ɣ, (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Resolution range (Å)	47.40-1.54 (1.57-1.54)	25.8-2.00 (2.05-2.00)
Total number of reflections	436883 (19459)	22530 (763)
Total number of unique reflections	30220 (1433)	12705 (5260)
R_{meas}	0.055 (0.355)	0.059 (0.0.364)
$I/\sigma I$	9.99 (8.7)	15.8 (1.3)
$CC_{1/2}$	0.999 (0.983)	0.995 (0.836)
Completeness (%)	99.4 (98.8)	91.36 (62.6)
Redundancy	14.5 (13.6)	1.8 (1.2)
Mosaicity	0.132	1.3

* Data in brackets represent the highest resolution shell

4.4.3 Expansion Trials

Initially, several expansion plates were conducted with the Top 96 E2 condition using different concentrations of PEG (19 to 31%), pH (5.6 – 8.4) and, protein concentration (15-17 mg/mL and 25-29.5 mg/mL), which included different ratios of stock protein solution to reservoir buffer in each plate created. Example layouts for expansion plates can be found in **Appendix B**. Attempts were made to recreate the successful conditions with incremental changes in all of the aforementioned variables. An example of a success can be seen in **Figure 4.27** where well H2 of the expansion plate contained the Top 96 E2 condition without any adjusted variable. While expansion crystals have not been sent for X-ray diffraction, they are being utilized for condition optimization and crystal seeding.

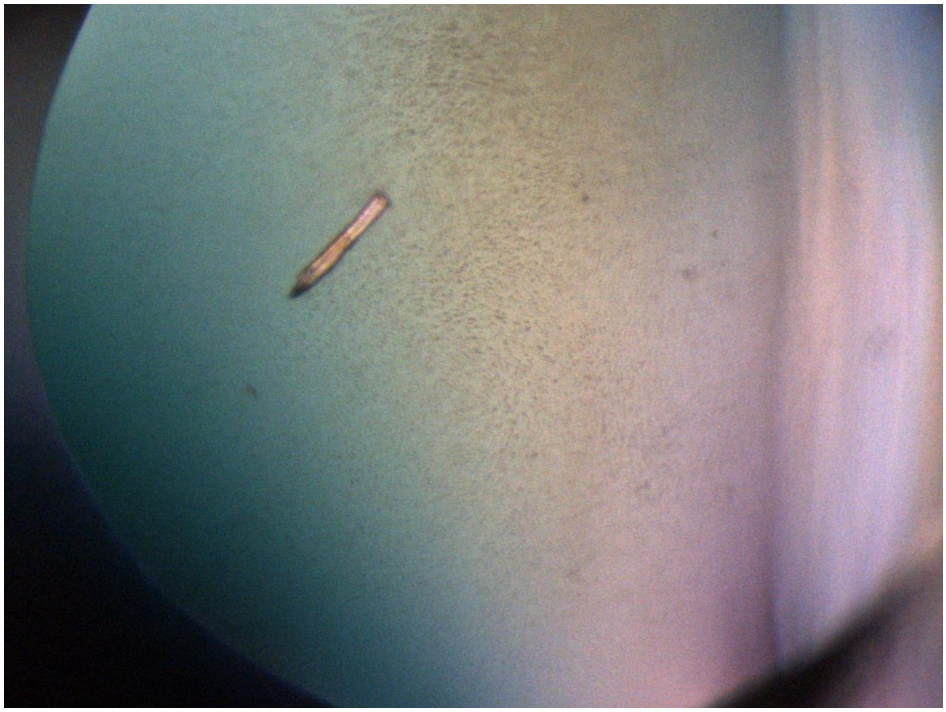


Figure 4.27 Crystal Images from Expansion Trials of WssF. A) A concentration of 15 mg/mL protein was mixed 1:1 with 25 % (w/v) PEG 3350 from well H2 (Top 96 E2 Condition) of expansion plate shown in **Appendix B (Figure G)**.

4.4.4 Molecular Replacement

Molecular replacement (MR) was initially attempted with the WssF data using automated molecular replacement programs (MRage in Phenix and BALBES, MrBump and MoRDa in CCP4 online), which scan the existing PDB databases for hits based on sequence identity to perform MR with the experimental data. Typical outputs from these automated MR trials refined with R_{work}/R_{free} values close to 0.5/0.5; thereby indicating that placement of the solution/model in the electron density is random and cannot be used. As an alternative MR route, the top five Phyre[2] results from the bioinformatics analysis (see **Table 4.2**) were used as models with the Phaser programs from the Phenix suite. However, none of these models were good candidates to pursue, given the poor fit of the model to the calculated electron density (R_{work}/R_{free} values close to 0.5/0.5) post refinement. Future experiments will have to focus on generating better replacement models or pursuing alternate methods to determine the phases and solve the structure.

4.4.5 Small Angle X-Ray Scattering

In parallel with protein crystallization and X-ray diffraction analysis, attempts were also made to structurally characterize WssF through SAXS analysis. The best quality SAXS model data, for both *Ai*WssF and *Pf*WssF were created from the highest concentration samples analyzed (10 mg/ml). Analysis of SAXS data was conducted using Frameslice to merge the data, followed by SCATTER to analyze and fit the data patterns. After merging and fitting the data patterns PRIMUS was used to fit the P(r) function for later processing with online servers (ATSAS and SIBYLS). The results in **Table 4.8** represent the manual data fitting done in SCATTER and automated online processing completed in ATSAS. The R_g value represents the radius of gyration of a particle and can be obtained from Guinier fitting both manually or using an auto R_g function

(129). The R_c value represents the cross-sectional radius, which is used in conjunction with the R_g value to obtain information about the shape of a particle. In the case of globular proteins both R_g and R_c values should be similar as the protein is almost as wide and it is long. This was the case for both *Ai*WssF and *Pf*WssF, with R_g values (30.5 and 25.1) close to their R_c values (18.4 and 16.1), which are predicted (Phyre2) to be globular. Optimization of the SAXS model in GNOM found an optimal max dimensional score (D_{max}) of 50.0 for *Ai*WssF and 53.0 for *Pf*WssF.

Table 4.8. SAXS Analysis Values from SCATTER and GNOM. R_g and R_c obtained from SCATTER and D_{max}(Å) obtained from GNOM.

Construct	R_g Value	R_c Value	D_{max} (Å)
*Ai*WssF	30.5	18.4	50.0
*Pf*WssF	25.1	16.1	53.0

Figure 4.28 is a comparison of the SAXS model of *Ai*WssF, shown in blue, to the Phyre 2 prediction model, shown in red, from **Figure 4.1**. The SAXS models were manipulated in PyMOL to overlay the structures with the Phyre2 prediction models using DAMAVER. Both models share similar generic characteristics of size and globular shape. The alignment of the two however was represented by a CRYSOL chi square value of 563.8. This value signified that the differences between the models were not due to random chance.

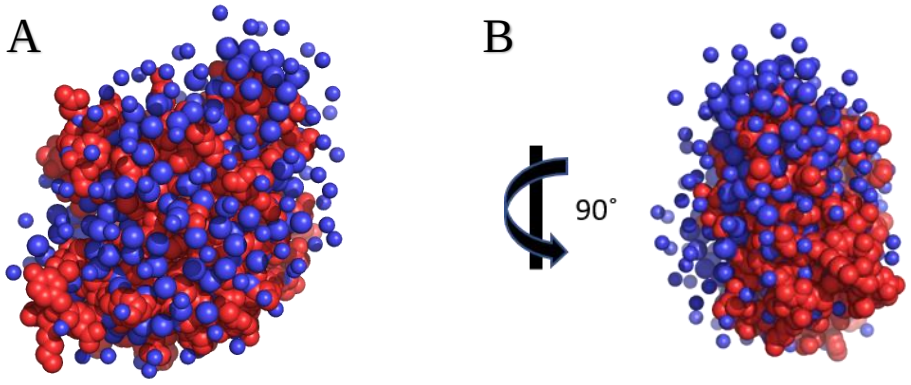


Figure 4.28 *Ai*WssF SAXS and Phyre2 *Ai*WssF Model Comparison Overlays. The Phyre2 prediction model *of* AiWssF is shown in red, while the SAXS model is depicted in blue. Panels A and B represent the same overlayed models rotated 90˚ relative to each other.

In the case of *Pf*WssF the same comparison revealed similar results. **Figure 4.29** compared the SAXS model of *Pf*WssF, shown in blue, to the Phyre 2 prediction model, shown in red, from **Figure 4.1**. Both models shared similar generic characteristics, such as, size and globular shape. The alignment of the two however was represented by a CRYSOL chi square value of 9999.9. This value signified that the differences between the models were not due to random chance.

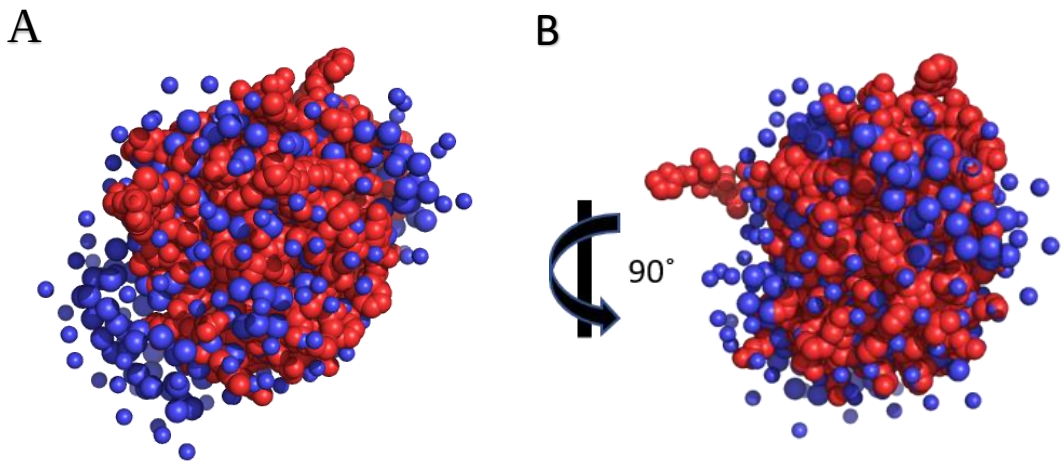


Figure 4.29 *Pf*WssF SAXS and Phyre2 *Pf*WssF Model Comparison Overlays. The Phyre2 prediction model *of Pf*WssF is shown in red, while the SAXS model is depicted in blue. Panels A and B represent the same overlayed models rotated 90˚ relative to each other.

A further comparison was made between the *Ai*WssF SAXS data model and the PDB structure of AlgX (PDB:4KNC)(69). This comparison revealed that the WssF model was approximately one quarter the size of the AlgX model **(Figure 4.30)**. When super imposed and aligned the two models did not share the generic characteristics of shape and size. However, the WssF model was imposed over the N-terminal SGNH hydrolase-like domain, where WssF displays the strongest homology. The fact that WssF was not predicted to contain a CBM domain like AlgX further supported this overlay and fitting orientation. The CRYSOL chi square value was found to be 712.4. As with the previous comparisons the difference between the models was not due to random chance.

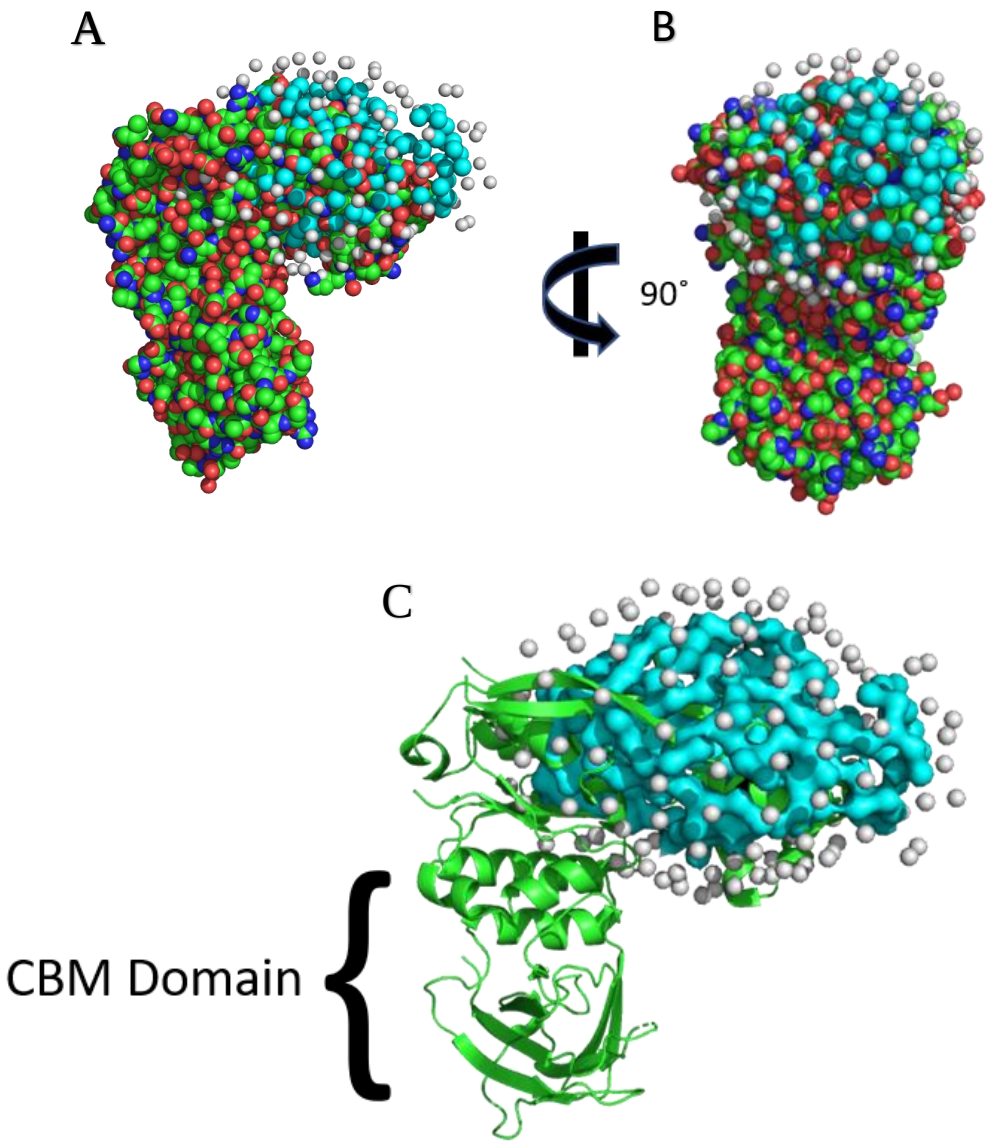


Figure 4.30 *Ai*WssF SAXS model and Solved AlgX structure comparison. The SAXS model *of* *Ai*WssF is shown in turquoise and white, while the AlgX model (PDB:4knc) is depicted in red, green and, blue. Panels A and B represent the same overlayed models rotated 90˚ relative to each other. Panel C represents a surface SAXS model of AiWssF while the cartoon model of AlgX depicted in green (PDB:4knc), the carbohydrate-binding module is labelled below the N-terminal SGNH hydrolase-like domain.

5. Discussion

5.1 Bioinformatics and Prediction models

Biofilms are a common protection mechanism for various bacteria (12, 64). For strains like *A. insuavis* and *P. fluorescens* acetylated cellulose comprises the primary structure of the biofilm. The *wss* operon is responsible for the synthesis and modification of the acetylated cellulose polymer. After synthesis, proteins WssG,H,I,F are responsible for the O-acetylation of cellulose before the polymer is exported from the periplasm (28). Based on sequence prediction **(Figure 5.1),** WssF is likely an acetyltransferase, utilizing a SGNH hydrolase like domain (148). In order to determine the best way to investigate the function of WssF, the amino acid sequences of the WssF constructs were analyzed using bioinformatic techniques **(Table 5.1)**. Sequence analysis from BLASTp predicted a potential catalytic triad formed by residues 17 (Serine), 188 (Aspartic acid) and, 191 (Histidine), which are characteristic of SGNH hydrolases. Furthermore residues 17 (Serine), 69 (Glycine) and, 103 (Threonine) are predicted to form an oxyanion hole for stabilization of tetrahedral intermediates that form during SGNH hydrolase reactions (143). Together these predicted residues may form an active site for WssF acetyltransferase function. Were these predictions to be correct, it would support the hypothesis that WssF functions as an acetyltransferase during polymer modification.

As an acetyltransferase, WssF would first need to remove acetate from an acetyl donor. Secondly, WssF would modify the cellulose polymer through O-acetylation in order to complete its function as an acetyltransferase. To achieve O-acetylation, WssF is most likely utilizing a double displacement reaction, as seen with peptidoglycan O-acetyltransferase B (PatB) (**Figure 5.1**), a protein with similar functionality (143). The double displacement reaction occurs in three

stages; 1) nucleophilic attack of the donor, 2) acid base catalysis of the acceptor and 3) nucleophilic attack of the enzyme substrate complex.

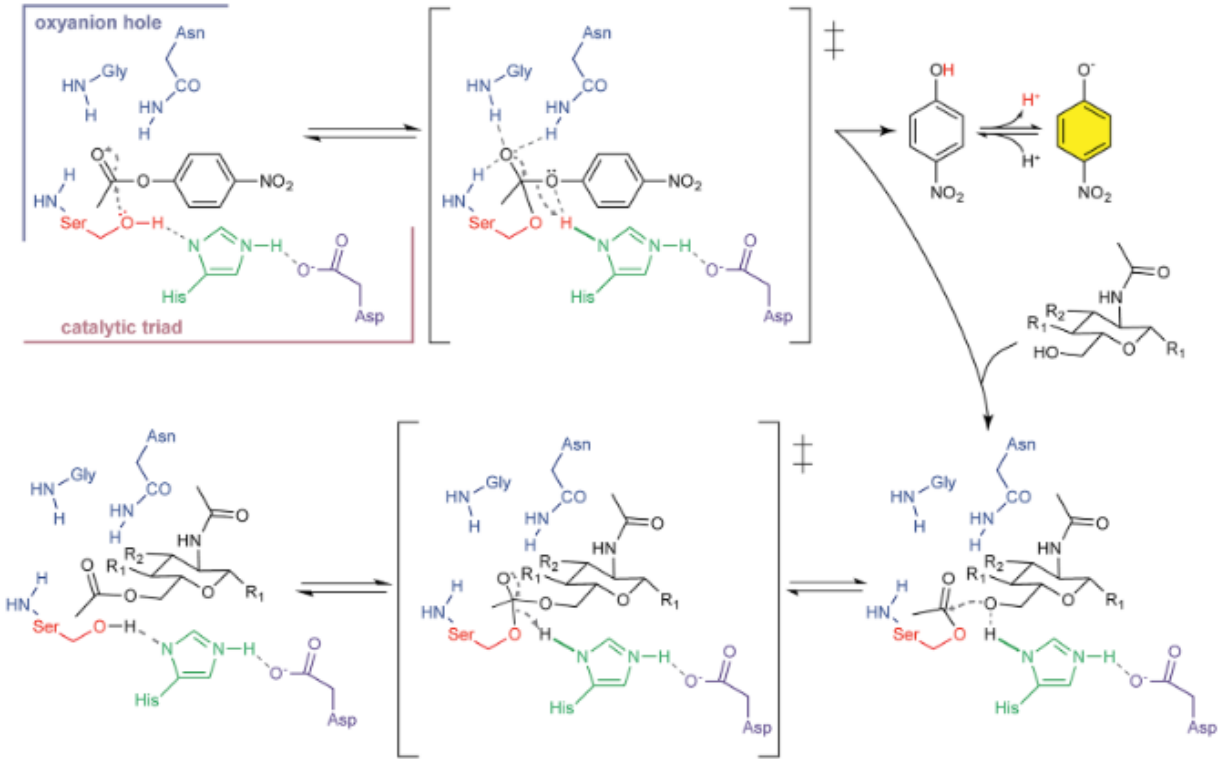


Figure 5.1. Proposed catalytic mechanism of PG O-acetyltransferases. The mechanism is presented with pNP-Ac as the acetyl donor, but *in vivo* presumably acetyl-PatA serves this function. R_1 is either a hydroxyl group or a β-1,4 linked sugar; R_2 is either a hydroxyl group of chito-oligosaccharides or a lactyl-peptide moiety of muramoyl-peptides in PG. The nitrophenolate ion is depicted in yellow to signify its detection at 405nm. Taken from (143)

To further test the predictions, sequence alignments were done to find potential homologues. This sequence alignment produced only low sequence identity matches, so only weak predictions could be made for how WssF folds **(Figures 4.2)**. However, the catalytic triad and oxyanion hole were found to be an exception. Most of the top 50 homologous proteins were some form of hydrolase, many of which were from the SGNH family. There was strong sequence coverage and confidence for the predicted active site. Hypothetical tertiary structure models predicted that low to moderately conserved residues form an α/β/α structure that is

common for members of the SGNH family (149). However, the *A.insuavis* construct formed a less common 4 parallel stranded β sheet compared to the common 5 parallel strands in the *P. fluorescens* construct. Mapping conservation scores onto the hypothetical structures reinforced the predictions that the active site was formed by a catalytic triad and oxyanion hole based on sequence alignment. Conservation scores for the amino acid sites closest to the predicted catalytic triad were the highest (142). High conservation scores indicated that those amino acid sites had the highest consistency among homologs. For an enzyme to retain its function and therefore usefulness the core catalytic residues must be retained. Empirical evidence of the functional and structural characteristics of WssF were required to confirm these predictions.

5.2 Profile of Optimal Protein Expression and Purification (Objective 1)

5.2.1 Expression Profile of WssF

To confidently work with WssF, stocks of pure protein were required. In order to facilitate this necessity, months were spent testing optimal conditions for the expression and purification of WssF. Previous work and assistance from Mr. Adam Miller and Ms. Emily Wilson (former Weadge lab members) allowed us to determine the optimal expression conditions for WssF. Various temperatures and lengths of expression were used to determine the largest yields of protein. For example, temperatures of 16 - 18 ºC were attempted as similar proteins in our lab, such as, WssI, have optimal expression patterns in this range. In parallel, expression at 37 ºC, the optimal growth temperature for *E. coli* was attempted. Choosing a temperature similar to that of a potential internal host may mimic the conditions that proteins like WssF are naturally induced at. However, neither expression produced sufficient protein yields, which led to further expression attempts at 30 ºC, the preferred growth temperature range of *P. fluorescens* and *A. insuavis*, as observed in lab (150). Strong bands on an SDS-PAGE gel were noted as

considerable yields when expressed at 30 ºC (**Figure 4.3**). Following this result, a longer expression window was attempted, and these expression trials revealed 16 - 18 h expression led to the best yields. Once a strong profile for expression was established trials for optimal purification were explored.

5.2.2 Purification Profile of WssF

To purify WssF, increasing imidazole concentrations were tested on an IMAC column to hone in on precise elution conditions. Analysis of imidazole gradient revealed that one wash without imidazole would remove some contaminates while a second with 50mM was the limit for removing additional contaminates without substantial protein elution **(Figure 4.4)**. The purer the eluted enzyme, the less likely secondary purification would be required, which would remove an entire day from the purification process. This reduction was important as an additional day was found to reduce esterase activity to 60 % of the previous day. Once sufficiently pure protein was produced for functional characterization, trials for storage of protein stocks began immediately. Optimal storage would allow for protein samples to be shipped for testing, such as, for HTS to minimize the need for repeated shipments and purifications. Successful storage trials at -80 ºC in 10 % (v/v) glycerol facilitated long-term storage greater than 35 days **(Figure 4.9)**. Concerns were raised when the presence of 10 % (v/v) glycerol led to a 2-fold increase in activity **(Figure 4.7)**. While the presence of glycerol may be stabilizing the protein in solution, its presence complicated HTS and the solubility of the compound screening. Analysis of the trials revealed that after dilution of protein stocks into final well conditions the final concentration was below the 1 % threshold for increased activity. Due to the chemical properties of the hydroxyl groups on glycerol it is possible that it may act as a better acceptor than water. In the case of the DltD enzyme a D-Ala- DltD acyl intermediate undergoes nucleophilic attack by

the 2-hydroxyl group of a glycerol attached to a lipoteichoic acid (LTA), which regenerates the enzyme producing D-Ala-LTA (74). The ability of glycerol to increase the esterase activity of WssF was not investigated further, but this represents an area of future functional characterization work.

For structural characterization purity of the sample was paramount, so secondary purification was required for these analyses. SEC produced a purer product but reduced the overall yield **(Figure 4.4)**. As a result, greater volumes of culture were expressed and purified when attempting to structurally characterize *Ai*WssF. Given the low yields from the tandem purifications, the amount of protein needed to create each crystal screen (96 wells) would generally take double the amount of time as a single step purification. While this created a bottleneck in work flow, successful crystal hits were generated **(Figure 4.26)**.

After optimal expression and purification conditions were determined, final protein yields of 3 mg/mL were typical from 1 L of culture. Protein yields after SEC varied from 2 mg/mL from 1 L of culture resulting in an almost 30 % loss. Overall Objective 1 was a success as sufficient yields of pure protein was obtained to work on functional (Objective 2) and structural characterization (Objective 3) simultaneously. Newer expressions with Rosetta cell lines were found to have 3 - 5 mg/mL from 1 L of culture and increased purity (**Figure 4.5**). Future work on N-terminal constructs and Rosetta cell lines could improve yields for further structural and inhibitor studies.

5.3 Functional Characterization of WssF (Objective 2)

5.3.1 Esterase Characterization of WssF

Characterizing the function of WssF was directed by the predictions made during bioinformatic analysis. If the hypothetical active site functions via a double displacement reaction, then there were two reactions that had to be assessed separately. Testing for esterase activity in the presence of a donor was the best place to start. Analysis of the profiles for both constructs revealed that increasing the concentration of protein within a reaction led to a concomitant increase in esterase product that was a linear relationship; thereby confirming enzyme dependent activity **(Figure 4.6)**. Specific activity values for both WssF constructs were comparable to AlgX (0.0068 ± 0.00057 µmol min^{-1} Mg^{-1}) and PatB (0.51 ± 0.01 µmol*min^{-1}*mg^{-1}) acetyltransferase enzymes found in the literature (69, 151). As AlgX is one of the closest homologs to WssF and an acetyltransferase that also works on an exopolysaccharide polymer, it is not surprising the rates are comparable. In order to test the reliability of measuring esterase activity of WssF Z' values were determined **(Figure 4.22)**. Measuring the average values for both the positive and control samples along with their corresponding standard deviations allowed for confidence that enzyme derived esterase activity was being measured. Z' values were consistently higher than 0.7 indicating that the negative control is significantly different from the positive reaction. The esterase was also used to determine kinetic parameters for this reaction **(Table 4.10)**. From the kinetic study *Ai*WssF was noted to have a higher maximum velocity than *Pf*WssF (V_{max} of 0.0063 and 0.0043 µmol/min, respectively). However, *Pf*WssF had a faster turnover rate (k_{cat} of 0.34±0.056 *s^{-1} vs. 0.24±0.033 *s^{-1}) and lower K_M (7.18±2.71 mM vs. 2.88±1.38 mM) and therefore a better catalytic efficiency (k_{cat}/K_M of 47 s^{-1}*M^{-1} vs. 83 s^{-1}*M^{-1}). In comparison to the literature, WssF has a k_{cat}/K_M comparable to AlgX both of which are only 3

fold different from PatB when reacting with the *p*NP acetyl donor analog (70, 151). The results of the kinetic studies were important as they revealed WssF was capable of efficiently catalyzing esterase activity. As previously mentioned, this hydrolysis would be the vital first step for any acetyltransferase reaction.

Further characterization of WssF esterase activity revealed the likelihood of an important histidine within the esterase reaction. Analysis of the pH profile **(Figure 4.11)** from pH 5.5-9.5 revealed esterase activity, for both constructs, that follows the ionization pattern of an imidazolium ring located on a histidine residue (pKa values of 9.44 and 6.04). A normal or "bell shaped" distribution indicates that protonation of the first nitrogen begins around pH 6.5 where esterase activity began to increase. Activity dropped at pH levels higher than 8.5, which is consistent with deprotonation of the second amino group of the histidine side chain residue. As the second nitrogen becomes fully deprotonated the structure loses the ability to shift its electron density, a vital step when priming the serine residue for catalysis. Based on the predicted catalytic triad, the pH profile is indicative that the conserved H191 is likely responsible for activating the S17 nucleophile between the pH values of 6.5 and 9. In the case of other esterases, like PatB, the same profile of deprotonation affected function of the catalytic triad (143, 148). While pH 8.5 was determined to be the pH where optimal esterase activity occurred, it does not take into account the biological relevance of the pH within the periplasm where WssF is predicted to be located (pH 5.5 - 7.5) (144). Although this would indicate that the enzyme is functioning below peak efficiency in the biological system, this may not matter as there may be other rate limiting steps or the conditions may allow the rate of activity for WssF to match that of the rest of cellulose synthesis. As a result of these physiological conditions, a pH 7 sodium phosphate buffer was chosen when maintaining WssF esterase activity for the routine assays.

5.3.2 Common inhibitor Profile

Among the esterase/hydrolase family there are many enzymes that function with different mechanisms. An easy way to test for functional differences is through the creation of an inhibitor profile **(Table 4.5)**. Common inhibitors for other hydrolases, such as, EDTA, PMSF and, MSF were unsuccessful at removing the esterase functionality of WssF even at biologically fatal concentrations. The retention of substantial esterase activity in the presence of EDTA provides support that WssF is not a metalloenzyme. Catalytic dependence on the presence of metals, such as, cobalt or magnesium, would have been disrupted when the EDTA metal chelator was introduced (152). However, the retention of activity in the presence of PMSF also suggested that a nucleophilic serine is not part of the reaction mechanism, as noted with serine proteases. Serine proteases utilize the same catalytic triad serine, histidine, and, aspartic acid predicted in WssF. Instead of transferring substrate molecules serine proteases bind and cleave peptides. PMSF inhibits serine proteases through sulfonylation of the hyperactive serine; thereby blocking the nucleophilic attack of the substrate (153). It is possible that the selective conditions of the active site for a cellulose polymer prevent the large hydrophobic ring structure of PMSF from binding to the active pocket. However, the use of MSF inhibitor, which is smaller than PMSF, did not lead to complete inhibition. PatB another O-acetyltransferase and Axe1 an acetyl xylan esterase also retained activity in the presence of PMSF while utilizing a catalytically active serine (143, 154). This provides evidence that WssF may have a similar active site that does use a nucleophilic serine but occludes these compounds by some unknown property. Additionally, there have been some unconventional serine proteases, which do not exhibit inhibition by PMSF (143, 155).

5.3.3 WssF Acceptor Specificity Profile

WssF was shown to exhibit esterase activity, a requirement for any acetyltransferase enzyme reaction. However, esterase activity alone was not enough to prove WssF was capable of catalyzing acetyltransferase activity. Thus, activity profiles in the presence of potential cellulose acceptors were used to test for the specificity of WssF **(Figures 4.12)**. Reactions with increasing lengths of cellulose oligomers produced increased esterase activity for *Ai*WssF in the presence of penta and hexamer polymers. Increased activity may be the result of selective conditions present within the active pocket of *Ai*WssF. Without cellulose, the esterase reaction relies on water to remove the acetyl group to regenerate the active site. With the presence of a preferred acceptor, instead of water, the activity was accelerated. The same increase was not seen for the *Pf*WssF construct, which suggests that there are subtle active site differences dependent on organismal background. The specificity of *Ai*WssF for cellulose oligomers was further confirmed when no changes in esterase activity were noted in the presence of a chito-oligomer acceptor instead of cello-oligomer **(Figure 4.13)**. Although structurally similar, the presence of the *N*-acetyl group appears to be enough to block binding of chito polymers. While the effect of cello-oligomers on esterase function is promising in the characterization of WssF as a cellulose acetyltransferase, there is still the possibility that the transferase reaction did not ensue after acetate release from the acetyl-donor.

5.3.4 Thin Layer Chromatography and Mass Spectrometry

To confirm the ability of WssF to act as an acetyltransferase, TLC was used to analyze reaction products. Analysis of the TLC experiments with increasing lengths of cellulose oligomers revealed secondary bands consistent with modified cellulose **(Figure 4.14)**. Longer lengths of cellulose more closely represent the natural cellulose polymer that undergoes O-

acetylation, whereas smaller cellulose oligomers may assume multiple orientations within the active site and interfere with proper binding. As the samples approach more biologically relevant acceptor lengths, the polymer may only assume a few orientations, and this may lead to more productive binding where activity ensues. Mass spectrometry **(Figure 4.15)** was used to confirm the secondary bands on the TLC plate were in fact acetylated cellulose. Analysis of the cellopentaose samples revealed that O-acetylated cellopentaose was present within the samples and that some polymers were O-acetylated twice. A follow up time course analysis **(Figure 4.16)** displayed acetylated cellopentaose in as little as 1 h and then increasing intensity over the course of the reaction. Mass spectrometry **(Figure 4.17)** again confirmed the TLC results and analysis of the MS spectra revealed acetylation profile differences between the enzymes from different organismal backgrounds. These results parallel the differences observed with the esterase assays in the presence of acceptor mentioned earlier. Under the conditions tested, mono-acetylated product from the *Ai*WssF reactions appeared as discrete bands on the TLC and could be detected by MS easily. In contrast, *Pf*WssF reactions analyzed by TLC had a smeared profile and analysis of the MS results noted the presence of a predominant mass of di-acetylated product and little to none of the mono-acetylated product. In retrospect, a lot of effort was spent attempting to improve the smeared bands on the *Pf*WssF TLC plated reactions, however, the smear may represent di-acetylated products that do not have a uniform acetylation profile and are therefore not easily resolved by TLC. To clarify, the cellopentaose consists of 5 sugar units with at least three positions per sugar unit (C2, C3 or C6 for all sugars and also C4 or C1 for the terminal sugars) that can potentially be modified by WssF. Future studies will focus on using NMR analysis of both *Ai*WssF and *Pf*WssF reaction products to elucidate the stoichiometry and regioselectivity of the acetyl modifications. However, even in the absence of this data, the

current TLC and MS results verify for the first time that WssF is a cellulose acetyltransferase and to the best of our knowledge, it is the first cellulose acetyltransferase to be functionally characterized.

After TLC and Mass spectrometry some questions remained about whether the *in vitro* conditions of the tests represent the *in vivo* conditions found within the periplasm of the cell. Within the assays the reaction was occurring without the removal of final product from the reaction mixture. Increasing the final concentration of the product should decrease the rate of reaction as the system reaches equilibrium. Within the periplasm this is not the case as any modified cellulose is cleaved and exported. So, while the reaction conditions within the periplasm are not isolated from external factors, the reaction is continually favouring the forward direction. In the case of kinetic testing these concerns are overcome by measuring initial rates that removes the complication of increasing product concentrations. In contrast, the TLC and Mass spectrometry analyses were testing for the presence or absence of modified cellulose, not specifically the rate of reaction. Future studies may aim at determining the kinetic parameters of the transferase reaction, but they were beyond the scope (and cost) of this thesis. Therefore, while not a perfect representation of the biological scenario, we are still confident that our tests approximate the function of WssF within the Wss biosynthetic apparatus.

5.3.5 Active Site Mutant Profile of *Ai*WssF

Following verification of WssF as an acetyltransferase, experiments were then conducted to understand the mechanism of modifying the cellulose polymer. Specifically, site direct alanine mutants **(Figure 4.18)** of the predicted catalytic triad were created and profiled to explore their role in catalysis similar to what has been accomplished for other homologous acetyltransferase enzymes (70, 143, 156). Determining successful site directed mutagenesis conditions was

difficult as two sets of primers were developed and only *Ai*WssF produced successful mutants. Mutant yield and purity were issues as AiWssF^{D188A} was only 50 % pure and had yields less than half the WT *Ai*WssF sample. Despite the low yields, specific activity assays were still able to be conducted. While individual alanine mutants of the catalytic triad residues did not completely abolish activity, the residual activity was lower than the complimentary inhibitor profiles that targeted the nucleophilic serine (*ie.* PMSF and MSF trials) (**Figures 4.19** and **Table 4.5**). Interestingly, mutation of serine had a smaller effect on esterase activity than histidine, when using pNP-AC as substrate **(Figure 4.19)**. This was surprising, as the loss of the catalytic serine should have hypothetically blocked the nucleophilic attack on the donor. With respect to the histidine, our pH profile analysis indicated the importance of this residue in catalysis. The dramatic loss in activity with this histidine mutant was thus expected, since the histidine can no longer participate in the proton relay that would render the serine nucleophilic. As an alternative substrate, each mutant was also assessed with 4-MUB Ac **(Figure 4.19)** as the acetyl donor. With this substrate, the residual esterase activity is reduced compared to the parallel *p*NPAc reactions. Even the unusually high activity retained by the Serine mutant with *p*NPAc is far lower with 4MUB-Ac. This difference may be explained by the fact that *p*NPAc has been noted to have non-specific hydrolysis occur in other *in vitro* esterase assays (157) due to the fact that the *p*NP is a good leaving group. While 4-MUB-Ac was a better substrate for these analyses, there was still residual esterase activity with each of the catalytic mutants. While puzzling, similar results were noted with the homologous AlgJ and AlgX enzymes from the alginate system and subsequent structural studies verified that these were in fact the catalytic residues for these proteins (70). In the absence of a structure of WssF, we elected to verify the role of these catalytic mutants by also assessing their ability to perform acetyl transfer onto cello-acceptors

(Figure 4.20 and 4.21). In contrast to our WT *Ai*WssF treated samples, no acetylated cellulose product was detected by MS for the catalytic triad mutants. This loss of transferase activity is important, as it verifies the critical nature of the three proposed active site residues to polymer modification.

5.3.6 Mechanism Prediction

The functional characterization of has provided concrete evidence that WssF, is a cellulose acetyltransferase. To achieve polymer modification WssF most likely utilizes a double displacement reaction similar to other O acetyltransferases, such as, PatB, OatA, AlgJ, AlgX and others (**Figure 5.1**) (70, 143). A double displacement reaction (**Figure 5.2**) occurs in three stages; 1) nucleophilic attack of the donor, 2) acid-base catalysis of the acceptor and finally 3) nucleophilic attack of the enzyme substrate complex. We predict that WssF follows a similar reaction mechanism. For the reaction to proceed the carboxyl group of D188 must form a salt bridge with the imidazolium ring of H191. As a result of this bridge, the electron density shifts allowing the second amide to remove the proton from S17; thereby priming it for nucleophilic attack of the donor. After priming S17, the newly formed alkoxide reacts with the carbonyl carbon of the acetate donor forming a tetrahedral intermediate state. The previously predicted oxyanion hole stabilizes the intermediate state before it decays into release of the donor molecule. When the donor molecule leaves it acquires the proton originally from the S17, which is bonded to H191, completing the first stage of the reaction. For the next stage, acid base catalysis, H191 once more removes a proton only this time from a hydroxy group located on the cellulose polymer. Now the alkoxide group on the acceptor repeats the process of the first stage, nucleophilic attack, to form a second tetrahedral intermediate. Once more the oxyanion hole stabilizes the intermediate before it either decays back into the enzyme substrate complex or

releases the acetylated cellulose polymer. After release of the modified cellulose polymer the predicted active site is once more primed for further reactions. Exporting the modified polymer reduces the concentration of product unbalancing the reaction. This exportation continually disrupts equilibrium favouring the reaction to move forward toward product.

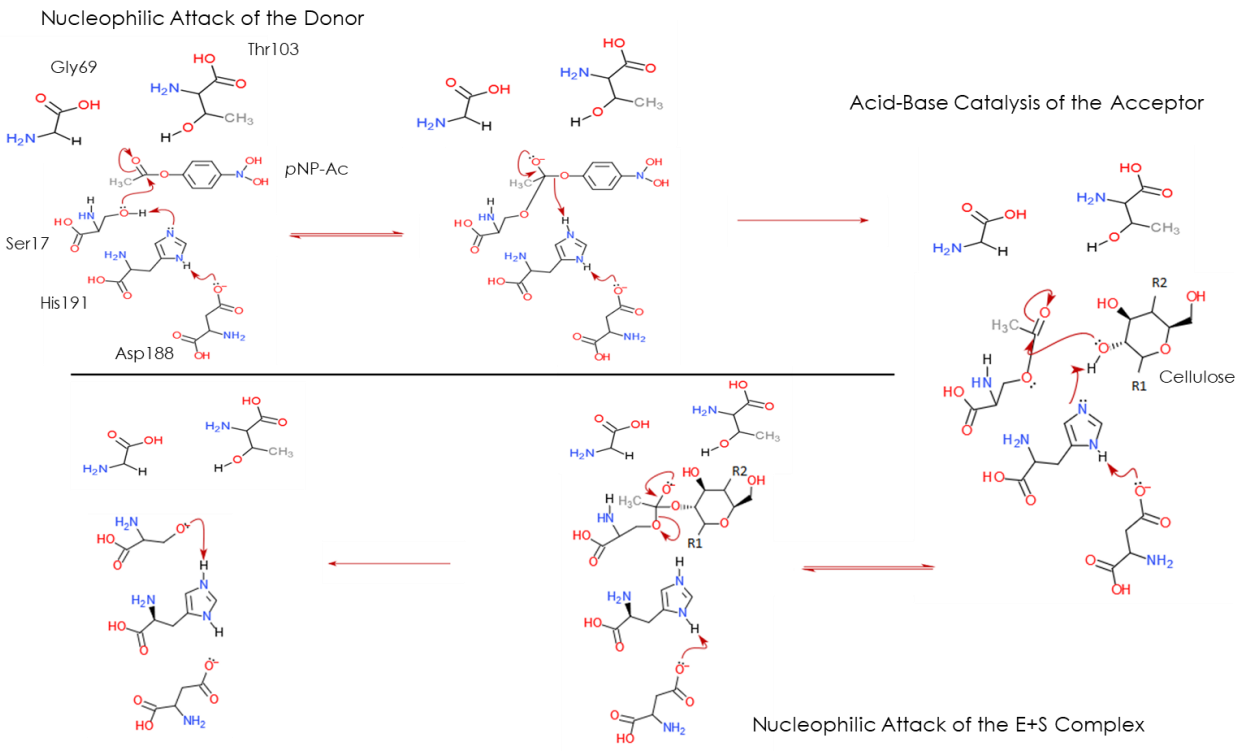


Figure 5.2 Hypothetical WssF reaction mechanism. The mechanism is presented with pNP Ac as the acetyl donor, representing the reaction in vitro. R1 and R2 represent the β1-4 linkages of the cellulose acceptor. G69, T103 and, a hydrogen from S17 are depicted as the oxyanion hole that stabilize the tetrahedral intermediates forming during the acetyltransferase reaction. D188 is depicted forming a salt bridge with H191.

5.3.7 High Throughput Screening Inhibitor Profile

Based on functional profiling and common inhibitor trails (**Table 4.5** and **Figure 4.22**) the WssF protein likely utilizes its esterase activity to serve as a critical acetyltransferase within the cellulose synthase complex. Therefore, after characterizing the esterase activity of WssF HTS was conducted on *Ai*WssF samples in hopes of determining potential inhibitors **(Figure 4.23)**.

HTS was done with the hopes of identifying compounds capable of blocking the esterase activity of WssF, which would subsequently arrest the acetyltransferase activity of the protein. Stopping the acetate modification of cellulose would affect the formation and architecture of the biofilm because without O-acetylation the biofilm will suffer from increased susceptibility to external pressures or may not be able to attach and proliferate to target surfaces in the same manner (158).

As previously mentioned, optimized storage conditions were a priority if HTS was to succeed. Under refrigerated conditions, 4 ˚C, WssF activity was not retained long enough for accumulation or transport of enzyme stocks. Storage at -80 ˚C in 10 % (v/v) glycerol allowed for strong Z′ values, > 0.7, after a month of storage. Successful long-term storage conditions allowed for confidence in the hit results from **Figure 4.23**. The HTS screening process also served to test the esterase activity of WssF for the reliability of measured activity. As previously mentioned Z′ values **(Figure 4.22)** were determined during the initial trials, which revealed that WssF esterase activity was consistently higher than negative controls. With confidence that false positives would not result from measuring WssF esterase activity compound screening could begin.

Promising compounds were selected for based on low IC_{50} values during HTS (**Figure 4.24**), however, analysis of the data from subsequent fluorescence quenching trials found these compounds to be unsuitable as their inhibition was the result of false positives. These false positives were the result of the compounds reducing the fluorescence of the product and not inhibition of esterase activity. Without any compounds from the initial quenching assay verified, 18 promising compounds from the primary and hit confirmation screens were selected for further analysis and these were found to not quench the fluorescence of the product.

From the top hit compounds, compounds A, B and, C were successfully assessed for IC_{50} values at Laurier by Adam DeReuchie-Tan (**Figure 4.25**). Initial testing with dose response

curves for these compounds determined IC_{50} values of 48 – 71 µM. Further testing with pNP-Ac and live cell assays are required to confirm the viability of these compounds as suitable inhibitors of biofilm formation/attachment.

Future inhibitor trials will attempt to find the greatest decrease in the esterase activity of WssF with the lowest concentration of inhibitor possible. Both these factors will be important when attempting to solve the problem of bioavailability. The ability to deliver the smallest effective dose of inhibitor to affected tissues will help to prevent reaching TD_{LO} (Toxic Dose Low**)** concentrations. TD_{LO} are the lowest dosages per unit body weight that produce signs of toxicity (159). As with any drug treatment preventing toxic side effects are paramount to health and safety. Starting with compounds that are highly potent and easily modifiable will help with achieving effective bioavailability within the TD_{LO} limitations.

5.3.8 Summary

Overall Objective 2 was accomplished due to the successful characterization of WssF as an acetyltransferase. Hypothesized to function through a double displacement reaction, the esterase activity of WssF was profiled through kinetic inhibitors and catalytic mutant assays. WssF was found to have strong esterase activity comparable to similar acetyltransferase enzymes (70, 143). Second, transfer activity was observed through the use of TLC and MS. Modified cellulose polymers were only observed in the presence of WssF. Secondary assays with chito-oligomers and the catalytic mutants verified that acetyl transfer is specific for cellulose and that the mechanism is consistent with a double displacement reaction that is typical of homologous enzymes (70, 143). Finally, the successful functional characterization of WssF allowed the initial steps of inhibitor screening to begin through HTS, which will involve future verification trials of biological efficacy.

Further characterization of WssF is possible through the creation of *Pf*WssF mutants and further site directed mutagenesis of the hypothetical oxyanion hole. Also, triple mutant genes would allow for further characterization of the predicted catalytic triad. In addition, WssF could be tested with more physiologically accurate donors, such as, acetyl-CoA or assessed against other Wss enzymes. Expanding the functional studies of WssF will strengthen the role of WssF as an acetyltranferase within the Wss cellulose synthase complex.

5.4 Structural Characterization of WssF (Objective 3)

5.4.1 *Ai*WssF Crystallization and Diffraction

Successful structural characterization of the protein WssF would confirm or correct the predictions made based on sequence analysis. Efforts to crystalize were met with success, as diffraction patterns with good resolution were produced **(Figure 4.27)**. Quality indicators, such as, $CC^{1/2}$, reported the data to be of high quality for the trimmed high-resolution shells. $CC^{1/2}$ measured the Pearson correlation coefficient by randomly splitting the unmerged data into halves and comparing the two data sets (160). Values below 0.3 were considered undesirable or insignificant correlations. Values for the successful diffraction patterns were significantly above 0.3, reaching 0.9 for the trimmed high-resolution shell. However, good diffraction data **(Table 4.7)** is not sufficient to solve the crystal structure, as the phase of the diffracted X-rays remains unknown. The issue known as the phase problem, stems from only the intensity of the collected X-ray being known (161). No information was measured related to where the collected X-ray was in its waveform, also known as the phase. With a strong diffraction pattern, the next logical step was to utilize molecular replacement to aid in determination of the phases.

5.4.2 Molecular Replacement

Molecular replacement relies on the solved structure of a homologue to generate assumptions that can be used to solve the phase of the unknown structure (162). This technique produces bias as it assumes that the unknown structure folds the same as the homologue based on sequence data alone. It is possible that, through chaperones or environmental conditions at the time of folding, the protein may not fold in the assumed pattern when solving the structure. Although not perfect, molecular replacement does aid in solving structures when no other technique is presently viable. However, as previously stated, WssF homologues have low sequence identity and this is likely why molecular replacement trials were met with no success. Without a solved structure to compare against there is no way to determine the missing phase information without further X-ray diffraction. One avenue that will be pursued is the use of heavy atom derivatives or selenomethione incorporation into the crystals to aid with solving the phases.

5.4.4 Small Angle X-Ray Scattering

Without a solved structure, complete structural characterization was not possible. Data collected from SAXS however was used to confirm some basic structural characteristics. SAXS measures density differences within the protein samples (129). This data is then used to create a representative space-filled model of the structure of WssF. These models would normally then be used to compare against solved structures. However, without a solved structure the models were instead compared against the predictive models **(Figures 4.28 and 4.29)** created during the bioinformatic analysis (namely Phyre2 outputs). When overlaid onto each other the models do not line up perfectly and the high chi square values indicated the models did not fit well, as the chi square value is commonly referred to as a measure of fit (163). The lower the value the

higher the likelihood the difference within the compared values are random and therefore significant. With both values significantly higher than the ideal value of one, the models matched poorly. This mismatching was most likely due to the creation of the predicted model from weak homologues with variable levels of residue conservation. To compound the problem further, the globular nature of the WssF proteins lowered the confidence of the SAXS models. Quality scoring around 65 % in the dimension of fit was observed, which indicated the level of agreement between the generated model and the curve that was analyzed. Although the models did not provide a clear pocket or any distinctive structural characteristics, they did further support the globular nature of the bioinformatics models, previously generated. Based on the SAXS models the overall size and globular shape of the WssF constructs roughly matched that of the predicted models despite sequence identity and overall size differences. For example, comparison between the SAXS data and a solved structure of AlgX **(Figure 6.30)** revealed that WssF was one quarter the size AlgX. This was expected as the molecular mass of WssF (24.5kDa) is roughly one quarter that of the reported molecular mass of AlgX (102kDa) due to the fact that AlgX has an appended carbohydrate binding module and some extended β-strand regions (69).

Interestingly, the max dimensional values (Rmax) of the SAXS models were 50 and 53. These values correspond within reason to the cell dimension values for the asymmetric unit of the diffracted crystals (between 42 and 77). While these data are not definitive, they do support each other in the structural elucidation of WssF and the predictive bioinformatics models.

5.4.5 Summary

Significant progress has been made towards accomplishing Objective 3, as high-quality diffraction data sets have been collected for *Ai*WssF, but the phase information still has to be

determined. After attempts at multiple solutions for the phase problem with molecular replacement were unfruitful, selenomethionine trials will be undertaken instead. Structural analysis of WssF was also achieved through the use of SAXS. Although SAXS provided more generic structural data than the previously attempted characterization techniques it was consistent with the predictive models generated as a globular monomeric protein. The generic parameters of WssF, such as, size and general shape match similar acetyltransferases endorsing the hypothetical models.

5.4.6 Future Directions

As part of the next steps to solving the structure of WssF, crystal trials will be made with selenomethionine substitution of the six naturally occurring methionine residues within the sequence of WssF. Substituting selenomethionine into the proteins structure allows for Multiwavelength Anomalous Diffraction (MAD), which is commonly used when solving the phase problem (139, 164). MAD relies on a protein to have at least 1 methionine residue per 75-100 amino acids and the sequence of WssF falls well within this limitation. Integration of selenomethionine residues should not significantly alter the protein structure. Once a selenomethionine crystal is obtained, structural characterization through MAD should be possible. If MAD techniques provide a high-resolution diffraction pattern, then incorporation of the previous diffraction patterns can be used to finally solve for the structure of *Ai*WssF.

In the future, efforts to crystalize and diffract the *Pf*WssF construct will need to be accomplished, which will strengthen the structural characterization already completed for *Ai*WssF. Based on functional complementarity and sequence conservation, we predict that the role of the active residues in the overall mechanism will be confirmed. The overall structure of

WssF, will likely follow the common trends observed for similar EPS acetyltransferases and their homologous SGNH hydrolase motifs.

6. Conclusion

The focus of this research was defining the function of WssF as a putative acetyltransferase. To achieve this goal, WssF was characterized both functionally and structurally. Work was broken down into three objectives, the first objective was profiles of predicted structure and function, which were created along with the development of optimal expression and purification conditions to inform how best to elucidate WssF in the last two objectives. The second objective was functional characterization of WssF as an acetyltransferase, which was achieved through various tests and analysis. Fluorescence and colourimetry spectrophotometry where used to access and profile the ability of WssF to catalyze esterase reactions. Following this, TLC and Mass spectrometry were used to assess and profile the ability of WssF to modify cellulose in the presence of an acetate donor. This data was then used to inform HTS screening and inhibitor assays in the hopes of identifying potential inhibitory compounds for future study. The third objective was structural characterization through X-ray diffraction and SAXS. Success in the production of diffraction quality crystals overcame this bottleneck in the research. However, efforts are still underway to solve the structure of WssF through selenomethionine replacement followed by MAD/SAD techniques. SAXS along with the functional characterization successfully endorsed our predicted model. Overall the first two objectives were completed successfully with the third providing significant insight and a clear path for the full characterization of WssF as an acetyltransferase within cellulose biofilm forming bacteria.

The benefits of understanding how to control biofilm development are broad, ranging from healthcare to agriculture (2, 158). As mentioned, healthcare would benefit directly from being able to inhibit biofilm synthesis or remove preexisting infections. Advances in antibiotics until the last decade or two was achieved through observing and testing substances used by other organisms to repel/kill bacteria (165). Newer antibiotic drugs rely on disrupting specific bacterial pathways to maximize efficacy and minimize side effects (166). With each new drug developed, resistance grows faster and creates more demand for diverse or combinatorial therapies to treat infections. Biofilms are one of the main pathways used by infections to survive antibiotic treatments (158). If the development of antibiotic resistance could be slowed, then the preventative task of eliminating diseases could replace the current priority of treating those already infected. Biofilms are commonly associated with infection in the case of agriculture this is not always the case (17, 18). Treatment of crop and dairy supplies are crucial to preventing disease and food spoilage (12). As previously mentioned even the presence of nonpathogenic biofilm forming bacteria within milk during pasteurization can have negative consequences. However, there are cases where eradication of bacteria is not the only goal. In the case of the radish and tomato crops discussed earlier the presence of some biofilm forming bacteria are immensely beneficial (17, 18). So, whether trying to remove an infection or promote a symbiotic relationship understanding to manipulate and control biofilms is critical. To do this an understanding of biofilm synthesis is important but also as important is the factors that give those biofilms strength. Understanding how cellulose biofilms are O-acetylated is crucial in this understanding, which is why research into proteins like WssF must be undertaken.

7. Integrative nature of this research

Determining the integrative nature of my research was not difficult. Studying the putative acetyltransferase WssF forced me to diversify not only how I thought about microbiology but also how I was going to approach my research. When I started like most people, I thought about my research in terms of how it would affect healthcare and the treatment of patients with pathogenic biofilm forming bacteria. From the very beginning the need to remove biofilms as a means of building antibiotic resistance was clear. What I came to see was that my research would affect other areas of society. As previously mentioned, biofilms are ubiquitous, in the case of agriculture a greater understanding of biofilms could lead to increased crop yields from both prevention of bacterial diseases but also promotion of beneficial microbial communities. As a result of this my research impacts both the areas of food consumption and the world's economies. On an individual's level there are added benefits, personal hygiene as previously discussed is directly impacted through the contamination of contact lens cases. Whether large or small in scope my research had greater possibilities than just helping those in need. While working on my research I discovered that microbiology would not be the only field I had to understand. Growing microbes and controlling their expression was only the beginning of my work. In order to understand how my protein functions, I had to expand my focus into organic and analytical chemistry not to mention biochemistry. Determining reaction kinetics and building pH profiles was a learning experience. More over my research around structural characterization moved my focus firmly into the areas of structural biology and crystallography. Creating crystallization conditions and changing small condition variables was only some of the new skills developed. It is not hard to see how my research is integrative when these are so many connections to healthcare, agriculture, organic chemistry, analytical chemistry, structural biology

and, crystallography. Compounding these connections are all of the disciplines of physics and mathematics required to develop the machines and apparatuses required to complete my research. These are just some of the reasons my research like all of biology is highly integrative in nature.

Acknowledgments

I would like to thank my supervisor Dr. Weadge, along with Dr. Horsman and Dr. Suits, my committee members, for their guidance and support. My external examiner Dr. McDonald and oral defence chair Dr. Culp. Alysha Burnett for her continued support throughout the entirety of the project. Alex Anderson and Will Scott for their guidance and assistance with the structural portion of my project. Past and present members of the Weadge Lab: Christopher Bartlett, Thomas Brenner, Becca Fievoli, Rutik Patel, Adam Deruchie Tan, Monica Garner, Adam Miller, Janice Prescott and, Emily Wilson. As well as Dave Sychantha, Leanne Wybenga-Groot, Rodolfo Gomez, Mark Nitz, members of the Clarke, Howell, Slawson, Suits and Horseman Labs. I would also like to thank Glyconet, NSERC and, Laurier for their funding and graduate services that have allowed me to achieve and continue my research.

References

1. Davey, M. E., and O'toole, G. A. (2000) Microbial Biofilms: from Ecology to Molecular Genetics. *Microbiol. Mol. Biol. Rev.* **64**, 847–867

2. Scharff, R. L. (2012) *Economic Burden from Health Losses Due to Foodborne Illness in the United States*, **75**, 123–131

3. Römling, U., Kjelleberg, S., Normark, S., Nyman, L., Uhlin, B. E., and Åkerlund, B. (2014) Microbial biofilm formation: A need to act. *J. Intern. Med.* **276**, 98–110

4. Costerton, J. W., Stewart, P. S., and Greenberg, E. P. (1999) Bacterial biofilms: a common cause of persistent infections. *Science*. **284**, 1318–22

5. Wolcott, R. D., Rhoads, D. D., Bennett, M. E., Wolcott, B. M., Gogokhia, L., Costerton, J. W., and Dowd, S. E. (2010) Chronic wounds and the medical biofilm paradigm. *J. Wound Care*. **19**, 45–53

6. Folkesson, A., Jelsbak, L., Yang, L., Krogh Johansen, H., Ciofu, O., Høiby, N., and Molin, S. (2012) Adaptation of Pseudomonas aeruginosa to the cystic fibrosis airway: an evolutionary perspective. 10.1038/nrmicro2907

7. Kilvington, S., Shovlin, J., and Nikolic, M. (2013) Identification and susceptibility to multipurpose disinfectant solutions of bacteria isolated from contact lens storage cases of patients with corneal infiltrative events. *Contact Lens Anterior Eye*. **36**, 294–298

8. International Organization for Standardization (2001) ISO 14729:2001(en), Ophthalmic optics — Contact lens care products — Microbiological requirements and test methods for products and regimens for hygienic management of contact lenses. [online] https://www.iso.org/obp/ui/#iso:std:iso:14729:ed-1:v1:en (Accessed October 15, 2017)

9. Sweeney, D. F., Jalbert, I., Covey, M., Sankaridurg, P. R., Vajdic, C., Holden, B. A., Sharma, S., Ramachandran, L., Willcox, M. D. P., and Rao, G. N. (2003) Clinical characterization of corneal infiltrative events observed with soft contact lens wear. *Cornea*. **22**, 435–42

10. Stapleton, F., Keay, L., Edwards, K., Naduvilath, T., Dart, J. K. G., Brian, G., and Holden, B. A. (2008) The Incidence of Contact Lens–Related Microbial Keratitis in Australia. *Ophthalmology*. **115**, 1655–1662

11. Dart, J. K. G., Radford, C. F., Minassian, D., Verma, S., and Stapleton, F. (2008) Risk Factors for Microbial Keratitis with Contemporary Contact Lenses. *Ophthalmology*. **115**, 1647–1654.e3

12. Marchand, S., De Block, J., De Jonghe, V., Coorevits, A., Heyndrickx, M., and Herman, L. (2012) *Biofilm Formation in Milk Production and Processing Environments; Influence on Milk Quality and Safety*, Blackwell Publishing Inc, **11**, 133–147

13. LIU, N. T., NOU, X., BAUCHAN, G. R., MURPHY, C., LEFCOURT, A. M., SHELTON, D. R., and LO, Y. M. (2015) Effects of Environmental Parameters on the

Dual-Species Biofilms Formed by Escherichia coli O157:H7 and Ralstonia insidiosa, a Strong Biofilm Producer Isolated from a Fresh-Cut Produce Processing Plant. *J. Food Prot.* **78**, 121–127

14. Cunty, A., Cesbron, S., Poliakoff, F., Jacques, M. A., and Manceau, C. (2015) Origin of the outbreak in France of Pseudomonas syringae pv. actinidiae biovar 3, the causal agent of bacterial canker of kiwifruit, revealed by a multilocus variable-number tandem-repeat analysis. *Appl. Environ. Microbiol.* **81**, 6773–6789

15. Scortichini, M., Marcelletti, S., Ferrante, P., Petriccione, M., and Firrao, G. (2012) Pseudomonas syringae pv. actinidiae: A re-emerging, multi-faceted, pandemic pathogen. *Mol. Plant Pathol.* **13**, 631–640

16. Takikawa, Y., Serizawa, S., Ichikawa, T., Tsuyumu, S., and Goto, M. The Causal Bacterium of Canker of Kiwifruit in Japan. [online] https://www-jstage-jst-go-jp.libproxy.wlu.ca/article/jjphytopath1918/55/4/55_4_437/_pdf (Accessed July 16, 2017)

17. Moretti, M., Gilardi, G., Gullino, M. U., and Garibaldi, A. (2008) Biological control potential of Achromobacter xylosoxydans for suppressing Fusarium wilt of tomato. *Int. J. Bot.* **4**, 369–375

18. de Boer, M., Bom, P., Kindt, F., Keurentjes, J. J. B., van der Sluis, I., van Loon, L. C., and Bakker, P. A. H. M. (2003) Control of Fusarium Wilt of Radish by Combining Pseudomonas putida Strains that have Different Disease-Suppressive Mechanisms. *Phytopathology*. **93**, 626–32

19. Netstate (1987) Arkansas State Fruit & Vegetable, South Arkansas Vine Ripe Pink Tomato, from NETSTATE.COM. [online] http://www.netstate.com/states/symb/vegetables/ar_tomato.htm (Accessed July 14, 2018)

20. Minor, Travis;Bond, J. (2016) USDA ERS - Tomatoes. *United States Dep. Agric.* [online] https://www.ers.usda.gov/topics/crops/vegetables-pulses/tomatoes.aspx (Accessed July 31, 2017)

21. Koza, A., Kuśmierska, A., McLaughlin, K., Moshynets, O., and Spiers, A. J. (2017) Adaptive radiation of Pseudomonas fluorescens SBW25 in experimental microcosms provides an understanding of the evolutionary ecology and molecular biology of A-L interface biofilm formation. *FEMS Microbiol. Lett.* **364**, 1581–96

22. Spiers, A. J., and Rainey, P. B. (2005) The Pdeudomonas fluorescens SBW25 wrinkly spreader biofilm requires attachment factor, cellulose fibre and LPS interactions to maintain strength and integrity. *Microbiology*. **151**, 2829–2839

23. Costa, A. M., Mergulhão, F. J., Briandet, R., and Azevedo, N. F. (2017) It is all about location: how to pinpoint microorganisms and their functions in multispecies biofilms. *Future Microbiol.* 10.2217/fmb-2017-0053

24. Tavernier, S., Crabbé, A., Tuysuz, M., Stuer, L., Henry, S., Rigole, P., Dhondt, I., and Coenye, T. (2017) Community composition determines activity of antibiotics against multispecies biofilms. *Antimicrob. Agents Chemother.* 10.1128/AAC.00302-17

25. Motta, J.-P., Allain, T., Green-Harrison, L. E., Groves, R. A., Feener, T., Ramay, H.,

Beck, P. L., Lewis, I. A., Wallace, J. L., and Buret, A. G. (2018) Iron Sequestration in Microbiota Biofilms As A Novel Strategy for Treating Inflammatory Bowel Disease. *Inflamm. Bowel Dis.* **24**, 1493–1502

26. Mayo Clinic Staff (2018) Inflammatory bowel disease (IBD) - Symptoms and causes - Mayo Clinic. [online] https://www.mayoclinic.org/diseases-conditions/inflammatory-bowel-disease/symptoms-causes/syc-20353315 (Accessed July 14, 2018)

27. Xiang, H., Cao, F., Ming, D., Zheng, Y., Dong, X., Zhong, X., Mu, D., Li, B., Zhong, L., Cao, J., Wang, L., Ma, H., Wang, T., and Wang, D. Aloe-emodin inhibits Staphylococcus aureus biofilms and extracellular protein production at the initial adhesion stage of biofilm development. 10.1007/s00253-017-8403-5

28. Spiers, A. J., Bohannon, J., Gehrig, S. M., and Rainey, P. B. (2003) Biofilm formation at the air-liquid interface by the Pseudomonas fluorescens SBW25 wrinkly spreader requires an acetylated form of cellulose. *Mol. Microbiol.* **50**, 15–27

29. Navarrete, F., and De La Fuente, L. (2014) Response of Xylella fastidiosa to zinc: decreased culturability, increased exopolysaccharide production, and formation of resilient biofilms under flow conditions. *Appl. Environ. Microbiol.* **80**, 1097–107

30. Flemming, H. C., and Wingender, J. (2010) The biofilm matrix. *Nat. Rev. Microbiol.* **8**, 623–633

31. Nodwell, J. (2016) Changing the way we look at antibiotics: Building up resistance is inevitable, so the need to look beyond penicillin is crucial - ProQuest. *Tor. Star*

32. European Union: Human and Animal Health - Commission continues fight against Antimicrobial Resistance with announcement of second Action Plan - ProQuest (2016) *Asia News Monit.*

33. Trancassini, M., Iebba, V., Citerà, N., Tuccio, V., Magni, A., Varesi, P., De Biase, R. V., Totino, V., Santangelo, F., Gagliardi, A., and Schippa, S. (2014) Outbreak of achromobacter xylosoxidans in an italian cystic fibrosis center: Genome variability, biofilm production, antibiotic resistance, and motility in isolated strains. *Front. Microbiol.* **5**, 1–8

34. Abreu, A. C., Saavedra, M. J., Simões, L. C., and Simões, M. (2016) Combinatorial approaches with selected phytochemicals to increase antibiotic efficacy against Staphylococcus aureus biofilms. *Biofouling*. **32**, 1103–1114

35. Lawrence, J. R., Neu, T. R., and Swerhone, G. D. W. (1998) Application of multiple parameter imaging for the quantification of algal, bacterial and exopolymer components of microbial biofilms. *J. Microbiol. Methods*. **32**, 253–261

36. Sutherland, I. W. (2001) The biofilm matrix - An immobilized but dynamic microbial environment. *Trends Microbiol.* **9**, 222–227

37. Mann, E. E., and Wozniak, D. J. (2012) Pseudomonas biofilm matrix composition and niche biology. *FEMS Microbiol. Rev.* **36**, 893–916

38. McDougald, D., Rice, S. A., Barraud, N., Steinberg, P. D., and Kjelleberg, S. (2012)

Should we stay or should we go: Mechanisms and ecological consequences for biofilm dispersal, **10**, 39–50

39. Nizalapur, S., Kimyon, O., Yee, E., Bhadbhade, M. M., Manefield, M., Willcox, M., Black, D. S., and Kumar, N. (2017) Synthesis and biological evaluation of novel acyclic and cyclic glyoxamide based derivatives as bacterial quorum sensing and biofilm inhibitors. *Org. Biomol. Chem.* 10.1039/C7OB01011G

40. Miller, M. B., and Bassler, B. L. (2001) Quorum Sensing in Bacteria. *Annu. Rev. Microbiol.* **55**, 165–199

41. Bassler, B. L. (1999) How bacteria talk to each other: regulation of gene expression by quorum sensing. *Curr. Opin. Microbiol.* **2**, 582–587

42. Povolotsky, T. L., and Hengge, R. (2012) "Life-style" control networks in Escherichia coli: Signaling by the second messenger c-di-GMP. *J. Biotechnol.* **160**, 10–16

43. Hengge, R. (2009) Principles of c-di-GMP signalling in bacteria. *Nat. Rev. Microbiol.* **7**, 263–273

44. Simm, R., Morr, M., Kader, A., Nimtz, M., and Römling, U. (2004) GGDEF and EAL domains inversely regulate cyclic di-GMP levels and transition from sessility to motility. *Mol. Microbiol.* **53**, 1123–1134

45. Boyd, C. D., and O'Toole, G. A. (2012) Second messenger regulation of biofilm formation: breakthroughs in understanding c-di-GMP effector systems. *Annu. Rev. Cell Dev. Biol.* **28**, 439–62

46. Weber, H., Pesavento, C., Possling, A., Tischendorf, G., and Hengge, R. (2006) Cyclic-di-GMP-mediated signalling within the ? ^{S} network of Escherichia coli. *Mol. Microbiol.* **62**, 1014–1034

47. Olsén, A., Jonsson, A., and Normark, S. (1989) Fibronectin binding mediated by a novel class of surface organelles on Escherichia coll. *Nature*. **338**, 652–655

48. Xavier, J. B., and Foster, K. R. (2007) Cooperation and conflict in microbial biofilms. *Proc. Natl. Acad. Sci.* **104**, 876–881

49. Costerton, J. W., Stewart, P. S., Greenberg, E. P., and Bacterial (1999) A common cause of persistent infections. *Science (80-.).* **284**, 1318

50. Stoodley, P., Lewandowski, Z., and Debeer, D. (1994) Liquid Flow in Biofilm Systems Liquid Flow in Biofilm Systems. *Appl. Environ. Microbiol.* **60**, 2711–2716

51. Tarafdar, A., Sarkar, T. K., Chakraborty, S., Sinha, A., and Masto, R. E. (2018) Biofilm development of Bacillus thuringiensis on MWCNT buckypaper: Adsorption-synergic biodegradation of phenanthrene. *Ecotoxicol. Environ. Saf.* **157**, 327–334

52. Kaplan, J. B. (2010) Biofilm Dispersal: Mechanisms, Clinical Implications, and Potential Therapeutic Uses. *J. Dent. Res.* 10.1177/0022034509359403

53. Lawrence, J. R., Scharf, B., Packroff, G., and Neu, T. R. (2002) Microscale evaluation of the effects of grazing by invertebrates with contrasting feeding modes on river biofilm

architecture and composition. *Microb. Ecol.* 10.1007/s00248-001-1064-y

54. Wong, H. C., Fear, A. L., Calhoont, R. D., Eichinger, G. H., Mayer, R., Amikam, D., Benziman, M., Gelfand, D. H., Meade, J. H., Emerick, A. W., Bruner, R., Ben-Bassat, A., Tal, R., Calhoon, R. D., Eichinger, G. H., Mayer, R., Amikam, D., Benziman, M., Gelfand, D. H., Meade, J. H., and Emerick, A. W. (1990) Genetic organization of the cellulose synthase operon in Acetobacter xylinum. *Proc. Natl. Acad. Sci. U. S. A.* **87**, 8130–8134

55. McNamara, J. T., Morgan, J. L. W., and Zimmer, J. (2015) A Molecular Description of Cellulose Biosynthesis. *Annu. Rev. Biochem.* **84**, 895–921

56. Omadjela, O., Narahari, A., Strumillo, J., Melida, H., Mazur, O., Bulone, V., Zimmer, J., Mélida, H., Mazur, O., Bulone, V., Zimmer, J., and Somerville, C. R. (2013) BcsA and BcsB form the catalytically active core of bacterial cellulose synthase sufficient for in vitro cellulose synthesis. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 17856–61

57. Morgan, J. L. W., McNamara, J. T., and Zimmer, J. (2014) Mechanism of activation of bacterial cellulose synthase by cyclic di-GMP. *Nat. Struct. Mol. Biol.* **21**, 489–96

58. Morgan, J. L. W., Strumillo, J., and Zimmer, J. (2012) Crystallographic snapshot of cellulose synthesis and membrane translocation. *Nature*. **493**, 181–186

59. Remminghorst, U., and Rehm, B. H. A. (2006) In vitro alginate polymerization and the functional role of Alg8 in alginate production by Pseudomonas aeruginosa. *Appl. Environ. Microbiol.* **72**, 298–305

60. Merighi, M., Lee, V. T., Hyodo, M., Hayakawa, Y., and Lory, S. (2007) The second messenger bis-(3'-5')-cyclic-GMP and its PilZ domain-containing receptor Alg44 are required for alginate biosynthesis in Pseudomonas aeruginosa. *Mol. Microbiol.* **65**, 876–895

61. Standal, R., Iversen, T. G., Coucheron, D. H., Fjaervik, E., Blatny, J. M., and Valla, S. (1994) A new gene required for cellulose production and a gene encoding cellulolytic activity in Acetobacter xylinum are colocalized with the bcs operon. *J. Bacteriol.* **176**, 665–72

62. Mazur, O., and Zimmer, J. (2011) Apo- and cellopentaose-bound structures of the bacterial cellulose synthase subunit BcsZ. 10.1074/jbc.M111.227660

63. Hu, L., Grim, C. J., Franco, A. A., Jarvis, K. G., Sathyamoorthy, V., Kothary, M. H., McCardell, B. A., and Tall, B. D. (2015) Analysis of the cellulose synthase operon genes, bcsA, bcsB, and bcsC in Cronobacter species: Prevalence among species and their roles in biofilm formation and cell-cell aggregation. *Food Microbiol.* **52**, 97–105

64. Spiers, A. J., Kahn, S. G., Bohannon, J., Travisano, M., and Rainey, P. B. (2002) Adaptive divergence in experimental populations of Pseudomonas fluorescens. I. Genetic and phenotypic bases of wrinkly spreader fitness. *Genetics*. **161**, 33–46

65. Whitney, J. C., and Howell, P. L. (2013) Synthase-dependent exopolysaccharide secretion in Gram-negative bacteria. *Trends Microbiol.* **21**, 63–72

66. Franklin, M. J., Nivens, D. E., Weadge, J. T., Howell, P. L., Frank, D., Wozniak, D., and Whitfield, C. (2011) Biosynthesis of the Pseudomonas aeruginosa extracellular polysaccharides, alginate, Pel, and Psl. 10.3389/fmicb.2011.00167

67. Chitnis, C. E., and Ohman, D. E. (1993) Genetic analysis of the alginate biosynthetic gene cluster of Pseudomonas aeruginosa shows evidence of an operonic structure. *Mol. Microbiol.* **8**, 583–93

68. Darzins, A., Wang, S. K., Vanags, R. I., and Chakrabarty, A. M. (1985) Clustering of mutations affecting alginic acid biosynthesis in mucoid Pseudomonas aeruginosa. *J. Bacteriol.* **164**, 516–24

69. Riley, L. M., Weadge, J. T., Baker, P., Robinson, H., Codée, J. D. C. C., Tipton, P. A., Ohman, D. E., and Howell, P. L. (2013) Structural and functional characterization of Pseudomonas aeruginosa AlgX: Role of algx in alginate acetylation. *J. Biol. Chem.* **288**, 22299–22314

70. Baker, P., Ricer, T., Moynihan, P. J., Kitova, E. N., Walvoort, M. T. C. C., Little, D. J., Whitney, J. C., Dawson, K., Weadge, J. T., Robinson, H., Ohman, D. E., Codée, J. D. C. C., Klassen, J. S., Clarke, A. J., and Howell, P. L. (2014) P. aeruginosa SGNH Hydrolase-Like Proteins AlgJ and AlgX Have Similar Topology but Separate and Distinct Roles in Alginate Acetylation. *PLoS Pathog.* **10**, e1004334

71. Franklin, M. J., Douthit, S. A., and McClure, M. A. (2004) Evidence that the algI/algJ gene cassette, required for O acetylation of Pseudomonas aeruginosa alginate, evolved by lateral gene transfer. *J. Bacteriol.* **186**, 4759–73

72. Franklin, M. J., and Ohman, D. E. (1996) Identification of algI and algJ in the Pseudomonas aeruginosa alginate biosynthetic gene cluster which are required for alginate O acetylation. *J. Bacteriol.* **178**, 2186–95

73. Franklin, M. J., and Ohman, D. E. (2002) Mutant analysis and cellular localization of the AlgI, AlgJ, and AlgF proteins required for O acetylation of alginate in Pseudomonas aeruginosa. *J. Bacteriol.* **184**, 3000–3007

74. McKay Wood, B., Santa Maria, J. P., Matano, L. M., Vickery, C. R., and Walker, S. (2018) A partial reconstitution implicates DltD in catalyzing lipoteichoic acid D-alanylation. *J. Biol. Chem.* **293**, 17985–17996

75. Gough, J., Karplus, K., Hughey, R., and Chothia, C. (2001) Assignment of homology to genome sequences using a library of hidden Markov models that represent all proteins of known structure. *J. Mol. Biol.* **313**, 903–919

76. Jain, S., Franklin, M. J., Ertesvåg, H., Valla, S., and Ohman, D. E. (2003) The dual roles of AlgG in C-5-epimerization and secretion of alginate polymers in Pseudomonas aeruginosa. *Mol. Microbiol.* **47**, 1123–1133

77. Whitney, J. C., Hay, I. D., Li, C., Eckford, P. D. W., Robinson, H., Amaya, M. F., Wood, L. F., Ohman, D. E., Bear, C. E., Rehm, B. H., Lynne Howell, P., Howell, P. L., and Lynne Howell, P. (2011) Structural basis for alginate secretion across the bacterial outer membrane. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 13083–8

78. Tan, J., Rouse, S. L., Li, D., Pye, V. E., Vogeley, L., Brinth, A. R., El Arnaout, T., Whitney, J. C., Howell, P. L., Sansom, M. S. P., and Caffrey, M. (2014) A conformational landscape for alginate secretion across the outer membrane of Pseudomonas aeruginosa. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 10.1107/S1399004714001850

79. Keiski, C.-L. L., Harwich, M., Jain, S., Neculai, A. M., Yip, P., Robinson, H., Whitney, J. C., Riley, L., Burrows, L. L., Ohman, D. E., and Howell, P. L. (2010) AlgK Is a TPR-Containing Protein and the Periplasmic Component of a Novel Exopolysaccharide Secretin. *Structure*. **18**, 265–273

80. Yu, B. J., Kim, J. A., Moon, J. H., Ryu, S. E., and Pan, J.-G. (2008) The diversity of lysine-acetylated proteins in Escherichia coli. *J. Microbiol. Biotechnol.* **18**, 1529–36

81. Zhang, J., Sprung, R., Pei, J., Tan, X., Kim, S., Zhu, H., Liu, C.-F., Grishin, N. V., and Zhao, Y. (2009) Lysine Acetylation Is a Highly Abundant and Evolutionarily Conserved Modification in *Escherichia Coli*. *Mol. Cell. Proteomics*. **8**, 215–225

82. Tan, W. H., Gilmore, E. C., and Baris, H. N. (2013) Human Developmental Genetics. in *Emery and Rimoin's Principles and Practice of Medical Genetics*, 10.1016/B978-0-12-383834-6.00018-5

83. Ma, Q. (2011) *Role of protein acetylation, formation and dispersal of biofilms, and their impact on insects - ProQuest*. Ph.D. thesis, Texas A&M University

84. Irsfeld, M., Prüß, B. M., and Stafslien, S. J. (2014) Screening the mechanical stability of Escherichia coli biofilms through exposure to external, hydrodynamic shear forces. *J. Basic Microbiol.* **54**, 1403–1409

85. Spiers, A. J. (2014) A mechanistic explanation linking adaptive mutation, niche change, and fitness advantage for the wrinkly spreader. *Int. J. Evol. Biol.* **2014**, 675432

86. Moynihan, P. J., Sychantha, D., and Clarke, A. J. (2014) Chemical biology of peptidoglycan acetylation and deacetylation. *Bioorg. Chem.* **54**, 44–50

87. Gille, S., and Pauly, M. (2012) O-Acetylation of Plant Cell Wall Polysaccharides. *Front. Plant Sci.* **3**, 12

88. Cleland, W. (1970) Steady State Kinetics. in *The Enzymes*, Volume 2 (Boyer, P. ed), pp. 1–65, Elsevier Inc

89. StatCan (2017) The Daily — Fruit and vegetable production, 2016. *StatCan*. [online] http://www.statcan.gc.ca/daily-quotidien/170201/dq170201c-eng.htm (Accessed July 31, 2017)

90. H.S.(Suncheon National Univ., J. (Korea R., Supanjani(McGill Univ., S.-A.-B. (Canada). D. of P. S., and Lee, K.D.(McGill Univ., S.-A.-B. (Canada). D. of P. S. co. (2016) Plant, soil and environment. *Plant, Soil Environ. - UZPI (Czech Republic)*. [online] http://agris.fao.org/agris-search/search.do?recordID=CZ2006000361 (Accessed July 31, 2017)

91. Al-Enazy, A.-A. R., Al-Oud, S. S., Al-Barakah, F. N., and Usman, A. R. (2017) Role of microbial inoculation and industrial by-product phosphogypsum in growth and nutrient

uptake of maize (*Zea mays* L.) grown in calcareous soil. *J. Sci. Food Agric.* **97**, 3665–3674

92. Maji, S., and Chakrabartty, P. K. (2014) Biocontrol of bacterial wilt of tomato caused by Ralstonia solanacearum by isolates of plant growth promoting rhizobacteria. *AJCS*. **8**, 208–214

93. Scales, B. S., Dickson, R. P., LiPuma, J. J., and Huffnagle, G. B. (2014) Microbiology, Genomics, and Clinical Significance of the Pseudomonas fluorescens Species Complex, an Unappreciated Colonizer of Humans. *Clin. Microbiol. Rev.* **27**, 927–948

94. Patel, S. K., Pratap, C. B., Verma, A. K., Jain, A. K., Dixit, V. K., and Nath, G. (2013) *Pseudomonas fluorescens* -like bacteria from the stomach: A microbiological and molecular study. *World J. Gastroenterol.* **19**, 1056

95. Dickson, R. P., Erb-Downward, J. R., Freeman, C. M., Walker, N., Scales, B. S., Beck, J. M., Martinez, F. J., Curtis, J. L., Lama, V. N., and Huffnagle, G. B. (2014) Changes in the Lung Microbiome following Lung Transplantation Include the Emergence of Two Distinct Pseudomonas Species with Distinct Clinical Associations. *PLoS One*. **9**, e97214

96. Wagner, J., Short, K., Catto-Smith, A. G., Cameron, D. J. S., Bishop, R. F., and Kirkwood, C. D. (2008) Identification and Characterisation of Pseudomonas 16S Ribosomal DNA from Ileal Biopsies of Children with Crohn's Disease. *PLoS One*. **3**, e3578

97. Stenhouse, M. A., and Milner, L. V (1992) A survey of cold-growing gram-negative organisms isolated from the skin of prospective blood donors. *Transfus. Med.* **2**, 235–7

98. WHO | Genes and human disease (2010) *WHO*

99. De Baets, F., Schelstraete, P., Van Daele, S., Haerynck, F., and Vaneechoutte, M. (2007) Achromobacter xylosoxidans in cystic fibrosis: Prevalence and clinical relevance. *J. Cyst. Fibros.* **6**, 75–78

100. Pereira, R. H. V, Carvalho-Assef, A. P., Albano, R. M., Folescu, T. W., Jones, M. C. M. F., Leão, R. S., and Marques, E. A. (2011) Achromobacter xylosoxidans: Characterization of strains in Brazilian cystic fibrosis patients. *J. Clin. Microbiol.* **49**, 3649–3651

101. Coward, A., Kenna, D. T. D., Perry, C., Martin, K., Doumith, M., and Turton, J. F. (2016) Use of nrdA gene sequence clustering to estimate the prevalence of different Achromobacter species among Cystic Fibrosis patients in the UK. *J. Cyst. Fibros.* **15**, 479–485

102. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S. S., Wilkins, M. R., Appel, R. D., and Bairoch, A. (2005) Protein Identification and Analysis Tools on the ExPASy Server. in *The Proteomics Protocols Handbook*, pp. 571–607, 10.1385/1-59259-890-0:571

103. Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* 10.1093/nar/25.17.3389

104. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., and Sternberg, M. J. E. E. (2015)

The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845–858

105. Celniker, G., Nimrod, G., Ashkenazy, H., Glaser, F., Martz, E., Mayrose, I., Pupko, T., and Ben-Tal, N. (2013) ConSurf: Using evolutionary data to raise testable hypotheses about protein function. *Isr. J. Chem.* 10.1002/ijch.201200096

106. Glaser, F., Pupko, T., Paz, I., Bell, R. E., Bechor-Shental, D., Martz, E., and Ben-Tal, N. (2003) ConSurf: Identification of functional regions in proteins by surface-mapping of phylogenetic information. *Bioinformatics*. 10.1093/bioinformatics/19.1.163

107. Edgar, R. C. (2004) MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 10.1093/nar/gkh340

108. Bateman, A., Martin, M. J., O'Donovan, C., Magrane, M., Apweiler, R., Alpi, E., Antunes, R., Arganiska, J., Bely, B., Bingley, M., Bonilla, C., Britto, R., Bursteinas, B., Chavali, G., Cibrian-Uhalte, E., Da Silva, A., De Giorgi, M., Dogan, T., Fazzini, F., Gane, P., Castro, L. G., Garmiri, P., Hatton-Ellis, E., Hieta, R., Huntley, R., Legge, D., Liu, W., Luo, J., Macdougall, A., Mutowo, P., Nightingale, A., Orchard, S., Pichler, K., Poggioli, D., Pundir, S., Pureza, L., Qi, G., Rosanoff, S., Saidi, R., Sawford, T., Shypitsyna, A., Turner, E., Volynkin, V., Wardell, T., Watkins, X., Zellner, H., Cowley, A., Figueira, L., Li, W., McWilliam, H., Lopez, R., Xenarios, I., Bougueleret, L., Bridge, A., Poux, S., Redaschi, N., Aimo, L., Argoud-Puy, G., Auchincloss, A., Axelsen, K., Bansal, P., Baratin, D., Blatter, M. C., Boeckmann, B., Bolleman, J., Boutet, E., Breuza, L., Casal-Casas, C., De Castro, E., Coudert, E., Cuche, B., Doche, M., Dornevil, D., Duvaud, S., Estreicher, A., Famiglietti, L., Feuermann, M., Gasteiger, E., Gehant, S., Gerritsen, V., Gos, A., Gruaz-Gumowski, N., Hinz, U., Hulo, C., Jungo, F., Keller, G., Lara, V., Lemercier, P., Lieberherr, D., Lombardot, T., Martin, X., Masson, P., Morgat, A., Neto, T., Nouspikel, N., Paesano, S., Pedruzzi, I., Pilbout, S., Pozzato, M., Pruess, M., Rivoire, C., Roechert, B., Schneider, M., Sigrist, C., Sonesson, K., Staehli, S., Stutz, A., Sundaram, S., Tognolli, M., Verbregue, L., Veuthey, A. L., Wu, C. H., Arighi, C. N., Arminski, L., Chen, C., Chen, Y., Garavelli, J. S., Huang, H., Laiho, K., McGarvey, P., Natale, D. A., Suzek, B. E., Vinayaka, C. R., Wang, Q., Wang, Y., Yeh, L. S., Yerramalla, M. S., and Zhang, J. (2015) UniProt: A hub for protein information. *Nucleic Acids Res.* 10.1093/nar/gku989

109. Bhagwat, M., and Aravind, L. (2007) PSI-BLAST tutorial. *Methods Mol. Biol.* **395**, 177–186

110. DeLano, W. L. (2002) The PyMOL Molecular Graphics System. *Schrödinger LLC wwwpymolorg*. **Version 1.**, http://www.pymol.org

111. Yoshida, N., and Sato, M. (2009) Plasmid uptake by bacteria: A comparison of methods and efficiencies. *Appl. Microbiol. Biotechnol.* **83**, 791–798

112. Guo, J., and Jia, R. A novel inducible expression system for the functional study of toxic gene in bacteria

113. Li, Z., Kessler, W., Van Den Heuvel, J., and Rinas, U. Simple defined autoinduction medium for high-level recombinant protein production using T7-based Escherichia coli

expression systems

114. Karakus, C., Uslu, M., Yazici, D., and Salih, B. A. (2016) Evaluation of immobilized metal affinity chromatography kits for the purification of histidine-tagged recombinant CagA protein. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **1021**, 182–187

115. Robinson, R. L., Neely, A. E., Mojadedi, W., Threatt, K. N., Davis, N. Y., and Weiland, M. H. (2016) Using an FPLC to promote active learning of the principles of protein structure and purification. *Biochem. Mol. Biol. Educ.* 10.1002/bmb.20980

116. Lawrence, A.-M., and Besir, H. U. S. (2009) Staining of proteins in gels with Coomassie G-250 without organic solvent and acetic acid. *J. Vis. Exp.* 10.3791/1350

117. Zhu, Y., Jia, H., Xi, M., Xu, L., Wu, S., and Li, X. (2017) Purification and characterization of a naringinase from a newly isolated strain of Bacillus amyloliquefaciens 11568 suitable for the transformation of flavonoids. *Food Chem.* **214**, 39–46

118. Savaryn, J. P., Toby, T. K., and Kelleher, N. L. (2016) A researcher's guide to mass spectrometry-based proteomics. *Proteomics*. 10.1002/pmic.201600113

119. Fodje, M., Janzen, K., Berg, R., Black, G., Labiuk, S., Gorin, J., and Grochulski, P. (2012) MxDC and MxLIVE: Software for data acquisition, information management and remote access to macromolecular crystallography beamlines. *J. Synchrotron Radiat.* 10.1107/S0909049511056305

120. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* 10.1016/S0076-6879(97)76066-X

121. Bunkóczi, G., Echols, N., McCoy, A. J., Oeffner, R. D., Adams, P. D., and Read, R. J. (2013) Phaser.MRage: Automated molecular replacement. in *Acta Crystallographica Section D: Biological Crystallography*, pp. 2276–2286, International Union of Crystallography, **69**, 2276–2286

122. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67**, 235–242

123. Long, F., Vagin, A. A., Young, P., and Murshudov, G. N. (2007) BALBES: A molecular-replacement pipeline. in *Acta Crystallographica Section D: Biological Crystallography*, pp. 125–132, International Union of Crystallography, **64**, 125–132

124. Keegan, R. M., and Winn, M. D. (2007) Automated search-model discovery and preparation for structure solution by molecular replacement. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **63**, 447–457

125. Lebedev, A., and Vagin, A. (2015) MoRDa, an automatic molecular replacement pipeline. *Acta Crystallogr. Sect. A Found. Adv.* **71**, 4

126. Classen, S., Hura, G. L., Holton, J. M., Rambo, R. P., Rodic, I., McGuire, P. J., Dyer, K.,

Hammel, M., Meigs, G., Frankel, K. A., and Tainer, J. A. (2013) Implementation and performance of SIBYLS: A dual endstation small-angle X-ray scattering and macromolecular crystallography beamline at the Advanced Light Source. *J. Appl. Crystallogr.* **46**, 1–13

127. Hura, G. L., Menon, A. L., Hammel, M., Rambo, R. P., Poole, F. L., Tsutakawa, S. E., Jenney, F. E., Classen, S., Frankel, K. A., Hopkins, R. C., Yang, S. J., Scott, J. W., Dillard, B. D., Adams, M. W. W., and Tainer, J. A. (2009) Robust, high-throughput solution structural analyses by small angle X-ray scattering (SAXS). *Nat. Methods*. **6**, 606–612

128. Dyer, K. N., Hammel, M., Rambo, R. P., Tsutakawa, S. E., Rodic, I., Classen, S., Tainer, J. A., and Hura, G. L. (2014) High-throughput SAXS for the characterization of biomolecules in solution: A practical approach. *Methods Mol. Biol.* **1091**, 245–258

129. Putnam, C. D., Hammel, M., Hura, G. L., and Tainer, J. A. X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. 10.1017/S0033583507004635

130. Förster, S., Apostol, L., and Bras, W. (2010) Scatter: Software for the analysis of nano- and mesoscale small-angle scattering. *J. Appl. Crystallogr.* **43**, 639–646

131. Grant, T. D., Luft, J. R., Carter, L. G., Matsui, T., Weiss, T. M., Martel, A., and Snell, E. H. (2015) The accurate assessment of small-angle X-ray scattering data. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 10.1107/S1399004714010876

132. Foster, S., and Apostol, L. (2010) Scatter. *Eur. Synchrotron Radiat. Facil.*

133. Svergun, D. I., Semyenuk, A., and Gajda, M. J. GNOM manual. [online] https://www.embl-hamburg.de/biosaxs/manuals/gnom.html (Accessed November 5, 2018)

134. Svergun, D. I., Petoukhov, M. V, and Koch, M. H. J. (2001) Determination of domain structure of proteins from x-ray solution scattering. *Biophys. J.* **80**, 2946–2953

135. Franke, D., Petoukhov, M. V., Konarev, P. V., Panjkovich, A., Tuukkanen, A., Mertens, H. D. T., Kikhney, A. G., Hajizadeh, N. R., Franklin, J. M., Jeffries, C. M., and Svergun, D. I. (2017) ATSAS 2.8: A comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J. Appl. Crystallogr.* **50**, 1212–1225

136. Volkov, V. V., and Svergun, D. I. (2003) Uniqueness of *ab initio* shape determination in small-angle scattering. *J. Appl. Crystallogr.* **36**, 860–864

137. Panjkovich, A., and Svergun, D. I. (2016) SASpy: A PyMOL plugin for manipulation and refinement of hybrid models against small angle X-ray scattering data. *Bioinformatics*. 10.1093/bioinformatics/btw071

138. Harrison, P. M., and Sternberg, M. J. E. (1996) The disulphide β-cross: From cystine geometry and clustering to classification of small disulphide-rich Protein folds. *J. Mol. Biol.* 10.1006/jmbi.1996.0664

139. Barton, W. A., Tzvetkova-Robev, D., Erdjument-Bromage, H., Tempst, P., and Nikolov,

D. B. (2006) Highly efficient selenomethionine labeling of recombinant proteins produced in mammalian cells. *Protein Sci.* **15**, 2008–2013

140. Marchler-Bauer, A., Derbyshire, M. K., Gonzales, N. R., Lu, S., Chitsaz, F., Geer, L. Y., Geer, R. C., He, J., Gwadz, M., Hurwitz, D. I., Lanczycki, C. J., Lu, F., Marchler, G. H., Song, J. S., Thanki, N., Wang, Z., Yamashita, R. A., Zhang, D., Zheng, C., and Bryant, S. H. (2015) CDD: NCBI's conserved domain database. *Nucleic Acids Res.* 10.1093/nar/gku1221

141. Marchler-Bauer, A., Bo, Y., Han, L., He, J., Lanczycki, C. J., Lu, S., Chitsaz, F., Derbyshire, M. K., Geer, R. C., Gonzales, N. R., Gwadz, M., Hurwitz, D. I., Lu, F., Marchler, G. H., Song, J. S., Thanki, N., Wang, Z., Yamashita, R. A., Zhang, D., Zheng, C., Geer, L. Y., and Bryant, S. H. (2017) CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res.* **45**, D200–D203

142. Mayrose, I., Graur, D., Ben-Tal, N., and Pupko, T. (2004) Comparison of Site-Specific Rate-Inference Methods for Protein Sequences: Empirical Bayesian Methods Are Superior. *Mol. Biol. Evol.* **21**, 1781–1791

143. Moynihan, P. J., and Clarke, A. J. (2014) Mechanism of action of peptidoglycan O-acetyltransferase B involves a Ser-His-Asp catalytic triad. *Biochemistry*. **53**, 6243–6251

144. Wilks, J. C., and Slonczewski, J. L. (2007) pH of the cytoplasm and periplasm of Escherichia coli: Rapid measurement by green fluorescent protein fluorimetry. *J. Bacteriol.* 10.1128/JB.00615-07

145. Kounnass, M. Z., Wolz, R. L., Gorbea, C. M., and Bond5, J. S. (1991) *THE JOURNAL OF BIOLOGICAL CHEMISTRY Meprin-A and-B CELL SURFACE ENDOPEPTIDASES OF T H E MOUSE KIDNEY**, [online] http://www.jbc.org/content/266/26/17350.full.pdf (Accessed December 20, 2018)

146. Okamoto, Y., Otsuka-Fuchino, H., Horiuchi, S., Tamiya, T., Matsumoto, J. J., and Tsuchiya, T. (1993) Purification and characterization of two metalloproteinases from squid mantle muscle, myosinase I and myosinase II. *Biochim. Biophys. Acta*. **1161**, 97–104

147. Evans, P. R., and Murshudov, G. N. (2013) How good are my data and what is the resolution? *Acta Crystallogr. Sect. D Biol. Crystallogr.* **69**, 1204–1214

148. Lee, L.-C., Lee, Y.-L., Leu, R.-J., and Shaw, J.-F. (2006) Functional role of catalytic triad and oxyanion hole-forming residues on enzyme activity of Escherichia coli thioesterase I/protease I/phospholipase L 1. *Biochem. J.* **397**, 69–76

149. Akoh, C. C., Lee, G.-C., Liaw, Y.-C., Huang, T.-H., and Shaw, J.-F. (2004) GDSL family of serine esterases/lipases. *Prog. Lipid Res.* **43**, 534–552

150. Quagliotto, L., Azziz, G., Bajsa, N., Vaz, P., Pérez, C., Ducamp, F., Cadenazzi, M., Altier, N., and Arias, A. (2009) Three native Pseudomonas fluorescens strains tested under growth chamber and field conditions as biocontrol agents against damping-off in alfalfa. *Biol. Control*. 10.1016/j.biocontrol.2009.05.006

151. Moynihan, P. J., and Clarke, A. J. (2014) Substrate specificity and kinetic characterization

of peptidoglycan O-acetyltransferase B from Neisseria gonorrhoeae. *J. Biol. Chem.* **289**, 16748–16760

152. Flora, S. J. S. S., and Pachauri, V. (2010) *Chelation in metal intoxication*, Multidisciplinary Digital Publishing Institute (MDPI), **7**, 2745–2788

153. Sharma, A., and Radha Kishan, K. V. (2011) Serine protease inhibitor mediated peptide bond re-synthesis in diverse protein molecules. *FEBS Lett.* **585**, 3465–3470

154. Margolles-clark, E., Tenkanen, M., Soderlund, H., and Penttila, M. (1996) contains an active-site serine residue and a cellulose-binding domain. *Eur. J. Biochem.* **237**, 553–560

155. Ekici, Ö. D., Paetzel, M., and Dalbey, R. E. (2008) Unconventional serine proteases: Variations on the catalytic Ser/His/Asp triad configuration. *Protein Sci.* 10.1110/ps.035436.108

156. Weadge, J. T., and Clarke, A. J. (2007) Neisseria gonorrheae O-acetylpeptidoglycan esterase, a serine esterase with a Ser-His-Asp catalytic triad. *Biochemistry*. 10.1021/bi700254m

157. De Yan, H., Zhang, Y. J., Liu, H. C., Zheng, J. Y., and Wang, Z. (2013) Influence of ammonium salts on the lipase/esterase activity assay using *p* -nitrophenyl esters as substrates. *Biotechnol. Appl. Biochem.* **60**, 343–347

158. Stewart, P. S., and William Costerton, J. (2001) Antibiotic resistance of bacteria in biofilms. *Lancet*. **358**, 135–138

159. *(Revised 1/03) TOXICOLOGY AND EXPOSURE GUIDELINES* [online] http://ehs.unl.eduhttp//ehs.unl.edu/ (Accessed December 17, 2018)

160. Karplus, P. A., and Diederichs, K. (2012) Linking crystallographic model and data quality. *Science*. **336**, 1030–3

161. Taylor, G. (2003) The phase problem. *Natl. Libr. Med.* **59**, 1991–1890

162. Boggon, T. J., and Shapiro, L. (2000) Screening for phasing atoms in protein crystallography. *Structure*. **8**, R143–R149

163. Zhu, J. (2017) *A Power Study of Gffit Statistics as Components of Pearson Chi-Square - ProQuest*. Ph.D. thesis, Arizona State University, [online] https://search-proquest-com.libproxy.wlu.ca/docview/1898267259/abstract/94DB43429A394C52PQ/2?accountid=15090 (Accessed October 25, 2018)

164. Hendrickson, W. A. (2014) Anomalous diffraction in crystallographic phase evaluation. *Q. Rev. Biophys.* **47**, 49–93

165. Clardy, J., Fischbach, M. A., and Currie, C. R. (2009) The natural history of antibiotics. *Curr. Biol.* 10.1016/j.cub.2009.04.001

166. Kohanski, M. A., Dwyer, D. J., and Collins, J. J. (2010) How antibiotics kill bacteria: From targets to networks. *Nat. Rev. Microbiol.* 10.1038/nrmicro2333

Appendix A.

Achromobacter insuavis AXX-A contig00047, whole genome shotgun sequence
TCATTTCTTCTGCAGCGCCTTGACCTGCGGCAGGTCGCCCAATTGCTTGGCAATGGCGTCGCCCCAGGC
TTTGTAGCCGGTGTTGGTGTAGTGCTGGCCGTCGGTGGTGGGCCACTGGCCGGGCGTCGAGAACTGGG
TCGAATCGATGTACGCGCAAGGCGCCACGTTGCTGCCCAGGAAGGCCGACACCTGCTTGACCCGCGCG
AAATTCTTGTTGTACTTGCCGCCCTCGGTGCCCCAGCTGGGACCGATCCAGACGCAGGCCGTGCCGGT
GCCGGAGATGGCCTTGGTCAGCGCGGTCGTCTGCTGCCAGGCCCAGGTCTTGGGAAAGGACGGATTGG
TGTAGCCGGCCATGGTGTCGCCGAGCACGATCACCACCAGGTTGGGCTTCTTGGCCGCGATCAGGTCG
GTCACCGGCGTGGTGCTGGCGTCCTTGCCCTTGAGCACGATCTGGCCGCCGTTCAGGCGCTCGGCGCC
GCCGCAGGTGCTGGGCGTGGCCTTGAGCCAGTCGCCGGCATTGCTGCCGCACACGCCGATGGTCTGCA
CCACCGCGCCCTGCTTGACCAGGTCGTCCTGCAGGGTCGTGATCAGGTAGCCCGGCGCGGCCAGATGG
CTGTCGCCGATCAGCAGGATGCTCATTCCGGCAAAAAGAGAGGTCAGCAT

Achromobacter insuavis AXX-A
MLTSLFAGMSILLIGDSHLAAPGYLITTLQDDLVKQGAVVQTIGVCGSNAGDWLKATPSTCGGAERLNGG
QIVLKGKDASTTPVTDLIAAKKPNLVVIVLGDTMAGYTNPSFPKTWAWQQTTALTKAISGTGTACVWIGPS
WGTEGGKYNKNFARVKQVSAFLGSNVAPCAYIDSTQFSTPGQWPTTDGQHYTNTGYKAWGDAIAKQLG
DLPQVKALQKK

Construct 1 (ordered from Genscript) – pET24-CtermHis
MLTSLFAGMSILLIGDSHLAAPGYLITTLQDDLVKQGAVVQTIGVCGSNAGDWLKATPSTCGGAERLNGG
QIVLKGKDASTTPVTDLIAAKKPNLVVIVLGDTMAGYTNPSFPKTWAWQQTTALTKAISGTGTACVWIGPS
WGTEGGKYNKNFARVKQVSAFLGSNVAPCAYIDSTQFSTPGQWPTTDGQHYTNTGYKAWGDAIAKQLG
DLPQVKALQKKlehhhhhh

Pseudomonas fluorescens SBW25
ATGCCTGTTTCTGCGATTGCCGGCCTAACCATGCTGGTATTGGGCGAAAGTCATATGAGCTTTCCCGAT
TCGTTGCTCAACCCGCTGCAAGACAACCTCACCAAGCAAGGCGCGGTGGTTCACTCCATCGGTGCGTG
CGGTGCCGGTGCGGCGGATTGGGTTGTGCCGAAAAAAGTCGAATGCGGCGGCGAACGCACGCCCACC
GGCAAGGCCGTGATCTATGGCAAAAACGCCATGAGCACCACGCCGATCCAGGAGCTGATCGCCAAAG
ACAAACCCGACGTGGTCGTGTTGATCATCGGCGACACCATGGGCTCCTACACCAACCCGGTGTTCCCT
AAAGCCTGGGCCTGGAAAAGCGTGACCTCGCTGACCAAAGCCATCACCGACACCGGCACCAAGTGCG
TGTGGGTCGGCCCGCCATGGGGCAAGGTCGGTTCGCAGTACAAGAAAGACGACACCCGCACCAAGCT
GATGTCGTCGTTCCTCGCCAGTAACGTGGCGCCGTGCACCTACATCGACTCGCTGACGTTCTCGAAGCC
AGGCGAGTGGATCACCACCGACGGCCAGCACTTCACCATCGACGGCTACCAGAAGTGGGCCAAGGCC
ATCGGTACAGCCTTGGGTGACCTGCCGCCATCGGCCTACGGTAAAGGAAACAAATAA

Pseudomonas fluorescens SBW25
MPVSAIAGLTMLVLGESHMSFPDSLLNPLQDNLTKQGAVVHSIGACGAGAADWVVPKKVECGGERTPTG
KAVIYGKNAMSTTPIQELIAKDKPDVVVLIIGDTMGSYTNPVFPKAWAWKSVTSLTKAITDTGTKCVWVGP
PWGKVGSQYKKDDTRTKLMSSFLASNVAPCTYIDSLTFSKPGEWITTDGQHFTIDGYQKWAKAIGTALGD
LPPSAYGKGNK

Construct 2 (ordered from Genscript) – pET24-CtermHis
MPVSAIAGLTMLVLGESHMSFPDSLLNPLQDNLTKQGAVVHSIGACGAGAADWVVPKKVECGGERTPTG
KAVIYGKNAMSTTPIQELIAKDKPDVVVLIIGDTMGSYTNPVFPKAWAWKSVTSLTKAITDTGTKCVWVGP
PWGKVGSQYKKDDTRTKLMSSFLASNVAPCTYIDSLTFSKPGEWITTDGQHFTIDGYQKWAKAIGTALGD
LPPSAYGKGNKlehhhhhh

Appendix B.

To help iterate the function of these plates the following figures were created in excel to visualize the layout of a standard 96 well microtiter plate. Each row of the plates contains a letter designation while the columns are designated by a column number. These prelabelled plates would allow for effective planning and efficient execution of the required assays.

	1	2	3	4	5	6	7	8	9	10	11	12
A	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12
B	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11	B12
C	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
D	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12
E	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12
F	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12
G	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11	G12
H	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12

Figure B.1 Example layout of an Assay Plate. Each well contains up to 300µL of liquid. Each plate containing 96 wells. Made from black plastics for Fluourescence spectrometry or clear plastics for colourimetric analysis.

	1	2	3	4	5	6	7	8	9	10
A	Blank	Blank	Blank	Blank	Blank	Blank	Blank	Blank	Blank	Blank
B	Blank	Blank	Blank	Blank	Blank	Blank	Blank	Blank	Blank	Blank
C	Blank	Blank	Blank	Blank	Blank	Blank	Blank	Blank	Blank	Blank
D	Blank	Blank	Blank	Blank	Blank	Blank	Blank	Blank	Blank	Blank
E	1.13µL	1.13µL	1.88µL	1.88µL	1.81µL	1.81µL	3.75µL	3.75µL	5.63µL	5.63µL
F	1.13µL	1.13µL	1.88µL	1.88µL	1.81µL	1.81µL	3.75µL	3.75µL	5.63µL	5.63µL
G	1.13µL	1.13µL	1.88µL	1.88µL	1.81µL	1.81µL	3.75µL	3.75µL	5.63µL	5.63µL
H	1.13µL	1.13µL	1.88µL	1.88µL	1.81µL	1.81µL	3.75µL	3.75µL	5.63µL	5.63µL

Protein Concentration	0µM	0.3µM	0.5µM	0.75µM	1µM	1.5µM

Figure B.2 WssF Esterase Activity Assay Layout. Blank wells represent negative control conditions (150µL of 100mM Sodium Phosphate pH 7, 15µL of 80mM 4-MUB Ac dissolved in DMSO, 135µL of Millipore filtered H_20). Samples are labelled with the volume of WssF protein added from a stock (2mg/mL). Sample wells replace a volume of H_2O with an equal volume of protein labelled on each well. Colour is used to denote the final well concentration of the protein WssF.

	1	2	3	4	5	6	7	8
A	Blank	Blank	Blank	Blank	WssF	WssF	WssF	WssF
B	Blank	Blank	Blank	Blank	WssF	WssF	WssF	WssF
C	Blank	Blank	Blank	Blank	WssF	WssF	WssF	WssF
D	Blank	Blank	Blank	Blank	WssF	WssF	WssF	WssF
E	Blank	Blank	Blank	Blank	WssF	WssF	WssF	WssF

Figure B.3 Z′ 4-MUB-Ac Fluorescence assay Layout. Blank wells represent negative control conditions (45µL of 50mM Sodium Phosphate pH 7, 0.5mM 4-MUB Ac and, 5 % (v/v) DMSO). Sample wells contained constant concentrations of protein across the plate. Multiple assays were conducted with various concentrations of protein (0.3µM, 0.5µM, 1µM).

	1	2	3	4	5	6	7
A	Blank	Blank	Blank	Blank	Blank	Blank	Blank
B	Blank	Blank	Blank	Blank	Blank	Blank	Blank
C	Blank	Blank	Blank	Blank	Blank	Blank	Blank
D	Blank	Blank	Blank	Blank	Blank	Blank	Blank
E	7.5µL	7.5µL	7.5µL	7.5µL	7.5µL	7.5µL	7.5µL
F	7.5µL	7.5µL	7.5µL	7.5µL	7.5µL	7.5µL	7.5µL
G	7.5µL	7.5µL	7.5µL	7.5µL	7.5µL	7.5µL	7.5µL
H	7.5µL	7.5µL	7.5µL	7.5µL	7.5µL	7.5µL	7.5µL

pNP Ac Concentration	0.25mM	0.5mM	1mM	2mM	4mM	6mM	8mM

Figure B.4 Substrate Kinetic Assay Layout. Blank wells represent negative control conditions (75µL of 100mM Sodium Phosphate, 7.5µL of pNP Ac, 67.5µL of Millipore H_2O). The concentration of pNP Ac stock solutions vary depending upon desired final well concentration. Sample wells contain 7.5µL of protein from a 1mg/ml stock resulting in a final well concentration of 2µM. Added protein volume replaces an equal volume of H_2O within the sample wells.

	1	2	3	4	5	6	7	8	9	10	11
A	Blank	Blank	Blank	Blank	Blank	Blank	Blank	Blank	Blank	Blank	Blank
B	Blank	Blank	Blank	Blank	Blank	Blank	Blank	Blank	Blank	Blank	Blank
C	Blank	Blank	Blank	Blank	Blank	Blank	Blank	Blank	Blank	Blank	Blank
D	Blank	Blank	Blank	Blank	Blank	Blank	Blank	Blank	Blank	Blank	Blank
E	7.5µL	7.5µL	7.5µL	7.5µL	7.5µL	7.5µL	7.5µL	7.5µL	7.5µL	7.5µL	7.5µL
F	7.5µL	7.5µL	7.5µL	7.5µL	7.5µL	7.5µL	7.5µL	7.5µL	7.5µL	7.5µL	7.5µL
G	7.5µL	7.5µL	7.5µL	7.5µL	7.5µL	7.5µL	7.5µL	7.5µL	7.5µL	7.5µL	7.5µL
H	7.5µL	7.5µL	7.5µL	7.5µL	7.5µL	7.5µL	7.5µL	7.5µL	7.5µL	7.5µL	7.5µL

Buffer	Mes	Mes	Mes	Phosphate	Phosphate	Phosphate	Tris	Tris	Tris	Tris	Tris
pH	5.5	6	6.5	6.5	7	7.5	7.5	8	8.5	9	9.25

Figure B.5 pH Profile Assay Layout. Blank wells represent negative control conditions (75µL of 100mM Sodium Phosphate, 7.5µL of 80mM pNP Ac, 67.5µL of Millipore H_2O). Utilized Buffers: Mes from pH 5.5-6.5, Sodium Phosphate from pH 6.5-7.5 and, Tris HCl from pH 7.5-9.25. Final well concentrations were 4µM protein and 6mM 4-Nitrophenyl acetate. Added protein volume replaces an equal volume of H_2O within the sample wells.

	1	2	3	4	5	6	7	8	9	10	11	12
A	Blank	Blank	Blank	Blank	*Ai*WssF	*Ai*WssF	*Ai*WssF	*Ai*WssF	*Pf*WssF	*Pf*WssF	*Pf*WssF	*Pf*WssF
B	Blank	Blank	Blank	Blank	*Ai*WssF	*Ai*WssF	*Ai*WssF	*Ai*WssF	*Pf*WssF	*Pf*WssF	*Pf*WssF	*Pf*WssF
C	Blank	Blank	Blank	Blank	*Ai*WssF	*Ai*WssF	*Ai*WssF	*Ai*WssF	*Pf*WssF	*Pf*WssF	*Pf*WssF	*Pf*WssF
D	Blank	Blank	Blank	Blank	*Ai*WssF	*Ai*WssF	*Ai*WssF	*Ai*WssF	*Pf*WssF	*Pf*WssF	*Pf*WssF	*Pf*WssF
E	Blank	Blank	Blank	Blank	*Ai*WssF	*Ai*WssF	*Ai*WssF	*Ai*WssF	*Pf*WssF	*Pf*WssF	*Pf*WssF	*Pf*WssF
F	Blank	Blank	Blank	Blank	*Ai*WssF	*Ai*WssF	*Ai*WssF	*Ai*WssF	*Pf*WssF	*Pf*WssF	*Pf*WssF	*Pf*WssF
G	Blank	Blank	Blank	Blank	*Ai*WssF	*Ai*WssF	*Ai*WssF	*Ai*WssF	*Pf*WssF	*Pf*WssF	*Pf*WssF	*Pf*WssF
H	Blank	Blank	Blank	Blank	*Ai*WssF	*Ai*WssF	*Ai*WssF	*Ai*WssF	*Pf*WssF	*Pf*WssF	*Pf*WssF	*Pf*WssF

Inhibitor (5mM)	EDTA	MSF	PMSF	None

Figure B.6 Inhibitor Profile Assay Layout. Blank wells represent negative control conditions (75µL of 100mM Sodium Phosphate, 7.5µL of 80mM pNP Ac, 67.5µL of Millipore H_2O). Final well concentrations were 4µM protein and 5mM 4-Nitrophenyl acetate.

		1	2	3	4	5	6	7	8	9	10	11	12
15mg/ml	A	20%	20%	23%	23%	25%	25%	27%	27%	30%	30%	N/A	N/A
17mg/ml	B	20%	20%	23%	23%	25%	25%	27%	27%	30%	30%	N/A	N/A
13mg/ml	C	20%	20%	23%	23%	25%	25%	27%	27%	30%	30%	N/A	N/A
29.5mg/ml	D	20%	20%	23%	23%	25%	25%	27%	27%	30%	30%	N/A	N/A
25mg/ml	E	20%	20%	23%	23%	25%	25%	27%	27%	30%	30%	N/A	N/A
29.5mg/ml	F	15%	15%	18%	18%	20%	20%	22%	22%	25%	25%	N/A	N/A
25mg/ml	G	15%	15%	18%	18%	20%	20%	22%	22%	25%	25%	N/A	N/A
	H	25%	25%	25%	25%	25%	25%	25%	25%	20%	20%	20%	20%

% (w/v) PEG 3350
0.1 M HEPES:NaOH pH 7.5, % (w/v) PEG 3350
0.1 M Bis-Tris:HCl pH 6.5, % (w/v) PEG MME 5000

Figure B.7 Protein Concentration Expansion Plate Layout. Example of expansion plates made from successful conditions. Shaded coded based on well condition. H1-H4, 15mg/mL; H5-H12, 25mg/mL. Percentages represent concentration of PEG used for each condition.

	1	2	3	4	5	6	7	8	9	10	11	12
A	5.6	6.2	6.4	6.8	7.4	7.4	N/A	7.6	7.8	8	8.2	8.4
B	5.6	6.2	6.4	6.8	7.4	7.4	N/A	7.6	7.8	8	8.2	8.4
C	5.6	6.2	6.4	6.8	7.4	7.4	N/A	7.6	7.8	8	8.2	8.4
D	5.6	6.2	6.4	6.8	7.4	7.4	N/A	7.6	7.8	8	8.2	8.4
E	5.6	6.2	6.4	6.8	7.4	7.4	N/A	7.6	7.8	8	8.2	8.4
F	5.6	6.2	6.4	6.8	7.4	7.4	N/A	7.6	7.8	8	8.2	8.4
G	5.6	6.2	6.4	6.8	7.4	7.4	N/A	7.6	7.8	8	8.2	8.4
H	5.6	6.2	6.4	6.8	7.4	7.4	N/A	7.6	7.8	8	8.2	8.4

PEG 3350 Concentration	19%	21%	23%	25%	25%	27%	29%	31%

Figure B.8 pH and PEG Expansion Plate Layout. Example of expansion plate based on successful well condition 15 mg/mL protein (Top) TOP96-E2: 25 % (w/v) PEG 3350, varied pH (5.6-8.4) and PEG concentrations (19-31 % (w/v)).

Table B.1 Molar Absorptivity Values. Molar absorptivity values used to calculate esterase activity of pNP Ac at various pH levels in **Figure 4.11**.

pH (Buffer)	Molar Absorptivity ($mmol^{-1}$ cm^{-1})
5.5 (MES)	3804
6 (MES)	12004
6.5 (MES)	28288
6.5 (NaP)	30100
7 (NaP)	57282.66
7.5 (NaP)	90276
7.5 (Tris-HCl)	83756
8 (Tris-HCl)	96040
8.5 (Tris-HCl)	103049.33
9 (Tris-HCl)	94561.33
9.25 (Tris-HCl)	108752
9.5 (Tris-HCl)	104952

TLC Activated carbon column preperation protocol.

1. The column was charged with one column volume (3 mL) of acetonitrile (100 % (v/v)).
2. The column was washed with two column volumes (6 mL) of water.

Filtration protocol

1. 150 µL of reaction mixture (Esterase activity assay) was diluted to 1 mL with ddH_2O.
2. Diluted sample was added to the column for 2 min.
3. The column was washed with two column volumes (6 mL) of ddH_2O.
4. Bound cellulose was eluted with 0.5 mL of 50 % (v/v) acetonitrile
5. Any remaining cellulose was eluted with 0.5 mL 100 % (v/v) acetonitrile in a separate sample

Column cleaning protocol

1. The column was washed twice with 2 mL of 50 % (v/v) Tetrahydrofuran or THF.
2. The column was washed with two column volumes (6 mL) of ddH_2O.

TLC protocol

1. Mobile phase was made and added to a glass tank, to a depth of 1 cm, 1 h before plates were to be added to the tank, which allowed for vapour diffusion.
2. 20 cm x 20 cm aluminum backed silica gel plate was cut to appropriate size.
3. A pencil line was drawn 1 cm from bottom of plate with marks every 1 cm for loading, outside marks were 1.5 cm from the edge of plate. Each mark was labeled with the sample to be loaded in that location.
4. Sample was loaded into 10 µL glass capillary tubes followed by absorption on to the plate. Each sample was absorbed onto their corresponding mark.
5. After each sample was loaded, the plate was left for 10 minutes to dry.
6. Steps 4 and 5 were repeated 5 times before the plate was added to the tank.
7. The plate was developed in the tank until the solvent front reached 1 cm from the top of the plate
8. The plate was removed from tank and pencil line was drawn along the upper solvent front before being left to dry for 20 minutes.
9. The plate was sprayed with Naphthol stain (contained within a fume hood) until slight discolouration appeared without completely dampening the plate.
10. The plate was placed on a hot plate (Contained within fume hood) until the stained bands appeared.
    *After addition of Naphthol stain all materials should remain in fume hood unless required.
    ** Mobile phase (8:3:4): 53.33 mL ethyl acetate, 20 mL ddH_2O, 26.67 mL Methanol.

*** Naphthol Stain (25:3:2): 208.33 mL Methanol, 25 mL H_2SO_4, 16.67 mL ddH_2O. (H_2SO_4 added slowly and last to avoid boiling the solution)