Protein Engineering of Cel6A from *Cellulomonas fimi* for Mechanistic Studies

by

Rebecca Catherine Whitney

A Thesis
presented to
The University of Guelph

In partial fulfilment of requirements
for the degree of
Master of Science
in
Molecular and Cellular Biology

Guelph, Ontario, Canada

ABSTRACT

PROTEIN ENGINEERING OF CEL6A FROM *CELLULOMONAS FIMI* FOR MECHANISTIC STUDIES

Rebecca Catherine Whitney
University of Guelph, 2014

Advisor:
Dr. Anthony J. Clarke

Cellulolytic enzymes commonly use one of two different mechanisms to achieve cellulose hydrolysis, differentiated by the stereochemical orientation of the anomeric carbon post reaction. The purpose of this study was to alter the reaction mechanism of Cel6A from *Cellulomonas fimi* using a catalytic cys variant in combination with chemical alkylation. To achieve this goal, the catalytic cys variant was purification under reducing conditions to preclude the 56% autooxidation that occurred to the introduced sulfhydryl under environmental oxygen levels. The inclusion of reducing agents in the purification buffers had a negative impact on the activity of the WT protein due to disruption of the native disulfide bonds. Additionally it was discovered that Tris-HCL has an inhibitory effect on the activity of Cel6A. Preliminary alkylation attempts indicated that the propionylation of the introduced cysteine may be able to rescue catalytic activity, but more definitive experimentation is required.

Acknowledgements

During the research and writing of this document I received support and assistance from many people. I would like to thank my advisor Dr. Anthony Clarke for all of his invaluable assistance and advice during the investigation and writing of this thesis, without which this would not have been possible. I would also like to thank the members of the Clarke lab, both past and present, who have aided me along this journey. The support and guidance from my lab mates was a great help to me and I will be eternally grateful for the wisdom they have imparted. I would also like to take this opportunity to thank my family and friends who have been encouraging me and supporting me throughout all of my academic endeavors over the years.

Table of Contents

List of Tables

Chapter 2: Methods

Chapter 3: Results

List of Figures

Chapter 1: Introduction

Chapter 2: Methods

Chapter 3: Results

List of Abbreviations

AA: Auxiliary Action

ACC: Advanced Analysis Center

ACN: Acetonitrile

AFEX: Ammonium Fiber Explosion

AGU: Anhydro Glucose Units

Asp: Aspartic acid

BCA: Bicinchoninic acid

BSA: Bovine serum albumin

CaZy: Carbohydrate Active enzyme

CBM: Carbohydrate binding module

CMC: Carboxymethyl cellulose

Cys: Cysteine

DS: Degree of Synergy

DP: Degree of Polymerization

DTNB: 5, 5`-2-dithiobis(2-nitrobenzoic acid)

DTDP: 2, 2` dithiodipyridine

FA: Formic Acid

Fn3: Fibronectin type-III

GH: Glycoside Hydrolase

Glu: Glutamic acid

HPLC: High-Performance Liquid Chromatography

IDT: Integrated DNA Technologies

IMAC: Immobilized metal affinity Chromatography

IPTG: Isopropyl-β-D-thiogalactoside

LB: Luria-Bertani

LC-MS: Liquid Chromatography- Mass Spectroscopy

LP: Leader Peptide

LPMOs: Lytic Polysaccharide Mono-Oxygenases

PAHBAH: p-hydroxybenzoic acid hydrazide

PDB: Protein Data Bank

PTP: Protein Tyrosine Phosphatase

WCL: Whole Cell Lysate

Chapter 1: Introduction

Cellulolytic enzymes play a key role both environmentally and industrially in the recycling of carbon trapped in cellulolytic biomass. As such they have been the subject of numerous scientific studies. Many of these investigations have been concerned with the creation or optimization of enzymes for use in providing sustainable fuel alternatives as reviewed in (Zinoviev et al., 2010). The environmental cost and sustainability of current fuel options is becoming increasingly harder to maintain. Therefore, alternative processes for generating fuel are being explored (Gowen and Fong, 2010; Solomon, 2010; Fushinobu, 2013). One of the more promising options includes utilizing cellulose as a renewable bio-ethanol source. In principal, cellulolytic enzymes are used to liberate glucose from cellulose chains which can be fermented into fuel ethanol. As such, developing uniquely active and efficient enzymes can impact the viability of cellulose as a competitive raw material for bio-fuel. One of the main cellulolytic enzymes utilized in the liberation of glucose from cellulose are endogluconases. This study was designed to expand the knowledge of the reaction mechanism of Cel6A, a model endoglucanase from *Cellulomonas fimi*. Cel6A has been utilized previously in studies concerning its reaction mechanism as well as rational design experiments.

1.1 Cellulose

Cellulose is the most abundant carbohydrate biopolymer in the world, being a constitutive structural component of plant cell walls. It is also found in certain algal and bacterial species as well as tunicates (Lynd et al., 2002; Klemm et al., 2005). Cellulose has a complex anisotropic structure composed of many chains. Cellulose chains contain

repeating units of the disaccharide cellobiose, which is composed of two β-D-glucose units linked by a β-(1→4)-glycosidic linkage (Figure 1.1). The cellulose chains are bundled together into cellulose microfibrils and then they aggregate into elementary

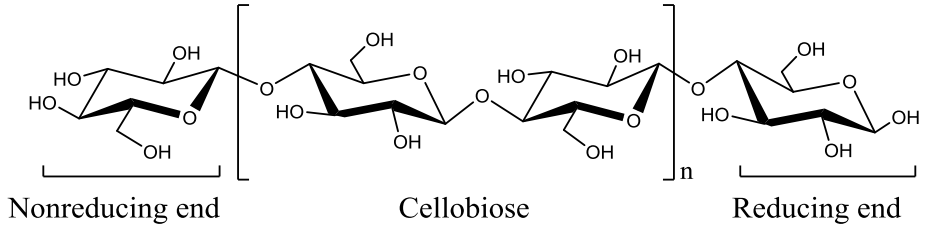


Figure 1.1: Chemical structure of cellulose. Cellobiose is the repeating unit of cellulose chains. The chains end in either a reducing sugar or a non-reducing sugar as indicated.

fibrils, due to the strong hydrogen bonds and van der Waal forces that occur between each chain (Ranby, 1969; Fan et al., 1980). Native cellulose, also known as cellulose I, contains parallel cellulose chains and it is further subdivided into α and β forms based on the hydrogen bonding pattern between the chains. The main physiological difference between the α and β forms is their solubility in alkaline solutions. The α form is insoluble and considered "true cellulose" while the β form is soluble (Nishiyama et al., 2003; Klemm et al., 2005). Each cellulose chain has a "reducing" and "non-reducing" end as indicated in Figure 1. The reducing end is terminated by the C1 hydroxyl group of the glucose moiety which is freely found in equilibrium with its aldehyde form in solution. This aldehyde can act as a reducing agent. In contrast, the chain end terminated in a C6 hydroxyl group is not able to form an aldehyde and has no reduction potential (Klemm et al., 2005).

Cellulose is often described using degree of polymerization (DP) and degree of crystallinity. The DP, which represents the number of constituent anhydroglucose units, is highly variable. As an illustration, wood pulp has a DP typically between 300 and 1700 depending on the post harvesting treatment, whereas cotton has a DP value within the 800-10 000 range depending on how it is treated (Klemm et al., 2005). This is an important characteristic of cellulose because the DP value correlates with the number of terminal and interior β-glycosidic bonds, which are the targets of either exo- or endo-acting enzymes (discussed further below). The arrangement and hydrogen bond formation between adjacent cellulose chains result in a spectrum of structural regions ranging from fully crystalline to fully amorphous. The exact measure of the ratio of crystalline to amorphous is referred to as the degree of crystallinity (Fan et al., 1980); the more crystalline the region, the higher the tensile strength of the strand (Al-Zuhair, 2008). The degree of crystallinity is often used as an indicator of reactivity, as well as mechanical strength of the cellulose molecule. The hydrolysis rates of cellulolytic enzymes acting on amorphous cellulose can be up to 30 times higher than the activity rate on highly crystalline cellulose (Zhang and Lynd, 2004).

Cellulose is rarely found on its own in natural systems. It is generally found within a composition of heterogeneous intertwined polysaccharide chains of hemicellulose and pectins surrounded by lignin. Hemicellulose, the alkali extract of plant cell walls, is a heterogeneous polymer of pentoses (for example, xylose and arabinose), as well as hexoses, like mannose, glucose and galactose, in combination with sugar acids (Saha, 2003). Lignin is the second most abundant polymer in the world. It is a complex molecule characterized by its racemic, aromatic, heteropolymeric composition. The main

constituents of lignin are hydroxycinnamyl alcohol monomers that differ in their degree of methoxylation. This includes *p*-hydroxyphenyl H, guaiacyl G and syringyl S phenylpropanoid units. The amount and composition of each component of the system varies depending on the species of origin (Boerjan et al., 2003).

1.2 Cellulolytic Enzymes

Classically it is has been shown that cellulose is naturally degraded by the concerted action of three different glycoside hydrolases: endoglucanases, exoglucanases, and β-glucosidases (Figure 1.2). Additionally, oxidative metalloproteins (AA9) and lytic polysaccharide mono-oxygenases (LPMOs) (CBM33, GH61) have been shown to aid in the digestion of cellulose (Horn et al., 2012; Fushinobu, 2013). These enzymes are found amongst bacteria and fungi that can use cellulose as their sole carbon source. Endoglucanases, exoglucanases, and β-glucosidases are able to hydrolyze the β-1,4-glycosidic bonds between glucose residues in cellulose. LPMOs are able to oxidize recalcitrant cellulose effectively "scratching" the surface of crystalline cellulose to produce an entry point for other cellulolytic enzymes (Fushinobu, 2013). This study is focused on the classical three enzyme cellulolytic system.

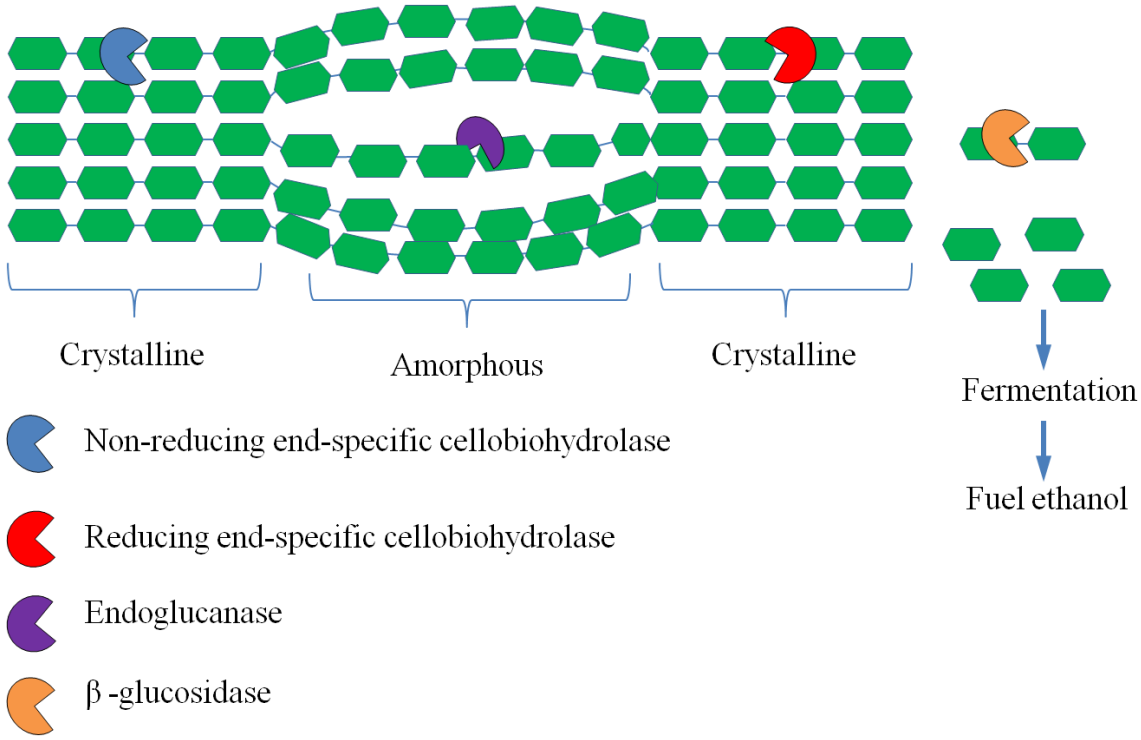


Figure 1.2: Specificity of cellulolytic enzymes. The three cellulolytic enzymes, endoglucanases, exoglucanases and β–glucosidases work in concert to degrade cellulose chains into single glucose molecules that can then be further fermented into ethanol for use as fuel.

1.3 Enzyme Nomenclature and Family Organization

In 1953, a system for organizing and grouping enzymes was proposed by Hoffmann-Ostenhof and it has been developed continually and updated ever since by the International Union of Biochemistry and Molecular Biology (IUBMB, http://www.chem.qmul.ac.uk/iubmb/). Enzymes that hydrolyze glycosidic bonds are assigned the number 3.2.1.x. where the 3 represents hydrolases, the 2 is an indicator of their glycosidase activity and the 1 classifies them as acting on *O*- or *S*- glycosidic linkages. The 'x' is used to indicate the particular substrate on which the enzyme acts as well as its mode of action. For example, endoglucanases are assigned the number 3.2.1.4,

while exoglucanases are 3.2.1.91 even though they both act on the same substrate due to the difference in their mode of action. This system is helpful when trying to identify the enzyme substrate and mode of activity. Unfortunately, it is difficult to determine any mechanical, structural or sequence based similarities between various enzymes using this system. Due to this difficulty a complementary system for organizing the glycoside hydrolases (GHs) has been developed that classifies the various enzymes based on their sequence similarity (Cantarel et al., 2009). There are currently over 15 000 GH modules organized into 132 families. As the enzymes are grouped based on primary and predicted secondary structure similarities, they generally have similar tertiary structures (Cantarel et al., 2009). The families that have similar three-dimensional folds are further grouped into clans (Cantarel et al., 2009). In this classification system, families of enzymes generally share a common catalytic mechanism, as well as conserved catalytic residues, even if they act on a variety of different substrates. All of this information, as well as an up to date database of all the enzymes and families currently classified under this system is maintained at the Carbohydrate Active Enzymes (CAZy) website http://www.cazy.org.

1.3.1 Endoglucanase

Endoglucanases (EC 3.2.1.4; cellulases) act to hydrolyze the β-(1, 4)-glycosidic linkages within cellulose chains. These enzymes thus create new chain ends and overall decrease the DP of cellulose. The cellulose chain binds to an active site cleft of the enzyme typically in a number of different binding sub-sites. For example, endoglucanase I from cellulolytic fungi has four sub-sites for cellulose to bind (Figure 1.3).

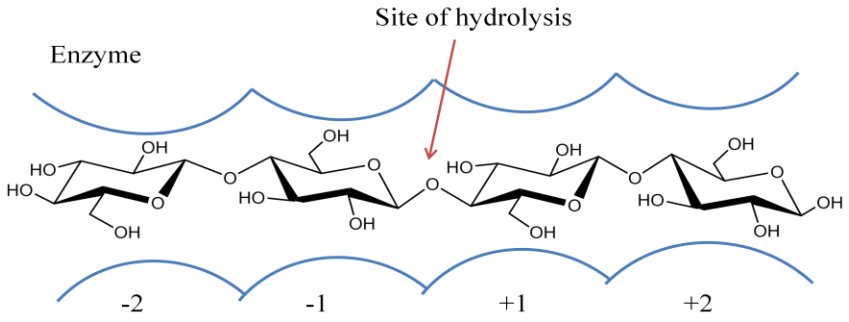


Figure 1.3: Subsite structure of endoglucanase I produced by cellulolytic fungi. The four kinetically significant subsites where glucose molecules bind are labelled from -2 to +2 (modified from (Davies et al., 1997).

Enzymes that hydrolyze internal bonds within soluble amorphous regions of cellulose, like endoglucanases, are more likely to have an open (to bulk solvent) substrate binding cleft or trough (Schülein, 2000). The absorption of endoglucanases to substrate is relatively fast compared to the hydrolysis of the cellulose chains (Zhang and Lynd, 2004) and it has recently been shown that during the act of hydrolysis, an initial decrease in cellulose fibril size is followed by a significant swelling event (Wang et al., 2012). In this way, it is postulated that endoglucanases allow for the introduction of water into more crystalline sections of cellulose.

1.3.2 Exoglucanases

Exoglucanases (EC 3.2.1.91; cellobiohydrolases) recognize and hydrolyze β-(1, 4)-glycosidic linkages from either the reducing or non-reducing end of a cellulose chain and release cellobiose as their reaction product. Exoglucanases act in a "processive" manner, and as such often have an enclosed tunnel shaped active site that allows them to separate a single glucan chain from the parent microfibril while hydrolyzing it. This

tunnel shape also acts to prevent the chain from re-adhering to the cellulose fiber it has been liberated from (Schülein, 2000). These tunnels often have flexible loops across them that play a pivotal role in pH dependence and the ring pucker formation believed to be involved in substrate binding (Bu et al., 2013) Exoglucanases are relatively slow compared to endoglucanases at decreasing the DP, however they are essential for the hydrolysis of microcrystalline regions of cellulose (Lynd et al., 2002).

1.3.3 β-Glucosidases

β-Glucosidases (EC 3.2.1.21; cellobiose) hydrolyze small oligosaccharides, like cellobiose, producing glucose. Therefore β-Glucosidases relieve end product inhibition, since cellobiose is a much more potent end product inhibitor of endo- and exo-glucanases (Holtzapple and Cognata, 1990). β-Glucosidases mechanistically function in the same fashion as exoglucanases and endoglucanases, employing two catalytic carboxyl groups provided by either Asp or Glu residues. Research into β-glucosidases has centered around their use in both the medical and environmental fields (Jäger and Kiss, 2005).

1.3.4 Reaction Mechanisms

The enzymatic hydrolysis of the β-(1,4)-glycosidic linkage is accomplished by employing either a retaining or an inverting mechanism of action as first proposed by (Koshland, 1952). These mechanisms of acid hydrolysis utilize a proton donor as well as either a nucleophile or base residue in the active site. The nomenclature of the two different mechanisms is based on the stereochemical orientation of the anomeric carbon of the product as compared to the substrate (Y.H. P. Zhang & Lynd, 2004).

The active site of an enzyme catalyzing an inverting reaction mechanism will consist of two carboxyl groups that are positioned approximately 10.5 Å apart. This is sufficient space to allow the substrate (in this case cellulose) as well as a water molecule to bind in between the two carboxyl groups. The reaction then proceeds through a single displacement step with one carboxyl group acting as an acid and the other as a general base catalyst (Figure 1.4). The end product has an inverted anomeric configuration compared to the substrate (reviewd in Zechel and Withers, 2000).

Enzymes using a retaining mechanism employ similar machinery with two carboxyl groups, but they are positioned characteristically 5.5 Å apart. The cellulose molecule will bind in between the two carboxyl groups and a double displacement reaction then occurs, with the first carboxyl involved acting first as an acid and the second carboxyl involved acting as a nucleophile to form a covalent adduct (Figure 1.4). The first carboxyl group then acts as a base to abstract a proton from a water molecule for attack by the resulting hydroxide ion on the oxocarbenium-like reaction intermediate to release product. This results in the retention of the original anomeric configuration by the product of the reaction.

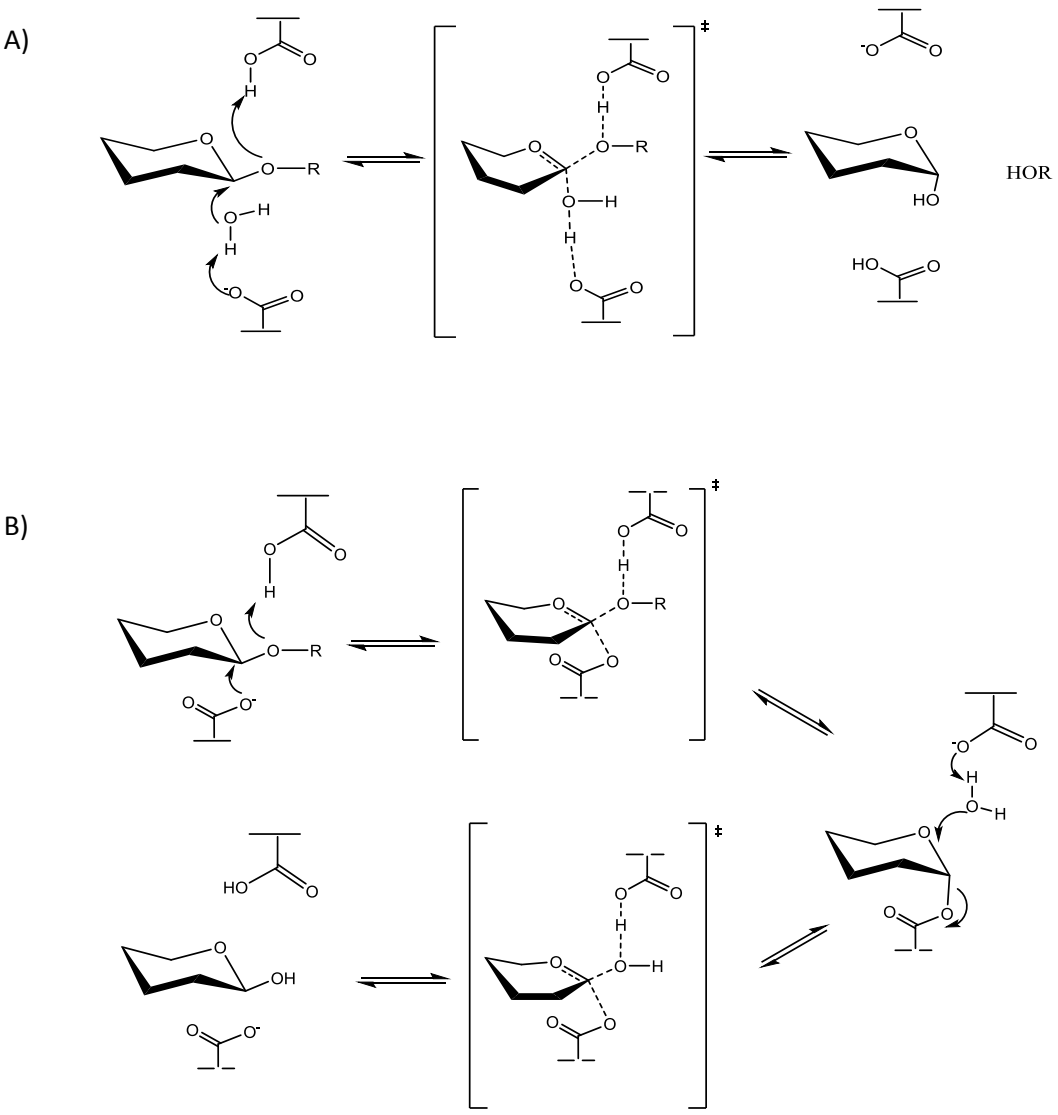


Figure 1.4: General reaction mechanisms of A) an inverting and B) a retaining enzyme. Oxocarbenium-like transition states are indicated by the square brackets. (modified from Zechel and Withers, 2000).

1.3.5 Cellulolytic Enzyme Modularity

Many cellulolytic enzymes have a characteristic modular structure that includes a catalytic domain separated from either one or more carbohydrate-binding module (CBM) *via* flexible linker regions. Fibronectin type III (Fn3) repeat modules are often found in cellulolytic enzymes as well but their function remains unknown. The number and type of CBMs present within a particular enzyme varies. To date, there have been over 32 000 CBM modules identified and these have been classified into 67 different families based on their sequence similarities, all of which have been collated and can be found online at www.CAZy.org.

The modularity of endogluconases is clearly demonstrated by examining the four different endogluconases produced by *C. fimi* (Figure 1.5), each of which contains a unique combination of modules.

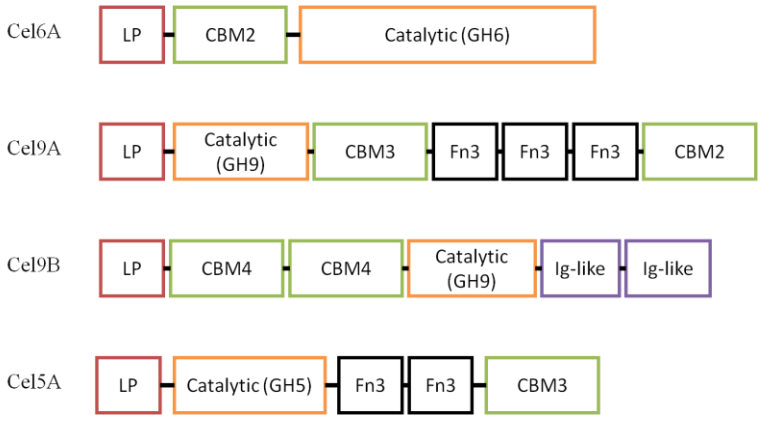


Figure 1.5: Modular structure of endoglucanases produced by *C. fimi*. Abbreviations: LP, leader peptides; CBMX, carbohydrate-binding module; Fn3, fibronectin type III repeats; and Ig-like, immunoglobulin-like fold.

The CBM modules aid in the binding of the enzymes to insoluble crystalline cellulose and as such increase the activity of cellulolytic enzymes on that particular substrate in an act of intramolecular synergy (Wang et al., 2008). As noted earlier, the role of the Fn3 repeats in bacterial cellulases has not been fully characterized. However, Fn3 repeats are widely distributed among eukaryotic proteins. In fact, the Fn3repeats found in bacteria were shown to be acquired from a eukaryotic source based on evolutionary studies (Campbell and Spitzfaden, 1994). In eukaryotes, Fn3 repeats facilitate protein to protein interactions and act as spacer modules in order to assure proper biological positioning of important residues and modules within the proteins they are located (Campbell and Spitzfaden, 1994).

1.4 Synergy

Cellulolytic systems display synergy throughout their digestion of cellulose. Synergy is defined generally in terms of degree of synergy (DS). This is quantified by the level of activity displayed by the components of a mixed reaction divided by the sum of activities of the separate components. Synergism has been displayed in a number of different combinations including; exoglucanase with endoglucanase, exoglucanase with exoglucanase, endoglucanase with endoglucanase, exoglucanase or endoglucanase with a β-glucosidase, as well as intramolecular synergism between catalytic domains and carbohydrate binding domains or other catalytic domains.

Synergy is displayed by complexed systems such as cellulosomes (see below), as well as non-complexed systems which are defined as discretely acting proteins whose

activity as a whole is improved when they are working in concert. Non-complexed systems are generally exploited by aerobic fungi and bacteria while complexed systems appear to be favoured by anaerobic microorganisms (Doi and Kosugi, 2004). The level of synergy itself can be affected by both the substrate properties, such as crystallinity of the cellulose, as well as the experimental conditions. For example, the amount of synergy observed during endo-exo activity can be increased by overloading enzyme, as long as the concentrations are maintained below saturation level. If the enzyme concentrations are oversaturated, a decrease in DS will be observed, at least *in vitro* (Zhang and Lynd, 2004). The synergy exhibited by endo- and exo-acting cellulolytic enzymes is the most extensively studied synergistic system.

Cellulosomes are large complexes of enzymes secreted by anaerobic bacteria (and highly suspected to be present in anaerobic fungi) that are composed of a number of enzymes linked together by scaffolding proteins (Doi and Kosugi, 2004). The "scaffoldin" protein is a fibrillar protein that contains numerous "cohesin" domains. These cohesion domains interact with "dockerin" domains located on cellulolytic enzymes in order to form the cellulosome. This "cohesion"/ "dockerin" interaction is species specific, and allows for plasticity in the quaternary structure of the cellulosome (Pages et al., 1997). The cellulosomes couple the enzymes required for efficient cellulose hydrolysis together into one package, allowing the enzymes to work synergistically. This includes cellulolytic enzymes, hemicellulases, pectinase, chitinase and other ancillary enzymes (Doi and Kosugi, 2004).

1.5 Cel6A from *C. fimi*

C. fimi is a free-living, Gram-positive, non-spore forming, facultative anaerobic rod. It can be isolated readily from soil samples, which is its most common habitat. *C. fimi* produces a variety of sugar active enzymes that can be found on CaZy including at least two cellobiohydrolases (Cel6B, and Cel48A), four endoglucanases (Cel6A, Cel9A, Cel9B and Cel5A) depending on the carbon source being used (Stackebrandt et al., 2006).

Of all of the endoglucanases produced by *C. fimi*, Cel6A (formerly CenA) has the simplest modular structure (Figure 1.5) and, as such, is a good candidate for mutational studies. It was discovered in 1984 (Gilkes et al., 1984) and its gene sequence was determined by the sequencing of unidirectional overlapping deletions and the amino acid sequence was subsequently deduced (Wong et al., 1986). The gene is 1350 bp long and is translated into a 449 amino acid polypeptide that, when folded, consists of an N-terminal CBM2 domain which is not required for catalytic activity, and a C-terminal catalytic domain which is classified as a CAZy GH6 family member. The two domains are connected by a 23 amino acid long Pro/Thr linker region (Wong et al., 1986). The first 31 N-terminal amino acids constitute a leader peptide that targets the protein to the periplasm. Once the protein is in the periplasm, the leader peptide is cleaved resulting in a mature enzyme of 418 amino acids (Wong et al., 1986) with a predicted molecular weight of 45.34 kDa (Gasteiger, 2003). The gene for Cel6A has been expressed recombinantly in *Escherichia coli, Rhodobacter capsulatus,* and *Klebsiella pneumoniae.* Asp 252 and Asp392 have been identified as the catalytic acid and base, respectively. As such, this endoglucanase employs an inverting mechanism of action relying on a

protonated acid and deprotonated base. While no crystal structure of Cel6A exists to date, four other GH6 enzymes have been crystallized, the enzymes from *H. insolens* (PDB: 1BVW), *T. fusca* (PDB: 1TML), *Hypocrea jecorina* (PDB: 1BVW) and *Mycobacterium tuberculosis* (PDB 1UOZ).

Cel6A has also been the target of various studies above and beyond the characterization of its catalytic acid (Wang et al.,1994) and identification of its catalytic base as Asp392 (Cockburn et al., 2010). Of relevance to the research presented in this thesis, catalytic base alterations have been made to modify the protein's function. The site-specific replacement of Asp392Cys ((D392C) Cel6A) was created and the engineered enzyme was shown to be inactive under the same assay conditions in which the WT enzyme was active. However, up to 52% activity could be restored by chemical alteration of the introduced Cys residue to cysteinesulfinic acid by oxidation utilizing hydrogen peroxide (Cockburn et al., 2010). Furthermore, the replacement of the catalytic Asp with Cys and its subsequent oxidation to cysteinesulfinic acid caused a shift in local pK_a and a consequent shift of pH optimum of enzyme activity. Thus, the oxidized (D392C) Cel6A showed increased activity below pH 5.5 compared to the WT enzyme culminating with the mutant enzyme displaying almost 300% of the activity displayed by the WT enzyme at a pH of 4.5 (Cockburn et al., 2010). Shifts in pH-activity have been demonstrated in non-cellulolytic enzymes, such as phytase (Kim et al., 2006), acid phosphatase (Makde et al., 2006), and α-amylase (Nielsen et al., 1999). Most of these types of studies involved changing residues in and around the catalytic site in order to shift the pK_a values of the catalytic residues, and as such change the pH optima for

activity. Unfortunately, however, many of these dramatic changes result in inactive enzymes.

Experiments have been conducted to study the replacement of surface residues on Cel6A with the goal of shifting its pH-optimum without affecting the catalytic mechanism or functionality of the enzyme (Cockburn and Clarke, 2011). In theory, the addition or removal of charged groups on the surface of an enzyme can work to increase the enzyme's stability at different pH values by altering its electrostatic interaction network. As such, increased enzyme activity can be accomplished for extreme pH values if the stability of the enzyme at those pH values is the limiting factor for activity. Thirty different mutant enzymes were made by substituting amino acids at 13 different positions. The amino acids were replaced with Asp, Glu, Ala, Arg, Lys, His as well as Cys which would later be oxidized for assays. Six amino acids were chosen based on structure similarity to the known structure of Cel6A from *T. fusca* and its reaction mechanism. Six other amino acids were chosen based on sequence alignments performed on Cel6A and other GH6 enzymes with known sequences and pH optima. The 13th position replaced was an eight amino acid long sequence starting at Gly218 that was universally conserved among enzymes with low pH optima. The activity of each mutant enzyme was systematically examined at three different pH values (5, 7 and 9). Four of the replacements caused an increase in activity (Tyr321Phe, Glu401Ala, Lys292Ala and Glu368Ala) (Cockburn and Clarke, 2011). This information was used to generate double mutants in order to assess any additive effect on activity. It was found that the Tyr321Phe/Glu407Ala mutant enzyme exhibited 162% of the wild type activity at pH 5, whereas the activities of Tyr321Phe and Glu407Ala were 141% and 120%, respectively,

on their own (Cockburn and Clarke, 2011). Interestingly, the activity of the double mutant Glu368Ala/Glu407Ala reverted back to WT levels while no difference was observed between a Lys292Ala mutant (148% WT activity) and its double mutant derivative Lys292/Glu407Ala (Cockburn and Clarke, 2011).

1.6 Altering the Mechanism of Glycoside Hydrolases

The first reported case of altering the catalytic mechanism of a glycosidic enzyme using site-directed mutagenesis was the replacement of the putative nucleophilic His95 residue to a glutamine in triosephosphate isomerase (EC 5.3.1.1) (Nickbarg et al., 1988). It was reported that the mutant triosephosphate had the same stereochemistry as the wild type (WT) enzyme but it had a drastically reduced k_{cat} value. This is in part due to the inability of the mutant enzyme to stabilize or neutralize the negative enediolate oxygen intermediate product of the reaction (Nickbarg et al., 1988). A series of studies have also examined the catalytic site of lysozymes, which function in a similar manner to endoglucanases. In one set of experiments, the catalytically active Thr26 residue of a T4 lysozyme was replaced with a His residue. The mutant lysozyme was crystallized isomorphously with the WT enzyme and compared for differences within the active site. Upon inspection, the WT enzyme had a water molecule placed in an ideal position to attack the α-side of the substrate and initiate a reaction that would result in an α-anomer product. In contrast, the Thr26His mutant enzyme had the N of the histidine side chain in the place formerly occupied by the water molecule. In this case, the His would act as the nucleophile in place of the water, resulting in an inverted product (Kuroki and Weaver, 1995).

Similar studies were performed on β-glucosidase Abg from *Agrobacterium faecalis.* Abg is classified as a retaining enzyme in its wild type form. In one study, the catalytic Glu358 residue was replaced with an Asp to increase the space within the active site. The activity was determined to be 2500-fold lower than the wild type enzyme, even though the mechanism was not altered. A site-directed replacement of Glu358Ala was then constructed. With the removal of the catalytically important Glu residue, the activity of the mutant enzyme was 10^7- fold lower, and what residual activity was present was attributed to contaminating wild-type enzyme. However, the activity of the enzyme was rescued to near wild type levels by the addition of azide or formate which could act as nucleophiles in the place of the Glu residue (Wang et al., 1994). This mutant Abg was further modified by replacing the Glu with a Ser residue and it is now used to synthetically create oligosaccharides (Mayer et al., 2000). The substitution of the Ser, like the Ala, interrupts the hydrolytic action of the enzyme while allowing it to keep the glycosyl transfer activity. The added polarity of the Ser side chain increased the reaction rate by providing a more favourable binding site for the substrate of the reaction, in this case a fluoride atom (Mayer et al., 2000).

Cellulases have also been the target of mutational studies to improve the understanding of their function. One common area of study is their function under different pH ranges. This is important due to the high or low pH required during pre-treatment of cellulose for fuel ethanol. Cellulases can display very different pH activity profiles while sharing high primary sequence homology. Such is the case between the family GH5 endoglucanase BSC found in *Bacillus subtilis* and the family GH5 endoglucanase NK1 found in alkalophilic *Bacillus sp.* N-4. These two enzymes have

significant sequence identity while having pH optima that range from 6 - 6.5 and 6 - 10.5 respectively for BSC and NK1 (Nakamura et al., 1991). The high alkaline activity was attributed to a group of 38 amino acids in the C-terminal section of the protein. Amino acids Ser287 and Ala296 were specifically identified as residues of importance (Nakamura et al., 1991). When a Ser287Asn and Ala296Ser mutant of the NK1 enzyme was created, the pH activity profile became comparable to the BSC enzyme from *B. subtilis*, suggesting the importance of those two residues for function in alkaline environments (Park et al., 1993). However, BSC was not active under conditions of high pH when those two residues were introduced into its sequence, suggesting that further replacements of residues are involved in the increased activity under alkaline conditions (Park et al., 1993).

Family GH7 cellulolytic enzymes have been found to act using a slightly modified catalytic method that includes a trio of carboxylates in their active center, in place of the standard two (Parsiegla et al., 1998). Mutational studies involving these carboxylate groups revealed two to be the catalytic nucleophile and acid/base respectively while the third was still required for proper function. The ability of exoglucanases to act on longer soluble substrates was also attributed to an increased number of sub-sites for cellulose binding (11 in exoglucanases compared to 4 - 5 in endoglucanases). This was confirmed by replacing selected residues in *Humicola insolens* Cel7B endoglucanase to Trp residues in order to mimic the sub-site binding structure of *Trichoderma reesei* Cel7A. This resulted in a decreased K_m on phosphoric acid swollen cellulose, indicating stronger binding to longer polymeric substrates by the modified Cel7B (Davies et al., 1997).

GH Family 45 enzymes have been utilized in studies concerning detergent compatibility. Detergents are an important component of industrial washes which are used to process cellulose; because of this, detergent compatibility among cellulolytic enzymes is a desirable trait. Cel45A from *H. insolens* has been modified by site-directed mutagenesis in order to increase stability with linear alkyl benzene sulphonate (C12-LAS) (Otzen et al., 1999).

The analysis of the catalytic sites of Cel6A from *C. fimi* was the first of its type carried out on family 6 enzymes and it has acted as a model study since. The potential catalytic residues of Cel6A were determined based on sequence similarity between Cel6A from *T. reesei*, *Thermobifida fusca* and *C. fimi*. The conserved catalytic Asp residues from the *T. reesei* and *T. fusca* enzymes were used to identify the residues of interest in the *C. fimi* enzyme. These residues were then subjected to site-directed replacement. The mutant enzymes were then characterized and Asp252 and Asp392 were identified as the catalytic acid and base, respectively (Wong et al., 1986). Site-directed mutagenesis studies have also been carried out on Cel6A from *T. fusca* demonstrating the importance of conserved 'non-catalytic' residues in the overall activity and specificity of the enzyme. Four residues were identified: His159, Arg237, Lys259 and Glu263. As well as affecting activity, some of these residues also seemed to affect the rate limiting step (Lys259) and the k_{cat}/K_m (His159). Unfortunately, varying results were seen depending on the substrates used (soluble cellodextrins, carboxymethyl cellulose (CMC), and crystalline cellulose) (Zhang et al., 2000).

The catalytic acid residue, as well as other conserved 'non-catalytic' residues, of Cel6A from *T. fusca* were identified by site-directed mutagenesis studies based on the

residues previously identified by Wong et al. (1986) using the highly homologous Cel6A from *C. fimi*. Despite having the same catalytic acid and other conserved 'non-catalytic' Asp residues, the putative base (Asp265) behaved unexpectedly. An Asp265Asn mutant enzyme was able to maintain up to 7% wild type activity, suggesting that Asp265 is not, in fact, the catalytic base in that particular enzymatic system. Family GH6 enzymes are the most extensively studied of the cellulolytic glycoside hydrolase families, yet despite the substantial analysis and protein engineering attempts, this family of glycoside hydrolases, in particular, continue to exhibit contrasting results from systems that seem closely related.

1.7 Chemical Modifications of GH Active Sites

The best way to understand and characterize an enzymes active site remains x-ray crystallography, however there are many instances where secondary options are explored or used to enhance the data obtained from x-ray crystallography. Often times, both competitive non-covalent as well as covalent inhibitors are employed (Kiss et al., 2002). Common inhibitors for GH active site studies include N-bromoacetyl-β-D-glycosylamines, glycosyl isothiocyanates, or glycosyl epoxide derivatives (Kiss et al., 2002; Jäger and Kiss, 2005; Brunzelle et al., 2008). A majority of these studies focus on β-D-glucosidases and β-D-xylosidases. These inhibitors are useful for trapping enzymes in their active sites and enzyme intermediates to discern useful information about active site conformation and catalytic residues (Brunzelle et al., 2008). Non-covalent fluorogenic substrates are also used in the study of GH activity. For example 4-methylumbelliferyl-β-cellobioside derivatives are used to characterize GH6 exoglucanases (Wu et al., 2013).

1.8 Biological Importance of Cys Residues

Cysteine is one of only two amino acids that contain a sulfur atom and the only one to possess a thiol. This thiol group is a nucleophile and is easily oxidized, particularly at physiological pH. This property allows Cys to easily form cystines and play a key structural role within protein tertiary structure. It is also a common scavenger residue believed to help deal with oxidative and nitrosative stress (Paget and Buttner, 2003; Poole et al., 2004). In recent years, Cys residues have been implicated in redox signalling pathways and H_2O_2 mediated stimulation of protein tyrosine phosphatases (PTP) transcription factors (Finkel, 2000; Rhee et al., 2000; Delaunay et al., 2002; Paget and Buttner, 2003). It appears that the physiological role of Cys residues relies heavily on its reactive nature. Cys can form cystines when bonded to another cysteine residue, as well as mixed disulfides with glutathione or other Cys derivatives (Schilling et al., 2004). Cys can also exist in a number of different oxidation states. A single oxidation event creates Cys-sulfenic acid, followed by Cys-sulfinic acid and culminating in Cys-sulfonic acid. Of the three different oxidation states, only Cys-sulfenic acid is reversible (Figure 1.6) (Reddie and Carroll, 2008).

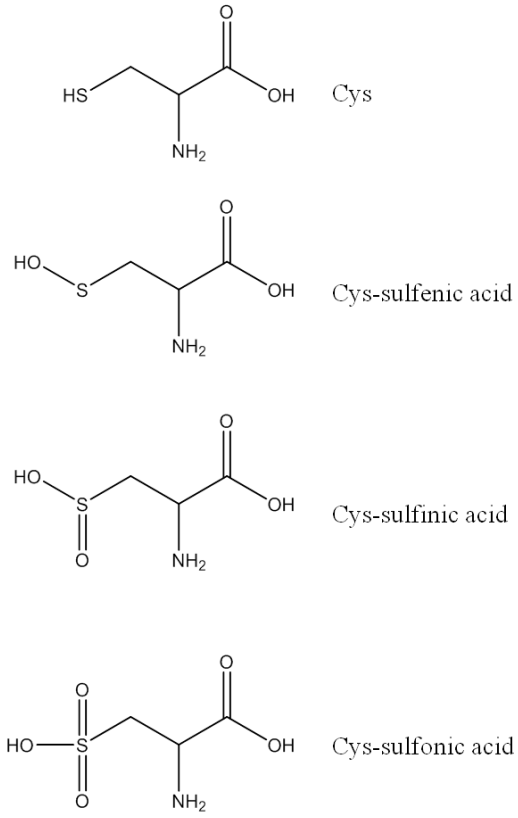


Figure 1.6: Potential oxidation states of Cys. Cys can be potentially oxidized as many as 3 times to produce a Cys-sulfenic, Cys-sulfinic, or Cys-sulfonic acid. Only the first oxidation to sulfenic acid is reversible.

Due to its important biological function, different techniques have been developed to identify and quantify the level of Cys oxidation within proteins. The most common and well documented way of quantifying free sulfhydryls is employing 5,5`-2-dithiobis( 2-nitrobenzoic acid) (DTNB), also known as Ellman`s reagent (Riener et al., 2002). Along with DTNB, another small reporter molecule 2,2` dithiodipyridine (DTDP) can be used. DTDP is more sensitive than DTNB due to its small amphiphilic nature, which allows it to more easily access buried free sulfhydryls (Riener et al., 2002). The detection limit of DTDP is ~0.2 nmol mL^{-1} whereas to approach that level of detection, DTNB must be incubated with cystamine as a mediator (Riener et al., 2002). Other methods of detection

include a papain amplification assay in which the active site of papain is reactivated by free sulfhydryls in the sample (Wright and Viola, 1998). Titration with bromobimane or maleimide derivatives of napthopyranones are useful fluorescent techniques, only hampered by cases of high background fluorescence (Wright and Viola, 1998).

There are a few techniques applied to differentially determine the oxidation state of Cys residues. In the case of Cys-sulfenic acid, dimedone based probes are utilized in order to trap Cys-sulfenic acids and allow for easy identification with gel analysis or mass spectrometry (Klomsiri et al., 2010). Differential alkylation is employed in order to determine the ratio of reduced to non-reduced cysteines or the levels of Cys-sulfinic and Cys-sulfonic species within a protein depending on how it is employed. Usually, differential alkylation is paired with mass spectroscopy or HPLC analysis (Hamman et al., 2002; Schilling et al., 2004).

1.9 Biofuel

The previously described mutational studies demonstrate that it is possible to alter the catalytic activity of an enzyme to tailor it to the reaction conditions required. This is important for the industrial application of cellulolytic enzymes in the production of bio-ethanol. Generally, incubation of untreated cellulosic materials with free enzyme systems or pure cultures of enzyme-degrading microbes results in hydrolysis yields of less than 20% of the theoretical yields (Zhang and Lynd, 2004). This is due to the general inaccessibility of the β-glycosidic bonds of the cellulose chains owing to the complex mixture of lignin, hemicellulose and cellulose that creates small pore sizes leaving the cellulose chains inaccessible to cellulolytic enzymes. Cellulose chains may

also prove difficult to digest depending on the degree of crystallinity. These problems are currently solved industrially using pre-treatment steps. These steps include: dilute-acid treatments, steam explosion at high solids concentration, hydrothermal processes, organic solvents with water, ammonium fiber explosion ("AFEX"), and strong alkali processes utilizing NaOH or lime (Zhang and Lynd, 2004). The goal of these steps is to partially or fully digest the associated hemicellulose and lignin, as well as reduce the crystallinity of the cellulose chains thereby increasing the efficiency of enzyme degradation (Cardona and Sánchez, 2007). Currently, fungal-derived enzymes from *T. reesei* are employed in the degradation of cellulose however the use of other enzyme systems, both bacterial and fungal is currently being investigated (Gowen and Fong, 2010; Fushinobu, 2013). Glucose produced by cellulolytic enzymes, as well as the sugars produced during hemicellulose degradation, can be fermented into ethanol by microorganisms for downstream use as fuels. It is important to identify and design more efficient systems in order to make the process of bio-fuel production more cost effective.

1.10 Hypothesis and Research Aims

Cel6A from *C. fimi* is an example of an inverting enzyme that can act as a model endoglucanase for cellulose degradation studies. It was hypothesized that it is possible to convert Cel6A from an inverting enzyme to a retaining enzyme by closing the gap between its catalytic residues through a combination of site specific replacement and chemical modification. Thus, I proposed to generate an Asp392Cys variant of Cel6A and then specifically alkylate the sulfhydryl group of the Cys residue with alkyl acetates to provide the catalytic base moiety within approx. 5 Å of the existing Asp252 catalytic acid residue (Figure 1.7).

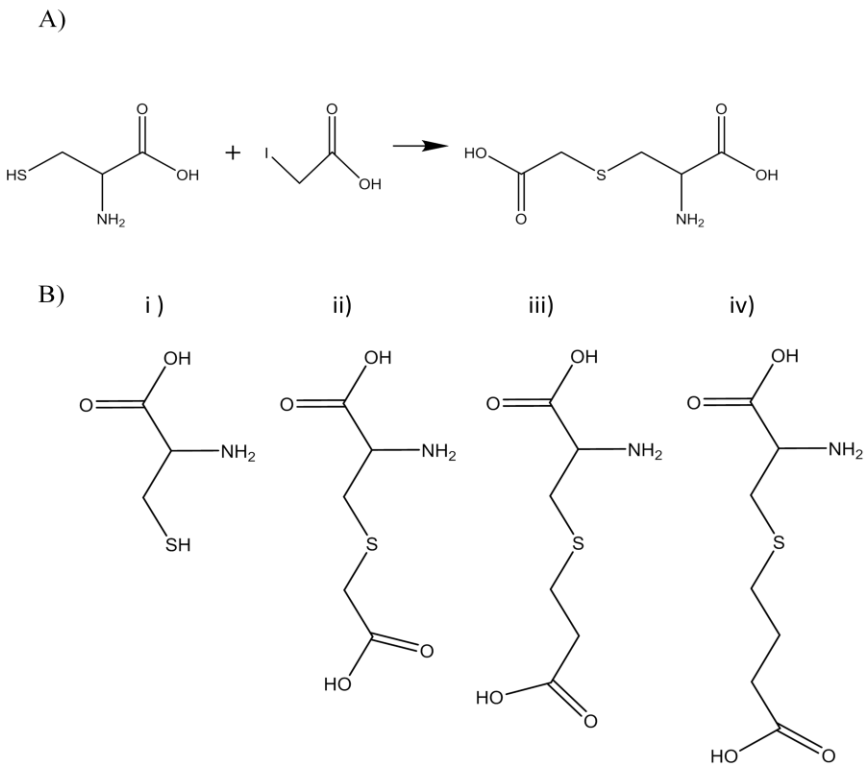


Figure 1.7: Extension of functional base catalyst by alkylation of Cys. A) Alkylation of Cys with iodoacetic acid as an example of the chemical modification possible to the Cys residue. B) A comparison of side chain lengths between i) an unmodified Cys residue ii) Cys-iodacetic acid iii) Cys-iodopropionic acid and iv) Cys-iodobutyric acid.

It is proposed that this introduced alkyl group will mimic the active site architecture of a retaining enzyme. Based on previous observations made by our lab (Cockburn et al. 2010), the introduced Cys is predicted to undergo approximately 20% autooxidation. Therefore, conditions will be explored to preclude some, if not all, of this oxidation during the purification and handling of the enzyme variant. To accomplish these goals, the major aims of this research were to: (i) generate recombinant enzyme lacking the leader peptide of both the wild-type and the Asp392Cys variant; (ii) characterize the oxidation state of the introduced Cys and test different reducing

conditions during purification to prevent auto-oxidation from occurring; (iii) alkylate the introduced Cys with various alkylating agents in an attempt to lengthen the catalytic residue chain (Figure 1.6); and (iv) determine the stereochemistry of any catalytic activity by ^{1}H-NMR analysis.

Chapter 2: Methods

2.1 *In Silico* Work

Primers were designed utilizing the plasmid editing program ApE (Davis, 2011). The FASTA sequence of Cel6A was analyzed using the SignalP 4.1 Server to determine any potential signal peptides (Peterson TN, Brunak S, Von Heijne G, 2011). The sequence was also analyzed using DISULFIND to determine potential disulfide bonds (Ceroni et al., 2006). The ProtParam tool available through Expassy was employed to determine predicted MW, PI, extinction coefficient and protein stability (Gasteiger, 2003). A Phyre2 model of *C. fimi* Cel6A structure was created using Cel6A from *Mycobacterium tuberculosis* as the template (Kelley and Sternberg, 2009).

2.2 General Experimental Procedures

2.2.1 DNA Preparation and Handling

All plasmid stocks were harvested using the PureLink® HiPure Plasmid Miniprep Kit (Invitrogen, Burlington, ON). All primers were ordered from Integrated DNA Technologies (IDT) and stored in 100 µg/mL stock solutions at -20 °C. For a list of all primers used throughout this study see Table 2.1. All DNA sequencing was performed at the Advanced Analysis Center (AAC) of the University of Guelph. All gel extractions and PCR clean up reactions were conducted using the PureLink™ Quick Gel Extraction and PCR purification combo kit (Invitrogen). All restriction enzyme digestions were performed using enzymes (New England BioLabs® Pickering, ON) for 3 hours at 37 °C. Ligation reactions used ligase (New England BioLabs®) with a 3:1 insert to vector ratio, stored at 4 °C for at least 24 hours prior to transformation according to the manufacturer's instructions. All DNA products were visualized using 1% agarose gel

electrophoresis. Agarose gels were prepared by adding 0.1% agarose (w/v) to TAE (40 mM Tris acetane, 1 mM EDTA, pH 8.0).

Table 2.1: PCR Primers Used in this Study

Designation	Description	Sequence[a]	Resulting Plasmid
Sequencing primers			
T7 Promoter	For sequencing pACDC003 and pACRW06,07,08, and 09	TAATACGACTCACTATAGGG	N/A
T7 Terminator	For sequencing pACDC003 and pACRW06,07,08, and 09	GCTAGTTATTGCTCAGCGG	N/A
Cloning Primers			
No His tag Forward	For sub-cloning *cel6A/cel6a D392C* into pET30a without the leader peptide	CCGCCGCG**CATATG**GCTCCC	PACRW-06/07
No His tag Reverse	For sub-cloning *cel6A/cel6a D392C* into pET30a without the leader peptide	TGCGGCCGC**AAGCTT**TCACCTG	pACRW-06/07
SUMO Forward	For sub-cloning *cel6A/cel6a D392C* into Champion^{tm} pET SUMO	GCTCCCGGCTGCCGCGTCGACTA	pACRW-08/09
SUMO Reverse	For sub-cloning *cel6A/cel6a D392C* into Champion^{tm} pET SUMO	TCACCACCTGGCGTTGCGCGCC	pACRW-08/09

[a] Introduced restriction sites are in bold.

All DNA samples were mixed with 5X Sample reaction buffer (Fermentas, Burlington, ON). The Sample buffer was diluted to a 1X working concentration. The samples were then loaded into the agarose gel wells and separated by electrophoreses at 100 mV for 30 minutes. The gels were then stained with SYBR®Safe (Invitrogen) for 10 minutes before being exposed to ultraviolet light to visualize the DNA banding pattern.

2.2.2 Cell Culture

All *E. coli* cells were cultured in LB media procured from Fisher Scientific (Ottawa, ON) and supplemented with the appropriate antibiotic. For a list of *E. coli* strains used in this study and their respective antibiotic resistance see Table 2.2. Overnight cultures were grown in 5 mL volumes at 37 °C with shaking. Cultures for protein expression and purification were grown in 1 L volumes. Freezer stocks of each *E. coli* strain, as well as *E. coli* DH5α containing each plasmid generated for this study, were made in duplicate by combining 500 µL of an overnight culture with 500 µL of 50% glycerol. They were then stored at -80 °C.

Table 2.2: Bacterial Strains Used in this Study

Strain	Genotype or Description	Source
E. coli DH5α	F– Φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17 (rK–, mK+) *pho*A *sup*E44 λ– *thi*-1 *gyr*A96 *rel*A1	Invitrogen
E. coli BL21 (λDE3)	F^- ompT gal dcm lon $hsdS_B(r_B^- m_B^-)$ λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5])	Novagen
E.coli Shuffle T7 Express	F´ *lac, pro, lacIQ* / *Δ(ara-leu)7697 araD139 fhuA2 lacZ::T7 gene1 Δ(phoA)PvuII phoR ahpC* galE (or U) galK λatt*::pNEB3-r1-*cDsbC* ($Spec^R$, $lacI^q$) *ΔtrxB rpsL150*(Str^R) *Δgor Δ(malF)3*	New England Biolabs

2.2.3 Calcium Chloride Competent Cell Derivation

E. coli strains used in this study were made chemically competent by the following method. A flask of LB (amount depending on how many competent cells needed) was inoculated with a 5 mL overnight culture which was then incubated at 37 °C with shaking until an OD_{600} of 0.4-0.6 was measured. For every 2 mL of the cell culture that was harvested the following procedure was completed. The cells were harvested by centrifugation at 21 000 *x g* for 30 seconds. The supernatant was discarded and the cells were re-suspended in 2 mL of chilled 0.1 M $MgCl_2$. This was stored on ice for 5 minutes and the cells were harvested by centrifugation at 21 000 *x g* for 30 seconds. The supernatant was once again discarded and the cells were re-suspended in 1 mL of chilled 0.1 M $CaCl_2$. These were then stored on ice for 30 minutes and after harvesting the cells in the same manner and discarded the supernatant the pellet was re-suspended in 300 µL of 0.1 M $CaCl_2$. Aliquots of 60 µL volumes were made and stored at -80 °C until required.

2.2.4 Transformations

Transformations were completed by adding either 1 µL of plasmid purification or 10 µL of ligation product to 60 µL of competent cells, thawed on ice. The cells would be mixed by gentle inverting the tube three times and then stored on ice for 30 minutes. The cells were then heat shocked at 42 °C for 30 seconds and allowed to recover on ice for 5 minutes before the addition of 500 µL LB and incubation for 1 hour at 37 °C with nutation. The cells were then pelleted and the supernatant discarded. The cells were re-suspended in 50 µL of LB and plated on agar plates supplemented with the appropriate

antibiotic. These plates were then grown over night at 37 °C. The following day an isolated colony was selected and streaked on a new plate, which was grown overnight at 37 °C to ensure an isolated colony was selected for creating the glycerol stocks.

2.2.5 SDS Polyacrylamide Gel Electrophoresis

Protein samples were analyzed by sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis (PAGE) using 12% gels (Sambrook and Russell, 2001). The protein samples were mixed with 10X Sample buffer in order to create a 1X working concentration. The samples were heated at 100 °C for ~10 minutes and then they were then loaded into gels adjacent to a PageRuler™ pre-stained protein ladder (Fermentas). The samples were separated using electrophoreses at 160 V for 1 hour. Gels not intended for Western immunoblot analysis were washed with water and microwaved for 30 seconds prior to being stained with a 1X working solution of Bradford reagent (Biorad, Hercules, CA). The gel was microwaved a second time post addition of the Bradford reagent and allowed to stain for 20 minutes before beginning the de-staining procedure. De-staining consisted of incubating the gel in dH_2O with a paper tissue present and gentle agitation. The water was changed and the paper tissue was replaced frequently until the de-staining was complete.

2.2.6 Western Immunoblotting

SDS-PAGE gels intended for Western immunoblot detection were analyzed using a method similar to that outlined by (Sambrook and Russell, 2001). The separated proteins from an SDS-PAGE were transferred to nitrocellulose paper *via* electrophoresis for 1 hour at 45 V. The nitrocellulose membrane was then washed using TBS (10 mM Tris-HCl, 150 mM NaCl, pH 7.5) for 5 minutes twice. After the washing procedure the

membrane was incubated with blocking buffer (3% w/v BSA in TBS) for at least 1 hour. The membrane was then washed twice for two minutes using T^3 buffer (20 mM Tris-HCl, 0.2% (w/v) Triton X-100, 0.05% (w/v) Tween-20, pH 7.5). Following the T^3 washes, the blot was washed for 5 minutes in TBS buffer. The blot was then incubated for 1 hour with the Primary antibody (20 µL of antibody in 20 mL of blocking buffer). Mouse Anti-His_6 antibodies (Biorad) were used to detect His_6 tagged proteins. After the primary antibody was removed, the blot was once again washed twice for 2 minutes with T^3 and once for 5 minutes with TBS. the same process was them completed using the secondary antibody. Goat anti-mouse secondary antibodies (Biorad) were used to detect the mouse antibodies. Once the secondary antibody was removed with the above mentioned washing procedure, NBT/BCIP detection solution (Thermo Scientific, Waltham Ma) was added to the blot. The blot was then developed according to the manufacturer's instructions. All washes and incubation times were complimented with gentle agitation.

2.2.7 Protein Quantification

Protein concentration was determined using one of three methods; BCA assay, BSA assay, or Beer's Law. The BCA and BSA assays were performed in 96 well microplates and read in BMG Fluorstar* Optima plate reader. Standards were made at concentrations ranging between 0-1 mg/ mL, using BSA purchased from BioRad for use in the BCA and BSA assays. Each standard was made using the same buffer as the protein sample being analyzed. For the BCA assay, 25 µL of each unknown or standard was mixed with 200 µL of a working reagent consisting of a 50 parts bicinchoninic acid and 1 part copper (II) sulfate. The samples were incubated at 37 °C for 30 minutes before being read at 595 nm. The BSA assay was completed by adding 5 µL of protein sample

to 195 µL of a 1 in 5 dilution of Bradford reagent (BioRad). The reaction was incubated at room temperature for 5 minutes before being read at 595 nm. All protein samples were compared to BSA standards. The BSA was used in cases where the protein was stored in the presence of a reducing agent. A predicted extinction coefficient for Cel6A was determined to be 2.66 A_{280} units per mg per cm using the Expasy server (Gasteiger, 2003). Protein samples were diluted in half using 8 M guanidine-HCl and their absorbance was measured at 280 nm. The protein concentration was then determined using Beer's Law.

2.3 Cloning

2.3.1 Conventional Cloning

PCRs, for cloning purposes, were conducted using the KAPA HiFi™ PCR Kit available through KAPA Biosystems (Boston, Massachusetts). Conditions were followed as outlined by KAPA Biosystems; see Table 2.3 for a list of reaction mixture components and Table 2.4 for PCR conditions. To create pACRW06/07, conventional cloning procedures were followed. Primers were designed to remove the 31 amino acid leader peptide and introduce a stop codon at the end of the sequence to remove the His_6-tag (Figure 2.1). All PCR products were subjected to a PCR cleanup and gel extraction procedure using the PureLink™ Quick Gel Extraction and PCR purification combo kit (Invitrogen). All restriction enzyme digests were completed using enzymes from New England BioLabs® for 3 hours at 37 °C. Ligations reactions used ligase available from New England BioLabs® with a 3:1 insert to vector ratio, stored at 4 °C for at least 24 hours prior to transformation.

Table 2.3: Reagents and Components of PCR Reactions

Component	Final concentration
5X KAPA HiFi GC Buffer (with 2.0 mM MG^{2+} at 1X)	1X
KAPA dNTP Mix	0.3 µM
Forward primer	0.3 µM
Reverse primer	0.3 µM
Template DNA	10 ng
KAPA HiFi DNA Polymerase	0.5 U/25 µL reaction
Total volume	25 µL

Table 2.4: PCR Reaction Conditions

Step	Temperature (°C)	Time (seconds)	
Heat activation of polymerase	95	300	
Denaturation	98	120	30x
Annealing	65	15	
Elongation	72	30	
Final elongation	72	60	

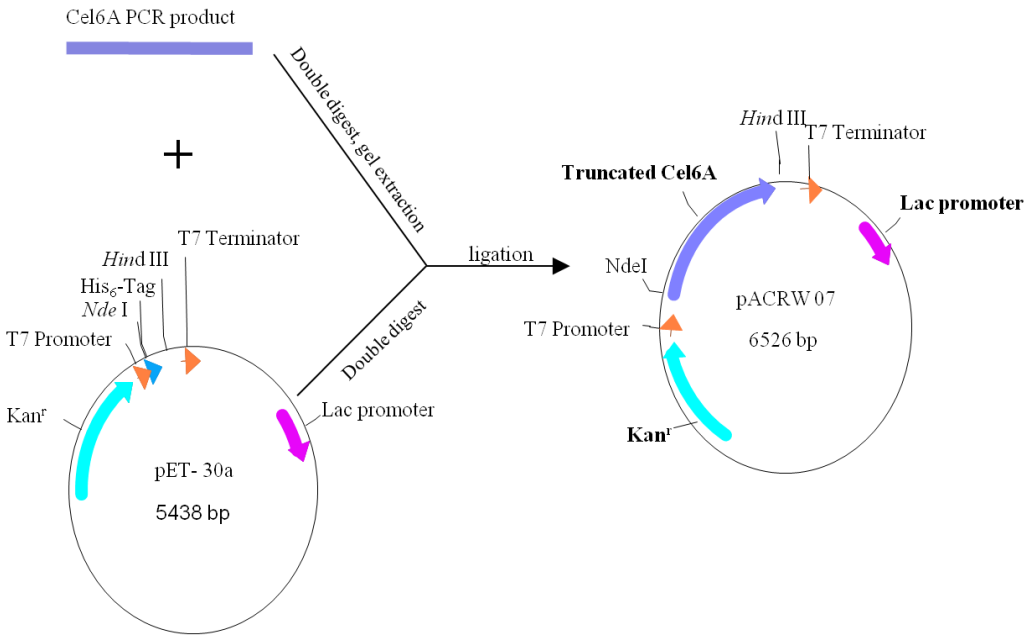


Figure 2.1: Cloning stragey for creating pACRW-06 and pACRW-07. The Cel6A and (D392C) Cel6A PCR product was engineered to contain *Nde* I and *Hind* III restriction sites for insertion into pET-30a. The inserted PCR product displaces the His_6-tag. The Lac promoter and Kan^r genes are useful for overexpression and selection.

2.3.2 Champion™ pET SUMO Cloning

Champion™ pET SUMO constructs (pACRW08/09) were created using the TA cloning system outlined in the Champion™ pET SUMO expression system (Invitrogen). Taq polymerase provided with the kit was utilized in order to introduce a deoxyadenosine overhang on each end of the PCR product. The PCR product was then ligated into the linearlized SUMO plasmid, which contained deoxythymidine overhangs (Weeks et al., 2007).

2.3.3 Recombinant Protein Screening

Ligation products were transformed into *E .coli* DH5α and grown overnight on a LB plate with the required antibiotic challenge. Each colony produced by this process

was then screened for the presence of the gene of interest using the following methods: analytical PCR, E-lyse, or analytical restriction digests. To confirm the gene was present and free of mutations, successfully screened colonies were sent for sequencing analysis.

2.3.4 Analytical PCR

To perform colony PCR, a small sample from the colony was mixed with 30 µL of dH_2O and boiled for 10 minutes. The insoluble fraction was separated from the soluble by centrifugation for 5 minutes at 21 000 *x g*. The supernatant was then used as the DNA template in a PCR using: 12.5 µL GoTaq® supermix (Promega), 2.5 µL of each of the primer used to generate the clones, 5 µL of the supernatant, and 2.5 µL of dH_2O. The reaction was completed following the protocol outlined in Table 2.4.

2.3.5 E-lyse

A sucrose solution was made by incubating a 25% (w/v) solution of sucrose with RNase A for 10 minutes at 90 °C and then adding lysozyme. A sample taken from freshly grown colony of each transformant was mixed with a 10 µL of TE buffer and 10 µL of the sucrose solution. The solution was then loaded into a 1% agarose gel with 0.2% SDS. The mixture was allowed to incubate for 15 minutes before the DNA was separated for 20 minutes at 20 V followed by 35 minutes at 90 V.

2.3.6 Analytical restriction digest

Potential colonies were used to inoculate overnight cultures. The plasmids from these cultures were harvested and subjected to restriction digestions. The same enzymes used in the creation of the construct were used in these digestions.

Descriptions of all plasmids created or used during the process of this study can be found in Table 2.5

2.4 Protein Over-Production and Purification

2.4.1 Expression Trials

All gene expression trials were performed in 5 mL volumes of either LB or Super Broth. Each 5 mL culture was inoculated with 100 µL of an overnight culture and grown to an OD_{600} between 0.4-0.6. The cultures were then induced with either 0.1 mM or 1 mM IPTG and grown at either 30°C, 37 °C, or 15°C with time points being taken at 1h, 3h, and O/N for the cultures growing at 30 °C and 37 °C. Samples were only taken after O/N expression for cultures growing at 15 °C. All samples obtained were 1 mL in size and the bacterial cells were collected by centrifugation at 21 000 *x g* for 1 minute. The samples were then treated with B-PER protein extraction reagent (Thermo Scientific) according to the manufacturer's instructions. The soluble and insoluble fractions of each sample were then analyzed using SDS-PAGE and Western immunoblotting to determine optimal expression conditions.

2.4.2 Over-Production of Proteins

Overnight cultures (5 mL x 3) grown in LB were sub-cultured into 1 L of LB supplemented with 50 mg/mL of Kanamycin. The culture was then grown at 37 °C with shaking until an OD_{600} of between 0.6-0.8 was reached. Gene expression was then induced with 1 mM IPTG and grown overnight at 15 °C with shaking at 200 rpm. The cells were harvested by centrifugation at 3000 *x g* for 15 minutes at 4 °C. The pellets were frozen between 1-24 hours. The pellets were re-suspended in 25 mL binding buffer (50 mM sodium phosphate (or 25 mM Tris-HCl), 300 mM NaCl, 10 mM imidazole, pH 8.0) per L harvested. One hundred µL of 5mg/mL DNAse and RNase solutions were added/25 mL of lysate. The cells were lysed using a French Press method with two passes

for each 25 mL aliquot. The soluble fraction was collected by centrifugation at 13 690 *x g* for 15 minutes at 4 °C.

2.4.3 Purification of Untagged Cel6A

Cel6A was expressed as described above. The whole cell lysate (WCL) was then passed through an anion-exchange column packed with Source 15Q resin (GE Healthcare Life Sciences, Baie d'Urfe, QC) previously equilibrated with 50 mM ammonium bicarbonate, pH 9.0. The flow through was collected and then a linear gradient to 1 M NaCl was applied over 60 minutes at a flow rate of 0.7 mL/min. One mL fractions were collected for further zymogram analysis. Attempts were made to further purify Cel6A by gel filtration with a 30 mL column packed with Superdex 200 (GE Healthcare Life Sciences). The collected fractions from the ion-exchange chromatography were lyophilized overnight and then rehydrated in 200 µL of dH_2O. Protein samples (100 µL at a time) were applied to column equilibrated in 50 mM sodium phosphate buffer, pH 8.0.

2.4.4 Purification of SUMO Tagged Cel6A and (D392C) Cel6A

The soluble fraction collected by centrifugation post over-expression was incubated with 1.5 mL of Ni^{2+}-NTA resin (Thermo Scientific) for 1.5 hours. The lysate was then placed in a re-usable IMAC column and allowed to settle for 10 minutes. Once settled, the lysate was allowed to flow through. The column was then washed with 150 mL of wash buffer (50 mM sodium phosphate, 300 mM NaCl, 50 mM imidazole, pH 8.0). The protein was eluted in 10 mL of elution buffer (50 mM sodium phosphate, 300 mM NaCl, 250 mM imidazole, pH 8.0). The SUMO tag was removed from the recovered protein by incubating it with 2 units of SUMO protease (provided with SUMO cloning kit) at 4°C during dialysis against 4 L of IMAC binding buffer (25 mM Tris-HCl, 300

mM NaCl, 10 mM imidazole, pH 8.0) overnight with three buffer changes. The enzyme mixture was separated by IMAC a second time following the above mentioned procedure. The protein of interest was collected in the flow-through component.

2.4.5 Purification of His_6-tagged Cel6A and (D392C) Cel6A

Recombinant Cel6A or (D392C) Cel6A was over-produced and collected as described above. The WCL was incubated with 0.5 mL/L of culture of Ni^{2+} resin (Thermo Scientific) 1.5 hours. The lysate was then placed in a re-usable IMAC column and allowed to settle for 10 minutes. Once settled, the lysate was allowed to flow through. The column was washed with 150 mL of wash buffer (50 mM sodium phosphate or 25 mM Tris-HCl), 300 mM NaCl, 50 mM imidazole, pH 8.0) prior to elution of the His_6-tagged proteins in 10 mL of elution buffer (50 mM sodium phosphate or 25 mM Tris-HCl), 300 mM NaCl, 250 mM imidazole, pH 8.0). When TCEP was present in the binding and elution buffers, the protein was eluted in two steps. The first elution was performed using 100 mM imidazole in two 2.5 mL steps and the second elution involved 2.5 mL additions of 250 mM imidazole until all the protein has been eluted.

Further purification was achieved by anion-exchange chromatography. The first elution from IMAC was dialyzed against 4L of 50 mM ammonium bicarbonate, pH 9.0 overnight with three buffer changes. This was then filtered and loaded onto a 2 mL Source 15Q column (GE Healthcare) equilibrated in 50 mM ammonium bicarbonate, pH 9.0. The protein of interest was collected in the flow through while the contaminating proteins remained bound to the column. The column was then washed with 1M NaCl and stored in 20% ethanol.

In the case of (D392C) Cel6A, the same purification procedure as described above was followed with the addition of 1 mM TCEP or 5 mM DTT to the binding, wash and elution buffers. One hundred mM TCEP stock solutions were made in 25 mM Tris-HCl and stored at 4 °C for no more than 2 weeks. The DTT solutions were made fresh for each use.

2.4.6 Zymogram analysis

Cel6A, prepared with no reducing agent or boiling, was separated using SDS-PAGE. The SDS was removed by incubating the gel for 2 hours in 50 mM disodium hydrogen phosphate, and 12.5 mM citric acid at pH 6.3 with the buffer being changed every 30 minutes. The first two 30 minute washes also contained 25% isopropanol. The gel was then placed in a 60 °C incubator for 10 minutes to remove excess liquid. During the wash time, an agarose replica gel was made consisting of 2% agar and 0.1% carboxymethyl cellulose (CMC) in 50 mM disodium hydrogen phosphate, and 12.5 mM citric acid at pH 6.3. The gel was allowed to cool at 4 °C prior to its use. Excess moisture was removed from the replica gel by incubation at 60 °C for 15 minutes. The SDS-PAGE gel was then laid on top of the agarose gel, avoiding the formation of air bubbles. The whole assembly was covered with parafilm and incubated for 3 hours at 60 °C. The gel was then cooled at 4 °C before being submerged in 0.1% Congo Red. The gel was stained for 30 minutes and then de-stained using 1 M sodium chloride, with frequent buffer changes. Gels were rinsed with 5% acetic acid as a counter stain. Zymogram analysis was also attempted in which in place of a replica gel, 0.2% CMC was incorporated into the SDS-PAGE. After separation of the proteins on the SDS-PAGE

zymogram gel, the same procedure was completed to remove the SDS and the gels were stained and de-stained in the same fashion as described above.

2.5 Free Sulfhydryl Detection

2.5.1 Dimedone Probe

To assay for the presence of Cys-sulfenic acid (D392C) Cel6A (6 µM) was reacted with 250 µM of DCP-Bio1 probe (KeraFAST) for 30 minutes on ice. The protein sample was then passed through a Millipore filter with a 30 kDa cut off to remove excess probe prior to SDS-PAGE and Western blot analysis. The Western blot was performed as described previously with an avidin-AP antibody (Biorad).

2.5.2 2,2' Dithiodipyridine (DTDP) Quantification of Free Thiols

DTDP was utilized to quantify the amount of free sulfhydryl in (D392C) Cel6A. DTDP (125 µL of 4 mM solution) was added to the unknown sample as well as to a series of known concentrations of L-cysteine and the absorbance of each measured at 324 nm. The amount of free sulfhydryl detected using this method was then compared to the amount of theoretical sulfhydryl to determine the percent oxidation of (D392C) Cel6A.

2.6 Activity Assays

Activity assays were performed using enzyme at a final concentration of 7.4 µg/ mL in 25 mM Tris-HCl buffer, pH 8.0. For determination of a suitable substrate, 1% w/v solutions of two different types of CMC were tested: low viscosity CMC (Sigma-Aldrich, St. Louis MO) and medium viscosity CMC (Megazyme, Burlington ON). Activity assays were incubated at 37 °C with nutation. At various time points (generally 0, 10, 20, and 60 min) 600 µl of solution was removed from reactions and mixed with 40 µl of 500

mM NaOH to stop the reaction. The time point samples were then frozen until all samples were collected and the amount of reducing sugars liberated could be determined.

Subsequent routine activity assays were conducted using either 1 or 0.1% w/v of the medium viscosity CMC. The assays may also have included 1 mM TCEP or 5 mM DTT, depending on the circumstances. Time points taken varied with specific assays, however, generally 0, 10, 60 and overnight time points were taken.

2.7 Reducing Sugar Assays

The amount of reducing sugars in a sample was determined using the alkaline p-hydroxybenzoic acid hydrazide (PAHBAH) method of determining reducing sugar content (Lever, 1972; Koziol, 1981). A standard curve using glucose was generated before the samples were assayed. A fresh 10 mM stock of glucose was made each time and then diluted to make 0.05, 0.1, 0.2, 0.4, 0.6, 0.8, and 1 mM standards. The standards were made in whichever buffer was used in the activity assays. Three aliquots of 200 µL of each standard (including a 0, buffer only, standard) had 400 µL of PAHBAH coloured solution (0.5 M PAHBAH, 0.05 M sodium sulfate, 0.01 M calcium chloride, 0.02 M sodium citrate and 0.3 M sodium hydroxide) added to them prior to vigorous boiling in a water bath for 5 minutes. They were immediately cooled in an ice water bath for 1 minute before being diluted based on the intensity of their reaction. The absorbance was then measured in a Beckman Coulter DU 530 UV/Vis spectrometer or a Nanophotometer p300 at 420 nm. Samples were diluted when necessary to obtain an accurate reading. For the samples collected during the activity assay, each 600 µL sample was thawed on ice and then divided into three aliquots of 200 µL. To each aliquot, 400 µL of PAHBAH colour solution was added and treated as described above. The absorbance values

obtained from were compared to a standard curve of glucose to determine the approximate amount of reducing sugars in each sample.

2.8 Determination of Specific Activity

Specific activity was determined using the following equation:

$$A_{420}/(m*p*t),$$

Where; m is the slope of the standard curve, p is the protein concentration and t is the time point resulting in a U/mg of protein. All specific activities were determined using time points that fell within the linear portion of the protein's activity curve.

2.9 Cysteine Oxidation

(D392C) Cel6A was purified in the presence of 1 mM TCEP through to IMAC elution. The IMAC eluant was dialyzed against 4L of 100 mM Tris-HCl buffer, pH 8.0. Oxidation reactions were set up with 12.5 µM (D392C) Cel6A and 100 µM H_2O_2 in 25 mM Tris-HCl. A control reaction was set up with no added H_2O_2. The reactions were incubated at 4 °C for 24 hours. Post incubation, the reactions were washed by ultrafiltration using Amicon Ultra 0.5 mL filters with 30 kDa cut offs (Millipore). The samples were washed seven times by centrifuging 400 µL of the sample for 5 minutes at 14 000 *x g* and then making the samples to 400µL with fresh buffer after each spin. The washed samples, as well as unwashed samples, were then used in activity assays with 0.1% CMC as described previously.

2.10 Cysteine alkylation

Cysteine alkylation trials were completed using iodoacetic acid, iodoacetamide, iodobutyric acid and iodopropionic acid. Each reagent was added at a 100 times molar excess to (D392C) Cel6A. The reaction buffer used was 100 mM Tris-HCl, pH 8.0 and

reactions were incubated at either 4 °C or 37 °C, with time points being taken after either 1 hour or 24 hours for further activity analysis. Prior to performing the activity assays, the enzyme samples were subjected to ultracentrifugation as described above using Amicon 30 kDa cut off filters (Millipore Billerica, MA) with at least six passes.

2.11 Mass Spectrometry

2.11.1 MALDI-TOF Mass Spectrometry

MALDI-TOF MS analysis was utilized in order to determine the size of the two protein species obtained during the purification of the full length His_6 tagged Cel6A from pACDC-003. Analysis was carried out by the Advanced Aanalysis Centre at the University of Guelph using a Bruker Reflex III MALDI-TOF with a sinnapinic acid matrix.

2.11.2 Trypsin Digests

For the identification of the cysteine oxidation state at position 392 in (D392C) Cel6A, a trypsin digest was performed for subsequent peptide mapping. (D392C) Cel6A (6 µg) was digested with 0.1 µg of trypsin for 3 hours at 37 °C. The samples were then analyzed by MS

Peptides for the identification of the propionic modification to C392 were prepared by in gel digestion following the protocol provided by the MS Facility of the AAC (University of Guelph). The protein of interest was resolved using SDS-PAGE as described above. After washing away the SDS with water, the bands of interest were excised and then de-stained with 25 mM NH_4HCO_3 in 50% acetonitrile (ACN) and periodic agitation for 10 minutes. This wash was repeated twice. The gel particles were incubated with 100% ACN until dehydrated (shrunken) and then reduced for 30 minutes

at 50 °C with 10 mM DTT. The gel particles were once again dehydrated with ACN prior to incubation with 55 mM iodoacetamide for 30 minutes at room temperature. Since iodoacetamide is light sensitive, the incubation was done in the dark. The gel pieces were then washed with 50 mM NH_4HCO_3 for 15 minutes with occasional agitation. The particles were once again dehydrated with ACN and then they were dried using a Speed Vac for 20 minutes. The gel slices were then digested with trypsin and incubated for 1 hour at room temperature before the addition of 50 mM NH_4HCO_3 for incubation at 37 °C overnight. The next day, 50 µL of H_2O was added to each sample and the sample was sonicated for 10 minutes. The sample was then briefly centrifuged and the supernatant was collected. The gel pieces were then exposed to 75 µL of 5% formic acid (FA) in ACN. They were then vortexed for 2 minutes, sonicated for 5 minutes, and briefly centrifuged. The supernatant was collected and this step was repeated once again with the exception of the sonication. The supernatant was then concentrated to approximately 10 µL using a Speed Vac. Once the sample was passed through a C18 Zip tip (Millipore), they were ready for MS analysis.

2.11.3 Liquid Chromatography-Mass Spectrometry (LC-MS)

LC-MS analyses were performed on a Dionex UHPLC UltiMate 3000 liquid chromatograph interfaced to anamaZon SL ion trap mass spectrometer (Bruker Daltonics, Billerica, MA) at the MS Facility of the Advanced Analysis Centre (University of Guelph). A C18 column (Phenomenex Aeris Peptide, 3.6 micron particle size, 150 mm x 2.1 mm) was used for chromatographic separation. The mobile phase gradient was as follows: initial conditions, 2% acetonitrile in 0.1% formic acid increasing to 55% acetonitrile (0.1% formic acid) over 35 min. The flow rate was maintained at 0.2 mL/min.

The mass spectrometer electrospray capillary voltage was maintained at 4.5 kV and the drying temperature at 220 °C with a flow rate of 10 L/min drying gas. Nebulizer pressure was 40psi. Nitrogen was used as both nebulizing and drying gas; helium was used as collision gas at 60 psi. The mass-to-charge ratio was scanned across the m/z range in enhanced resolution positive-ion auto MS/MS mode. The Smart Parameter Setting (SPS) was used to automatically optimize the trap drive level for precursor ions. The instrument was externally calibrated with the ESI TuneMix (Agilent). The sample injection volume was 20 µl. Data processing was performed using DataAnalysis software from Bruker where the 200 most intense peaks that triggered Auto MS/MS were tabulated and then Bruker BioTools software was used to compare peptides fragmentations to the protein sequence with and without modification.

Chapter 3: Results

3.1 *In Silico* Work

The SignalP algorithm determined the signal sequence cleavage site in Cel6A to be between amino acid 31 and 32 (Figure 3.1). Disulfind predicted that all six native

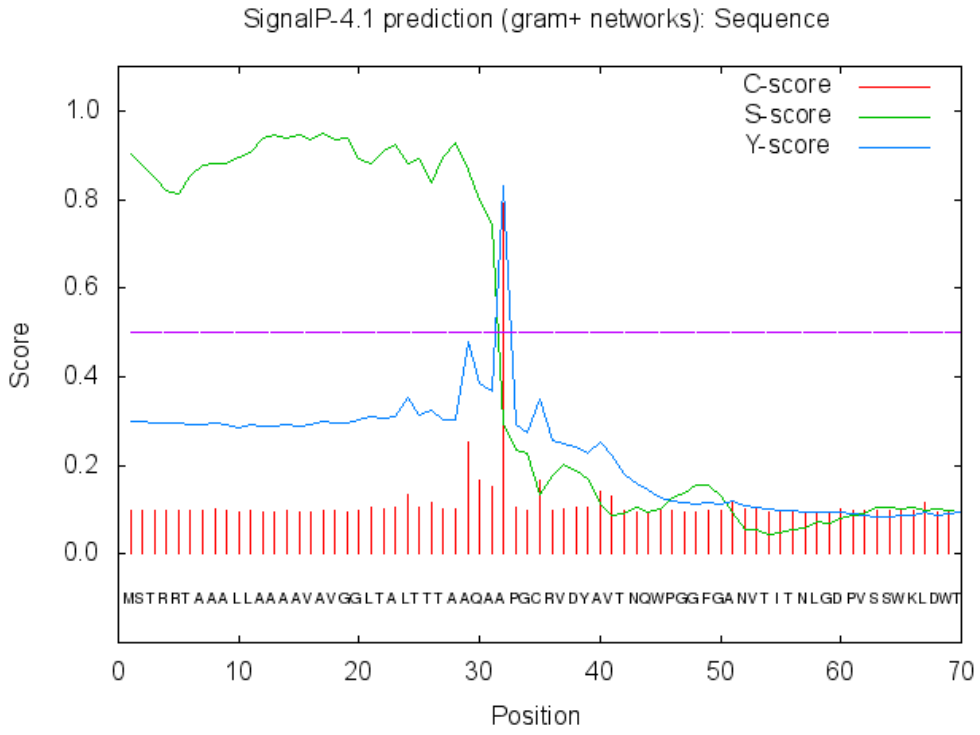


Figure 3.1: SignalP prediction of signal cleavage site within Cel6A amino acid sequence. The sequence was compared to known Gram-positive bacterial signal peptides. The analysis revealed the cleavage site to be between amino acid 31 and 32.

cysteines formed disulfide bonds with a confidence of 0.666374. The disulfide partners were predicted to be: C35 - C390, C134 - C291, and C248 - C426. The ProtParam tool predicted the MW to be 46730.9 kDa, the PI to be 7.56, the extinction coefficient to be 107745 M^{-1} cm^{-1}, and estimated half life to be greater than 10 hours in *E. coli* for the WT

protein. For (D392C) Cel6A, the MW was predicted to be 47541.8 Da, the theoretical PI as 8.01, an extinction coefficient of 107745 M^{-1} $cm^{-1}$and the estimated half life of over 10 hours in *E. coli*. The Phyre2 structural model of Cel6A was based on the known structure of *M. tuberculosis* Cel6A and the two catalytic residues within the active site are predicted to be 9.8 Å apart (Figure 3.2).

`

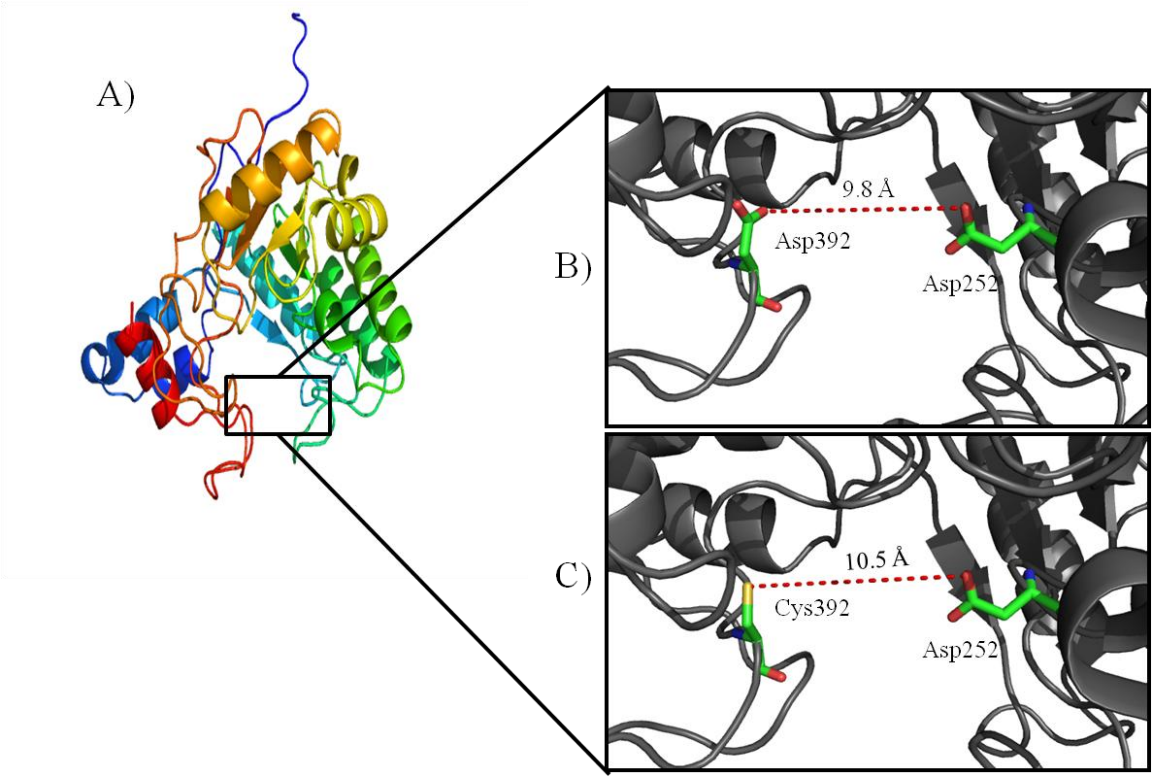


Figure 3.2: Threaded structure of truncated Cel6A predicted by Phyre2. A) Predicted structure of Cel6A using Cel6A from *Mycobacterium tuberculosis* (d1ouza) as the template. Calculated distances between (B) the two catalytic Asp residues in the WT enzyme, and (C) the catalytic Asp252 and Cys392 residues of (D392C) Cel6A.

3.2 Cloning

3.2.1 Conventional Cloning

The *cel6A* and *(D392C) cel6A* fragments were successfully amplified using the PCR method outlined in Table 2.4 using pACDC-003 and pACDC-006 as templates, respectively (Figure 3.3). The genes were excised and subjected to restriction digestions

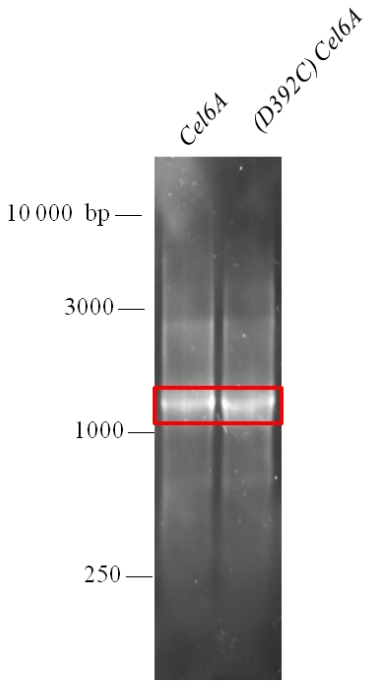


Figure 3.3: Representative agarose gel electrophoresis of PCR products. The red box indicates the PCR products of a reaction amplifying *cel6A* and *(D392C) cel6A* for insertion into the SUMO vector. Electrophoresis was conducted at 100 V for 30 minutes with a GeneRuler™ 1kb DNA ladder, and detected using SYBR® safe.

using *Nde*I and *Hind*III. The PCR products were then inserted into a pET30a plasmid which had also been digested with *Nde*I/*Hind*III. The subsequent colonies were screened using analytical PCR (Figure 3.4), analytical restriction digests (Figure 3.5) and E-lyse (Figure 3.6).

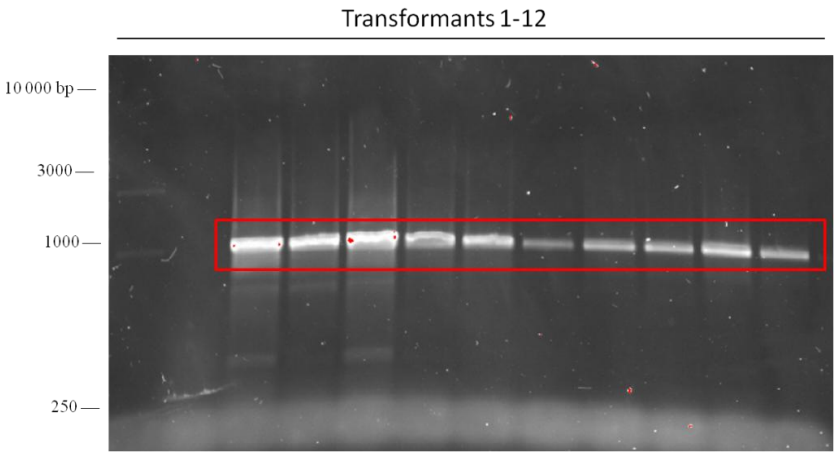


Figure 3.4: Representative analytical PCR agarose gel. Lanes 1-12 contain colony PCR products from 12 recombinant Cel6A clones. Clones 3-12 tested positive with a band appearing at the expected 1200 bp, indicated by a red box. Electrophoresis was conducted at 100 V for 30 minutes with a GeneRuler™ 1kb DNA ladder and detected using SYBR® safe.

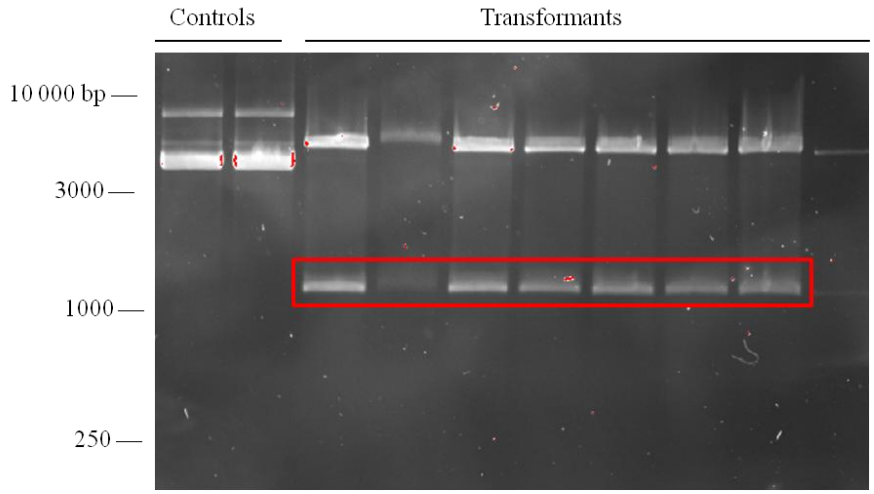


Figure 3.5: Representative agarose gel electrophoresis of analytical restriction digest. Controls were undigested pET30a and undigested recombinant plasmid. Recombinant plasmids were digested with *Hind*III and *Nde*I. This particular gel was used to analyze recombinant plasmids during the creation of pACRW-07. Plasmids in lanes 4-10 are positive for the presence of *(D392C) cel6A* as indicated by the red box. Electrophoresis was conducted at 100 V for 30 minutes with a GeneRuler™ 1kb DNA ladder and detected using SYBR® safe.

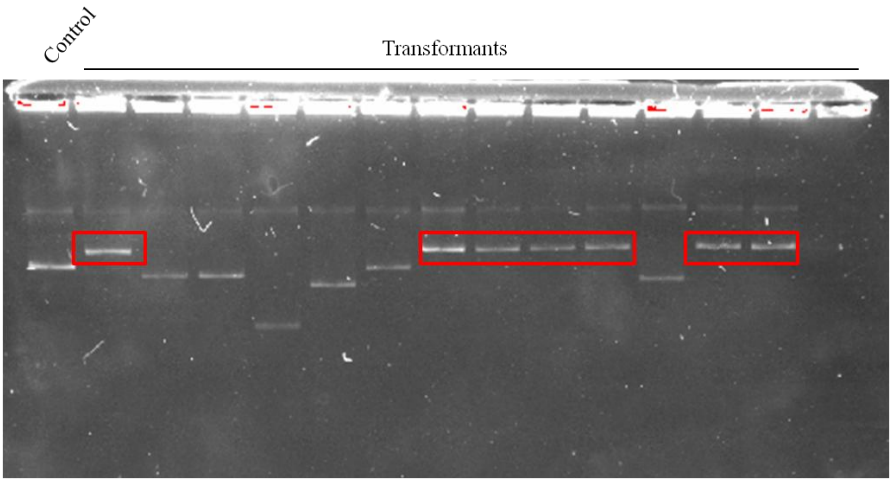


Figure 3.6: e-lyse gel electrophoresis analyzing *cel6A* transformants during the creation of pACRW-06. An empty pET30a plasmid was used as a control to visualize the migration pattern of a plasmid without an insert. Transformants that potentially contain an insert are indicated by a red box. Electrophoresis was conducted at 100 V for 30 minutes and detected using SYBR® safe.

Colonies that screened positive were sent for sequencing analysis which confirmed that the inserted genes contained the correct sequence, with the exception of one silent point mutation in which a cytosine at position 786 was replaced with a thymine. From the positive sequences one of the colonies was chosen for all future work.

3.2.2 Champion™ pET SUMO Cloning

The genes of interest (*cel6A* and *(D392C) Cel6A*) were amplified from pACDC-003 and pACDC-006, respectively using primers found in Table 2.1. The PCR products were gel extracted and ligated into the pET SUMO vector. Successful transformants were screened using colony PCR. Plasmids from the colonies that screened positive were confirmed to have the proper insert by nucleotide sequence analysis. One of the plasmids was then chosen for all future work.

3.3 Protein Purification

3.3.1 Cel6A With No His_6 Tag or Leader Peptide

Expression trials involving the construct encoding Cel6A lacking both its leader peptide and a His_6 tag indicated that expression was best at 37 °C for 3 hours using 1 mM IPTG. However, zymogram analysis of these samples showed very little enzymatic activity compared to the full-length WT enzyme (Figure 3.7). Regardless, attempts to purify this form of Cel6A were made. Whole cell lysate from 1 L cultures of transformed cells were collected. These crude enzyme preparations were subjected to anion-exchange chromatography using Source 15Q resin. Under the conditions employed, a protein with the appropriate MW eluted in a number of fractions after application of the linear NaCl gradient (approx. 15 mM to 40 mM NaCl) (Figure 3.8). These fractions were analyzed by zymography but no cellulolytic activity was observed in any of the fractions eluted from the SourceQ resin. This is likely because the protein of interest was located in the flow through, since its predicted pKa does not correspond with its measured pKa. The fractions were nonetheless pooled, concentrated, and applied to a gel filtration column for further purification. However, as before no fraction with cellulolytic activity was recovered.

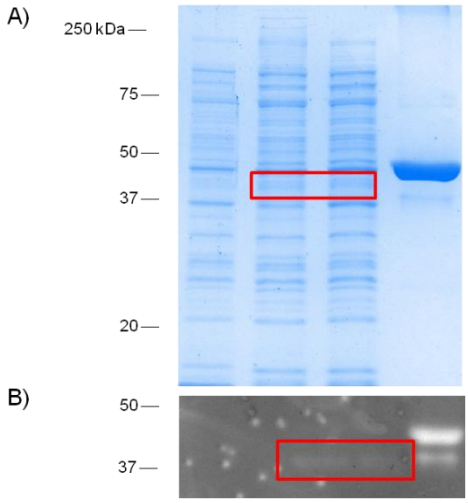


Figure 3.7: SDS-PAGE and zymogram analysis of truncated Cel6A from expression trials. A) SDS PAGE analysis of cell lysates obtained before (lane 1) and after induction with 1 mM IPTG at 37 °C for 3 hours which had not been boiled (lane 2) and had been boiled (lane 3) during the sample preparation. Lane 4, Purified full length Cel6A as a positive control. B) Zymogram analysis using 1% CMC as substrate. The red box denotes weak zones of clearing that were observed in the original gel at the expected size of the truncated Cel6A. MW markers (kDa) are indicated on the left of the respective gels.

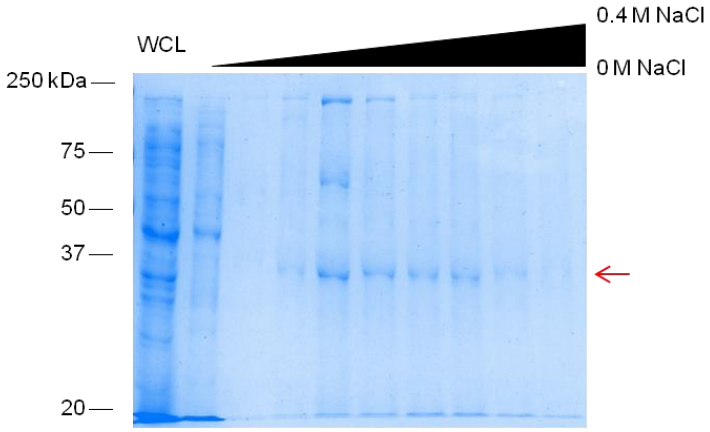


Figure 3.8: SDS-PAGE analysis of Cel6A purification by Source 15Q anion-exchange chromatography. The whole cell lysate (WCL) of *E. coli* transformed with pACRW-06 was loaded onto the Source 15Q column, previously equilibrated in 50 mM ammonium bicarbonate buffer, pH 9. Cel6A was recovered with the application of a linear gradient to 1.0 M NaCl and 1 mL fractions were collected. A protein with the predicted MW of Cel6A is indicated by the red arrow, which eluted between 0.2 M and 0.4 M NaCl. Samples eluted post 0.4 M are not shown. MW markers (kDa) are indicated on the left.

3.3.2 SUMO Fusions

Given the failure to recover active Cel6A lacking its leader peptide, attempts were made to produce the enzyme fused to SUMO. Expression trials with *E. coli* Bl21(DE3) transformed with pACRW-08 indicated that the SUMO construct was produced primarily in an insoluble form. Nevertheless, Cel6A was most successfully expressed with a 1 mM IPTG induction overnight at 15 °C (Figure 3.9). These conditions were used for large scale production of the fusion protein. The overproduced protein was released from the cells using the French Press method at 1000 psi and then purified using a two step IMAC procedure. Thus, IMAC affinity chromatography was used to isolate the fusion protein from the cell digests and then used again to bind (remove) the His-tagged SUMO protein following its cleavage from Cel6A (Figure 3.10). Cleavage of the fusion protein was achieved using the SUMO protease during dialysis at 4 °C overnight. This protocol led to the release, recovery, and purification of the majority of Cel6A. However, it is evident from Figure 3.10 that the digestion was not complete. Regardless, sufficient amounts of Cel6A could be prepared by this protocol for subsequent experimentation.

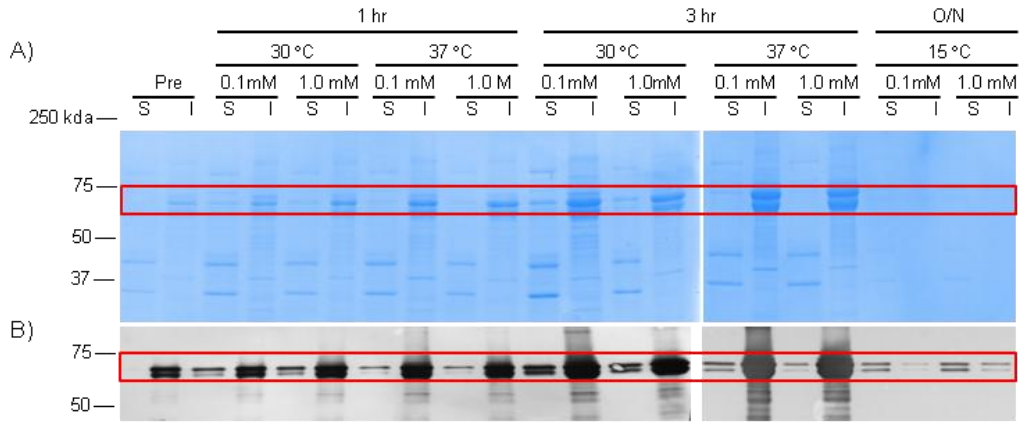


Figure 3.9: Expression of *cel6A* from pACRW-08 for soluble protein product. Cultures (5 mL) of *E. coli* BL21 (DE3) transformed with pACRW-08 were grown to an OD_{600} of 0.8 at the temperatures shown and then induced for gene expression with the addition of either 0.1 mM or 1.0 mM IPTG and incubated for a further 1, 3, or 15 (O/N) hours. The soluble (S) and insoluble (I) proteins were collected and subjected to *-SDS PAGE analysis with (A) staining with Bradford reagent and (B) Western immunoblotting with an anti-His_6 antibody. Electrophoresis was conducted at 160 V for 60 minutes. The red boxes denote the expected migration of Cel6A.

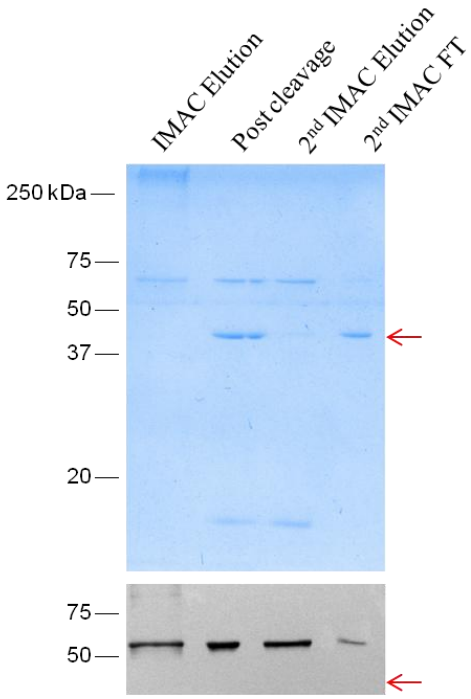


Figure 3.10: SDS-PAGE and corresponding Western immunoblot of SUMO cleavage. Cleavage was performed by incubating the 1st IMAC elution with the SUMO protease during dialysis. The native Cel6A was recovered in the flow through (FT) of the 2nd IMAC. Red arrows indicate the expected migration pattern of Cel6A with the SUMO tag removed. As expected it does not appear on the Western blot as the His_6 tag is removed with the SUMO tag. All protein samples were separated by electrophoresis at 160 V for 60 minutes and detected using 1X Bradford reagent. Western Blot was performed as outlined in the method section using mouse-anti-His_6 primary antibodies and goat anti-mouse secondary antibodies.MW markers (kDa) are indicated on the left.

Production and purification of (D392C) Cel6A-SUMO was attempted as described above for WT Cel6A. Thus, the encoding gene in *E. coli* Bl21(DE3) transformed with pACRW-09 was expressed with a 1 mM IPTG induction overnight at 15 °C and the overproduced protein was released from the cells using the French Press method at 1000 psi. Unfortunately however, extensive degradation of the protein was observed during the initial IMAC purification (Figure 3.11) and no intact (D392C) Cel6A could be recovered.

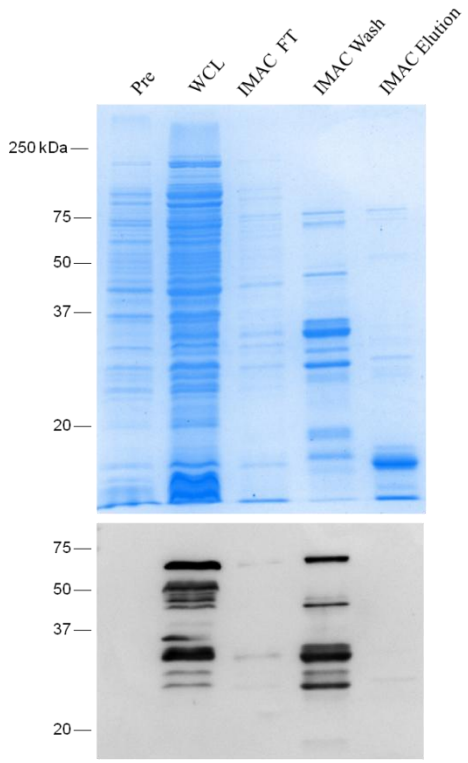


Figure 3.11: Purification of (D392C) Cel6A. (D392C) Cel6A was produced as a SUMO fusion encoded on pACRW09. Pre-induced cells (Pre) were induced for protein production and the whole cell lysate (WCL) was subjected to IMAC on Ni^{2+}-NTA agarose. The flow through (FT) fraction was collected and the resin was washed (Wash) with 30 mM imidazole before the fusion protein was eluted with 250 mM imidazole (Elution). Electrophoretic conditions of the SDS-PAGE analysis (top) and Western immunobloting (bottom) were as described in the legend to Fig. 3.10. MW markers (kDa) are indicated on the left.

3.3.3 His_6 Tags

The WT and (D392C) forms of Cel6A possessing C-terminal His_6 tags were overproduced in either *E. coli* Bl21 (De3) or *E. coli* Shuffle® T7 transformed with pACDC-03 and pACDC-06, respectively. Optimal expression and overproduction conditions were found to be using a 1 mM IPTG induction at 15°C overnight. The overproduced proteins were successfully harvested using the French Press method of cell lysis at a pressure of 1000 psi followed by centrifugation at 3800 *x g* for 15 minutes at 4 °C. The proteins were purified from the soluble the whole cell lysate using a combination of IMAC affinity chromatography followed by anion-exchange chromatography (Figure 3.12). Both forms of Cel6A were recovered from the IMAC resin by the addition of 250 mM imidazole to the elution buffer. Anion-exchange chromatography on Source 15Q led to the further purification of the two enzymes where both were recovered in the flow through fractions while contaminating proteins remained bound. Under the conditions employed, typically 20 mg of purified enzyme was recovered per L of cells.

In an attempt to minimize any auto-oxidation of the free Cys392 residue, (D392C) Cel6A was also purified from cell lysate using the protocol described above but in the presence of either 1 mM TCEP or 5 mM DTT. The inclusion of either reducing agent did not interfere with the initial IMAC affinity chromatography, but the 1 mM TCEP interfered with the anion-exchange chromatography with Source 15Q. In this case, (D392C) Cel6A was purified by only IMAC affinity chromatography which involved step gradients. Thus, after the first wash step containing 30 mM imidazole, a second step was completed using 50 mM imidazole. A small amount of (D392C) Cel6A eluted from

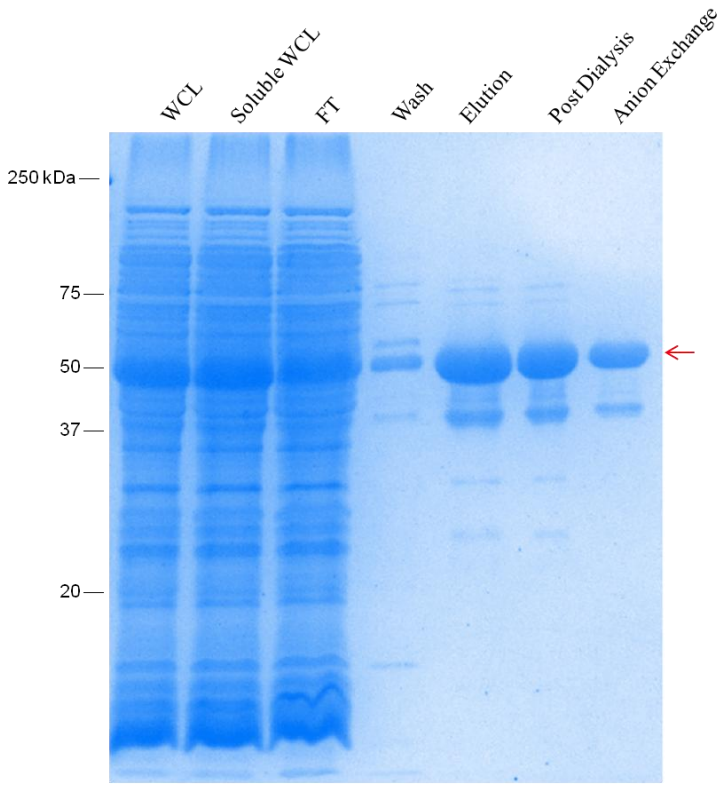


Figure 3.12: A representative SDS-PAGE analysis of a typical (D392C) Cel6A purification. Whole cell lysates (WCL) obtained by French Press were subjected to centrifugation and the soluble WCL was to Ni^{2+}-NTA agarose. The flow through (FT) fraction was collected and the resin was washed (Wash) 30 mM imidazole before the (D392C) Cel6A was eluted with 250 mM imidazole (Elution). The recovered enzyme was dialyzed before anion-exchange chromatography on SourceQ. Expected migration of Cel6A is indicated by the red arrow. Electrophoretic conditions and staining were as described in the legend to Fig. 3.9. MW markers (kDa) are indicated on the left.

the column together with contaminants, but the majority was recovered in the third step involving the addition of 250 mM imidazole.

3.4 Characterization of Recombinant Forms of Cel6A

Zymogram analysis of the full length Cel6A indicated that both it as well as the large degradation product typically formed are both active against CMC (Figure 3.13). MALDI-TOF MS analysis of these two forms of Cel6A provided masses close to those expected, 47,044 Da for the full length enzyme (theoretical mass, 47 553 Da) and 45,674 Da for the truncated digestion product. Western immunoblot analysis indicated that the large degradation product still retained the C-terminal His tag suggesting that truncation occurred at the N-terminus. Analysis of the predicted N-terminal sequence of Cel6A suggests that the difference of 1,370 Da would involve the first 27 residues, Met to Thr located within the leader peptide. Intriguingly however, zymogram analysis of Cel6A with an engineered truncation, removing the leader peptide showed minimal activity (Figure 3.7).

The specific activity of the purified WT Cel6A was determined to be 23.3 ± 0.08 U/mg using 1% CMC (medium viscosity) as substrate in 50 mM sodium phosphate buffer, pH 8.0. An attempt was made to confirm the pH-activity optimum of Cel6A. This experiment involved the use of sodium acetate buffers for pH 4.5 and 5, and Tris-HCl buffers for activity measurements made at pH 6.2 to 8 (Figure 3.14). However, it was subsequently determined that Cel6A is inhibited by Tris which thus compromised the

data collected earlier (Figure 3.15). Thus, the specific activity of Cel6A in 50 mM Tris-HCl, pH 8.0 was 8.5 ± 0.3 U/mg, only 37% of that determined in phosphate buffer.

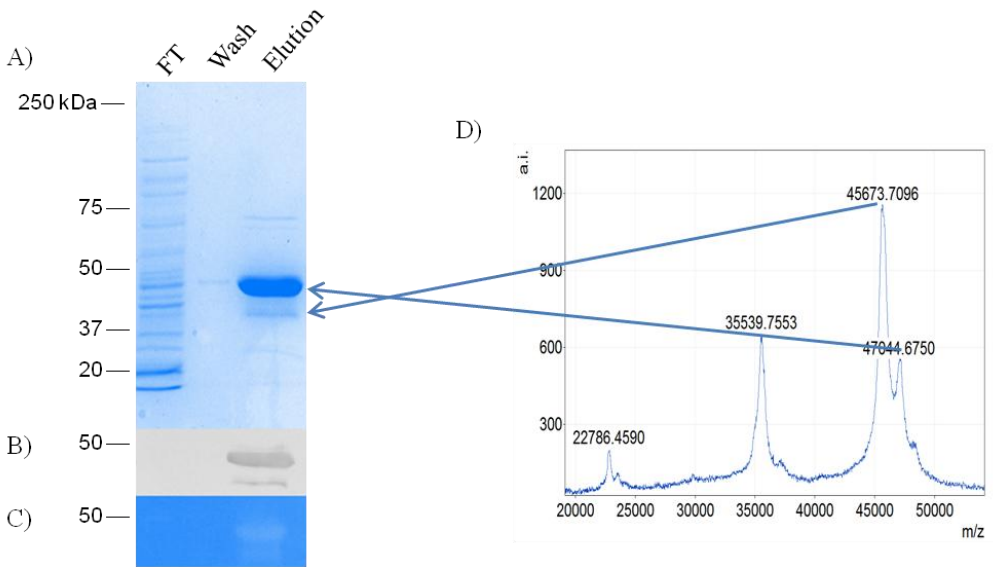


Figure 3.13: MALDI TOF analysis of Cel6A. The flow through (FT), wash, and imidazole eluted (Elution) fractions of an IMAC on N^{2+}-NTA agarose were subjected to SDS-PAGE analysis with Commassie Blue staining (A) and corresponding Western immunoblot (B) a as well as zymography (C). MALDI-TOF MS analysis was performed on protein extracted from the SDS-PAGE as indicated by the arrows. Electrophoretic conditions and Western immunobloting were as described in the legend to Fig. 3.10. Zymogram analysis was performed as described in the methods section. MW markers (kDa) are indicated on the left.

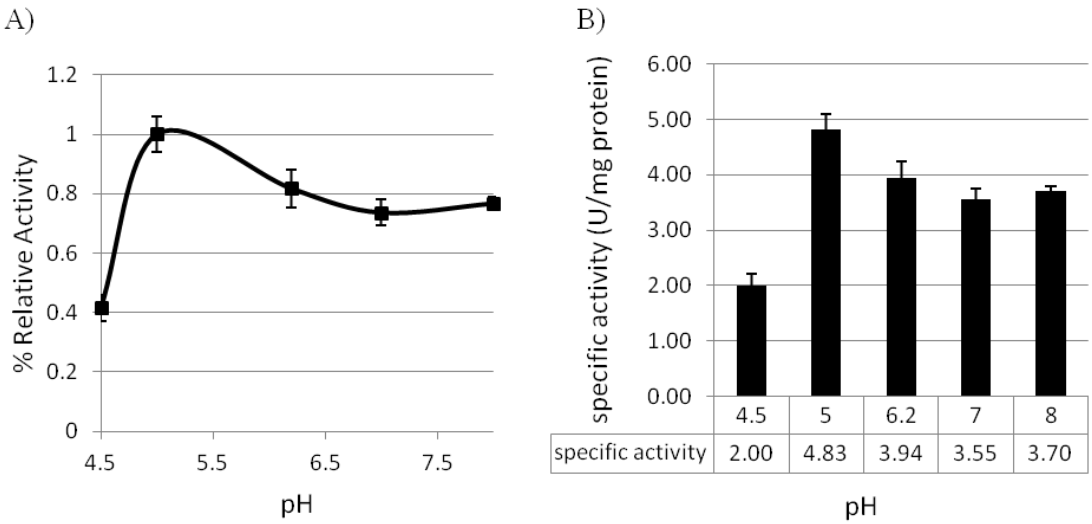


Figure 3.14: Dependence of Cel6A activity on pH. A) pH curve showing relative specific activity of Cel6A. B) The specific activity of Cel6A at various pH values. The specific activities of Cel6A using 0.1% medium viscosity CMC as substrate were determined in sodium acetate buffers, pH 4.5 and 5, and Tris-HCl buffer, pH 6.2, 7 and 8. Concentrations of reducing sugar product were determined using the PAHBAH reducing sugar assay. A) scatter plot; B) bar graph.

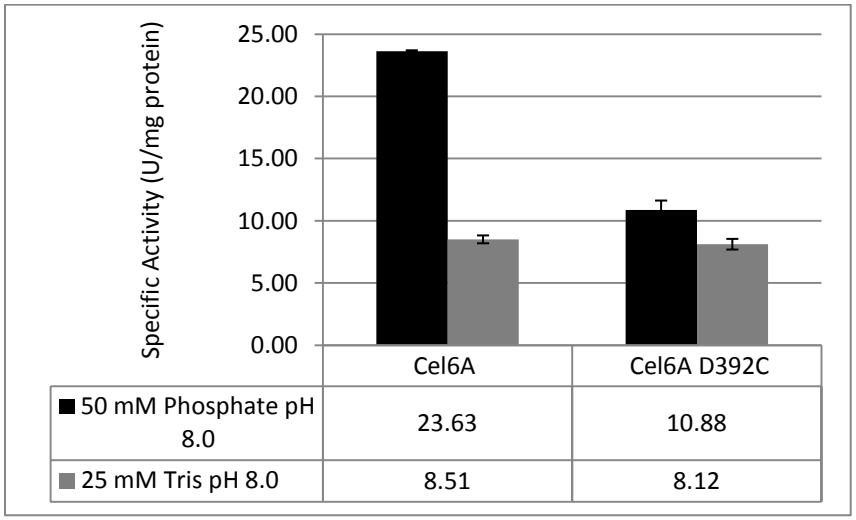


Figure 3.15: Effect of buffer composition on Cel6A and (D392C) Cel6A activity. The specific activity of the enzymes was measured using either 50 mM sodium phosphate buffer, pH 8.0 or 25 mM Tris.HCL buffer, pH 8.0using 1% medium viscosity CMC as substrate.

3.5 Characterization of (D392C) Cel6A

Previous studies conducted by former graduate student D. Cockburn indicated that the introduced sulfhydryl of Cys at position 392 was prone to auto-oxidation which resulted in the restoration of some catalytic activity. (D392C) Cel6A was assayed for cellulolytic activity against medium viscosity CMC and a specific activity of 10.9 U/mg protein at pH 8.0 was determined, a value of 46% of WT. This finding suggested that autooxidation of Cys392 had indeed occurred. LC-MS/MS analysis of the enzyme purified in the absence of reducing agents showed that the Cys392was indeed oxidized to sulfinic acid during purification (Table 3.1). Three separate analyses were performed to

Table 3.1: LC-MS analysis of tryptic peptides from (D392C) Cel6A that contain C392*.

M/Z	Charge	Intensity	Mass (Da)		Sequence[a]	Modification(s)
			Calculated	Predicted		
984.2	3+	10563	2949.578	2950.292	A	6: Sulfinic Acid 26: Met oxidation
973.53	3+	359722	2917.568	2918.302	A	26: Met oxidation
973.5	3+	114824	2917.478	2918.302	A	26: Met oxidation
1159.76	3+	7347	3476.258	3477.552	B	6: Sulfinic Acid 26: Met oxidation
1149.18	3+	20692	3444.518	3445.562	B	26: Met oxidation

[a] The sequence of the peptide were: (A) **LPGESCGACNGGPAAGQWWQEIALEMAR**
(B) **LPGESCGACNGGPAAGQWWQEIALMARW**
With both peptides, C392 is at position 6 and a Met residue (M412) is at position 26.

quantify the extent of this auto-oxidation. The first involved reaction with dimedone followed by Western immunoblot analysis using streptavidin-AP conjugate. The Western immunoblot was positive for both the sample washed previous to the addition of SDS sample buffer as well as the sample incubated directly with SDS sample buffer with no additional washing (Figure 3.16). However, the washed sample reacted less intensely to the antibody compared to the unwashed sample suggesting non-specific reaction of the probe. In support of this conclusion, incubation with the dimedone probe appeared to have no effect on the activity of (D392C) Cel6A (Figure 3.17).

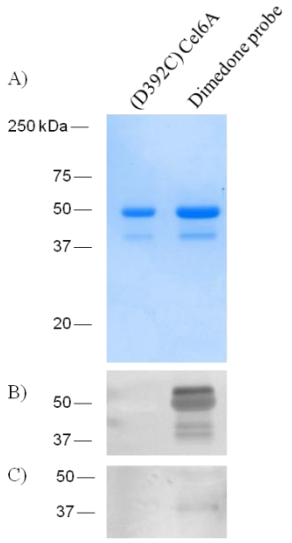


Figure 3.16: Dimedone-based probe analysis of (D392C) Cel6A. A) SDS-PAGE analysis of (D392C) Cel6A as a negative control and Cel6A D392C reacted with a dimedone probe. Electrophoretic conditions and staining were as described in the legend to Fig. 3.9. Western immuno blot analysis using a streptavidin-AP conjugate was performed on (D392C) Cel6A, which had been incubated with a dimedone-biotin conjugate and (B) not washed and (C) washed three times by ultrafiltration prior to electrophoresis. MW markers (kDa) are indicated on the left.

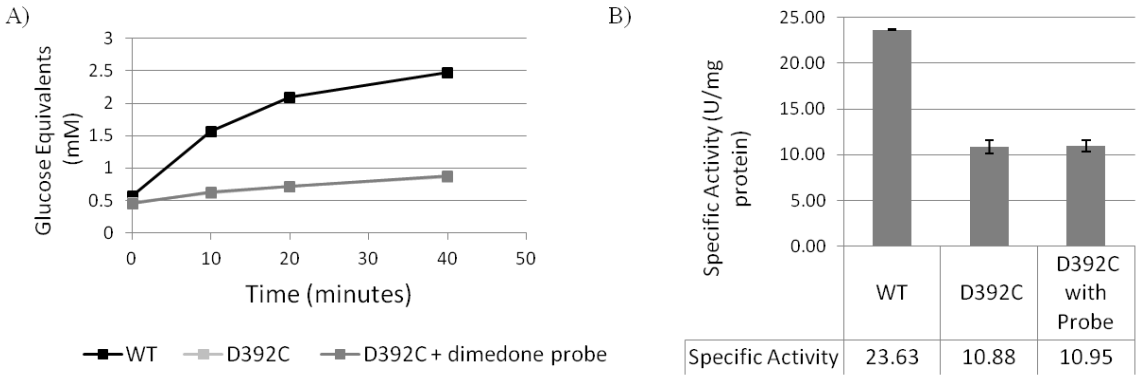


Figure 3.17: The effect of the dimedone probe on (D392C) Cel6A activity. A) Activity of Cel6A and (D392C) Cel6A incubated prior in the absence and presence of dimedone. The activity was measured in 50 mM sodium phosphate at a pH of 8.0 with 1% medium viscosity CMC as a substrate. B) The specific activities of Cel6A, Cel6A D392C and Ce l6A D392C after being incubated with dimedone.

An attempt was made to quantify the amount of free sulfhydryl present in (D392C) Cel6A by titration with the group-specific reagent DTNB. Titrations were performed on an enzyme preparation in 25 mM Tris-HCl, pH 8.0 immediately following its purification and overnight dialysis. However, measurements made in the absence and presence of 1 M guanidinium chloride failed to detect any free sulfhydryl group.

On the other hand, success with sulfhydryl quantification was achieved with DTDP titrations. Reaction of 24 nmol of freshly prepared (D392C) Cel6A resulted in the detection of 10.4 nmol of sulfhydryl. Therefore, it appeared that 56.7% of the introduced cysteine was oxidized. This extent of auto-oxidation increased to 75% after the enzyme was stored for three days post anion-exchange chromatography.

3.5.1 Protection From Auto-oxidation

Attempts to protect (D392C) Cel6A from auto-oxidation were made by including reducing agents in all buffers used to isolate, purify, and store the enzyme. Both Cel6A and (D392C) Cel6A were purified successfully using the standard purification methods with either 5 mM DTT or 1 mM TCEP added to the buffers. WT Cel6A was handled in this manner as a control for the presence of the reducing agents to determine the effect on enzymatic properties. The extent of auto-oxidation protection was determined by DTDP titration. (D392C) Cel6A isolated and purified in the presence of 1 mM TCEP was found to be protected from auto-oxidation but not completely as the enzyme possessed between 0.82 and 0.86 molar equivalents of free Cys (*i.e.*, 14 % to 18% oxidation).

This protection from auto-oxidation was reflected by the lower amount of cellulolytic activity associated with (D392C) Cel6A handled in the presence of reducing agent; it is expected that fully reduced Cys392 would not serve in a catalytic role . (D392C) Cel6A purified in 25 mM Tris-HCl buffer displayed 95% of the activity of the Cel6A purified under the same conditions. However, (D392C) Cel6A purified in the presence of TCEP had a specific activity of only 1.6 ±0.064 U/mg against CMC with a DS of 0.6 under the conditions employed compared to 8.1 ± 0.43 U/mg for the unprotected enzyme. Thus, this preparation of enzyme had only approx. 20% activity of the unprotected enzyme. (D392C) Cel6A purified in the presence of 5 mM DTT was even less active as it had only 15% residual activity compared to control enzyme handled and assayed under the same conditions (Figure 3.18). However, the inclusion of reducing

agents had a negative effect on the activity of the WT enzyme with only 41% residual activity in comparison to untreated enzyme (Figure 3.18).

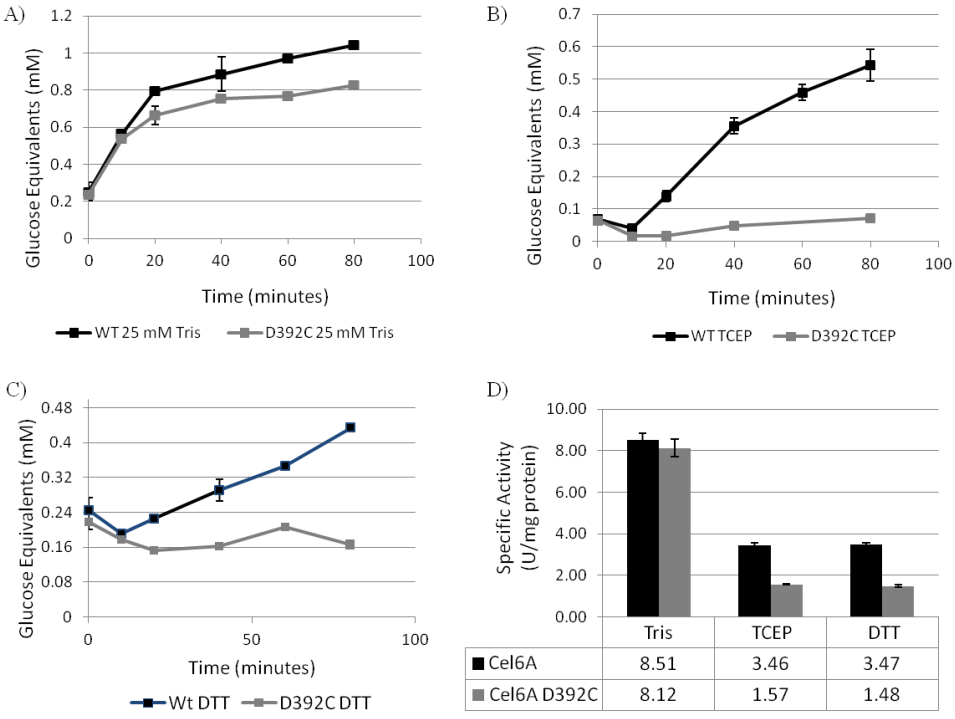


Figure 3.18: Effect of reducing agents on Cel6A and (D392C) Cel6A activity. A) Cel6A and (D392C) Cel6A activity assays were conducted with 1% medium viscosity CMC as substrate in 25 mM Tris-HCl, pH 8.0, and reducing sugar products were quantified using the PAHBAH assay. Separate standard curves were created for each buffer condition involving the reducing agents. Enzymes purified and assayed in the, A) absence of reducing agents; B) presence of 1 mM TCEP; and C) presence of 5 mM DTT. D) Summary of all data.

Furthermore, the exposure to reducing agents post purification to Cel6A and (D392C) Cel6A did not have the same inhibitory results as purification in the presence of reducing agents. There was a 24% drop in specific activity of the WT protein when exposed to TCEP and 41% drop when exposed to DTT. The same negative effect was not observed with (D392C) Cel6A with only 3% and 21% drop in activity after exposure to TCEP and DTT respectively (Figure 3.19). It was also demonstrated that (D392C) Cel6A activity could be restored with oxidation post purification in the presence of reducing agents, once those agents have been removed via dialysis (Figure 3.20).

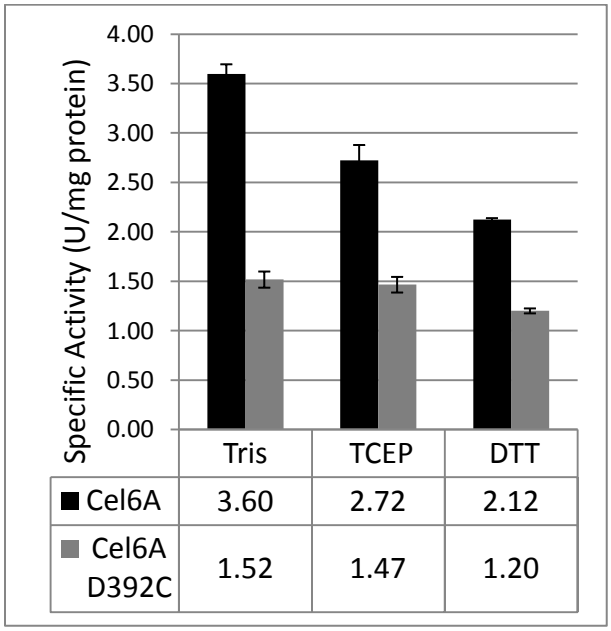


Figure 3.19: Specific activities of Cel6A and (D392C) Cel6A exposed to reducing agents after purification. Assay conditions as described in the legend to Figure 3.18.

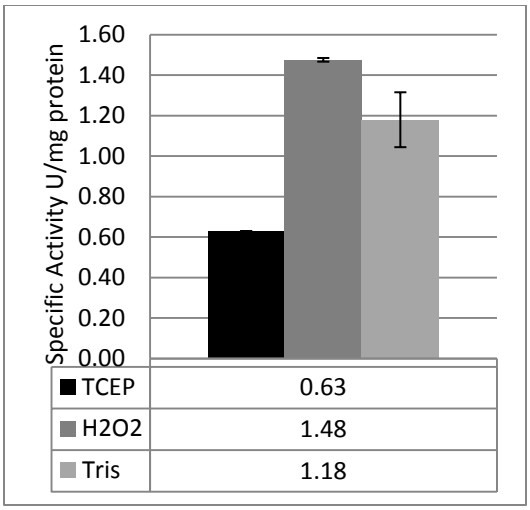


Figure 3.20: Restoration of catalytic activity with oxidation of Cys392. The specific activities of (D392C) Cel6A with 0.1% medium viscosity CMC as substrate in 25 mM Tris-HCl, pH 8.0 after purification in the presence of TCEP, and then incubated in the absence (Tris) and presence of H_2O_2 (1:8 ratio enzyme to reagent) for 24 hours at 4 °C.

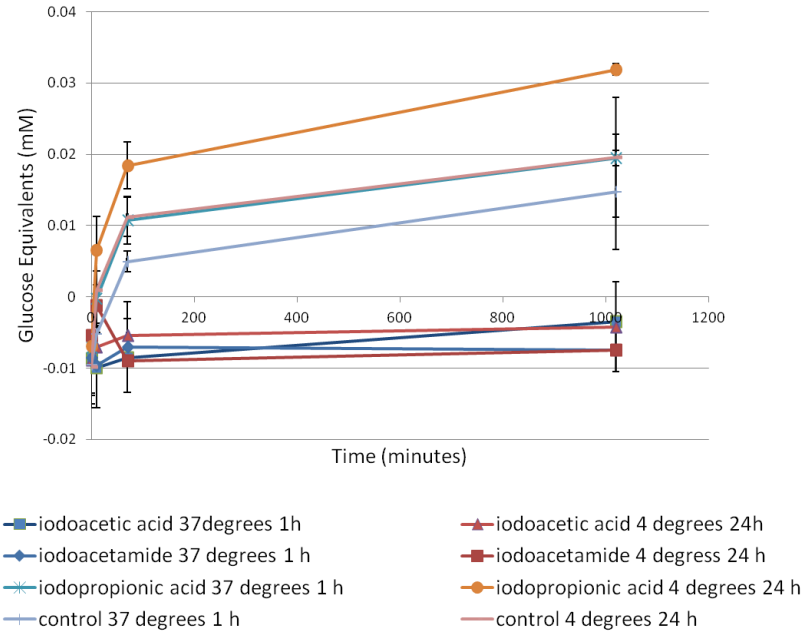
3.6 Cysteine Alkylation

Initial alkylation experiments involved the separate incubations of (D392C) Cel6A with iodopropionic acid, iodoacetic acid, and iodoacetamide at 37 °C for 1 hour, and at 4 °C for 24 hours. The specific activity of the control enzymes lacking reagent were 0.42 ± 0.04 U/mg after 1 hour at 37 °C and 0.5 ± 0.04 U/mg after 24 hours at 4 °C. As seen in Figure 3.21, no change in specific activity was observed with alkylation by either acetic acid or acetamide under either reaction condition but a 20% increase in activity relative to the control enzyme was seen with propionylation at 4 °C for 24 hours. Given this slight increase in activity observed with propionylation at 4 °C, a second experiment was conducted in which (D392C) Cel6A was incubated for extended periods of time with iodopropionic acid. The results, presented in Figure 3.22, indicate that incubation to 48

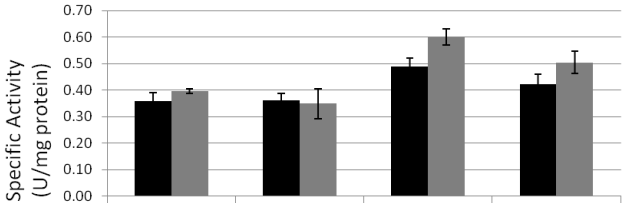
hours led to a further increase in recovery of specific activity. However, a similar increase in activity was observed with the control enzyme following incubation for the extended period such that propionylation led to an overall 29 % rate enhancement (Figure 3.21). LC-MS/MS analysis was performed on (D392C) Cel6A alkylated with propionic acid to confirm alkylation had indeed occurred. As seen in Table 3.2, propionyl groups were found on Cys392 but additional alkylation was observed on other Cys residues in the protein.

Finally, alkylation with iodobutyric acid was included in this second experiment. However, this treatment led to only a 7.7% increase in specific activity over that of control enzyme under the conditions employed (Figure 3.22).

A)



B)



	Acetic Acid	Acetamide	Propionic acid	Control
37 degrees celsius	0.36	0.36	0.49	0.42
4 degrees celsius	0.40	0.35	0.60	0.50

Figure 3.21: Effect of alkylation on (D392C) Cel6A on specific activity. (D392C)Cel6A was purified in the presence of 1 mM TCEP and then the reducing agent was removed by dialysis before incubation with a 100X molar excess of each alkylation agent as indicated for either 1 hour at 37 °C or 24 hours at 4°C. Specific activities were determined with 0.1% medium viscosity CMC as substrate in 25 mM Tris-HCl buffer, pH 8.0 after ultrafiltration of each enzyme preparation through 30 kDa cut off Amicon filters with four washes to remove any excess iodine or iodo-compounds. A) Time course of activities for enzymes treated as indicated; B) Summary of specific activity data.

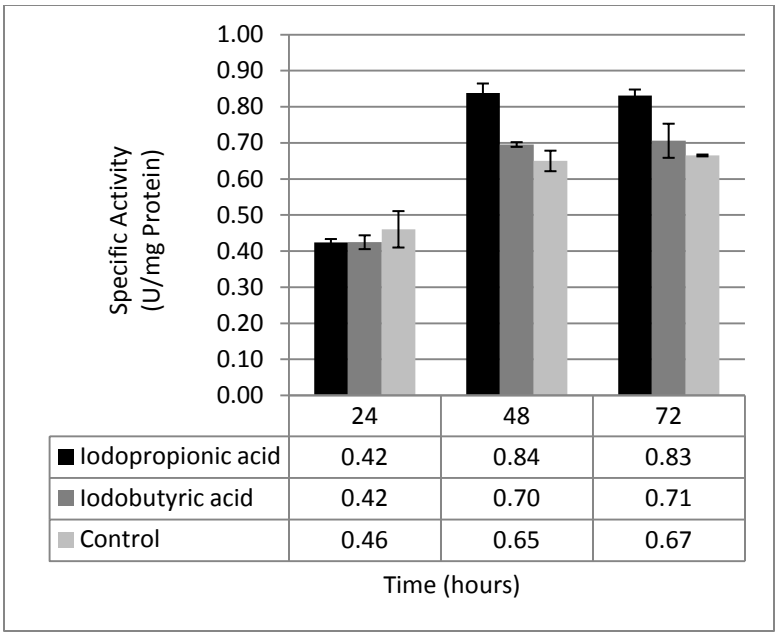


Figure 3.22: Effect on specific activity of extended incubation of (D392C) Cel6A alkylation with iodoalkyl acids. (D392C) Cel6A was purified in the presence of 1 mM TCEP and then the reducing agent was removed by dialysis before incubation with a 100X molar excess of each alkylation agent as indicated for either 24, 48, or 72 hours at 4°C. Specific activities were determined with 0.1% medium viscosity CMC as substrate in 25 mM Tris-HCl buffer, pH 8.0 after ultrafiltration of each enzyme preparation through 30 kDa cut off Amicon filters with four washes to remove any excess iodine or iodo-compounds.

Table 3.2: LC-MS/MS analysis of propionylated tryptic peptide.

N-Term	Ion	A	B	Y	C-Term	Ion
1	L	86.096	114.091	175.119	28	R
2	P	183.149	211.144	246.156	27	A
3	G	240.171	268.166	377.197	26	M
4	E	369.213	397.208	506.239	25	E
5	S	456.245	484.24	619.323	24	L
6	C*	632.283	660.278	690.36	23	A
7	G	689.305	717.3	803.444	22	I
8	A	760.342	788.337	932.487	21	E
9	C*	936.38	964.375	1060.546	20	Q
10	N	1050.423	1078.418	1246.625	19	W
11	G	1107.445	1135.439	1432.704	18	W
12	G	1164.466	1192.461	1560.763	17	Q
13	P	1261.519	1289.514	1617.784	16	G
14	A	1332.556	1360.551	1688.821	15	A
15	A	1403.593	1431.588	1759.858	14	A
16	G	1460.614	1488.609	1856.911	13	P
17	Q	1588.673	1616.668	1913.933	12	G
18	W	1774.752	1802.747	1970.954	11	G
19	W	1960.832	1988.827	2084.997	10	N
20	Q	2088.890	2116.885	2261.035	9	C*
21	E	2217.933	2245.928	2332.072	8	A
22	I	2331.017	2359.012	2389.094	7	G
23	A	2402.054	2430.049	2565.132	6	C*
24	L	2515.138	2543.133	2652.164	5	S
25	E	2644.181	2672.176	2781.207	4	E
26	M	2775.211	2803.216	2838.228	3	P
27	A	2846.258	2874.253	2935.281	2	P
28	R	3002.359	3030.354	3048.365	1	L

C* indicates cysteine with propionic acid modifications.

Chapter 4: Discussion

4.1 Production of recombinant Cel6A

Whereas a number of studies have been conducted in the past on recombinant forms of Cel6A from *C. fimi,* several problems regarding its overproduction in *E. coli* warranted addressing. Thus, a construct with the leader peptide removed was created. The decision to remove the LP was based on the consistent purification of two Cel6A products when a construct containing the entire WT sequence (pACDC-003) was used for expression. Both species purified are active. The size difference between the two Cel6A species purified must be from degradation on the N-terminus, since both species can be visualized using Western immunoblot. Therefore removing the non-essential LP from the N-terminus should have theoretically resulted in the purification of one stable, active species of Cel6A. Additionally, this accounts for any potential inefficient processing of the Gram positive LP in the Gram negative expression host. The removal of the His_6 tag was performed simultaneously in order to limit any effect on activity the tag may have.

During sequence analysis, a single point mutation was noted at nucleotide 786 where a thymine occurs in place of a cytosine. This replacement results in the codon GGC being translated as GGT. Fortuitously both codons encode for glycine, resulting in the same primary amino acid sequence. Both pACRW-06 and 07 encode this error indicating that the initial mutation occurred during previous cloning experiments and that it has been propagated through the use of pACDC-003 and pACDC-006 as templates.

However native purification of either Cel6A or (D392C) Cel6A was not achieved. The expression levels were too low, and Cel6A could not be isolated. Further trouble

shooting of the over-expression protocol in order to increase the amount of soluble expression should be completed. Additionally, during purification attempts isolates from each step should be tested for reducing sugar capabilities to determine the success of each step.

To address the solubility and expression levels the champion™ SUMO expression system was chosen to replace the untagged native purification. The champion™ SUMO expression system offers advantages to native purification due to its ability to increase the fusion protein's stability and the potential to increase soluble yields, all while generating a native, untagged protein (Satakarni and Curtis, 2011).

However, expression/production trials revealed that much of the Cel6A-SUMO protein was insoluble, despite the fusion to SUMO which in the past has been proven to increase solubility (Satakarni and Curtis, 2011). Expression at 15 °C over night with 1 mM IPTG resulted in the largest ratio of soluble to insoluble production. The level of expression at 15 °C was far below that observed at warmer temperatures, but the warmer temperatures resulted in a majority of insoluble protein (Figure 3.9). The insolubility of this form of Cel6A could also explain the low cellulolytic activity associated with the purification of the WT, untagged form. Despite the low yield, a pilot purification attempt of Cel6A-SUMO was completed to test the double IMAC purification and subsequent SUMO cleavage. This resulted in successful purification and cleavage of the SUMO tag. However the cleavage did not go to 100% completion (Figure 3.10). More extensive expression trials could be conducted in the future to trouble shoot the solubility issues coupled with testing different cleavage conditions. With more exhaustive optimization this construct could provide a good source of native Cel6A.

On the other hand each purification attempt of (D392C) Cel6A-SUMO resulted in unstable protein as extensive degradation during over-production was observed (Figure 3.11). Therefore native (D392C) Cel6A was not purified using the champion™ SUMO tag system

The SUMO protein tag is attached to the N-terminus of protein, whereas the construct previously utilized was tagged at the C-terminus. Initially during purification of the C-terminal His_6 tagged protein there was degradation occurring that resulted in two proteins identifiable *via* Western immunoblotting. In this fashion it was postulated that the degradation was occurring at the N-terminus. With the (D392C) Cel6A-SUMO construct there are multiple degradation products visualized on Western immunoblots. This infers that degradation is occurring on the C-terminal in this instance. The Cys replacement of the catalytic Asp is located near the C-terminal of the sequence and it is possible that the Cys substitution plays a role in the instability of the protein when there is no stabilizing tag present. Neither of the two (D392C) Cel6A constructs lacking a tag on the C-terminal (the construct with no His_6 tag and the N-terminally tagged SUMO protein) were successfully purified.

Therefore, due to time constraints, solubility issues, and stability issues it was decided that the best constructs to proceed with for this study were pACDC-003 and pACDC-006 containing the full length His_6-tagged Cel6A and (D392C) Cel6A respectively. Both Cel6A and (D392C) Cel6A were purified based on a previously designed protocol (Cockburn et al., 2010). The purifications of both Cel6A and Cel6A (D392C) Cel6A were successful and resulted in larger yield of protein than either the champion™ SUMO constructs and the native untagged constructs (Figure 3.12).

4.2 Characterization of Recombinant Cel6A and (D392C)Cel6A

The initial investigation of the dependence of activity on pH indicated that the pH optimum for both Cel6A and the cysteine catalytic variant was at 5.0 (Figure 3.15). This is contrary to previous findings that indicated the pH optimum for Cel6A was around pH 7 (Cockburn et al., 2010). The difference between the previous and current studies was the buffers used in activity assays, and more specifically the use of phosphate and Tris-HCl buffers, respectively, for the pH range 6 to 8. This prompted an investigation of the effect of Tris on Cel6A activity and it was discovered that there are a number of reported cases in the literature of Tris being an enzyme inhibitor. Of these cases, a few have concerned different glycoside hydrolases. For example, Tris was shown to be a pH dependant inhibitor of a β-mannanse (Burke and Khan, 2000). The pH dependence is likely due to the protonation state of the Tris and its interaction with the catalytic base and a pivotal His residue at a pH above 6 (Burke and Khan, 2000). Cel6A contains three His residues and one of these, His321, is predicted to be in the active cleft (Figure 4.1). The position of this His residue is similar to that within the β-mannanse noted above (Burke and Khan, 2000) and so it is tempting to speculate that association of Tris with His321 and the Asp 392 caused the observed inhibition of activity. On the other hand, Bis-Tris propane, which is two conjoined Tris molecules, has been shown to inhibit a β-D-xylosidase by compromising its catalytic acid and catalytic base (Brunzelle et al., 2008). Hence, it is equally possible that Tris may act similarly with Cel6A.

The dramatic decrease in Cel6A activity in the presence of 25 mM Tris compared to 50 mM phosphate can be attributed to the inhibitory properties of the buffer. Given this, it was surprising that the same dramatic decrease is not observed with (D392C)

Cel6A. There was a 26 % decrease in activity in the presence of Tris, compared to the 64 % decrease in activity of WT Cel6A. Both the phosphate and Tris activity assays were completed at a pH of 8.0, well above the reported optimum of 5.0 for (D392C) Cel6A. The presence of Cys-sulfinic acid in place of the physiological Asp

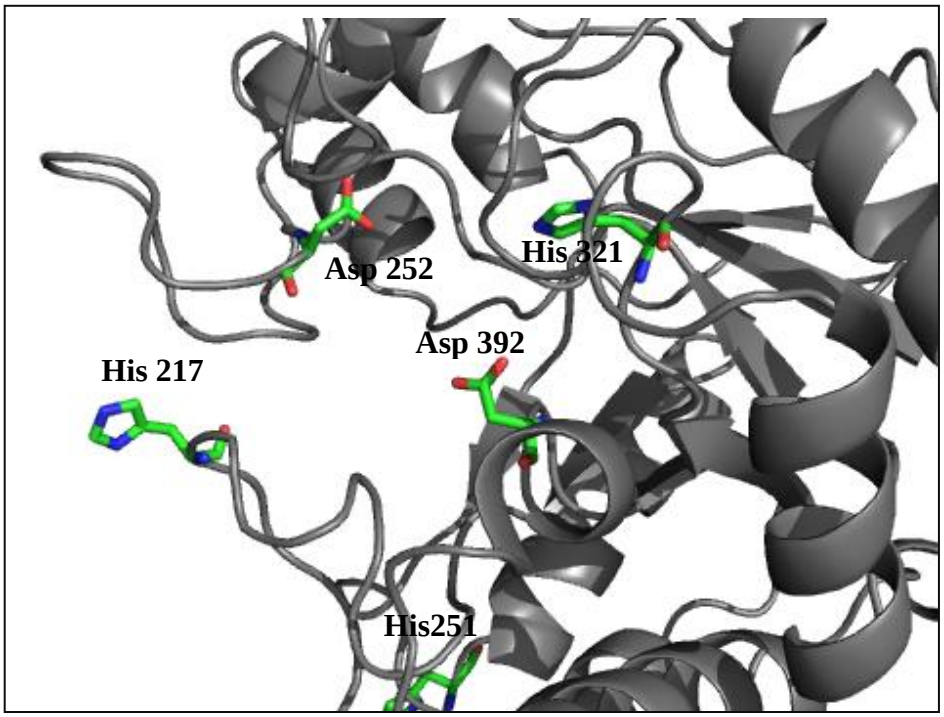


Figure 4.1: His locations in predicted active site region of Cel6A. Catalytic site of Cel6A with all three native His residues shown including His321 which is located deep within the active site.

residue already results in a dramatic decrease of activity, especially at a pH of 8 and it appears that the inclusion of Tris does not compound this inhibition. This could be due to the difference in pKa's of Cys-sulfinic acid relative to Asp and the possible influence the two residues have on other residues of the active site. The pKa of the introduced sulfinate is expected to be approximately 1.5 compared to 3.8 of the carboxyl group in an Asp residue. Thus, it is not unreasonable to expect that the greater than 2 units lower pKa value of Cys-sulfinate would have a significant effect on the local environment of the

active site residues to in turn influence the inhibitory effect of Tris, which may be interacting with it.

4.2.1 Cys392 Characterization

The rescued activity of (D392C) Cel6A is imparted by oxidation that can occur spontaneously if precautions are not made. In order to identify the oxidation state of the introduced Cys residue that results from auto-oxidation, a number of analyses were performed. This characterization of the Cys oxidation state was crucial for determining the protein's viability for the planned chemical modification experiments. Since any oxidation past Cys-sulfenic acid occurring at the introduced Cys would preclude it from chemical alkylation, considering that only oxidation to Cys-sulfenic acid is known to be reversible (Paget and Buttner, 2003).

A dimedone based probe was used to assay for the presence of Cys-sulfenic acid. Cys-sulfenic acid is the first oxidation state of Cys and is one of the potential outcomes of the oxidation that occurs during purification. A strong reaction was observed in the protein that was not washed prior to the addition of SDS sample buffer, however there was a weaker reaction observed when protein sample was washed to remove excess probe prior to the addition of the SDS sample buffer. Additional activity assays indicated that incubation with the dimedone probe has no effect on the protein's activity (Figure 3.19). Therefore, it would appear that a small proportion of the "catalytic" Cys may be oxidized to Cys-sulfenic acid but it is likely that the majority is being oxidized further than sulfenic acid. Alternatively, it is possible that the retention of the dimedone probe was non-specific and that all of the auto-oxidation results in the formation of the higher state of sulfinic acid. Evidence for the latter scenario was provided by the LC-MS/MS

analysis of the oxidized protein which revealed only the presence of Cys-sulfinic acid; no Cys-sulfenic acid at position 392 was detected. Along with Cys-sulfinic acid, un-oxidized Cys at this position was also detected which corroborates with the partial oxidation determined by DTDP titration experiments.

The first attempt at quantifying the level of free sulfhydryl was conducted with the commonly used reagent DNTB. However, the amount of free Cys present in the samples tested was at the detection limit of the assay without the addition of a mediator protein (Riener et al., 2002). The inclusion of guanidine-HCL in the protein sample to perturb its structure and thereby possibly better expose the Cys did not help with detection. However, that titrations with DTDP, a sulfhydryl reagent with nM sensitivity (Reiner et al. 2002), did provide informative data suggests that reagent accessibility was not the issue with DTNB. Finally, the DTDP titrations revealed that (D392C) Cel6A is prone to continued auto-oxidation after its manipulation during purification procedures if precaution is not taken to maintain a reducing environment.

In order to maintain a reducing environment reducing agents were included in the purification buffers. Both Cel6A and (D392C) Cel6A were successfully purified in the presence of 1 mM TCEP and 5 mM DTT. The inclusion of TCEP had a negative impact on anion exchange chromatography. Generally at a pH of 9 Cel6A flows through the anion exchange resin and is collected in the FT, however when TCEP is present it appears to retard this process. The same amount of protein takes approximately 30 more minutes to pass through the column. TCEP has a positive charge and it is likely that this is interfering with the protein/resin interaction. This necessitated purification using a double IMAC elution process.

The activity of both Cel6A and (D392C) Cel6A was reduced by the inclusion of reducing agents in the purification buffers as seen in Figure 3.18. The reduced activity of Cel6A is likely due to the disruption of the disulfide bonds that occur within the protein's native conformation, and the use of Tris-HCl as the purification and activity buffer. The decrease in activity of (D392C) Cel6A is not proportional to the decrease in activity of Cel6A. This indicates that the reducing agents at least partially prevent spontaneous oxidation of the introduced cysteine. This is reflected in the reduction of activity that would normally have been imparted by the Cys-sulfinic acid. Therefore, the inclusion of reducing agents during the purification process prevents some oxidation from occurring. However they also disrupt the native disulfide bonds. This in combination with the suggested Tris-HCl buffer for use with TCEP results in the drastic decrease in specific activity observed in (D392C) Cel6A.

Both Cel6A and (D392C) Cel6A were purified in the absence of reducing agent and then exposed to either 1 mM TCEP or 5 mM DTT before undergoing activity assays (Figure 3.19). The addition of a reducing agent post purification had a significant effect on the activity of the WT protein however had little effect on the activity of (D392C) Cel6A, with DTT having a larger negative effect than TCEP. This indicates that including the reducing agent during the purification process is necessary to reduce the amount of cysteine oxidation, further demonstrating that the cysteine is oxidized past sulfenic acid and is therefore irreversibly oxidized and that the reduction of activity of (D392C) Cel6A is in fact due to the prevention of auto-oxidation during purification, rather than the presence of Tris.

TCEP was determined to be a better candidate for the prevention of oxidation because it is a more potent reducing agent that is more stable than DTT. The structure of TCEP is also amenable to the downstream modification of cysteines since, unlike DTT it does not contain thiols. Overall the purification of Cel6A in the presence of a reducing agent is not ideal due to the important disulfide bonds within its structure. These bonds are not essential for catalytic activity, but their disruption has a significant impact on the efficiency of catalysis.

4.3 Rescuing Catalytic Activity

The activity of (D392C) Cel6A can be rescued after purification in the presence of TCEP by incubation with H_2O_2. This indicates that the introduced cysteine remains amenable to modification once the reducing agents are removed, implying it will be available for chemical modifications. Therefore chemical alkylation was attempted to rescue catalytic activity. The initial alkylation trials showed that the protein incubated with the propionic acid had a marginally higher specific activity than the control (Figures 3.21 and 3.22). However the activity displayed was only about one half of the specific activity of (D392C) Cel6A when oxidized. In each case, the protein used in the reaction was previously purified under reducing conditions to prevent oxidation. The control represents activity restored by introduced oxidation during the incubation and handling of the protein during the alkylation reactions.

The specific activity of the protein when incubated with either iodoacetic acid or iodoacetamide was significantly lower than the specific activity of the protein control and the protein incubated with iodopropionic acid. This indicates that in the case of iodoacetic

acid and iodoacetamide the compounds were able to react with the catalytic cysteine and preclude any oxidation from occurring; as such there was less activity than the control. Iodoacetic acid provides a carboxyl group that could rescue activity however that was not observed. This implies that the addition of an acetic acid does not place the carboxyl group appropriately for efficient catalysis; the gap between the two acidic groups would be too far for a single-displacement mechanism and too short for a double-displacement reaction. The modification with iodoacetamide was performed as a negative control, as an introduced amide group would not provide a potential catalytic moiety.

The enzyme alkylated with propionic acid displayed a slightly higher specific activity than the control. These experiments hint at the possibility that activity could be restored with alkylation, however much needs to be optimized and clarified. The minimal level of rescued activity suggests that the reaction is either not efficient, or is not going to completion or that the reaction interferes with the stability of the protein. Whether or not the reactions were going to completion needs to be investigated in order to determine the efficacy of the reaction conditions. Alkylation reactions can be pH, temperature and time dependant (Hale et al., 2004). Therefore these conditions should be tested, wherein the completeness of the reaction can be monitored by MALDI-TOF analysis as described by Hale et al., 2004. This would ensure that the modifications are in fact taking place and that the negative modifying group is not somehow being repelled or in excluded from the catalytic site during the reaction. The propionic acid modification was identified on the catalytic cysteine through LC-MS/MS. However the sample analyzed was one in which the protein sample was not washed prior to the addition of SDS sample buffer. Therefore the propionic addition could have occurred either before or after the addition of the SDS

sample buffer. In combination with Mass spec analysis stability curves can be generated of the enzyme (Rees and Robertson, 2001). This will give insight into whether the low activity is caused by an incomplete reaction, stability issues, or inefficient catalysis.

4.4 Conclusions

The use of Tris-HCl buffer in activity assays is detrimental due to enzyme inhibition. Therefore many of the activity assays could be repeated with phosphate buffer to obtain a more accurate depiction of the effects each experiment had on enzyme activity.

Despite the inhibition of Tris there was still measurable activity of both Cel6A and (D392C) Cel6A. The inclusion of reducing agents during the purification process is detrimental to the activity of Cel6A, both due to the buffer of choice; TCEP is more stable in Tris-HCl buffer, and the disruption of native disulfide bonds. However the activity of (D392C) Cel6A can be rescued by oxidation, after purification in the presence of reducing agents.

Overall the alkylation experiments suggest that it is possible to rescue protein activity by chemical modification, specifically with the addition of a propionic group. It is yet to be determined whether or not the protein's reaction mechanism is altered in the process of restoring activity. However the use of Cel6A from *C. fimi* is not an ideal model protein. The proposal of changing the reaction mechanism of a GH using chemical alkylation would be simplified by using a protein that does not contain a number of structurally important disulfide bonds. For example, endo-β-1,4-galactosidase (EC 3.2.1.21) is an inverting GH from family 98 that does not contain any Cys and its

structure has been solved (PDB: 2WMI). The family GH15 Glucodextranase (EC 3.2.1.70) is also an inverting GH that does not contain any Cys residues and its structure is known (PDB: 1UG9). These two enzymes could be potential candidates for future catalytic chemical modification experiments. Using either of these enzymes would simplify any attempts to utilize an introduced Cys for chemical modifications of the reaction mechanism.

References

Al-Zuhair, S. (2008). The effect of crystallinity of cellulose on the rate of reducing sugars production by heterogeneous enzymatic hydrolysis. Bioresour. Technol. *99*, 4078–4085.

Boerjan, W., Ralph, J., and Baucher, M. (2003). Lignin biosynthesis. Annu. Rev. Plant Biol. *54*, 519–546.

Brunzelle, J.S., Jordan, D.B., McCaslin, D.R., Olczak, A., and Wawrzak, Z. (2008). Structure of the two-subsite beta-d-xylosidase from Selenomonas ruminantium in complex with 1,3-bis[tris(hydroxymethyl)methylamino]propane. Arch. Biochem. Biophys. *474*, 157–166.

Bu, L., Crowley, M.F., Himmel, M.E., and Beckham, G.T. (2013). Computational investigation of the pH dependence of loop flexibility and catalytic function in glycoside hydrolases. J. Biol. Chem. *288*, 12175–12186.

Burke, M.D., and Khan, S. a (2000). Triggered enzymatic degradation of a water-soluble polymer solution using a novel inhibitor. Biomacromolecules *1*, 688–695.

Campbell, I.D., and Spitzfaden, C. (1994). Building proteins with fibronectin type III modules. Structure *2*, 333–337.

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

Cardona, C. a, and Sánchez, O.J. (2007). Fuel ethanol production: process design trends and integration opportunities. Bioresour. Technol. *98*, 2415–2457.

Ceroni, A., Passerini, A., Vullo, A., and Frasconi, P. (2006). DISULFIND: a disulfide bonding state and cysteine connectivity prediction server. Nucleic Acids Res. *34*, W177–81.

Cockburn, D.W., and Clarke, A.J. (2011). Modulating the pH-activity profile of cellulase A from *Cellulomonas fimi* by replacement of surface residues. Protein Eng. Des. Sel. *24*, 429–437.

Cockburn, D.W., Vandenende, C., and Clarke, A.J. (2010). Modulating the pH-activity profile of cellulase by substitution: replacing the general base catalyst aspartate with cysteinesulfinate in cellulase A from *Cellulomonas fimi*. Biochem. *49*, 2042–2050.

Davies, G.J., Ducros, V., Lewis, R.J., Borchert, T.V., and Schulein, M. (1997). Oligosaccharide specificity of a family 7 endoglucanase: insertion of potential sugar-binding subsites. J. Biotechnol. *57*, 91–100.

Davis, W. ApE: A Plasmid Editor.

Delaunay, A., Pflieger, D., Barrault, M.B., Vinh, J., and Toledano, M.B. (2002). A thiol peroxidase is an H2O2 receptor and redox-transducer in gene activation. Cell *111*, 471–481.

Doi, R.H., and Kosugi, A. (2004). Cellulosomes: plant-cell-wall-degrading enzyme complexes. Nat. Rev. Microbiol. *2*, 541–551.

Fan, L., Lee, Y.H., and Beardmore, D. (1980). Major chemical and physical features of cellulosic materials as substrates for enzymatic hydrolysis. Adv. Biochem. Eng. *14*, 101–117.

Finkel, T. (2000). Redox-dependent signal transduction. FEBS Lett. *476*, 52–54.

Fushinobu, S. (2013). Metalloproteins: A new face for biomass breakdown. Nat. Chem. Biol. *10*, 88–89.

Gasteiger, E. (2003). ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res. *31*, 3784–3788.

Gilkes, N.R., Langsford, M.L., Kilburn, D.G., Miller, R.C., and Warren, R. a (1984). Mode of action and substrate specificities of cellulases from cloned bacterial genes. J. Biol. Chem. *259*, 10455–10459.

Gowen, C.M., and Fong, S.S. (2010). Exploring biodiversity for cellulosic biofuel production. Chem. Biodivers. *7*, 1086–1097.

Hale, J.E., Butler, J.P., Gelfanova, V., You, J.-S., and Knierman, M.D. (2004). A simplified procedure for the reduction and alkylation of cysteine residues in proteins prior to proteolytic digestion and mass spectral analysis. Anal. Biochem. *333*, 174–181.

Hamman, M., Zhang, T., Hendrich, S., and Thomas, J. (2002). Quantiation of protein sulfinic and sulfonic acid, irreversibly oxidized protein cysteine sites in cellular proteins. Methods Enzymol. *348*, 146–156.

Holtzapple, M., and Cognata, M. (1990). Inhibition of Trichoderma reesei cellulase by sugars and solvents. Biotechnol. Bioeng. *36*, 275–287.

Horn, S.J., Vaaje-Kolstad, G., Westereng, B., and Eijsink, V.G. (2012). Novel enzymes for the degradation of cellulose. Biotechnol. Biofuels *5*, 45.

Jäger, S., and Kiss, L. (2005). Investigation of the active site of the extracellular β-D-glucosidase from Aspergillus carbonarius. World J. Microbiol. Biotechnol. *21*, 337–343.

Kelley, L.A., and Sternberg, M.J.E. (2009). Protein structure prediction on the Web: a case study using the Phyre server. Nat. Protoc. *4*, 363–371.

Kim, T., Mullaney, E.J., Porres, J.M., Roneker, K.R., Crowe, S., Rice, S., Ko, T., Ullah, A.H.J., Daly, C.B., Welch, R., et al. (2006). Shifting the pH profile of *Aspergillus niger* PhyA phytase to match the stomach pH enhances its effectiveness as an animal feed additive. Appl. Environ. Microb. *72*, 4397–4403.

Kiss, T., Erdei, A., and Kiss, L. (2002). Investigation of the active site of the extracellular beta-D-xylosidase from Aspergillus carbonarius. Arch. Biochem. Biophys. *399*, 188–194.

Klemm, D., Heublein, B., Fink, H.-P., and Bohn, A. (2005). Cellulose: fascinating biopolymer and sustainable raw material. Angew. Chem. Int. Ed. Engl. *44*, 3358–3393.

Klomsiri, C., Nelson, K.J., Bechtold, E., Soito, L., Johnson, L.C., Lowther, W.T., Ryu, S.-E., King, S.B., Furdui, C.M., and Poole, L.B. (2010). Use of dimedone-based chemical probes for sulfenic acid detection evaluation of conditions affecting probe incorporation into redox-sensitive proteins. (Elsevier Inc.).

Koshland, D. (1952). Effect of Catalysts on the Hydrolysis of Acetyl Phosphate. Nucleophilic Displacement Mechanisms in Enzymatic Reactions. J. Am. Chem. Soc. *0*, 2286–2292.

Koziol, M. (1981). An evaluation of the alkaline p-hydroxybenzoic acid hydrazide procedure for the determination of reducing sugars. Anal. Chim. Acta *128*, 195–205.

Kuroki, R., and Weaver, L. (1995). Structure-based design of a lysozyme with altered catalytic activity. Nat. Struct. Biol. *2*, 1007–1011.

Lever, M. (1972). A new reaction for colorimetric determination of carbohydrates. Anal. Biochem. *47*, 273–279.

Lynd, L.R., Weimer, P.J., Van Zyl, W.H., and Pretorius, I.S. (2002). Microbial cellulose utilization: fundamentals and biotechnology. Microbiol. Mol. Biol. Rev. *66*, 506–577.

Makde, R.D., Dikshit, K., and Kumar, V. (2006). Protein engineering of class-A non-specific acid phosphatase (PhoN) of *Salmonella typhimurium* : modulation of the pH-activity profile. Biomol. Eng. *23*, 247–251.

Mayer, C., Zechel, D.L., Reid, S.P., Warren, R. a, and Withers, S.G. (2000). The E358S mutant of *Agrobacterium* sp. beta-glucosidase is a greatly improved glycosynthase. FEBS Lett. *466*, 40–44.

Nakamura, A., Fukumori, F., and Horinouchi, S. (1991). Construction and characterization of the chimeric enzymes between the *Bacillus subtilis* cellulase and an alkalophilic *Bacillus* cellulase. J. Biol. Chem. *266*, 1579–1583.

Nickbarg, E.B., Davenport, R.C., Petsko, G.A., and Knowles, J.R. (1988). Triosephosphate isomerase: removal of a putatively electrophilic histidine residue results in a subtle change in catalytic mechanism. Biochem. *27*, 5948–5960.

Nielsen, J.E., Beier, L., Otzen, D., Borchert, T.V., Frantzen, H.B., Andersen, K.V., and Svendsen, A. (1999). Electrostatics in the active site of an α-amylase. E. J. Biochem. *264*, 816–824.

Nishiyama, Y., Sugiyama, J., Chanzy, H., and Langan, P. (2003). Crystal structure and hydrogen bonding system in cellulose I(alpha) from synchrotron X-ray and neutron fiber diffraction. J. Am. Chem. Soc. *125*, 14300–14306.

Otzen, D.E., Christiansen, L., and Schülein, M. (1999). A comparative study of the unfolding of the endoglucanase Cel45 from *Humicola insolens* in denaturant and surfactant. Prot. Sci. *8*, 1878–1887.

Pages, S., Belaich, A., and Belaich, J. (1997). Species-specificity of the cohesin-dockerin interaction between *Clostridium thermocellum* and *Clostridium cellulolyticum* : prediction of specificity determinants of. Proteins *29*, 517–527.

Paget, M.S.B., and Buttner, M.J. (2003). Thiol-based regulatory switches. Annu. Rev. Genet. *37*, 91–121.

Park, J.S., Hitomi, J., Horinouchi, S., and Beppu, T. (1993). Identification of two amino acids contributing the high enzyme activity in the alkaline pH range of an alkaline endoglucanase from a *Bacillus* sp. Prot. Eng. *6*, 921–926.

Parsiegla, G., Juy, M., Reverbel-Leroy, C., Tardif, C., Belaïch, J.P., Driguez, H., and Haser, R. (1998). The crystal structure of the processive endocellulase CelF of *Clostridium cellulolyticum* in complex with a thiooligosaccharide inhibitor at 2.0 A resolution. EMBO J. *17*, 5551–5562.

Peterson TN, Brunak S, Von Heijne G, N. h. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods *8*, 785–786.

Poole, L.B., Karplus, P.A., and Claiborne, A. (2004). Protein sulfenic acids in redox signaling. Annu. Rev. Pharmacol. Toxicol. *44*, 325–347.

Ranby, B. (1969). Recent progress on the structure and morphology of cellulose. Adv. Chem. Ser. *95*, 139–151.

Reddie, K.G., and Carroll, K.S. (2008). Expanding the functional diversity of proteins through cysteine oxidation. Curr. Opin. Chem. Biol. *12*, 746–754.

Rees, D.C., and Robertson, A.D. (2001). thermostability of proteins. Prot. Sci. *10*, 1187–1194.

Rhee, S.G., Bae, Y.S., Lee, S.-R., and Kwon, J. (2000). Hydrogen Peroxide: A Key Messenger That Modulates Protein Phosphorylation Through Cysteine Oxidation. Sci. Signal. *2000*, pe1–pe1.

Riener, C., Kada, G., and Gruber, H. (2002). Quick measurement of protein sulfhydryls with Ellman's reagent and with 4,4'-dithidipyridine. Anal. Bioanal. Chem *373*, 266–276.

Saha, B.C. (2003). Hemicellulose bioconversion. J. Ind. Microbiol. Biotechnol. *30*, 279–291.

Sambrook, J., and Russell, D.W. (2001). Molecular cloning: a laboratory manual (Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press).

Satakarni, M., and Curtis, R. (2011). Production of recombinant peptides as fusions with SUMO. Protein Expr. Purif. *78*, 113–119.

Schilling, B., Yoo, C.B., Collins, C.J., and Gibson, B.W. (2004). Determining cysteine oxidation status using differential alkylation. Int. J. Mass Spectrom. *236*, 117–127.

Schülein, M. (2000). Protein engineering of cellulases. Biochim. Biophys. Acta. *1543*, 239–252.

Solomon, B.D. (2010). Biofuels and sustainability. Ann. N. Y. Acad. Sci. *1185*, 119–134.

Stackebrandt, E., Schumann, P., and Prauser, H. (2006). The Prokaryotes. 983–1001.

Wang, J., Quirk, A., Lipkowski, J., Dutcher, J.R., Hill, C., Mark, A., and Clarke, A.J. (2012). Real-time observation of the swelling and hydrolysis of a single crystalline cellulose fiber catalyzed by cellulase 7B from Trichoderma reesei. Langmuir *28*, 9664–9672.

Wang, L., Zhang, Y., and Gao, P. (2008). A novel function for the cellulose binding module of cellobiohydrolase I. Sci China C Life Sci. *51*, 620–629.

Wang, Q., Graham, R., Trimbur, D., Warren, R., and Withers, S. (1994). Changing enzymic reaction mechanisms by mutagenesis: conversion of a retaining glucosidase to an inverting enzyme. J. Am. Chem. Soc. *116*, 11594–11595.

Weeks, S.D., Drinker, M., and Loll, P.J. (2007). Ligation independent cloning vectors for expression of SUMO fusions. Protein Expr. Purif. *53*, 40–50.

Wong, W.K., Gerhard, B., Guo, Z.M., Kilburn, D.G., Warren, A. J., and Miller, R.C. (1986). Characterization and structure of an endoglucanase gene cenA of *Cellulomonas fimi*. Gene *44*, 315–324.

Wright, K., and Viola, R. (1998). Evaluation of methods for the quantitation of cysteines in proteins. Anal. Biochem. *265*, 8–14.

Wu, M., Nerinckx, W., Piens, K., Ishida, T., Hansson, H., Sandgren, M., and Ståhlberg, J. (2013). Rational design, synthesis, evaluation and enzyme-substrate structures of improved fluorogenic substrates for family 6 glycoside hydrolases. FEBS J. *280*, 184–198.

Zechel, D.L., and Withers, S.G. (2000). Glycosidase mechanisms: anatomy of a finely tuned catalyst. Acc. Chem. Res. *33*, 11–18.

Zhang, S., Barr, B.K., and Wilson, D.B. (2000). Effects of noncatalytic residue mutations on substrate specificity and ligand binding of *Thermobifida fusca* endocellulase cel6A. E. J. Biochem. *267*, 244–252.

Zhang, Y.-H.P., and Lynd, L.R. (2004). Toward an aggregated understanding of enzymatic hydrolysis of cellulose: noncomplexed cellulase systems. Biotechnol. Bioeng. *88*, 797–824.

Zinoviev, S., Müller-Langer, F., Das, P., Bertero, N., Fornasiero, P., Kaltschmitt, M., Centi, G., and Miertus, S. (2010). Next-generation biofuels: Survey of emerging technologies and sustainability issues. ChemSusChem *3*, 1106–1133.

Appendix A: Standard Curves

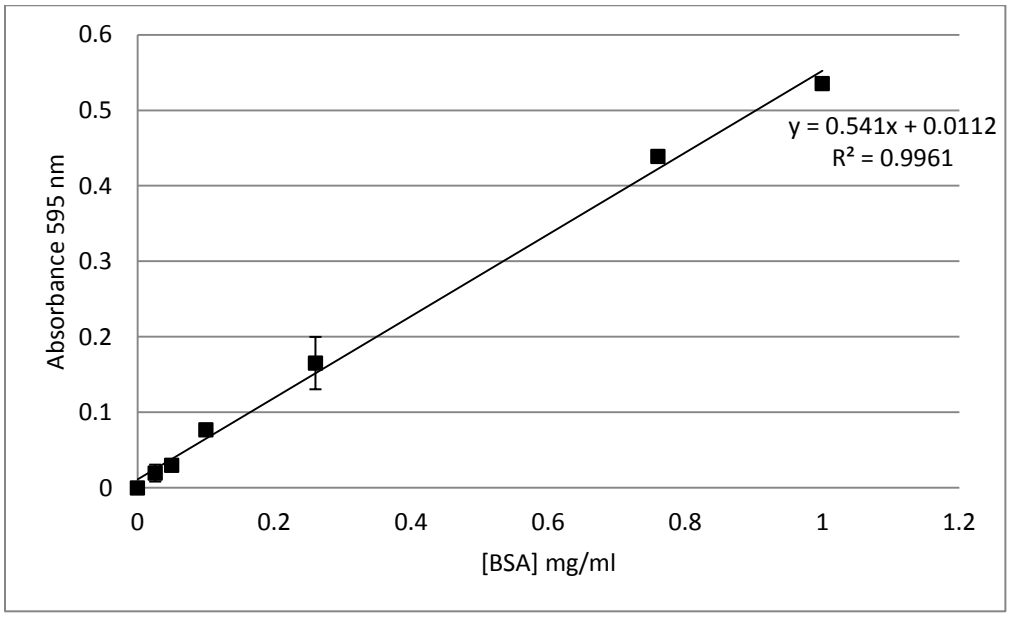


Figure 5.1: Example BCA standard curve. Standard curve generated using BSA standards in 25 mM Tris-HCL buffer.

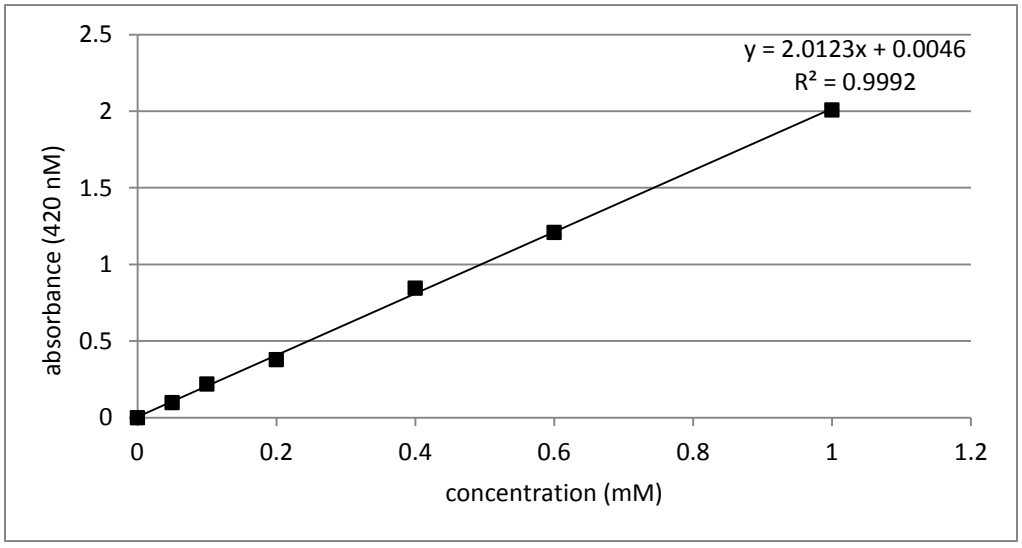


Figure 5.2: Example glucose standard curve. This particular curve was generated in the presence of 1 mM DTT.

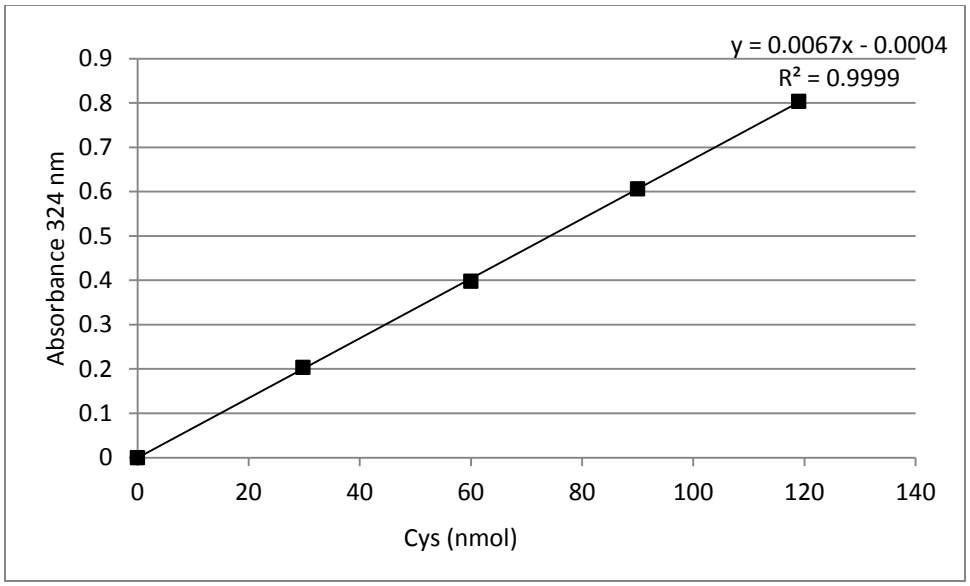


Figure 5.4: Example Cys standard curve. This curve was generated using DTDP and nmol quantities of L-Cys.