

FERULATE CROSS-LINKING AND CONJUGATIONS: THE ROLE OF FERULATE
IN THE GRASS CELL WALL AND SPECIALIZED METABOLISM

By

Emily Frankman

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Biochemistry and Molecular Biology—Doctor of Philosophy

2018

ABSTRACT

FERULATE CROSS-LINKING AND CONJUGATIONS: THE ROLE OF FERULATE IN THE GRASS CELL WALL AND SPECIALIZED METABOLISM

By

Emily Frankman

Members of the *Poaceae* include many of the most economically important plants, such as maize, rice, wheat, and sorghum. These grasses not only account for approximately half of the human diet, they are also a likely source of renewable bioenergy. An impediment to using biomass for the production of biofuels and other products is the recalcitrance of the plant cell to enzymatic digestion. The cell wall is vital for plant strength in that it allows the plant to grow upright, and acts as protection against pests and the environment. Grasses have a unique cell wall that contains glucuronoarabinoxylan instead of xyloglucan and has a large amount of esterified hydroxycinnamic acid. Hydroxycinnamates are derived from phenylalanine in the phenylpropanoid pathway and serve many roles in the plant, including cell wall biosynthesis and specialized metabolism. One of these hydroxycinnamates, namely ferulate, can modify the arabinose side-chain of xylan. These feruloylated arabinoxylan moieties are able to dimerize via radical coupling to form diferulates, resulting in cross-links between xylan polymers and between xylan and lignin. Cross-linking in monocots likely contributes to cell wall strength, but the lack of identified genes involved in this pathway limit our understanding of this monocot-specific modification. The enzyme that adds ferulate to arabinoxylan is currently unidentified preventing alteration of the amount of ferulate in the wall by reverse genetic methods. In order to better understand ferulate-mediated cross-linking in grasses, our goal has been to find the gene encoding the enzyme responsible for adding ferulate onto arabinose residues in arabinoxylan, namely the arabinoxylan ferulate acyltransferase (AraFAT). To accomplish this goal we selected

candidates from a grass-specific BAHD acyltransferase clade that were highly differentially expressed in monocots compared to dicots, produced proteins from these genes, and then assayed these proteins with arabinose-containing substrates and feruloyl-CoA. We also explored wheat seedling protein extractions as a method of determining a testable assay for AraFAT.

During our search for the AraFAT gene, we discovered that one of our candidate genes, Bradi1g36980, is able to add ferulate to phenylamines to create phenylamides or hydroxycinnamic acid amide conjugates. These phenylamides have several functions in plants that include defense responses to pathogens and wounding. Phenylamides also play an important role in the plant cell wall, providing crosslinks between polymers that add rigidity and strength. The Bradi1g36980 enzyme has activity with donor substrate feruloyl-CoA and the acceptors tyramine, 2-phenyl-ethylamine, tryptamine, and serotonin. To our knowledge, this is the first time a BAHD acyl-transferase has been shown to use tryptamine or serotonin as a substrate. The true substrates of this enzymes have not been established by mutational analysis. Because these hydroxycinnamic acid amides are known to increase in plants exposed to jasmonate, a hormone involved in wound response pathways, we subjected *Brachypodium* seedlings to methyl-jasmonate and measured the relative expression level of Bradi1g36980 at various time points. We observed a slight increase in Bradi1g36980 expression occurred from methyl-jasmonate exposure, and are investigating the role of the phytohormone ethylene in its regulation.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Curt, for his guidance and willingness to take me on as his graduate student. Thank you to my committee members for their feedback and advice at every turn. Many thanks to my family, especially my parents, for their unconditional love and support. I am grateful to my friends for their support and kindness during an arduous time. To my fiancé, Nick, thank you for loving me and for always being there when I needed someone. Without all of you, this dissertation and degree would not have been possible.

TABLE OF CONTENTS

LIST OF TABLES	vii
LIST OF FIGURES	viii
KEY TO ABBREVIATIONS	x
INTRODUCTION	1
The importance of liquid fuel and plants as feedstocks	1
Background on plant cell wall and diferulate cross-linking in grasses	1
Significance of feruloylation and cross-linking in grasses	4
Literature review of current AraFAT research	7
Possible pathways of arabinoxylan feruloylation in grasses	11
Objectives of the graduate research project.....	14
CHAPTER 1.....	16
INTRODUCTION ON BIOCHEMICAL APPROACHES TO DISCOVERING AraFAT ...	16
MATERIALS AND METHODS.....	19
Selection of candidate genes	19
Enzymatic synthesis of CoA-substrates using 4CL2	19
Cloning, expression, and purification of BAHD acyltransferases in <i>E. coli</i>	20
Expression of genes using Cell-Free Protein Synthesis	21
Protein purification with HisTrap and MBPTrap columns using Fast Protein Liquid Chromatography (FPLC).....	21
Enzymatic assays	22
Reversed Phase Liquid Chromatography (RPC) analysis	22
Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS) with C18 column ..	22
Wheat seedling growing conditions	23
Wheat seedling protein extraction.....	23
FPLC desalting of wheat protein extract	24
FPLC anion exchange to separate proteins by charge	24
Assaying extracted protein for AraFAT activity	24
Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS) with Amide column	25
RESULTS.....	25
Bioinformatics approach to selecting AraFAT candidates.....	25
Cloning, expressing, and purifying candidates in <i>E. coli</i> and in a cell-free system	29
Characterization of Bradi1g36990 in an effort to find a testable assay for AraFAT	36
Wheat seedling protein extraction and arabinose substrate assays	39
DISCUSSION.....	40

CHAPTER 2.....	48
INTRODUCTION ON HYDROXYCINNAMATE CONJUGATIONS IN GRASSES.....	48
Background on jasmonate response pathway in grasses	51
MATERIALS AND METHODS.....	52
Acyltransferase identification and selection	52
Collection of compounds known to be hydroxycinnamated	52
Enzymatic synthesis of CoA-substrates using 4CL2	53
Cloning, expression, and purification of Bradi1g36980 in <i>E. coli</i>	53
Protein purification with MBPTrap column using FPLC.....	53
Enzymatic assays	53
UPLC RPC analysis	54
Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS) with C18 column ..	54
Time course enzyme kinetic assay methods and analysis	55
Exposure of Brachypodium to methyl-jasmonate (MeJA).....	55
Quantitative Real-Time PCR (qRT-PCR) of MeJA-treated plants.....	56
Tobacco <i>Agrobacterium</i> -infiltration of Bradi1g36980	56
Analysis of <i>Agrobacterium</i> -infiltrated <i>N. benthamiana</i> leaves	56
RESULTS.....	57
Characterization of Bradi1g36980 with feruloyl-CoA and substrates	57
Michaelis-Menten kinetic constants for substrates	63
Brachypodium seedling exposure to methyl-jasmonate.....	69
<i>Agrobacterium</i> -infiltrated tobacco plants with overexpressed gene.....	72
DISCUSSION.....	74
REFERENCES	80

LIST OF TABLES

Table 1: Cloned constructs of AraFAT candidates in various vectors heterologously expressed in <i>E. coli</i>	29
Table 2: CoA-acyl substrate preference of AraFAT candidates.....	31
Table 3: List of compounds known to be hydroxycinnamated in plants.....	59
Table 4: Gene enrichment table of the ratio of Replicates A to B of Bradi1g36980 RNA-seq data.....	70

LIST OF FIGURES

Figure 1: Lignin biosynthetic pathway from phenylalanine to the monolignols <i>p</i> -coumaroyl, sinapyl, and coniferyl alcohols.....	3
Figure 2: Ferulate addition on an arabinose side chain of xylan.....	5
Figure 3: Diferulate cross-linking between xylan polymers and between xylan and lignin.....	6
Figure 4: PF02458 clade of BAHD acyltransferases.....	9
Figure 5: Simplified schematic of the plant cell and the possible localization of arabinoxylan feruloylation.....	12
Figure 6: Graph of saponifiable ferulic and <i>p</i> -coumaric acid contents (% w/w) of the cell wall in growing wheat seedlings.....	18
Figure 7: Cell wall composition and RNA-seq dataset of Brachypodium developmental stages.....	27
Figure 8: Expression patterns of the top ten Brachypodium AraFAT candidates.....	28
Figure 9: RPC separation of a methanol assay with feruloyl-CoA and <i>p</i> -coumaroyl-CoA.....	31
Figure 10: Arabinose-containing substrates used in the AraFAT assays.....	33
Figure 11: Schematic cartoon of the cell-free protein synthesis bilayer reaction translating the mRNA into protein.....	34
Figure 12: SDS-PAGE gels of CFPS-WGX synthesized protein of AraFAT candidates.....	36
Figure 13: Expression pattern of Bradi1g36990 during different developmental stages in Brachypodium.....	38
Figure 14: RPC separation of methanol assays with Bradi1g36990 show activity with both <i>p</i> -coumaroyl- and feruloyl-CoA.....	38
Figure 15: Simplified structures of hydroxycinnamate conjugates.....	49
Figure 16: SDS-PAGE gels of Bradi1g36980 MBP-tag affinity purification.....	58
Figure 17: RPC separation of Bradi1g36980 assay products with feruloyl-CoA and phenylamines.....	59

Figure 18: Amino-containing compounds assayed with Bradi1g36980 resulting in no activity.....	61
Figure 19: Mass spectra of hydroxycinnamic acid amides made by Bradi1g36980 enzyme.....	62
Figure 20: Fragmentation patterns of feruloyl-amides in LC-MS/MS in positive ion mode.....	62
Figure 21: Saturation kinetics of Bradi1g36980 with varied feruloyl-CoA concentrations, measuring feruloyl-tryptamine product formation versus time.....	64
Figure 22: Saturation kinetics of Bradi1g36980 with varied tryptamine concentrations, measuring feruloyl-tryptamine product formation versus time.....	65
Figure 23: Saturation kinetics of Bradi1g36980 with varied serotonin concentrations, measuring feruloyl-serotonin product formation versus time.....	66
Figure 24: Michaelis-Menten and asymptotic fits of feruloyl-CoA by Bradi1g36980 enzyme....	67
Figure 25: Potential errors in data series of kinetic assays of feruloyl-CoA by Bradi1g36980 enzyme	67
Figure 26: Preliminary data on initial velocity of product formed versus substrate concentration for tryptamine and serotonin substrates by Bradi1g36980	68
Figure 27: Graph of three biological replicates of RNA-seq data of Bradi1g36980 at different developmental stages of Brachypodium	69
Figure 28: Relative gene expression in response to methyl-jasmonate (MeJA) over time analyzed by qRT-PCR.....	71
Figure 29: LC-MS/MS data of soluble extracts of transiently overexpressed Bradi1g36980 and untreated leaves in tobacco leaves.....	73
Figure 30: Mass spectra of feruloyl-containing peaks in Bradi1g36980 overexpression in tobacco leaves.....	74

KEY TO ABBREVIATIONS

4CL – 4-coumarate:CoA ligase

(6x)His-tag – six Histidine residues (for affinity tag)

A1P – arabinose-1-phosphate

A2XX – 2³- α -L-arabinofuranosyl-xylotriose

A3X – 3²- α -L-arabinofuranosyl-xylobiose

ACC – 1-aminocyclopropanecarboxylic acid

ACN – acetonitrile

AIR – alcohol insoluble residue

Ara f – arabinofuranose

AraFAT – arabinose ferulate acyltransferase

Arap – arabinopyranose

AraPAT – arabinose *p*-coumarate acyltransferase

AT – acyltransferase

AX – arabinoxylan

BAHD – class of acyltransferases; stands for the first four enzymes characterized in this class

BS – Brachypodium stem

C3H – *p*-coumarate 3-hydroxylase

C4H – cinnamate 4-hydroxylase

CAD – cinnamyl alcohol dehydrogenase

CCoA-OMT – caffeoyl-CoA *O*-methyltransferase

CCR – cinnamoyl-CoA reductase

CESA – cellulose synthase

CFPS-WGX – cell-free protein synthesis wheat germ extract

CoA – Coenzyme A

COMT – caffeic acid *O*-methyltransferase

CSC – cellulose synthase complex

DASF – day(s) after seed filling, typically about 7 days after flowering for *Brachypodium*

DHFR – dihydrofolate reductase, a control protein used in CFPS-WGX

DTT – dithiothreitol

eV – electron volt

F5H – ferulate 5-hydroxylase

fA – ferulate or ferulic acid

FPLC – Fast Protein Liquid Chromatography

GAX – glucuronoarabinoxylan

GNAT – GCN5-related *N*-acyltransferase

GT – glycosyl transferase

HCA – hydroxycinnamate

HCT – *p*-hydroxycinnamoyl-CoA:quinic/shikimate *p*-hydroxycinnamoyltransferase

HPLC – High Performance Liquid Chromatography

JA – jasmonate or jasmonic acid

K_M – Michaelis-Menten constant

LC-MS/MS – Liquid Chromatography Tandem Mass Spectrometry

MBP – maltose-binding protein (for affinity tag)

Me- α -Ara – methyl- α -L-arabinofuranoside

Me- β -Ara – methyl- β -L-arabinofuranoside

MeJA – methyl-jasmonate

MLG – mixed-linked glucan

MS – Mass Spectrometry

m/z – mass-to-charge ratio

PAL – phenylalanine ammonia-lyase

pCA – *p*-coumarate or *p*-coumaric acid

PMT – *p*-coumarate monolignol transferase

qRT-PCR – quantitative real-time polymerase chain reaction

RGP – reversibly glycosylated protein

RNA-seq – RNA sequencing

RNAi – RNA interference

RPC – reversed phase chromatography

RPKM – reads per Kilobase of transcript, per million mapped reads

SDS-PAGE – sodium dodecyl sulfate polyacrylamide gel electrophoresis

TEV – tobacco etch virus

TFA – trifluoroacetic acid

THT – tyramine:*N*-hydroxycinnamoyl transferase

TIC – total ion chromatogram

UAM – UDP-arabinopyranose mutase

UDP – uridine diphosphate

UDP-Araf – uridine diphosphate arabinofuranose

UDP-Arap – uridine diphosphate arabinopyranose

UPLC – Ultra-high Performance Liquid Chromatography

USP – UDP-sugar pyrophosphorylase

V_{\max} – maximum enzyme reaction rate at saturating substrate concentration

XyG – xyloglucan

INTRODUCTION

The importance of liquid fuel and plants as feedstocks

Fossil fuels are the major source of energy for powering most of the world's technologies. The United States alone gets roughly 80% of its total energy from oil, natural gas, and coal (NASEM, 2018). With the inevitable depletion of fossil fuels drawing near, the demand to find alternative sources of renewable and sustainable fuels is growing. Many scientists believe that renewable sources of fuels can be obtained from plant sources. While plants play a vital role in nutrition for humans and animals, the non-edible parts of plants can be utilized for conversion to liquid biofuels, such as ethanol. Most sources of ethanol are made from easily extractable starches and sugars, but a massive pool of potential renewable energy is found elsewhere in the cell. Two of the most abundant organic polymers on Earth are found in plants: lignin and cellulose, components of the plant cell wall (Liu, 2010). Although easily overlooked, the cell wall holds much potential as a source of biofuels and bioproducts (Perlack *et al.*, 2005).

Background on plant cell wall and diferulate cross-linking in grasses

The plant cell wall is essential to plant development and morphology. The wall, supported by turgor pressure, allows plants to grow vertically to compete more effectively for light (Neutelings, 2011). Cell walls not only play a role in cell-to-cell communication, as shown through hormone transport and metabolism in the extracellular cell wall space (Tameshige *et al.*, 2015), they also provide a physical barrier against pathogens and herbivorous insects (Keegstra, 2010). This vital component of the plant cell acts as a protective layer and strengthens the plant, but its recalcitrance to enzymatic digestion or chemical processing limits human utilization (Simmons *et al.*, 2010; Foster *et al.*, 2010). Fiber is also a major source of nutrition for ruminant animals, but its poor digestibility allows for less than 50% utilization (Hatfield *et al.*, 1999a).

A factor that affects the usage of cell walls as a feedstock for the production of biofuels, is the type of cell wall. There are two types of cell wall: the primary cell wall, which is synthesized during the early stages of cell growth that allows for cell expansion while resisting turgor pressure, and the secondary cell wall, designed to provide mechanical strength to organs such as stems and to waterproof surfaces to the xylem as well as resistance to pathogens (Keegstra, 2010). The secondary cell wall is formed inside the primary cell wall after the cell is fully elongated (Gibeaut and Carpita, 1994).

The main components of the primary cell wall include cellulose, hemicellulose, and pectin. Cellulose is synthesized by a class of enzymes referred to as cellulose synthase (CESA) proteins, which form cellulose synthase complexes (CSCs). These complexes occur as hexagonal rosettes, and each rosette subunit contains six CESA proteins. These 36 CESA proteins simultaneously synthesize 36 parallel β -(1,4)-glucan chains, which crystallize via hydrogen bonding to form cellulose microfibrils (Doblin et al., 2002; Somerville, 2006; Li *et al.*, 2014). Cellulose acts as the structural framework on which matrix polysaccharides build (Geisler et al., 2008; Lerouxel *et al.*, 2006).

A major component of the secondary cell wall is lignin, a hydrophobic, phenolic polymer that typically comprises a large portion of the secondary cell wall. This component waterproofs the cell and accounts for much of the recalcitrance of the wall to enzymatic digestion (Vanholme *et al.*, 2008; Vanholme *et al.*, 2010; Whetten and Sederoff, 1995). The lignin biosynthetic pathway involves the conversion of the amino acid phenylalanine to three major monolignol subunits: syringyl (S), guaiacyl (G), and *p*-hydroxyphenyl (H) (**Figure 1**) (Boerjan *et al.*, 2003).

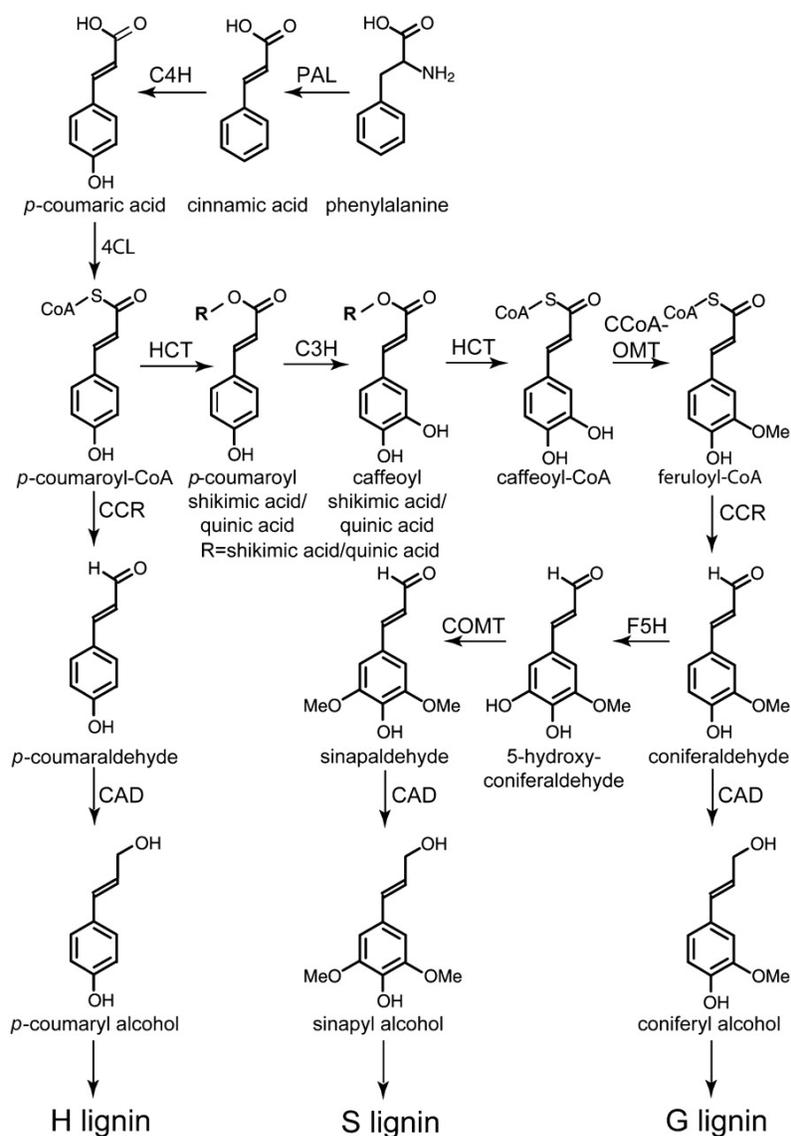


Figure 1: Lignin biosynthetic pathway from phenylalanine to the monolignols *p*-coumaroyl, sinapyl, and coniferyl alcohols. PAL, phenylalanine ammonia-lyase; C4H, cinnamate 4-hydroxylase; 4CL, 4-coumarate:CoA ligase; HCT, *p*-hydroxycinnamoyl-CoA:quinic acid/shikimate *p*-hydroxycinnamoyltransferase; C3H, *p*-coumarate 3-hydroxylase; CCoA-OMT, caffeoyl-CoA O-methyltransferase; CCR, cinnamoyl-CoA reductase; F5H, ferulate 5-hydroxylase; COMT, caffeic acid O-methyltransferase; CAD, cinnamyl alcohol dehydrogenase (Figure from Vanholme *et al.*, 2010; material is copyright by the American Society of Plant Biologists)

Hemicelluloses are polysaccharide chains made of β -(1,4)-linkages that include xylans, glucans, xyloglucans, and β -(1,3;1,4)-glucans, known as mixed-linked glucans (MLG). MLG is unique to monocots and is a good source of soluble dietary fiber (Wood, 2007). This group of polysaccharides is likely required for interactions between cellulose microfibrils to establish a strong cell wall. Alterations in the hemicelluloses *in muro* are likely necessary for the remodeling of the wall during growth and development (Scheller and Ulvskov, 2010). Pectin is the most structurally diverse polysaccharide, made up of covalently linked galacturonic acid-rich polysaccharides that act in a number of functions, including growth and development, cell-cell adhesion, and cell expansion (Mohnen, 2008). The cell wall also has various structural proteins, such as extensins and expansins, which assist in the assembly and expansion of the cell wall (Cassab, 1998).

Significance of feruloylation and cross-linking in grasses

Members of the *Poaceae*, a family of monocotyledon grasses, contain some of the most economically important plants. Grasses are extremely promising candidates for biofuel crops because they are highly productive, and some are able to grow on marginal lands (Escamilla-Treviño *et al.*, 2013; Hatfield *et al.*, 2009). Despite the dietary and economic importance of grasses, to date most studies of plant cell walls have been on dicotyledonous plants in an effort to reduce costs of lignin recalcitrance for pulping and paper production (Grabber *et al.*, 2004). Past research has shown that the cell walls of grasses differ significantly from that of other monocots and of dicots (Gibeaut and Carpita, 1994; Carpita, 1996; Vogel, 2008). The primary cell wall of *Poaceae* is known as Type II, and its composition differs greatly from Type I primary cell walls. Some of these differences include walls made of glucuronoarabinoxylan (GAX) rather than xyloglucans (XyG), large quantities of MLG, and high levels of hydroxycinnamates (Vogel,

2008; Faik, 2010). Type II cell walls are very low in pectin, having only about 2-10% compared to about 35% in Type I (Mohnen, 2008). Monocots also have a reduced amount of lignin in their secondary cell walls compared to eudicots (Vogel, 2008; Engels and Jung, 1998; Grabber, 2005).

A key distinguishing feature of the grasses is the relatively high amounts of ferulic acid in the cell wall, up to 4% in both the primary and secondary cell walls (Carpita, 1996; Saulnier *et al.*, 1999; Hatfield *et al.*, 1999b; Vogel, 2008). Ferulate is added to the arabinose side chain of xylan, on the C5 carbon of arabinose (**Figure 2**). This hydroxycinnamate is capable of forming dimeric structures via free-radical coupling, leading to a variety of diferulate isomers such as 5-5 and 8-5 coupling (Ralph *et al.*, 1994; Vismeh *et al.*, 2013). These diferulates can then cross-link between and within xylan polymers (Hatfield *et al.*, 1999c).

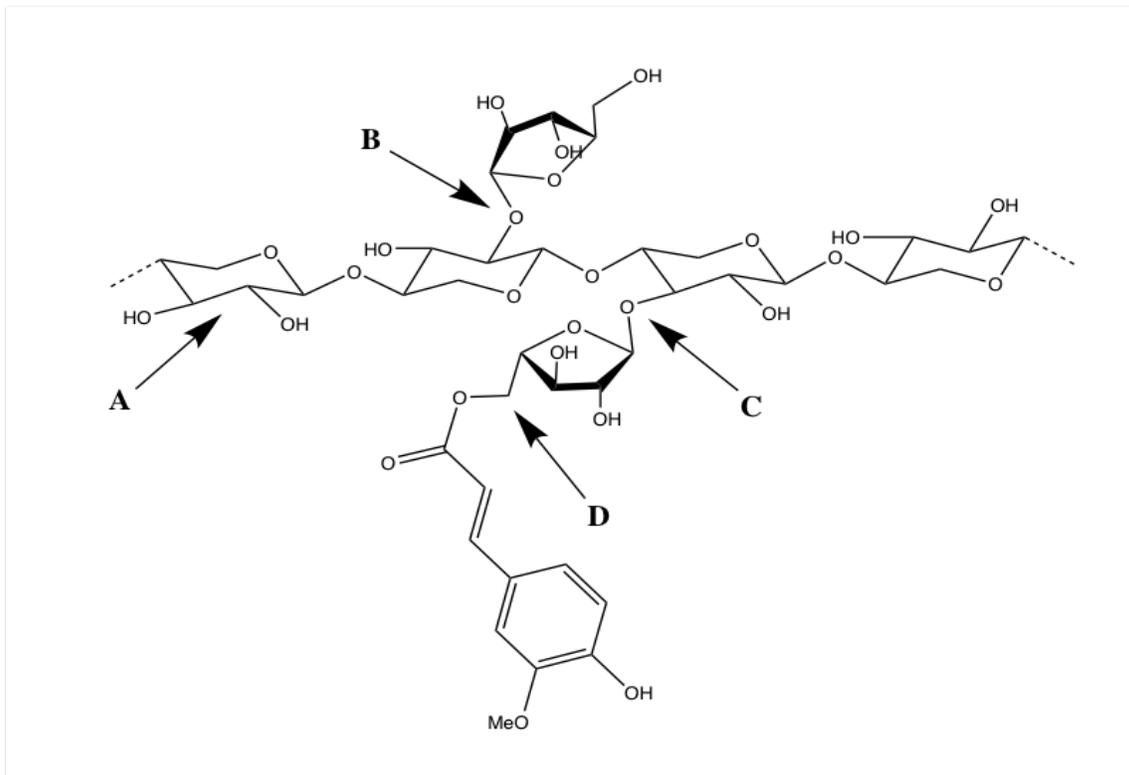


Figure 2: Ferulate addition on an arabinose side chain of xylan. A is the β-(1,4)-xylose backbone of xylan. B is an arabinofuranose (Araf) addition on the α-(1,2) position of xylose. C highlights the α-(1,3)-linked Araf, with D, a ferulate addition on the C5-carbon of arabinose. The α-(1,3) Araf is also able to be substituted with *p*-coumarate or xylose (usually at the C2 of arabinose).

Feruloyl esters participate as substrates in the free-radical polymerization of lignin, leading to cross-links between xylan and lignin (**Figure 3**) (Ralph *et al.*, 1992; Ralph *et al.*, 1994; Grabber *et al.*, 1995; Ralph, 2010). Although the significance of ferulate cross-linking in the plant cell wall is yet to be determined, research has suggested that this modification impacts the digestibility of cell walls due to the polysaccharide cross-linkages (Fry and Miller, 1989; Ralph *et al.*, 1994). The enzymes involved in this unique grass addition of ferulate to arabinoxylan are also unknown. Understanding the enzymes involved in this modification will allow for scientists to study the impact of cross-linking in grasses, and potentially design plant cell walls to be more readily digestible (Mottiar *et al.*, 2016; Ralph, 2010).

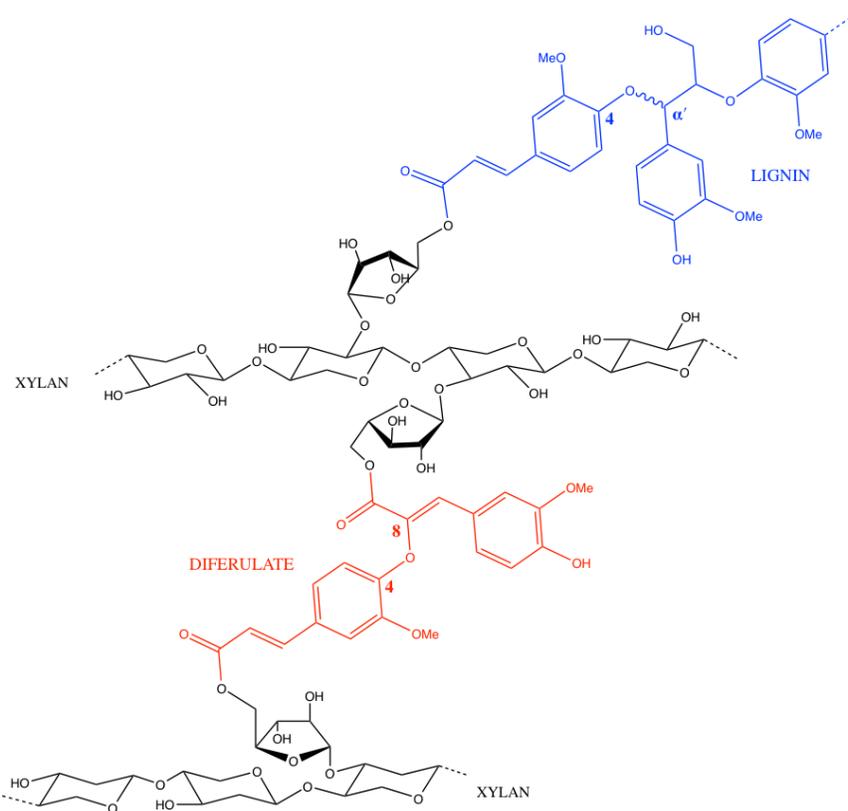


Figure 3: Diferulate cross-linking between xylan polymers and between xylan and lignin. In red, an 8-O-4 diferulate cross-links between two xylan chains (black). Ferulate acts as a nucleation site for lignin polymerization, and is able to form ester (not shown) and ether linkages with lignin, as shown in blue with the 4-O- α' ether coupling, a result of trapping intermediate quinone methides (Ralph *et al.*, 1992; Grabber *et al.*, 1995).

Literature review of current AraFAT research

Diferulate cross-linking in grasses has become a topic of great interest due to its potential for altering the digestibility of these walls. Initial studies on *Poaceae* addressed the role of ferulic acid in cell wall function. Data showing that diferulates were forming both ester and ether linkages led to the speculation that ferulates were involved in cross-linking of cell wall components (Scalbert *et al.*, 1985; Smith and Hartley, 1983). It was not until 1991 that evidence of xylan cross-linking *in vivo* was published from research on cell wall composition of young bamboo shoots. Using the fungal enzyme preparation Driselase, the group was successful at isolating a diferuloyl arabinoxylan hexasaccharide from the growing cell wall (Ishii, 1991). Work performed by Iiyama *et al.* showed that ferulic acid residues were simultaneously esterified (to arabinoxylan) and etherified (to lignin), providing further evidence that ferulates form cross-links in the cell wall (Iiyama *et al.*, 1990).

Some approaches to studying ferulate cross-linking and its impact on grasses use genetic approaches. One study used high performance liquid chromatography (HPLC) analysis of maize internodes throughout development and found evidence that ferulate deposition occurs in both primary and secondary cell walls. This effort also found that ferulate esters were being deposited in the secondary cell wall of maize, contradicting a previous held hypothesis that these cross-links only occurred in primary cell walls of grasses (Jung, 2003). In a separate experiment studying putative maize mutants with reduced ferulate esters, this group found an association between the reduction in ferulate esters and a reduction in ferulate ethers (Jung and Phillips, 2010). These data suggest the ferulate esters were involved in lignin cross-linking. In addition, plants with lowered ferulate levels had improved cell wall degradability. These results demonstrate a role for ferulate esters in wall function (Jung and Phillips, 2010).

There have been many efforts to find genes involved in arabinoxylan biosynthesis, including the gene responsible for the addition of ferulate to the arabinose moieties of arabinoxylan (Mitchell *et al.*, 2007; Piston *et al.*, 2010; Anders *et al.*, 2012; Chiniqy *et al.*, 2012; Bartley *et al.*, 2013). In one study by Mitchell *et al.*, researchers developed a bioinformatics approach in an attempt to discover this enzyme. Because xylan makes up a larger portion of the cell wall in grasses than it does in dicots, the hypothesis was that they would find genes involved in arabinoxylan biosynthesis through observing genes in grasses highly differentially expressed in monocots compared to dicots. They were able to conclude that three families of glycosyl transferases (GTs), GT43, GT47 and GT61, are most likely involved in arabinoxylan biosynthesis. GT61 has been shown to encode the xylosyl transferase that adds arabinose to the xylan backbone (Anders *et al.*, 2012). It is the most highly differentially expressed gene between dicots and cereals. It also has statistically significant correlation with the group of related proteins (Pfam PF02458) that are BAHD acyltransferases (Mitchell *et al.*, 2007). For this reason, the PF02458 clade of BAHD acyltransferases is likely to contain the gene for arabinoxylan ferulate transferase (**Figure 4**).

Additional evidence that the PF02458 group might contain AraFAT is through reverse genetic approaches. In one experiment, the expression of four rice genes in this clade were knocked-down in rice using RNA-interference (RNAi) and a significant reduction in ester-linked ferulate was measured (Piston *et al.*, 2010). Although this study indicates that one of these genes could be AraFAT, the simultaneous suppression of these genes made it impossible to identify which, if any, were responsible for AX feruloylation. One of the genes studied in that paper, Os06g39390, was further investigated by Bartley *et al.* using activation-tagged overexpression lines of rice. This gene, named OsAT10, when overexpressed in rice showed an increase in cell

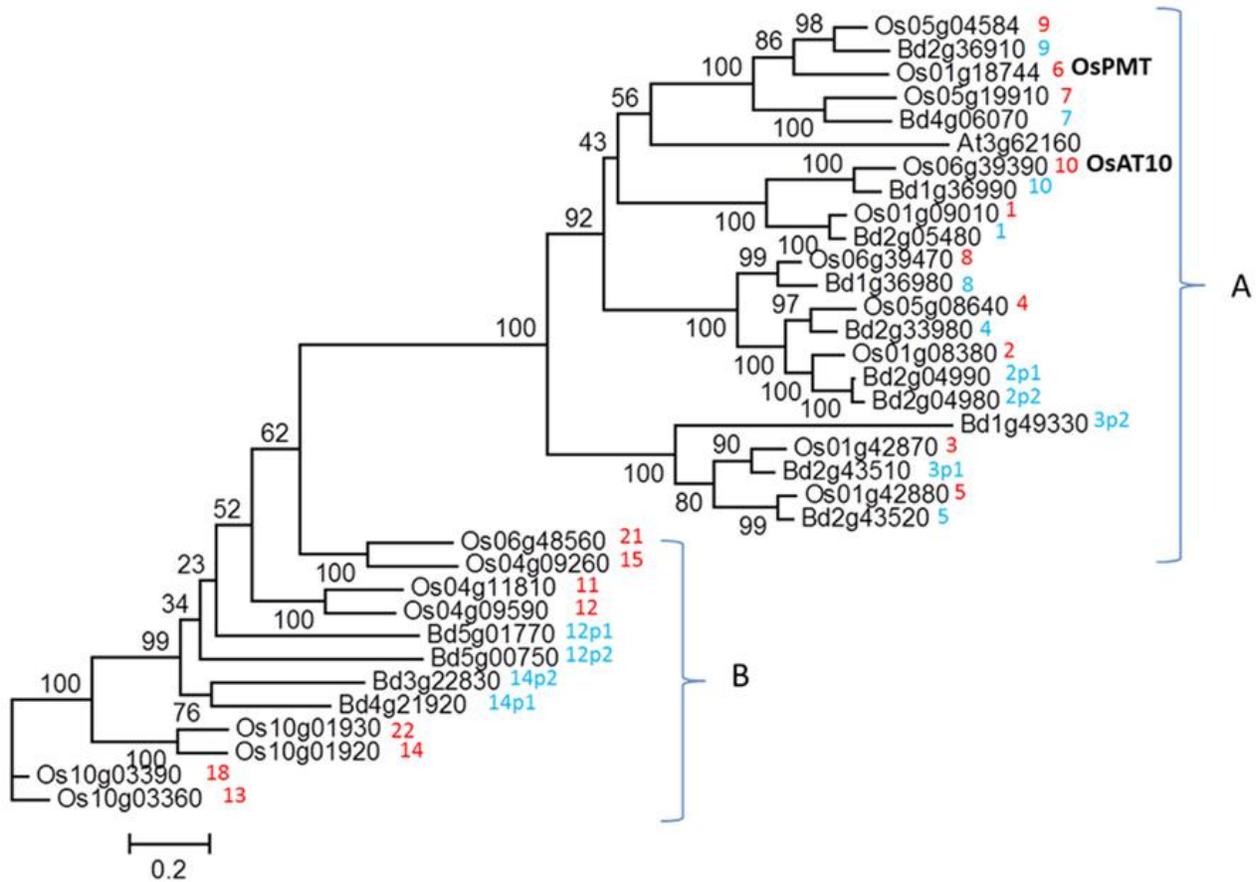


Figure 4: PF02458 clade of BAHD acyltransferases. This grass-specific clade belongs to the PF02458 group identified as having potential for holding the gene responsible for feruloylation of arabinoxylan (Mitchell *et al.*, 2007). The clade above is borrowed from the work by Molinari *et al.* (2013) and was organized into two subclades based on the gene coexpression with xylan biosynthesis genes. Clade A belongs to those highly coexpressed with xylan pathway genes, and Clade B genes are not highly expressed or coexpressed and have no clear rice or Brachypodium orthologs. The genes most likely to be involved in arabinoxylan feruloylation are in Subclade A (upper right). OsPMT is the rice *p*-coumarate monolignol transferase (Withers *et al.*, 2012). OsAT10, when overexpressed in rice, had an increase in *p*-coumarate on arabinoxylan, genetic evidence that it is the arabinoxylan *p*-coumarate acyltransferase (AraPAT) (Bartley *et al.*, 2013). (Figure from Molinari *et al.*, 2013 published in *Frontiers in Plant Science*)

wall *p*-coumarate levels by up to 300%, and a reduction of about 60% in ferulate levels. They verify that the *p*-coumarate added is predominantly bound to GAX and not lignin because it was measured mostly in the trifluoroacetate (TFA)-soluble matrix fraction, which corresponds to hemicellulose contents of the cell wall, excluding cellulose and lignin. Their experiments provide strong genetic evidence that the OsAT10 gene was responsible for the addition of *p*-coumarate

onto arabinoxylan. This study gave more confidence that the related *ferulate* acyltransferase would be in this same clade (Bartley *et al.*, 2013).

Our lab has identified the function of a member of the PF02458 clade . We discovered the *p*-coumaroyl-CoA monolignol transferase (PMT) gene, which encodes an enzyme that produces a monolignol conjugate found in grass lignin. (Withers *et al.*, 2012). From the kinetic assays, this enzyme showed strong affinity for *p*-coumaroyl-CoA, and served as the donor substrate for sinapyl and *p*- coumaroyl alcohols. A study by Petrik *et al.* verified that PMT is specifically a lignin biosynthetic enzyme and is not involved in arabinoxylan *p*-coumaroylation using genetically altered PMT in *Brachypodium* plants (Petrik *et al.*, 2014). While this is not the gene responsible for producing feruloylated arabinoxylan, the grass-specific nature of the enzyme is consistent with Mitchell *et al.*'s bioinformatics (Mitchell *et al.*, 2007; Withers *et al.*, 2012).

Further evidence that the PF02458 group contains genes involved in grass-specific cell wall biosynthesis can be found in two recent publications. The first one found that the *Brachypodium distachyon* gene, Bradi1g36980, upon overexpression in Arabidopsis, had increased levels of alkali-releasable *p*-coumarate from cell walls to that of grass lignin levels. They postulate that this gene is a second *p*-coumarate monolignol transferase, labeled PMT2 (Sibout *et al.*, 2016). However, Chapter 2 of this dissertation shows that Bradi1g36980 is *not* PMT2 and is in fact responsible for a different chemical reaction.

The second publication to explore members of this clade discovered genes involved in cell wall feruloylation (de Souza *et al.*, 2018). RNAi was utilized to silence the gene Sevir.5G130000 in the model grass *Setaria viridis* and found a 60% reduction of feruloylated arabinoxylan in stems. Silencing ortholog Bradi2g05480 in *Brachypodium* also showed a

decrease in feruloylation, but to a lesser degree (about 10-20%). This study provides strong genetic evidence that Bradi2g05480 is a gene encoding the enzyme that adds ferulate to arabinoxylan in grasses (de Souza *et al.*, 2018). In Chapter 1 of this dissertation, I present the biochemical approaches we used to determine if this gene was able to perform the predicted AraFAT function, which we were ultimately unsuccessful at proving.

Possible pathways of arabinoxylan feruloylation in grasses

Much is still unknown about the pathway of arabinoxylan feruloylation. There are two major considerations that must be addressed in order to discover the enzyme required to catalyze this reaction: the location of the ferulate addition and the substrates required to make this addition. Understanding where in the cell ferulate is added to arabinose (or arabinoxylan) would provide a lot of information about what enzymes are involved. For example, if the reaction occurred in the Golgi, it could be assumed that BAHD acyltransferases would not be the enzymes responsible for the reaction, since they are localized in the cytoplasm (D'Auria, 2006). There are three possible locations where arabinose feruloylation might occur: 1) in the cytosol, 2) in the Golgi, or 3) *in muro* (at the cell wall) (**Figure 5**).

There are a few reasons to believe that this reaction occurs in the cytosol. One, the likely candidate group of BAHD acyltransferases (PF02458) is cytosolic since they lack a signal peptide. If a gene in this group is responsible for feruloylation of arabinose, the reaction would have to occur cytosolically. Another reason that this location is attractive is that the likely acceptor substrate, UDP-arabinofuranose (UDP-Araf) is made in the cytosol. A majority of UDP-arabinopyranose (UDP-Arap) is made from UDP-glucose in the Golgi, then transported out of the Golgi where it is converted to UDP-Araf by the UDP-arabinopyranose mutase (UAM)

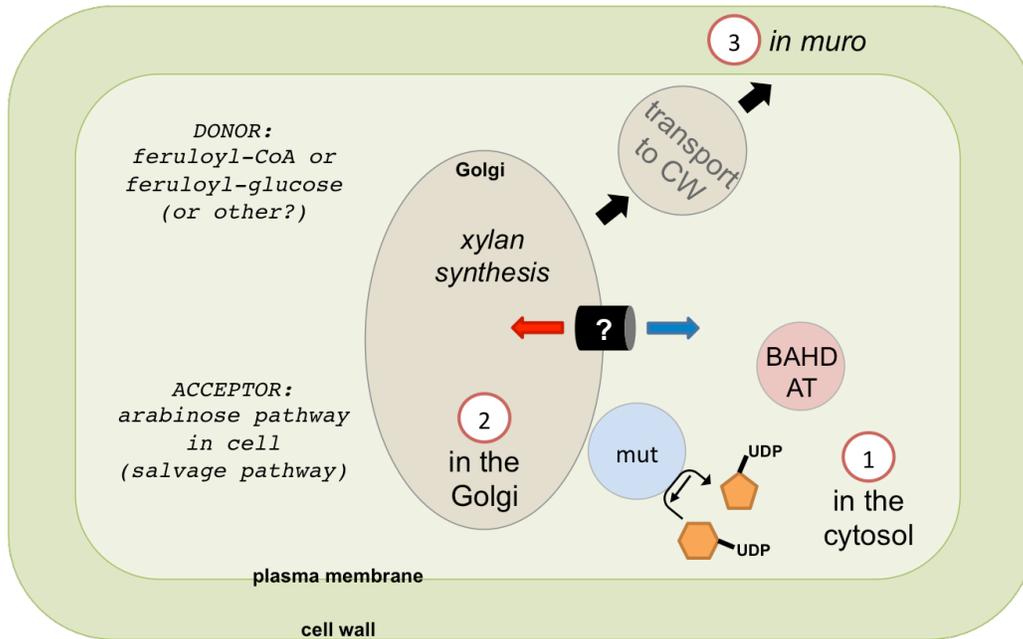


Figure 5: Simplified schematic of the plant cell and the possible localization of arabinoxylan feruloylation. The three possible locations of ferulate to be added to arabinose (or AX) are 1) in the cytosol, 2) in the Golgi (the location of xylan synthesis), or 3) *in muro*, at the cell wall. Here, I highlight the cytosolic localization as the most likely possibility, as the predicted acceptor substrate UDP-Araf (orange pentagon) is newly converted (from UDP-Arap, orange hexagon) and BAHD ATs are cytosolic proteins. UAMs (labeled “mut” in blue) could potentially play a role in this reaction, being cytosolic proteins that often associate with the Golgi and play a role in the conversion to the likely acceptor substrate UDP-Araf. The black cylinder with “?” represents the unknown transporters responsible for UDP-arabinose transport across the Golgi membrane, and if there is one specific to feruloylated arabinose. The question of donor and acceptor substrates is another unknown part of this pathway. Feruloyl-CoA or -glucose are two possible donor substrates, however the latter is stored in the vacuole, and would make it difficult to occur in the substrate pathway to the Golgi.

(Rautengarten *et al.*, 2017). There is a minor amount of UDP-Arap that is made by the sugar salvage pathway cytosolically, using arabinose-1-phosphate (A1P) and a UDP-sugar pyrophosphorylase (USP) (Geserick and Tenhaken, 2013). Either pool of UDP-Arap is present in the cytosol before conversion to UDP-Araf, making the cytosol an attractive location for arabinose feruloylation.

For feruloylation on arabinoxylan to occur, there needs to be a donor substrate and an acceptor substrate. The donor is a ferulate compound, but the activated form used for this

reaction is unknown. The most likely candidate is feruloyl-CoA, as it is synthesized in the cytosol and is used as a substrate in many other reactions. Another possibility is feruloyl-glucose, which a number of studies have suggested as a possible substrate (Bokern *et al.*, 1991; Obel *et al.*, 2003; Lenucci *et al.*, 2009). Obel *et al.* (2003) fed radiolabeled ferulic acid to a cell suspension culture of wheat cells and observed a rapid conversion to feruloyl-glucose, which they correlated to the resulting arabinoxylan feruloylation. However, they assumed that the measured feruloylated arabinose is associated with glycoproteins, which is incorrect as glycoproteins make up a minor component of the cell (Obel *et al.*, 2003; Buanafina, 2009). Further, the rapid conversion of ferulic acid into feruloyl-glucose does not eliminate feruloyl-CoA as the donor substrate. Feruloyl-glucose may act as a reserve pool of feruloyl units until used for turn-over to feruloyl-CoA, or other feruloyl-form (Lenucci *et al.*, 2009). As for the acceptor substrates, there are several arabinose forms, any of which could be feruloylated. The reaction may take place on arabinose-1-phosphate (A1P). It could also occur as UDP-arabinopyranose (UDP-Arap) or UDP-Araf, two interconverting forms of arabinose. One reason why UDP-Araf is the most promising candidate for acceptor substrate is that it is the substrate used for xylan synthesis.

The mutases responsible for interconversion of UDP-arabinose, UAMs, could play an active part in this reaction as well. These cytosolic proteins, known as Reversibly Glycosylated Proteins (RGPs) often associate with the Golgi membrane, despite having no known transmembrane domains nor signal peptides (Konishi *et al.*, 2007; Rautengarten *et al.*, 2011). The method of catalytic ring interconversion requires a conserved DXD motif, which serves to coordinate the manganese cation and the phosphate groups on the UDP-arabinose, and an arginine residue in the mutase, which links to arabinose and acts as the site for reversible

glycosylation (Konishi *et al.*, 2010; Singh *et al.*, 1995). There are three known rice mutases, known as UAM1, UAM2, and UAM3, but only UAMs1 and 3 have any mutase activity (Konishi *et al.*, 2007). Concomitant knock-downs of the two active mutases in Arabidopsis cause severe growth defects and result in plants almost entirely deficient in cell wall-derived arabinose (Rautengarten *et al.*, 2011).

The role of UAM and its subcellular localization is likely a factor in this reaction, since the predominant form of arabinose is the furanose form but requires the transport of UDP-Araf from inside the Golgi to the cytosol in order to convert to UDP-Araf (Rautengarten *et al.*, 2017). One possibility is that the mutase plays a role in the transfer of ferulate onto arabinose. Once converted to the furanose form, UDP-Araf can rapidly be used as a substrate. It is also possible that ferulate is added while arabinose is bound to the arginyl residue on the mutase during interconversion, although this has not been proven. Another interesting phenomenon of the UAMs is their formation of homo- and heteroprotein complexes (Langeveld *et al.*, 2002; Drakakaki *et al.*, 2006; De Pino *et al.*, 2007; Rautengarten *et al.*, 2011). The purpose of these complexes is not yet understood. In one study, Konishi *et al.* (2010) showed evidence that the nonmutase UAM2 in complexes with UAM1 or UAM3 both showed minor increases in mutase activity. One consideration that, to our knowledge, has yet to be studied is if these UAM complexes interact with other proteins, such as BAHD ATs or UDP-arabinose transporters. Perhaps understanding the function of these homo- and heterocomplexes of UAMs would give us new insights into the pathway of arabinoxylan feruloylation in the cell.

Objectives of the graduate research project

My goal was to find the enzyme(s) responsible for the addition of ferulate on arabinoxylan using mainly biochemical approaches. In this dissertation, I detail our efforts to

find the gene encoding AraFAT and our approaches to elucidating the AraFAT pathway in grasses. In Chapter 1, I present the work we did to characterize the AraFAT candidates that we selected for study. I also examine the potential for using wheat seedling protein extracts as a means to find and develop an assay for AraFAT activity. In Chapter 2, I introduce results from a project that arose during our search for AraFAT, where one of the candidates showed activity with phenylamines. This experiment, to our knowledge, is the first instance of a BAHD acyltransferase *N*-acylating serotonin and tryptamine. While we were ultimately unsuccessful at finding the AraFAT enzyme, we believe this work has contributed to the overall understanding of grass-specific cell wall modifications. I further hope this work convinces the reader the importance of using biochemistry to further characterize members of the PF02458 BAHD acyltransferase clade, and that genetic approaches alone to gene discovery are not enough to predict protein function in the cell.

CHAPTER 1

INTRODUCTION ON BIOCHEMICAL APPROACHES TO DISCOVERING AraFAT

As stated in the introduction, the goal of my research was to discover the gene encoding the enzyme responsible for adding ferulate to arabinoxylan in grasses. Our first step was to determine an approach finding candidate genes. Several labs have been using genetic approaches, such as overexpression or knock-down of a specific gene, in order to discover gene function in the plant (Bartley *et al.*, 2013; de Souza *et al.*, 2018). While these methods have been successful at measuring changes in the cell wall as a result of manipulating a single gene, it is difficult to know if these changes come directly from the gene's function or as a result of substrate pool depletion. For example, if overexpressing a gene leads to increased ferulate on lignin, is this gene directly responsible for the addition of ferulate onto lignin, or is it using a different substrate (such as *p*-coumaroyl-CoA) and modifying other parts of the cell with *p*-coumarate, leaving feruloyl-CoA free to be added to lignin.

Ideally, our project would involve both biochemical and genetic approaches to understanding and discovering the AraFAT pathway. However, *Agrobacterium* transformation into *Brachypodium distachyon* is a lengthy process, taking about 20-31 weeks before harvesting transgenic seeds (Bragg *et al.*, 2014). Having had prior success with purifying BAHD acyltransferase proteins and characterizing their activities *in vitro*, we chose to use a similar approach to find AraFAT. Cloning, expressing, and purifying our candidates followed by *in vitro* assays will give us evidence that these enzymes are performing reactions with specific substrates, something that we are unable to visualize in the plant.

Over the course of this project, a great deal of effort was made to clone, express, and assay candidate genes for AraFAT. Unfortunately, using both bacterial expression and cell-free

protein synthesis using wheat germ extract (CFPS-WGX), we were unable to find a candidate with the ability to ferulate arabinose. We cannot determine if this failing is from a lack of an assay, not having the correct BAHD AT candidates, or if our proposed AraFAT reaction is altogether incorrect. It is possible that we have been assaying our candidates with the correct substrates, but we have yet to find the correct protein(s) involved. To reiterate our dilemma, the three unknowns in this reaction are 1) the enzyme, 2) the donor substrate, and 3) the acceptor substrate. One way to account for one variable in this reaction is by extracting protein from a grass species, which has feruloylated AX and will therefore contain our protein of interest, and assaying the total protein extract with various substrates. This approach would allow us to assay the soluble proteins all at once and determine an assay for the AraFAT reaction. It will also give us insight into the proper form of arabinose used in the plant. Once we have an assay, we can use proteomics to identify enzymes that were present in the successful reaction mixture and narrow down the list of potential AraFAT candidates.

There are a number of reasons why using wheat seedlings for protein extraction is an advantageous approach to finding AraFAT activity. As mentioned in the CFPS-WGX section, wheat is a member of *Poaceae*, a monocot family that modifies arabinoxylan with ferulate. During the first week of growth and development, wheat seedlings contain a large amount of saponifiable ferulate between days 3 and 7 of growth (**Figure 6**). Further, these seedlings are at a stage of development when the cells are rapidly growing and expanding, and the amount of arabinoxylan is increasing (Obel *et al.*, 2002; Lu *et al.*, 2006). This information gives us confidence that during the first week of growth, feruloylation of arabinoxylan is occurring, and therefore, the proteins responsible for this modification are likely present in the wheat cell.

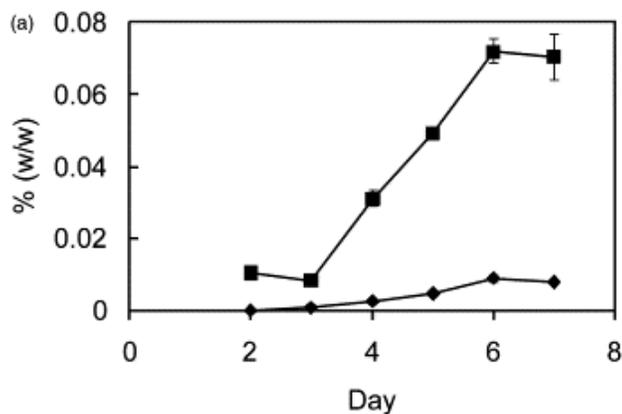


Figure 6: Graph of saponifiable ferulic and *p*-coumaric acid contents (% w/w) of the cell wall in growing wheat seedlings. The content of ferulic acid (■) and *p*-coumaric acid (♦) from 2-7 days after germination. (Figure from Obel *et al.*, 2002, published in *Phytochemistry*)

Given that wheat seedlings are synthesizing feruloylated arabinoxylan, protein extracted from wheat seedlings very likely contains the enzyme responsible for addition of ferulate on arabinoxylan. Our hypothesis is that the enzyme responsible for arabinoxylan feruloylation is a BAHD acyltransferase, a cytosolic protein. Therefore, we hypothesized that AraFAT would be present in the soluble fraction of our protein extract. If we extract the total soluble protein from wheat and assay it with feruloyl-CoA, the donor substrate used by BAHD acyltransferases, alongside a number of arabinose substrates, then we could determine three critical things: 1) an assay for determining AraFAT, 2) the acceptor substrates for the AraFAT reaction, and 3) we could identify the protein(s) responsible for this reaction using proteomics.

In this chapter, I present our biochemical approaches with the aim of determining the gene responsible for the arabinoxylan ferulate acyltransferase (AraFAT). We successfully cloned, expressed, and purified several BAHD acyltransferase candidates using heterologous expression in *E. coli* and in CFPS. We were able to characterize a number of these candidates using an artificial acceptor substrate, methanol, which acts as an acceptor by BAHD ATs, resulting in methyl-ferulate and methyl-*p*-coumarate conjugates. This “methanol assay” gives us

insight into CoA-acyl donor preferences and informs us about individual BAHD AT activity. Assaying these proteins with a variety of arabinose substrates proved unsuccessful, as we did not find a feruloylated arabinose product. However, our experiments using wheat protein extract lacked extensive investigation, and given more time, we believe this approach could be a powerful tool to discover AraFAT activity and the AraFAT gene.

MATERIALS AND METHODS

Selection of candidate genes

The *Brachypodium* genes selected for study were grass-specific BAHD acyltransferases, some of which belonged to the PF02458 clade (Mitchell *et al.*, 2007). An RNA-seq gene expression dataset, collected in our lab by Dr. Jacob Jensen, was used to find genes highly expressed during elongating internode stages BS04-BS05, developmental stages where saponifiable ferulates are highly present in the cell.

Enzymatic synthesis of CoA-substrates using 4CL2

Feruloyl- and *p*-coumaroyl-CoA were enzymatically synthesized with 4-coumaroyl:Coenzyme A ligase 2 (4CL2), using a modified method from Beuerle and Pichersky (2002). 4CL2 was codon optimized and synthesized using DNASTrings (GeneArt by Thermo Fisher Scientific) and cloned into the pMAL-c5x vector, which adds a maltose-binding protein (MBP) affinity tag to the protein of interest. Purified 4CL2 (0.5 mg) was incubated with 3 mg ferulic or *p*-coumaric acid, 2 mg Coenzyme A hydrate, and 6.9 mg ATP in 50 mM Tris-HCl, pH 7.5, a 10 mL reaction volume for five hours, after which an additional 6.9 mg ATP, 2 mg CoA hydrate, and 0.5 mg 4CL2 was added and incubated overnight. Products were purified using Sep-Pak C18 1cc Vac Cartridges (Waters) preconditioned with 4% ammonium acetate.

The reaction mixture was adjusted to 4% ammonium acetate, and loaded onto the cartridges (3-4 cartridges per 10 mL). Aromatic CoA esters bind to C18 columns in the presence of ammonium acetate, and are able to be recovered using water (Beuerle and Pichersky, 2002). The cartridges were washed with 5 CVs of 4% ammonium acetate and the absorbance at 260nm was below 0.050 AU. Feruloyl- and *p*-coumaroyl-CoAs were eluted from the column using 1 mL water. The resulting product was evaluated for purity using reversed-phase chromatography (RPC) (method shown below). Concentrations were measured using UV spectrophotometry and molar absorption coefficient value of 19000 M⁻¹ cm⁻¹ (Gross and Zenk, 1974).

Cloning, expression, and purification of BAHD acyltransferases in *E. coli*

Primers were designed to amplify some genes from *Brachypodium distachyon* ecotype Bd21-3 cDNA developmental stages 2-5. Some genes (Bradi1g36990, Bradi3g12497, Bradi4g14210, and Bradi5g01240) were not successfully amplified from cDNA by PCR. These genes were codon optimized and synthesized using DNAStrings (GeneArt by Thermo Fisher Scientific). The BAHD AT genes were cloned into several different vectors, including pDEST17 (N-terminal 6xHis-tag), pMAL-c5x (N-terminal MBP-tag), pETITE (N-terminal SUMO-tag), and pET-DEST42 (a C-terminal 6xHis-tag vector, but can be used as a tagless construct if stop codon is kept). *E. coli* containing the above plasmids were grown in 2 L baffled flasks containing 500 mL LB media with appropriate antibiotics at 37°C, 200 rpm. Once the cultures reached an OD₆₀₀ of between 0.4 and 0.6, expression of the protein encoded by the plasmid was induced by adding IPTG to a concentration of 0.2 mM and incubating overnight at 15°C with shaking (200 rpm). Cells were pelleted (4000 rpm for 10 minutes at 4°C) and stored at -80°C until used for purification. Cells were lysed using a Constant Systems Ltd. TS benchtop cell disruptor at 20

kPSI and soluble protein was purified using affinity columns on the FPLC (method described below).

Expression of genes using Cell-Free Protein Synthesis

The wheat germ cell-free protein synthesis system Premium PLUS Expression Kit was obtained from CellFree Sciences. The candidate genes were cloned into the vector pEU-E01-MCS using InFusion cloning at the XhoI and Sall sites. The clones were sequenced to verify correct DNA sequence. The DNA was then transcribed into mRNA using their Transcription Premix LM, followed by translation into protein using the WEPRO9240 (wheat germ extract containing creatine kinase) and SUB-AMIX SGC translation buffer. Protein sizes were analyzed on a TGX gel (Bio-Rad).

Protein purification with HisTrap and MBPTrap columns using Fast Protein Liquid Chromatography (FPLC)

Affinity chromatography was performed using an FPLC and the affinity columns HisTrap and MBPTrap for His-tagged and MBP-tagged proteins, respectively. For His-tag fusion proteins, the buffers used were Binding (20 mM Tris-HCl, pH 7.5, 500 mM NaCl), Wash (20 mM Tris-HCl, pH 7.5, 500 mM NaCl 20 mM imidazole) and Elution (20 mM Tris-HCl, pH 7.5, 150 mM NaCl, 500 mM imidazole). After binding to the HisTrap column with Binding Buffer, the column was washed for 2 column volumes (CVs) with Wash Buffer. Still in Wash Buffer, elution of the His-tag protein was performed using a gradient of 0 to 100% of Elution Buffer over 20 mL. For MBP-tag fusion proteins, buffers used were Binding (20 mM Tris-HCl, pH 7.5, 200 mM NaCl) and Elution (20 mM Tris-HCl, pH 7.5, 200 mM NaCl, 10 mM maltose). After binding to the column, elution of the MBP-fusion protein occurred by an immediate ramp to 100% Elution Buffer for 5 CVs. Peaks corresponding to eluted protein were run on SDS-PAGE for analysis.

Enzymatic assays

Assays were performed in 50 μ L volumes in 50 mM Tris-HCl, pH 7.5. CoA substrates and acceptor substrates were added to 1 mM final concentrations. Incubations were performed at room temperature at varying lengths of time, but typically for 1 hour (unless stated otherwise). To stop the reactions, 1 μ L of 10 M HCl was added to each tube and then brought up to 50% methanol in 200 μ L volumes. The reactions were filtered through Amicon 10K 0.5 mL filters for 10 minutes on maximum speed and 10 μ L of the flowthrough was injected into the UPLC for analysis.

Reversed Phase Liquid Chromatography (RPC) analysis

All assays were analyzed on a Waters UPLC H-series using a BEH C18 column, 1.7 μ M, 2.1 x 100 mm. Flow rate of 0.300 mL/min, Solvent A: ddH₂O with 0.1% TFA and Solvent B: 100% ACN. The linear gradient began with 87% Solvent A, moving to 60% Solvent A from 0 to 7 minutes. The gradient goes to 100% B from 7 to 8 minutes, and at 11 minutes returns then to 87% A for 4 minutes. Compounds were detected using TUV detector at 262 and 340 nm wavelengths.

Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS) with C18 column

Enzymatic assay reaction products were analyzed at the Mass Spectrometry Facility at Michigan State University using a G2-XS QToF instrument. Separation was performed on the a Waters Acquity BEH C18 column, 1.7 μ M, 2.1 x 100 mm, with a slightly modified RPC gradient: Solvent A is ddH₂O with 0.1% formic acid and Solvent B is 100% ACN. The flow rate is 0.300 mL/min, and begins with 87% A. From 0 to 7 minutes, the Solvent A goes from 87% to 60%, and between 7 and 8 minutes ramps up to 100% B (until 10 minutes). At 10.01 minutes, it returns to 87% A and equilibrates until 12 minutes. Mass spectrometry analysis was performed

in positive and negative electrospray ionization (ESI) modes searching for a mass range between 50 and 1500 Da, with ramped collision from 20-80 eV to fragment the compounds. Capillary was set at 2.0 kV and lockspray capillary at 2.5 kV. Scan duration was 0.200 seconds, with source and desolvation temperatures at 100 and 350 degrees Celsius, respectively. Cone gas flow and desolvation gas flow rates were at 25.0 and 600 L/Hr, respectively.

Wheat seedling growing conditions

Triticum wheat seeds were soaked in water for 3-4 hours before planting to improve germination. To keep them aerated during this time, we inserted a hose with air into the beaker. Presoaked seeds were then planted in medium grade vermiculite and transferred to a growth chamber with 16/8 hour light conditions.

Wheat seedling protein extraction

Six to seven days after planting, above-ground tissue of wheat seedlings was collected into a pre-chilled glass beaker on ice. Protein extraction buffer, 20 mM Tris-HCl, pH 7.5, 150 mM NaCl, 1% ascorbic acid, and protease inhibitor tablet (1 per 100 mL; from Sigma), was added to the beaker until seedlings were submerged (150-200 mL). Seedlings were sheared into smaller pieces with the scissors before using the Thermo Fisher immersion blender to disintegrate cells and homogenize tissue. After homogenization, Miracloth was used to separate liquid containing protein and fibrous plant material.

To remove cell wall debris, intact cells, and larger organelles, the Miracloth-filtered extract was spun in a Beckman-Coulter Allegra 25R table-top centrifuge with fixed-angle rotor TA10.250 for 10 minutes at 10,000 x g at 4°C. The supernatant was measured and used for ammonium sulfate protein precipitation. Using EnCor Biotechnology, Inc.'s calculator (<http://www.encorbio.com/protocols/AM-SO4.htm>), the extract was brought to 80% ammonium

sulfate saturation and equilibrated on ice for 2 hours with constant stirring. The ammonium sulfate mixture was then centrifuged for 10 minutes at 10,000 x g at 4°C to pellet the protein. Once the supernatant was decanted, the pelleted protein was resuspended in 5 mL of sodium phosphate buffer (50 mM sodium phosphate, pH 7, 150 mM NaCl, protease inhibitor tablet). To separate the soluble from insoluble protein, this extract was centrifuged in a Beckman-Coulter Optima L-90K Ultracentrifuge with fixed-angle rotor 50 TI for 30 minutes at 50,000 x g at 4°C.

FPLC desalting of wheat protein extract

The soluble protein was separated from small molecules using a HiTrap desalting column (four 5 mL columns) on the FPLC using a buffer of 50 mM sodium phosphate, pH 7.5, 150 mM NaCl. Fractions were pooled together, and an aliquot was assayed for AraFAT activity.

FPLC anion exchange to separate proteins by charge

Desalted protein was further separated by charge using a CaptoQ anion exchange column (1 mL, GE Healthcare). Solvent A was 20 mM Tris-HCl, pH 7.5 and Solvent B was 20 mM Tris-HCl, pH 7.5, 1 M KCl. A total of 30 fractions were collected in 0.5 mL volumes over 15 mL of an elution gradient from 10% to 100% B, and every other fraction was assayed for AraFAT activity.

Assaying extracted protein for AraFAT activity

Protein samples at different steps (desalted fractions and anion exchange fractions) were used in protein assays to determine if predicted AraFAT activity was found. Several arabinose substrates were used (see Chapter 1 **Figure 10**) to incubate with the protein and feruloyl-CoA. We also performed assays that included UTP and ATP, as well as with and without TCEP. After various incubation times (1-2 hours to overnight), the reaction was stopped by adding ACN containing 0.1% TFA to final concentration of 66%. Samples were separated and analyzed using RPC using LC-MS/MS in both positive and negative modes.

Liquid Chromatography-Tandem Mass Spectrometry (LC/MS/MS) with Amide column

Some enzymatic assay reaction products were analyzed at the Mass Spectrometry Facility at Michigan State University using a G2-XS QToF instrument. Separation was performed on a BEH Amide column, 1.7 μ M, 2.1 x 100 mm, with the gradient: Solvent A, 10 mM aqueous ammonium acetate (NH₄OAc) and Solvent B is 100% ACN. The flowrate is 0.300 mL/min, and begins with 5% A. From 1 to 9 minute, the Solvent A goes from 5% to 65%, and remains until 10 minutes. At 10.01 minutes, it returns to 5% A and equilibrates until 14 minutes. Detection was performed using electrospray ionization in both positive and negative electrospray ionization (ESI) modes searching for a mass range between 50 and 1500 Da, with ramped collision from 20-80 eV to fragment the compounds. Capillary was set at 2.0 kV and lockspray capillary at 2.5 kV. Scan duration was 0.200 seconds, with source and desolvation temperatures at 100 and 350 degrees Celsius, respectively. Cone gas flow and desolvation gas flow rates were at 25.0 and 600 L/Hr, respectively.

RESULTS

Bioinformatics approach to selecting AraFAT candidates

The bioinformatics approach to finding enzymes involved in arabinoxylan biosynthesis as shown in Mitchell *et al.* (2007) provided a great starting point for the search for the arabinoxylan ferulate acyltransferase. Their hypothesis was that putative glycosyl transferase (GT) genes that are highly differentially expressed in monocots compared to eudicots would lead to discovering genes involved in arabinoxylan biosynthesis, the major hemicellulose component in monocots. Another unique feature of the monocot cell wall is the feruloylation of arabinoxylan. When looking at the putative GT families likely involved in arabinoxylan biosynthesis, one family of

BAHD acyltransferases, PF02458, was highly coexpressed with these GTs. Because BAHD acyltransferases utilize CoA-acyl substrates, the hypothesis was that this could be the enzyme adding ferulate (in the form of feruloyl-CoA) to arabinose/arabinoxylan. Further, this is a grass-specific clade of acyltransferases, which implies that these enzymes might be more likely than other protein families to be performing grass-specific modifications.

One approach to finding the AraFAT gene would be to clone all the genes in the PF02458 clade and assay them for AraFAT activity. Given the large number of genes in the clade, this approach would require cloning, protein expression, and assaying of a large number of candidates. In an effort to reduce the number of candidate genes, as well as broaden our search for potential BAHD acyltransferases outside of the PF02458 clade, we decided to look at a gene expression dataset produced by RNA-seq analysis of different developmental stages of *Brachypodium distachyon* internodes (**Figure 7**). We also measured cell wall-bound saponifiable monoferulates from these samples in an effort to understand amounts of ferulate available during each developmental stage of *Brachypodium* (**Figure 7C**). We hypothesized that the gene responsible for adding ferulate to arabinose(-xylan) would be expressed as ferulate is present in the cell. Analyzing the amount of ferulate versus developmental stage, there is a significant increase in ferulate from stem stage BS03 to stage BS04, where it remains fairly level across the older stem internodes. We therefore decided to look at BAHD acyltransferase genes with an expression pattern that has high expression in elongating stems corresponding to stages BS04-05 (**Figure 8**).

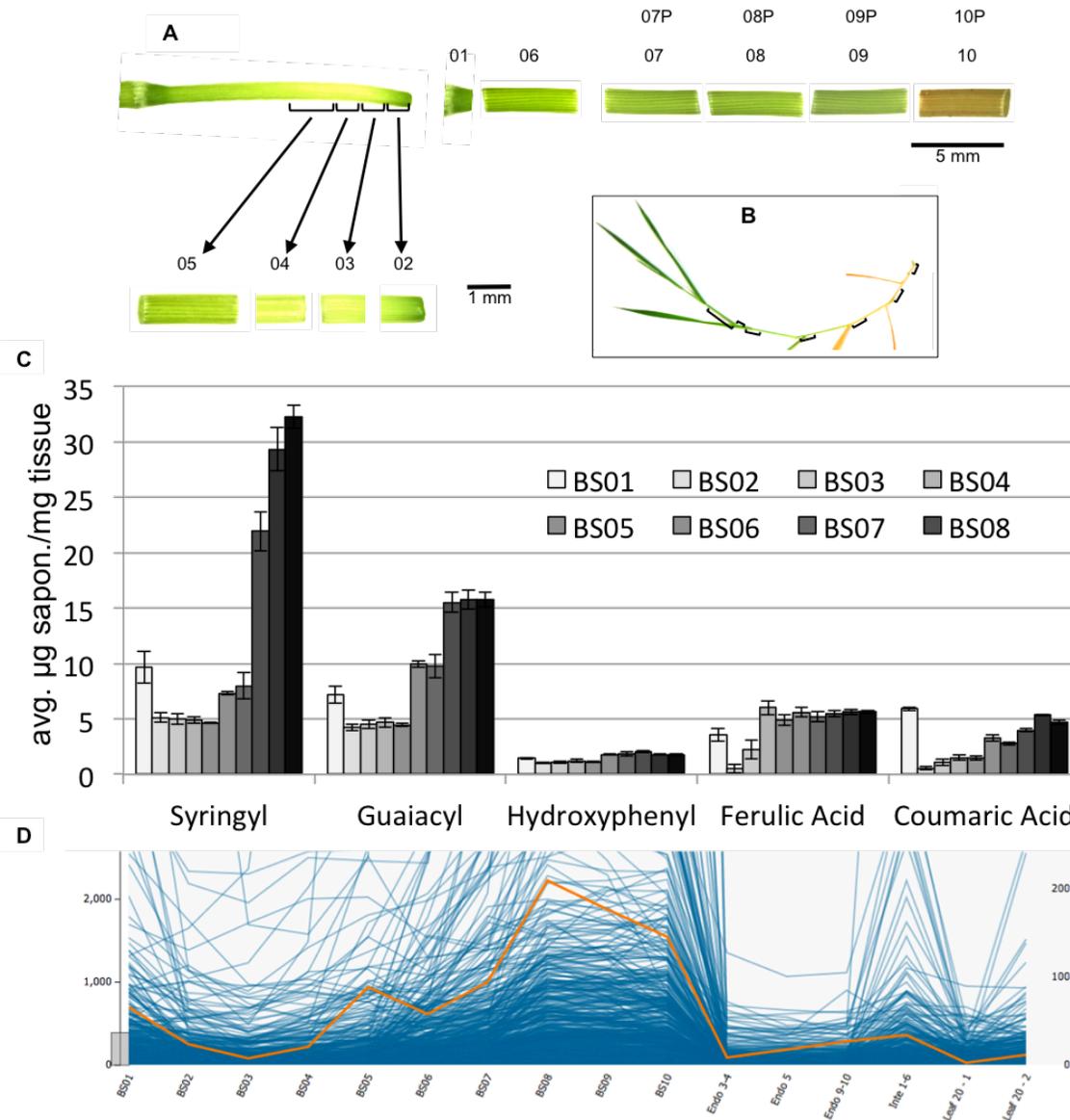


Figure 7: Cell wall composition and RNA-seq dataset of *Brachypodium* developmental stages. Panel A shows the fractionated internodes of *Brachypodium* stem. Section 01 is a node, containing both the youngest and oldest stem tissue. From stages 02-10, the developmental stages go from youngest to oldest internode. Stages 02-05 represent the elongating internode as it grows into a fully elongated internode, shown by stages 06-10. The full stem with leaves is pictured in panel B. These stem and tissue sections were used for RNA-seq experiments to gather expression data at each stage. Tissue was also gathered to analyze the cell wall components shown in panel C. Lignin Composition in the form of extractable syringyl (S), guaiacyl (G), and hydroxyphenyl (H) lignin monomers was performed as described in Foster *et al.* (2010). The pattern shows the lignification that occurs in later stages of stem development. Saponifiable hydroxycinnamates show the cleavable ester-linked (mono)ferulic and *p*-coumaric acids of each developmental stage. BS is *Brachypodium* Stage and corresponds to the number system shown in panel A. The figure in panel D is from the GLBRC Genome Suite coexpression viewer that shows genes coexpressed with a gene of choice. In this image, the orange line corresponds to

Figure 7 (cont'd): Bradi2g33980, a PF02458 clade BAHD acyltransferase. This is an example of a gene in the PF02458 clade that would not be the best candidate based on its expression pattern (higher in later stages). The Y-axis is the normalized reads. BS, Brachypodium stem stages 01-10 (see Figure 7A); Endo 3-4, endosperm 3-4 days after seed filling (DASF); Endo 5, endosperm 5 DASF; Endo 9-10, endosperm 9-10 DASF; Inte, integument 1-6 DASF; Leaf 20-1, leaf of 20 day old plants set 1; Leaf 20-2, leaf of 20 day old plants set 2. Y-axis is in normalized RPKM.

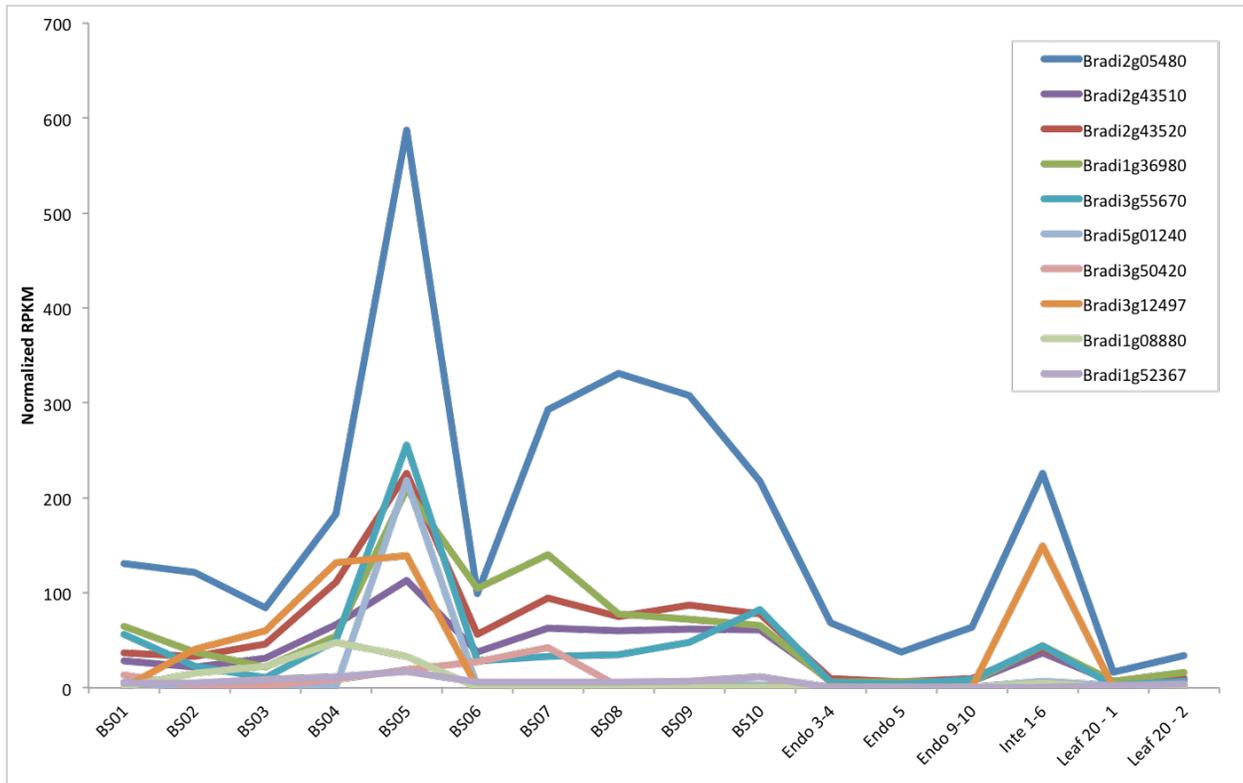


Figure 8: Expression patterns of the top ten Brachypodium AraFAT candidates. Most of the candidates had highest expression in stages BS-04-05, when ferulate was present in the cell. BS, Brachypodium stem stages 01-10 (see Figure 7A); Endo 3-4, endosperm 3-4 days after seed filling (DASF); Endo 5, endosperm 5 DASF; Endo 9-10, endosperm 9-10 DASF; Inte, integument 1-6 DASF; Leaf 20-1, leaf of 20 day old plants set 1; Leaf 20-2, leaf of 20 day old plants set 2. Y-axis is in normalized RPKM.

Cloning, expressing, and purifying candidates in *E. coli* and in a cell-free system

AraFAT candidate genes were cloned into a number of vectors with various affinity tags (Table 1). Based on our prior success with 6xHis-tags, we began to clone genes into the pDEST17 Gateway vector. We also cloned a number of candidate genes into the pMAL-c5x vector designed with a tobacco etch virus (TEV) recognition sequence downstream of the MBP-tag, which created an N-terminal MBP-tag fusion protein with a cleavage site to remove MBP. MBP-tags are advantageous because they assist in protein folding and solubility, as well as a highly specific way of purifying the protein. The end product is an MBP-fusion protein with very little other proteins co-purified. After expression and purification of our candidate proteins, they were then assayed for enzyme activity.

<i>location of tag</i>	pMAL-c5x (MBP-tag) (N-term)	pDEST17 (6xHis-tag) (N-term)	FLAG-tag (N-term)	C-tag (C-term)	pET-DEST42 (no tag) (n/a)	pETite (SUMO tag) (N-term)	pEU-E01 (for CFPS) (n/a)
Brachy BAHD AT							
Bradi2g05480	—	—	—	—	—	—	✓
Bradi2g43510	—	—	—	—	—	—	✓
Bradi2g43520	✓	—	✓	—	—	—	✓
Bradi1g36980	✓	✓	—	—	—	—	✓
Bradi3g55670	—	—	—	—	—	—	✓
Bradi5g01240	—	✓	—	✓	✓	✓	✓
Bradi3g50420	—	—	—	—	—	—	✓
Bradi3g12497	✓	—	✓	—	—	—	—
Bradi1g08880	—	—	—	—	—	—	UC
Bradi1g52367	—	—	—	—	—	—	✓
Bradi1g36990	✓	✓	✓	✓	✓	✓	✓

Table 1: Cloned constructs of AraFAT candidates in various vectors heterologously expressed in *E. coli*. Constructs that were successfully cloned and expressed protein are labeled with a check. UC is an unsuccessful clone. Dashes show that clones were not attempted with the gene and/or vector. The location of tag signifies if the tag is on the N- or C-terminus of the protein. n/a signifies that there is no tag. In the case of Bradi1g08880, cloning attempts into the pEU-E01 vector were unsuccessful. Bradi1g36990 is the gene studied by Bartley *et al.* and shown to have genetic evidence it adds *p*-coumarate to arabinoxylan (Bartley *et al.*, 2013). All genes and constructs cloned were sequenced to verify the correct DNA sequence.

During our work to characterize the ferulate monolignol transferase (FMT) enzyme, we discovered that this enzyme used methanol as an acceptor (Wilkerson *et al.*, 2014). Prior to RPC analysis, FMT reaction mixtures were brought up to 50% methanol by addition of methanol. This step was performed to solubilize certain feruloyl-monolignol compounds and improve peak separations on the C18 column of the UPLC. Once samples were separated by RPC, it was discovered that the remaining feruloyl-CoA was consumed, and a novel peak occurred, identified using standards as methyl-ferulate. It appears that FMT readily uses methanol as an acceptor substrate. We tested to determine if other BAHD ATs are capable of using methanol as an artificial acceptor. Many of the BAHD AT candidates we have studied are able to conjugate methanol with CoA-acyl substrates, including *p*-coumarate monolignol transferase (PMT)

The reaction of FMT with methanol to make methyl-ferulate taught us two things: first, all enzyme reactions should be stopped by either by acidification or by first precipitating the protein by the addition of ACN. Second, the discovery that methanol can be used as an artificial acceptor allowed us to test the CoA-substrate preferences for BAHD acyltransferases. Determining which candidates used feruloyl-CoA as a substrate allowed us to further restrict the proteins we needed to assay with various arabinose substrates.

Our first step after purifying a BAHD AT was to characterize the enzyme by assaying it with methanol and feruloyl-CoA or methanol and *p*-coumaroyl-CoA. These methanol assays were analyzed by RPC and the retention time of the product peaks were compared with methyl-ferulate and methyl-*p*-coumarate standards (**Figure 9**). The CoA-acyl donor substrates that each enzyme used are listed in **Table 2**.

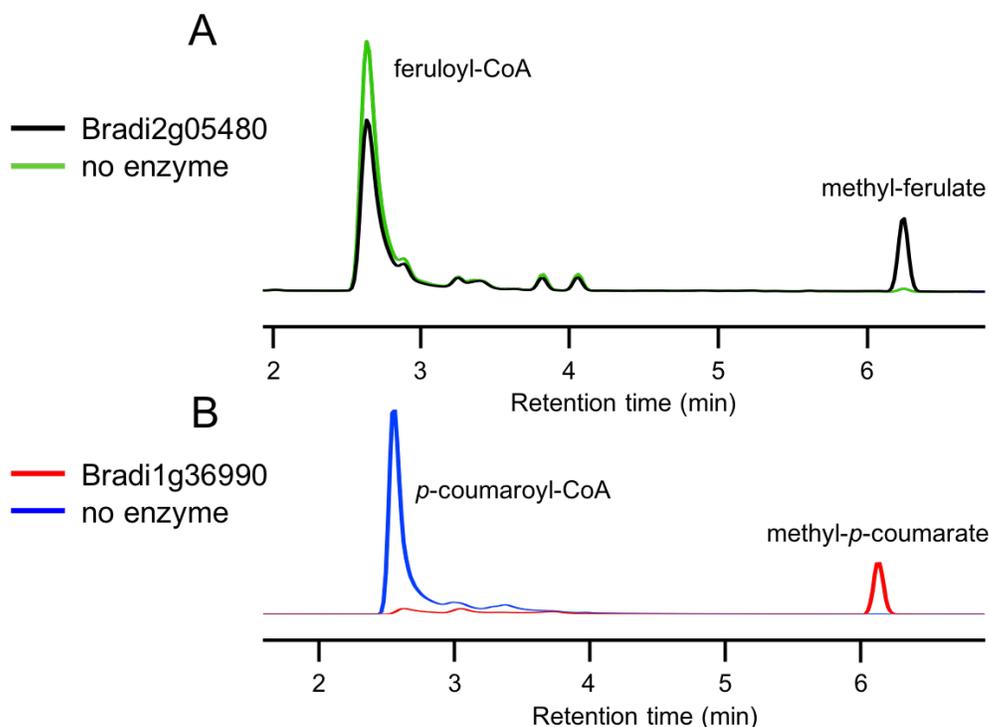


Figure 9: RPC separation of a methanol assay with feruloyl-CoA and *p*-coumaroyl-CoA. Bradi2g05480 was incubated with feruloyl-CoA (A) and Bradi1g36990 was incubated with *p*-coumaroyl-CoA (B) for 1 hour each before the reaction was analyzed.

Candidate	CoA Activity
Bradi2g05480 *	ferulate
Bradi2g43510 *	none
Bradi2g43520 *	none
Bradi1g36980 *	ferulate
Bradi3g55670	ferulate and <i>p</i> -coumarate
Bradi5g01240	ferulate and <i>p</i> -coumarate
Bradi3g50420	ferulate and <i>p</i> -coumarate
Bradi3g12497	none
Bradi1g08880	n/a
Bradi1g52367	none
Bradi1g36990 *	ferulate and <i>p</i> -coumarate

Table 2: CoA-acyl substrate preference of AraFAT candidates. Each BAHD AT CoA substrate activity and preference was determined with the methanol activity assay. For the enzymes labeled none, no activity with any CoA-substrate was detected. n/a for Bradi1g08880 is due to an unsuccessful cloning. Asterisks (*) depict genes in the PF02458 clade.

The enzymes having activity with feruloyl-CoA were then assayed with feruloyl-CoA and a variety of arabinose substrates. We hypothesized that one of these candidate enzymes would use feruloyl-CoA and an arabinose-containing compound. As stated in the introduction, the substrates used in the AraFAT reaction are unknown, and therefore we chose to assay our candidate enzymes with a series of arabinose substrates (**Figure 10**). With the transfer of ferulate to arabinose, we would expect to see a consumption of feruloyl-CoA and the occurrence of a novel peak. Since Feruloyl-CoA and ferulate conjugates absorb at 340 nm, we monitored our RPC separations of the reaction products at this wavelength. These enzyme assays were typically performed using 0.5-1 µg of protein and 1 mM of each substrate (feruloyl-CoA and an arabinose acceptor). In all assay conditions and with all forms of arabinose, we were unsuccessful at finding any AraFAT activity with our candidates. In order to eliminate the possibility that the protein we purified was inactive due to the presence of the affinity tag, we removed the N-terminal MBP-tag using TEV protease, but were still unable to observe AraFAT activity.

We also cloned some genes into vectors using alternative affinity tags or placed the tag on the C-terminus. We also cloned some candidates fused to the following tags: C-tag, a C-terminal affinity tag consisting of only four amino acids (EPEA) able to be purified using CaptureSelect Affinity Matrix (Thermo Fisher); FLAG-tag, an N-terminal affinity tag consisting of eight amino acids (DYKDDDDK) with mild elution conditions; and SUMO-tag, a Small Ubiquitin-like Modifier protein that is similar to MBP in that it improves protein solubility and includes a protease cleavage recognition site with highly efficient SUMO (ULP-1) protease to remove SUMO-tag from our protein of interest (Champion pET SUMO Expression System from Thermo Fisher).

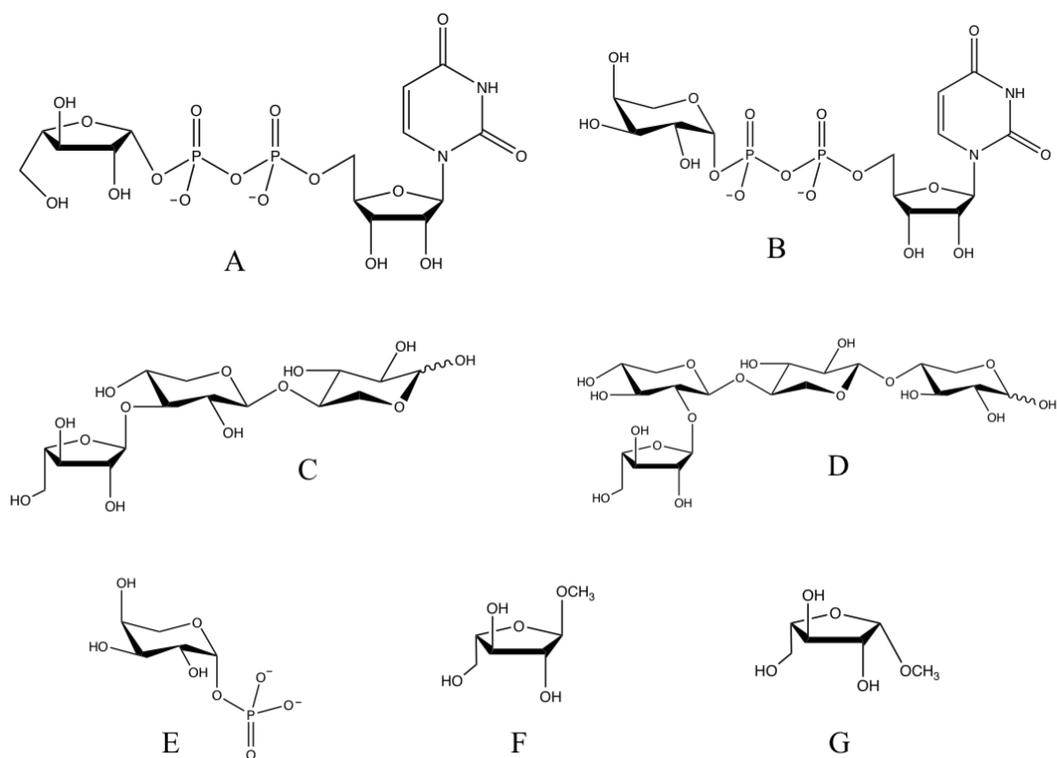


Figure 10: Arabinose-containing substrates used in the AraFAT assays. **A**, UDP-arabinofuranose (UDP-Araf); **B**, UDP-arabinopyranose (UDP-Arap); **C**, 3²-α-L-arabinofuranosyl-xylobiose (A3X); **D**, 2³-α-L-arabinofuranosyl-xylotriose (A2XX); **E**, arabinose-1-phosphate (A1P); **F**, methyl-α-L-arabinofuranoside (Me-α-Ara); **G**, methyl-β-L-arabinofuranoside (Me-β-Ara).

After expressing these constructs in *E. coli*, and purifying the recombinant protein, we performed a series of arabinose-containing assays. None of the enzymes we assayed had AraFAT activity. A few possibilities for these results include: 1) we were assaying with the wrong substrates; 2) we were looking at the wrong enzymes; 3) we were missing other proteins involved in the reaction; or 4) the affinity-tags could be interfering with the reaction. While the first two possibilities are likely, we wanted to exhaust all of our options before abandoning our candidates. There were ways to address the latter two possibilities, one of which was to use a cell-free protein synthesis (CFPS) system prepared using wheat germ (WGX).

Cell-free protein synthesis (CFPS) is a tool used for producing large amounts of protein that has been developed and is now being commercialized for use in research labs. As its name implies, CFPS produces protein outside of a cell. The cell machinery is still required, provided in the form of protein extracts, in our case, wheat germ extract. It works using a ‘bilayer system’: the bottom layer contains the reaction mixture and the wheat germ extract, slightly heavier than the substrate solution that forms on top (**Figure 11**) (Takai *et al.*, 2010). This bilayer acts to remove byproducts and inhibitors formed by the reaction over time by diffusing between lower and upper layers. In the upper layer, the buffered substrate solution contains amino acids, ATP, and GTP, required for the production of protein. The bilayer system also includes creatine kinase and creatine phosphate, both used to regenerate ATP for continued protein expression (Anderson *et al.*, 2015; Sawasaki *et al.*, 2002; Takai *et al.*, 2010).

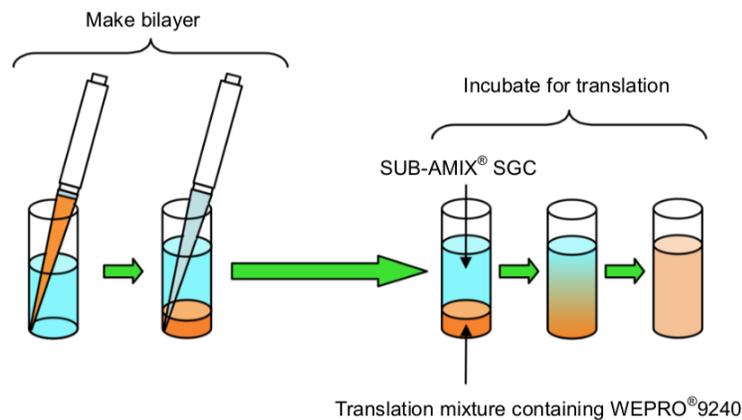


Figure 11: Schematic cartoon of the cell-free protein synthesis bilayer reaction translating the mRNA into protein. On the left side, it shows how to create the bilayer by carefully pipetting the RNA and WEPRO-9240, wheat germ extract plus creatine kinase, mixture (depicted in orange) to the bottom of the well containing the SUB-AMIX-SGC translation buffer (blue). The green arrow represents the translation that occurs in a 20 hour incubation at 15°C, with *substrates* passing through the permeable bilayer to replenish depleted *substrates*. (Figure from CellFree Sciences Co., Ltd., Premium PLUS Expression Kit_E_ver.1.2, Feb 8, 2013).

There are a number of benefits to using CFPS-WGX as a means of expressing and assaying our AraFAT candidates. For one, the transcription from DNA to mRNA, followed by translation into protein, altogether takes about a day compared to the several days for producing *E. coli* clones containing the expression plasmid protein, with additional time for protein expression and purification. Because of the ease in expressing protein and the small volume reaction mixture (~200 μ l), we were able to perform several CFPS reactions in a day. This allowed us to assay a number of candidates in a short amount of time without the need to optimize their expression and purification in *E. coli* (Takai *et al.*, 2010). Finally, the benefit of assaying our proteins in the wheat germ reaction. Wheat is a member of the *Poaceae* family and is known to have an abundance of arabinoxylan in its germ, with measurable feruloyl esters bound to AX (Saulnier *et al.*, 2007) This means that it is very likely its wheat germ extract includes the enzyme(s) involved in this reaction. If there are other proteins that assist in AX feruloylation, such as the UDP-arabinopyranose mutase (UAM), they are likely present in the wheat germ extract as well. Using CFPS-WGX was successful at producing a large amount of protein in a short amount of time. We were able to clone and express several of our candidates using this expression system (**Figure 12**).

Upon assaying the genes for activity using the methanol assay in the CFPS-WGX system, we found that most of them showed activity with either feruloyl- or *p*-coumaroyl-CoA, or both (see **Table 2**). Two candidates, Bradi2g43510 and Bradi2g43520, did not show any activity with either CoA-substrates, so it is possible they use different CoA-acyl donors, or they are examples of BAHD ATs that are not active with methanol. For all of the synthesized proteins generated by CFPS, we performed assays with our list of arabinose candidates, and were still unsuccessful at finding any AraFAT activity.

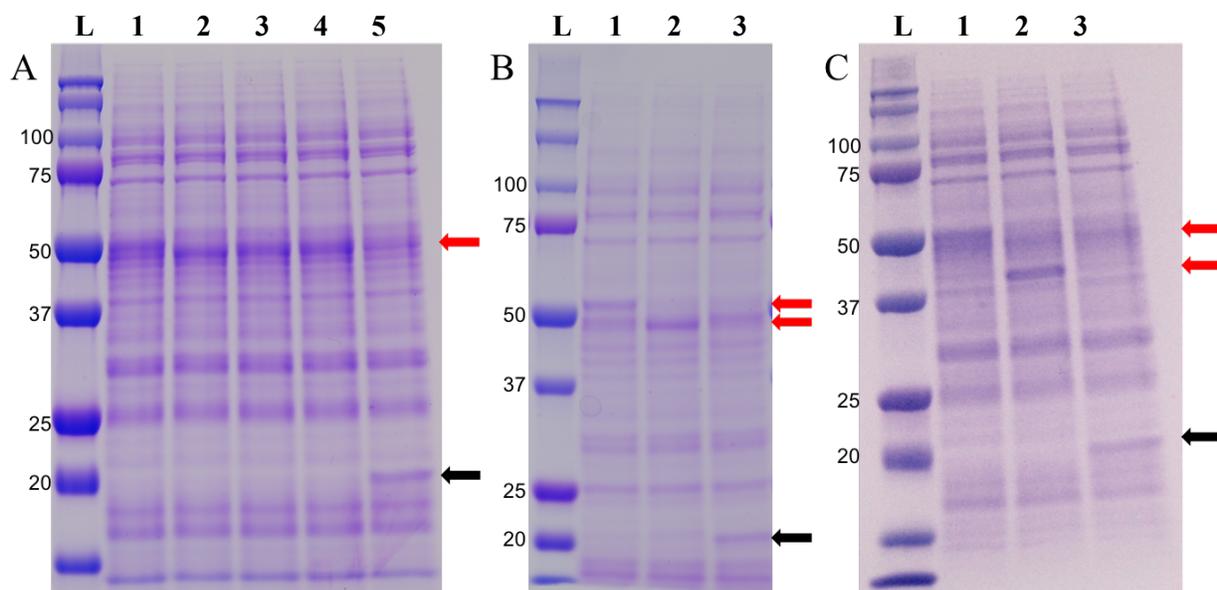


Figure 12: SDS-PAGE gels of CFPS-WGX synthesized protein of AraFAT candidates. In gel **A**, the proteins synthesized are Bradi2g05480 (1), Bradi2g43510 (2), Bradi2g43520 (3), Bradi1g36980 (4), and the control DHFR (5). In **B**, Bradi1g36990 (1), Bradi5g01240 (2), and the control DHFR (3). In **C**, Bradi3g55670 (1), Bradi3g50420 (2), and the control DHFR (3). BAHD ATs range from 45-55 kDa, and their bands are indicated by the red arrows on the right side of each SDS-PAGE. DHFR runs at ~20 kDa, and is indicated on the gels by the black arrows. The ladder (L) is Precision Plus Dual Color Protein Ladder (Bio-Rad).

Characterization of Bradi1g36990 in an effort to find a testable assay for AraFAT

As described in the previous section, our attempts to clone, express, and purify candidates and assay them with arabinose substrates proved unsuccessful in finding the enzyme responsible for the AraFAT reaction. One concern was that our candidates contained the AraFAT gene, but that we lacked a testable assay to demonstrate AraFAT activity. In an effort to elucidate a testable assay for AraFAT, we selected the gene Bradi1g36990, ortholog to the OsAT10 gene studied by Bartley *et al.* (2013) that showed strong genetic evidence of being responsible for *p*-coumarate on AX, hereafter referred to as arabinose *p*-coumarate acyltransferase (AraPAT). It was advantageous that a group had previously studied the gene in rice and found an increase in *p*-coumarate on arabinoxylan when overexpressing it. Another advantage is that this reaction is likely performing an almost identical reaction to AraFAT, only instead of ferulate, it adds *p*-

coumarate. This knowledge gave us an opportunity to test an enzyme that is likely AraPAT, develop an assay to prove *p*-coumarate transfer onto arabinose, and then use that same approach to test for AraFAT activity with our candidate genes and feruloyl-CoA.

The expression pattern of Bradi1g36990 as measured in our Brachypodium RNA-seq dataset shows a relatively low expression throughout all developmental stages except for a peak of expression during stem stage BS05. This is a stage that has elongating internodes, which our assays show to have a large pool of ferulate (see **Figure 7C**) (Jung, 2003; Obel *et al.*, 2002). This gives us confidence that our hypothesis that an enzyme utilizing ferulate and/or *p*-coumarate for arabinoxylan substitution would have large gene expression in stem stages 04-06 (**Figure 13**).

Bradi1g36990 was characterized using the “methanol assay” to determine enzyme activity and CoA-substrate preference (**Figure 14**). Bradi1g36990 is able to use both feruloyl- and *p*-coumaroyl-CoAs with methanol to make methyl-ferulate and *p*-coumarate, respectively. The preference between these two CoA-substrates, using a saturation of methanol, was not measured or determined.

Assaying Bradi1g36990 with various arabinose substrates (see **Figure 10**), gave no activity. We would anticipate seeing consumption of feruloyl-CoA and a new product peak forming (at 340 nm UV), but no new products formed. As shown in **Table 1**, this gene was cloned into several different vectors with and without affinity tags, and heterologously expressed in *E. coli*. It was also cloned and expressed tagless in the CFPS-WGX system. In all situations, this protein showed activity with methanol as an acceptor and feruloyl- and *p*-coumaroyl-CoA as

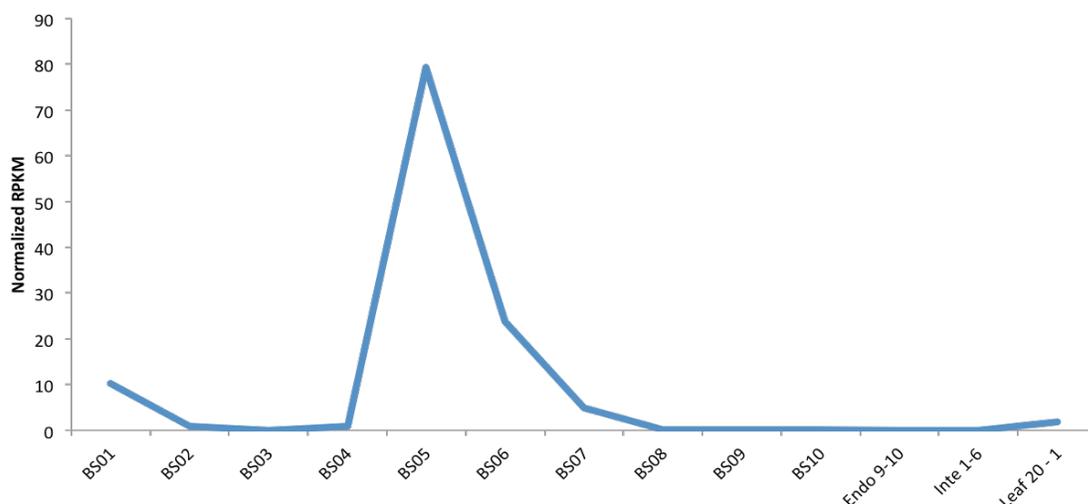


Figure 13: Expression pattern of Bradi1g36990 during different developmental stages in Brachypodium. This gene's expression peaks at elongating internode section BS05, and does not have much expression in other tissues. BS, Brachypodium stem stages 01-10 (see Figure 7A); Endo 3-4, endosperm 3-4 days after seed filling (DASF); Endo 5, endosperm 5 DASF; Endo 9-10, endosperm 9-10 DASF; Inte, integument 1-6 DASF; Leaf 20-1, leaf of 20 day old plants set 1. Y-axis is in normalized RPKM.

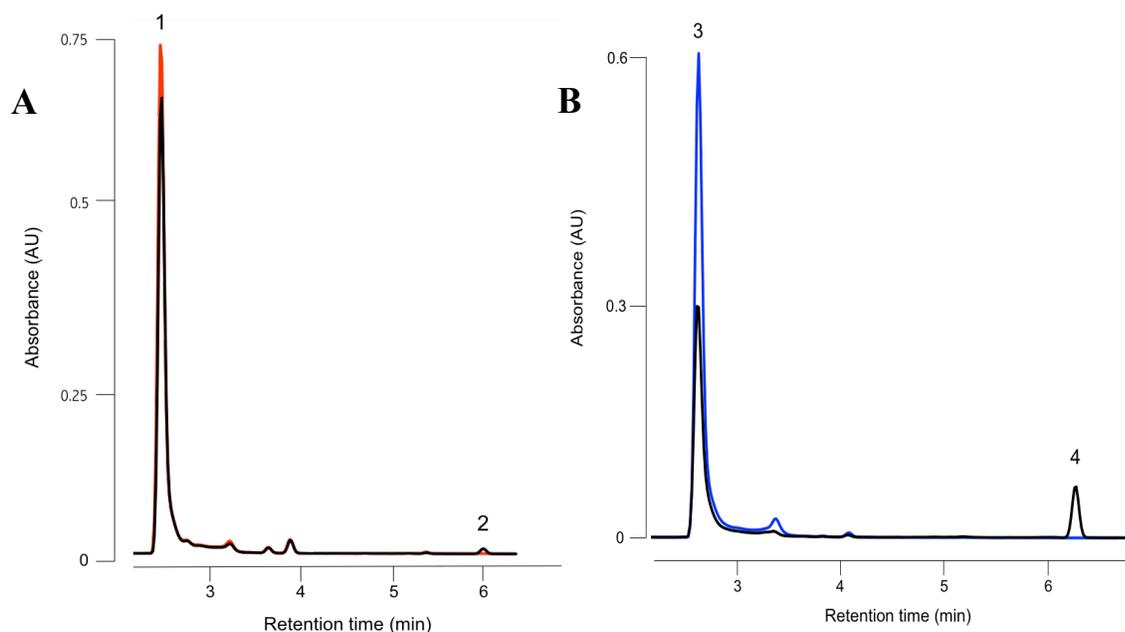


Figure 14: RPC separation of methanol assays with Bradi1g36990 show activity with both *p*-coumaroyl- and feruloyl-CoA. In panel A is \pm enzyme with *p*-coumaroyl-CoA and methanol, and panel B is \pm enzyme with feruloyl-CoA and methanol. Peak 1 is *p*-coumaroyl-CoA, 2 is methyl-*p*-coumarate, 3 is feruloyl-CoA, and 4 is methyl-ferulate. Black lines represent assays with addition of enzyme. The formation of conjugated products proves that the Bradi1g36990 enzyme is active and is able to use both CoA substrates. Minor peak in blue at \sim 3.4 min in B is *cis*-feruloyl-CoA. Other peaks between 3-4 minutes are compounds were not identified, but were commonly found in CFPS-WGX reactions. Detection performed at 340 nm.

donor substrates, leading us to believe it was properly folded and active. However, the enzyme had no activity with any arabinose acceptor substrates. This protein was also assayed with coniferyl, sinapyl, and *p*-coumaryl alcohols to determine if it had PMT or FMT activity.

Wheat seedling protein extraction and arabinose substrate assays

Protein extraction was performed on 6- to 7-day old wheat seedlings and the soluble fraction was desalted to remove small molecules. This desalted fraction was then used to test for AraFAT activity with arabinose substrates, UDP-Araf, UDP-Arap, A1P, and L-arabinose. The donor substrates we used were feruloyl-CoA and feruloyl-glucose, in case the predicted ferulate donor substrate was incorrect. In many assays, we found a number of feruloyl esterases or feruloyl-CoA-utilizing enzymes that consumed feruloyl-CoA to produce ferulic acid. This hydrolysis of feruloyl-CoA suggests that the protein extract contained enzymes that consumed feruloyl-CoA in a non-AraFAT reaction.

After optimization of our protein extraction protocol using 1% ascorbic acid as an antioxidant, we found that there was less conversion or break-down of feruloyl-CoA to ferulic acid. We then repeated our assays with UDP-Araf, UDP-Arap, and A1P and analyzed using RPC. One assay showed all three had small unidentified peaks, close in retention time. It also looked like feruloyl-CoA was almost completely consumed, but there was no break-down to ferulic acid. It appeared that the flow through peak at the beginning of the trace at 262 nm was larger in all three cases, so we thought the product might not be binding to the column. We analyzed these reactions using LC-MS/MS and a BEH Amide column. This separation using liquid chromatography did not show us any peaks containing feruloyl compounds, as we looked for fragment ion *m/z* 193 in negative ion mode. Unfortunately, we did not also run these reactions using RPC on the LC-MS/MS QToF instrument, so we cannot conclude whether or not

the peaks we saw on the UPLC were feruloylated arabinose substrates or not. We plan to re-run this assay and analyze it using chromatography and RPC with LC-MS/MS and identify those peaks.

Our assays with wheat protein extract using feruloyl-CoA or -glucose and arabinose substrates did not result in a positively identified peak corresponding to a feruloyl-arabinose compound. However, the lack of time and extensive research to optimize a method for analyzing these protein assays can account for our failure to find a testable AraFAT assay and product.

DISCUSSION

In this chapter, I have presented our work in search of finding the gene encoding arabinose ferulate acyltransferase (AraFAT), the enzyme responsible for adding ferulate to arabinoxylan. Using a variety of vectors, affinity tags, and protein expression techniques, including the cell-free protein synthesis system, we were able to successfully clone, express, and purify 11 BAHD acyltransferase (AT) candidates. Once purified, these candidates were assayed using an artificial acceptor substrate, methanol, which allows us to determine protein activity and CoA-substrate preference for some BAHD ATs. We determined that two of our candidates, Bradi2g05480 and Bradi1g36980, only used feruloyl-CoA as a donor substrate with methanol. Four candidates, Bradi3g55670, Bradi5g01240, Bradi3g50420, and Bradi1g36990 (Bartley's AraPAT gene), were able to produce both methyl-ferulate and *p*-coumarate, showing activity with both feruloyl- and *p*-coumaroyl-CoAs, respectively. We did find genes, namely, Bradi2g43510, Bradi2g43520, Bradi3g12497, and Bradi1g52367 that had no activity with either CoA-acyl substrates. This suggests that they do not use either of these CoA-substrates, they were misfolded or inactive proteins, or that not all BAHD ATs use methanol as an artificial acceptor.

For a majority of our BAHD ATs, we were able to determine protein activity and CoA preference using methanol as an artificial acceptor. We are interested to learn if other BAHD ATs are able to be characterized as well using this methanol assay.

Our results from assaying these AraFAT candidates with feruloyl-CoA and arabinose-containing substrates did not show any evidence of ferulate addition on arabinose. None of the enzyme reactions analyzed by RPC showed any novel peaks that would correspond to feruloyl-arabinose compounds, leading us to believe that either these enzymes were not responsible for AX feruloylation or that we were missing something in the assay reaction, such as a chaperone protein. We hypothesized that assaying these candidates in the CFPS mixture would allow for incubation with other wheat germ proteins and would therefore address this concern. Unfortunately, this approach also showed us no sign of feruloyl-arabinose conjugation.

One approach to developing an assay for AraFAT was by testing Bradi1g36990, the gene, studied by Bartley *et al.* (2013), thought to encode the arabinose *p*-coumarate acyltransferase, AraPAT. The protein showed activity with both feruloyl- and *p*-coumaroyl-CoAs in the assays with methanol as an artificial acceptor substrate. We hypothesized that assaying the active protein with arabinose substrates would lead us to a direct, translatable method of testing and finding the AraFAT gene. Unfortunately, every construct of Bradi1g36990 heterologously expressed in *E. coli* as well as expression in the CFPS system failed to produce any positive results upon assaying with our arabinose-containing substrates. This outcome was the first point where we considered that perhaps Bradi1g36990 was not the gene encoding AraPAT; it was possible that this genetic approach was causing changes in the cell wall indirectly. For instance, if the gene Bradi1g36990 encoded a protein that uses feruloyl-CoA in a process separate from cell wall biosynthesis, then we might not see any changes in feruloyl-CoA

on arabinoxylan or lignin, but we might see an increase in *p*-coumarate content in the cell wall, as shown in the reverse genetic experiments by Bartley *et al* (2013). This would be described as an indirect modification of Bradi1g36990.

Another instance where we were faced with this direct-versus-indirect-modification conundrum was with Bradi2g05480, the gene previously studied and published by de Souza *et al.* (2018), in which the RNAi line in *Brachypodium* measured a decrease in feruloylated AX by ~10-20%, and in *S. viridis* about a 60% reduction in feruloylated AX. This publication gave us another gene in the PF02458 clade with strong genetic evidence that it is responsible for a hydroxycinnamate transfer onto arabinoxylan. As one of our candidate genes for AraFAT, we cloned and expressed the protein in CFPS, then assayed it for AraFAT activity. The enzyme is able to use only feruloyl-CoA in the methanol assay, indicating that it is active and prefers feruloyl-CoA to *p*-coumaroyl-CoA. However, this protein did not produce any feruloyl-arabinose products. It is considerably more difficult to explain an indirect effect as a result of a gene suppression line, as knocking-down the gene's expression in plant and subsequent reduction in protein expressed does not consume a pool of hydroxycinnamates. After these results, the lack of activity in our *in vitro* assays made us wonder if we were missing an important component in our assays for AraFAT and AraPAT activities.

One studied enzyme that may help to explain the issue of indirect effect is with PMT, the *p*-coumarate monolignol transferase. This gene was biochemically characterized in our lab and found to add *p*-coumarate onto monolignols (Withers *et al.*, 2012). After knowing its function *in vitro*, Petrik *et al.* (2014) overexpressed and knocked-down this gene (Bradi2g36910) in *Brachypodium*. Upon measuring the *p*CA and fA released from lignin and hemicelluloses, they found a dramatic increase of *p*CA content (~220%) released from BdPMT OX stem tissue

biomass compared to wild-type. The point of interest in this experiment is the HCA content released from arabinoxylan using mild acidolysis. For the PMT OX line, no increase in *p*CA was found, but a significant increase in fA was measured (~140%) compared to WT plants (Petrik *et al.*, 2014). While the researchers did not speculate as to why this phenomenon occurred, I believe this is likely due to an increased amount of PMT protein that used *p*-coumaroyl-CoA for lignin acylation. While PMT was using *p*-coumarate to add to lignin, it possibly left arabinoxylan free to be decorated by available ferulate. There does appear to be a slight decrease in *p*CA on AX in the OX line, but not a significant measurement.

Another explanation as to why both Bradi1g36990 and Bradi2g05480 were not able to transfer *p*-coumarate and ferulate, respectively, onto arabinose substrates in our assays could be that we did not perform the assays with the correct substrates or that the CFPS wheat germ extract was missing required proteins involved in this reaction. It is curious that both of these previously studied genes showing strong genetic evidence for hydroxycinnamating AX do not show any activity with potential substrates using *in vitro* assays. Perhaps this problem is giving us a clue that these modifications of arabinoxylan are more complex than previously thought, and that future studies will need to be creative in elucidating the pathways to feruloylated and *p*-coumaroylated arabinoxylan.

The second part of this chapter discussed using wheat seedling protein extracts to find the AraFAT gene. The goal of this experiment was to determine an assay of the AraFAT activity by eliminating one unknown factor in the AraFAT reaction: the protein. Through several trials of this extraction, we were able to optimize a protocol for extracting soluble proteins from 6-7 day old wheat seedlings and for separating out nonspecific feruloyl esterases from our protein extract. Assaying our protein extracts, both total desalted protein and protein separated by anion

exchange chromatography, with feruloyl-CoA and arabinose-containing substrates and analyzed using RPC and LC-MS/MS, did not find any feruloylated arabinose products resulting from protein reactions. However, this failure must be qualified with a lack of extensive investigation.

In order to fully determine if these wheat protein extracts contain the AraFAT enzyme, more experiments need to be conducted with substrate feeding and searching for potential feruloyl-arabinose products. One of the areas in which we failed was realizing the amount of substrate needed for the enzyme reactions, specifically feruloyl-CoA. A large portion of our feruloyl-CoA was cleaved by feruloyl esterases or nonspecific esterases that utilize CoA-substrates. I believe an abundance of feruloyl-CoA and arabinose-containing substrates (at least 1 mM each), along with addition of nonspecific esterase inhibitors such as sodium fluoride, would need to be incubated with the proteins and analyzed with both chromatography and RPC using LC-MS/MS and thoroughly searched for feruloyl-containing substrates. Because there likely would be a number of nonspecific reactions with non-AraFAT proteins, we would do our best to ensure that the protein was completely removed of small molecules by desalting methods before assaying with these substrates, as feruloyl conjugates have been known to form. For example, when assaying our candidates in CFPS-WGX, we found a peak that was identified using LC-MS/MS as a feruloylated dithiothreitol (DTT) compound.

An assumption we had was that the reaction by AraFAT to produce a feruloyl-arabinose conjugate would be abundant enough to see a significant product peak on the UPLC. With other BAHD ATs we have assayed, the product peak was easy to detect. It is possible that this reaction is not abundant, and in that case, we would need to use LC-MS/MS to analyze and identify feruloyl-containing peaks. Using LC-MS/MS allows us to search for compounds based on their predicted mass-to-charge ratio. Since we are looking for feruloyl-containing conjugates, we

already know the characteristic fragmentation pattern of ferulate in both positive and negative ion modes. In positive ion mode, the fragmentation patterns would be m/z 177, 145, 117, and 89. In negative ion mode, an ester-bound ferulate would fragment at m/z 193. Given what we now know about these AraFAT assays, we are potentially looking for a very small product peak within a sample that has other feruloyl-containing product peaks. However, we also know the predicted masses of our feruloyl-arabinose products which should help us determine the correct products.

One experiment that would be critical for this investigation is assaying other subcellular components for AraFAT activity. With the hypothesis that the AraFAT enzyme is a BAHD AT and therefore cytosolic, we are overlooking the Golgi apparatus, an organelle that potentially holds or is involved in the AraFAT reaction. As previously stated, the Golgi is the site of xylan synthesis and a majority of UDP-Arap synthesis (Rautengarten *et al.*, 2017; Scheller and Ulvskov, 2010). The cytosolic face of the Golgi is where UAMs tend to associate. These mutases convert UDP-Arap to UDP-Araf (Konishi *et al.*, 2007; Rautengarten *et al.*, 2011). Although not studied, UAMs and their homo- and heterocomplexes may play a role in the transfer of ferulate onto arabinose or assist in the transport of feruloyl-arabinose into the Golgi. A proposed experiment would be to fractionate and isolate intact Golgi by differential centrifugation and assay this fraction with feruloyl-CoA and arabinose substrates (Morré and Mollenhauer, 1964). We would also propose assaying the Golgi fraction with soluble protein, in case a cytosolic protein (such as a BAHD AT) contains AraFAT and/or necessary proteins for the AraFAT reaction to occur. This would also ensure that the UAMs are in the assay.

Golgi reaction products can be analyzed by using LC-MS/MS, but this would require a few considerations. First, determining where this product would be located: outside of or within

the Golgi. In order to isolate the products outside of the Golgi, the Golgi and proteins could be removed by using Amicon Ultra centrifugal filter units and analyzing the product-containing flow-through. For analyzing the products within the Golgi apparatus, a detergent would be required to denature membranes and lyse the Golgi. One issue is that most detergents are incompatible with LC-MS/MS systems. RapiGest SF Surfactant (from Waters) is a mild detergent able to be hydrolyzed at low pH, making it compatible with LC-MS/MS.

Another possibility is that the reaction takes place within the Golgi apparatus, and we would need to lyse the Golgi in order to assay the Golgi-localized proteins with our substrates. Again, we would propose using the RapiGest reagent to perform this denaturation step, followed by a series of assays with our arabinose substrates and feruloyl-CoA and feruloyl-glucose. We could then solubilize RapiGest and analyze using LC-MS/MS to look for feruloyl-arabinose compounds. We believe that these series of assays with isolated Golgi could help us to determine AraFAT activity by looking in a subcellular component that has yet to be fully investigated.

Given that the pathway of AX feruloylation is still not understood in terms of its cellular localization and substrates used, further information of these key factors is needed to predict which proteins are involved in and responsible for this grass-specific modification.

Understanding ferulate pools and pathways within the cell using radioactive tracing techniques may help us elucidate the AraFAT pathway. Prior work has been performed to determine if feruloyl-glucose is the donor substrate in this pathway, but it is unclear if feruloyl-glucose was used as the donor in these experiments or acted as a reserve pool for conversion to feruloyl-CoA (Obel *et al.*, 2003; Lenucci *et al.*, 2009). While reverse genetic experiments are indispensable to understanding gene in plants, it is important that we also find an *in vitro* assay to characterize these proteins as well. The wheat protein extraction method is a potentially powerful tool to

determine an AraFAT assay and find the AraFAT protein. With additional time, isolation and assaying of Golgi, and more thorough searching of feruloyl-arabinose products, we believe this approach may lead the way to discovery of AraFAT activity and the AraFAT gene.

CHAPTER 2

INTRODUCTION ON HYDROXYCINNAMATE CONJUGATIONS IN GRASSES

Plants have developed many strategies for surviving and being successful in a potentially hostile environment. One of these strategies is the development of secondary metabolites from phenolic compounds, such as hydroxycinnamic acids (HCAs). These HCA-derived secondary metabolites are involved in number of plant defense responses: abiotic stresses, wounding, ultraviolet (UV) irradiation, and pathogen attacks. HCAs and their conjugates also play important roles in plant growth and development and signaling pathways.

HCAs are a class of aromatic acids that derive from the amino acid phenylalanine in the lignin biosynthetic pathway (**Figure 1**, in Introduction). They have a C6-C3 skeleton and undergo a series of hydroxylation and *O*-methylation steps to create a variety of structures, the simplest and most common being ferulate, *p*-coumarate, and caffeate. These compounds are usually activated in the form of Coenzyme A (CoA) thioesters or 1-*O*-acylglucosides, which can then be used to make several thousands of phenolic compounds, most of them flavonoids or lignin (Grisebach, 1981; Strack, 2001). HCA-CoA compounds produce the lignin subunits guaiacyl (G), syringyl (S), and *p*-hydroxyphenyl (H), as described in the Introduction of this dissertation. HCAs are also able to conjugate with other compounds to create amides, esters, and glycosides (**Figure 15**) (Strack, 2001).

Phenylamides serve many functions in plants. In grasses, they are known to play an important role in defense responses and disease resistance. When barley plants were subjected to powdery mildew fungus, levels of *p*-coumaroyl-hydroxyagmatine increased (Von Röpenack, Parr, and Schulze-Lefert, 1998). Maize showed induction of tyramine conjugates when subjected to wounding (Ishihara *et al.*, 2000) and insect herbivory (Marti *et al.*, 2013).

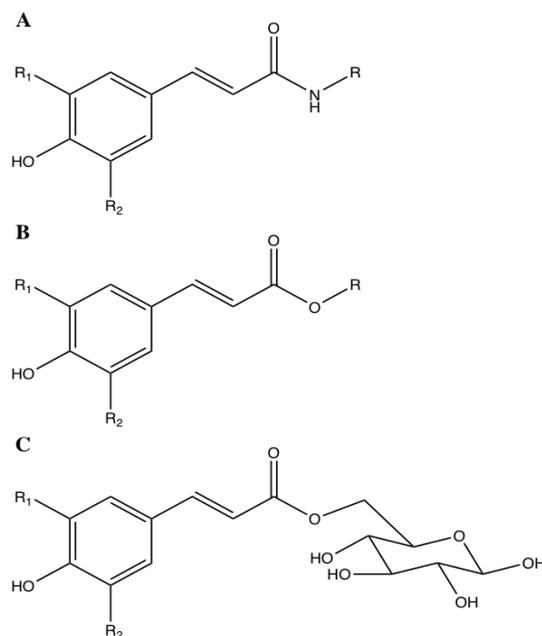


Figure 15: Simplified structures of hydroxycinnamate conjugates. HCA conjugations in the forms of amides (**A**), esters (**B**), and glycosides (**C**). For ferulate, $R_1 = \text{OMe}$, $R_2 = \text{H}$; *p*-coumarate, R_1 and $R_2 = \text{H}$; caffeate, $R_1 = \text{OH}$, $R_2 = \text{H}$. The R following the amide and ester linkages refers to the conjugates formed on HCA.

It has been shown that some of these metabolites conjugated with ferulate are incorporated via ether linkages into the cell wall following wounding, possibly as a way to repair the wounds or to create a barrier to incoming pathogens. Understanding the enzymes that make these phenylamide conjugates and determining their pathways in the plant will help us understand how plants detect external danger and protect themselves from pathogen and insect attacks. This knowledge could allow scientists to develop plant-made treatments for discouraging pests without the use of potentially harmful pesticides.

There are three types of enzymes that perform this *N*-acylation conjugation in plants: GCN5-related *N*-acyltransferase (GNAT) proteins, serine-carboxypeptidase-like (SCPL) proteins, and BAHD acyltransferases (D'Auria, 2006; Kang, 2005; Stehle, 2009). GNAT enzymes specifically catalyze *N*-acylation condensation reactions with hydroxycinnamoyl CoA

thioesters (Bassard, 2010). The GNAT proteins have molecular masses that range from 20 to 25 kDa, and contain two conserved regions, RGF₂GIG and FYXR₂XG, which are required for enzyme activity (Kang, 2006). One of the first genes in this group to be characterized was the tyramine:*N*-hydroxycinnamoyl transferase (THT), found in tobacco (Hohlfeld, 1995; Lu, 1996; Négrel and Martin, 1984).

SCPL proteins are 1-*O*-acylglucose ester-dependent enzymes that differ from true serine carboxypeptidases by their lack of peptidase ability. They perform transacylation using a catalytic triad of serine (Ser), histidine (His), and aspartic acid (Asp) with 1-*O*-β-glucose esters (Bontpart *et al.*, 2015; Liu, 2010). One example of an SCPL protein is sinapoylglucose:malate sinapoyltransferase (SMT). SMT performs the acyltransferase of the donor, sinapoylglucose, to the acceptor, malate, to make sinapoyl malate, which accumulates in the leaves and protects the plant from UV damage (Strack, 1982; Steffens, 2000). The product and the SMT enzyme are localized in the vacuole (Strack and Sharma, 1985). SCPL proteins contain N-terminal signal peptides that suggest the secretory pathway is directing them towards the vacuole (Lehfeldt *et al.*, 2000; Mugford *et al.*, 2009).

BAHD acyltransferases are named after the first four characterized genes from this family (BEAT, AHCT, HCBT, and DAT). These monomeric enzymes range in size from 48 to 55 kDa and are localized in the cytosol, as they lack transit or signal peptides (D'Auria, 2006). One key feature of the BAHD ATs is their conserved motifs, HXXXD and DFGWG. The former motif is important for histidine deprotonation of the oxygen or nitrogen on the acceptor substrate, which then allows for a nucleophilic attack on the carbonyl carbon of the CoA donor. The DFGWG motif is not located near the active site, so likely serves a structural integrity purpose. The first BAHD AT to show an ability to transfer an acyl group to a nitrogen atom was HCBT,

an *N*-hydroxycinnamoyl/benzoyltransferase ((D'Auria, 2006; Ma *et al.*, 2005). While using very different acyl donors, the SCPL and BAHD proteins are able to perform both *O*- and *N*-acylation and have some overlapping functions in which metabolites they modify (Liu, 2010). It can be difficult to discern which class of enzyme is responsible for what conjugation in the plant.

Background on jasmonate response pathway in grasses

Phenylamides have been shown to increase in plants in response to jasmonate (Gális *et al.*, 2006; Ishihara *et al.*, 2008a; Ishihara *et al.*, 2008b; Lee *et al.* 1997). The jasmonate (JA) response pathway is involved in a number of plant functions, including response to wounding and insect herbivory (Koo and Howe, 2009). When plants are wounded or attacked by insects or pathogens, JA is released into the cell (Schilmiller and Howe, 2005). Jasmonate is a hormone that occurs in a number of forms, including the volatile methyl-jasmonate (MeJA) and jasmonate-isoleucine (JA-Ile). JA-Ile is directly involved in the interaction between jasmonate ZIM-domain (JAZ) and coronatine insensitive 1 (COI1) proteins, which leads to the degradation of JAZ proteins. JAZ acts as a repressor of MYC, a group of transcription factors which interact with the MED25 mediator co-activator complex. This interaction recruits RNA polymerase II and turns on the jasmonate responsive genes (Thines *et al.*, 2007; Zhang *et al.*, 2017). In summary, when JA is detected in the cell, a chain of reactions leads to the transcription of responsive genes that assist in defense against external attacks.

In this chapter, I present the research that we conducted on Bradi1g36980, a gene encoding a BAHD acyltransferase that has shown activity with feruloyl-CoA and a number of phenylamines, including serotonin and tryptamine. To our knowledge, this is the first time a BAHD acyltransferase has shown activity with either serotonin or tryptamine. We have found that it specifically uses feruloyl-CoA as a donor substrate with a calculated K_M value of 2.5 μ M

and has no activity with *p*-coumaroyl-CoA. We have gathered some data on tryptamine and serotonin, calculating K_M values of 493 and 893 μM , respectively. These K_M values are relatively high for acceptor substrates, which may indicate that these are not the natural substrates used by the enzyme in the plant. When transiently overexpressing this gene in tobacco leaves to analyze metabolites resulting from Bradi1g36980, we see a significant change in the amount of ferulates in the overexpression sample compared with untreated leaves. we find that it is indeed feruloylating compounds, although these compounds have not yet been identified.

MATERIALS AND METHODS

Acyltransferase identification and selection

Bradi1g36980 was selected from a grass-specific clade in the PF02458 family of BAHD acyltransferases for AraFAT activity.

Collection of compounds known to be hydroxycinnamated

Several compounds known to be hydroxycinnamated in plants were obtained to assay with our BAHD acyltransferases. 16-hydroxyhexadecanoic acid, 2-phenylethylamine, agmatine, dopamine, kaempferol, kuromanin, norepinephrine, putrescine, quercetin, quercitrin, D-(-)-quinic acid, rutin, serotonin, beta-sitosterol, spermidine, spermine, sucrose, tryptamine, and tyramine were all purchased from Sigma-Aldrich. Octopamine was purchased from Santa Cruz Biotechnology. Hordenine was obtained from Abcam Biochemicals, and beta-sitostanol was purchased from Matreya, LLC. A2XX and A3X were purchased from Megazyme.

Enzymatic synthesis of CoA-substrates using 4CL2

Feruloyl- and *p*-coumaroyl-CoA were enzymatically synthesized with 4-coumaroyl:Coenzyme A ligase 2 (4CL2), using a modified method from Beuerle and

Pichersky (2002). 4CL2 was synthesized by GeneArt and cloned into the pMAL-c5x vector, which adds a maltose-binding protein (MBP) affinity tag to the protein of interest. Purified 4CL2 was incubated with 3 mg ferulic or *p*-coumaric acid, 2 mg Coenzyme A (CoA), and 6.9 mg ATP in 50 mM Tris-HCl, pH 7.5, a 10 mL reaction volume.

Cloning, expression, and purification of Bradi1g36980 in *E. coli*

Primers were designed to amplify Bradi1g36980 from *Brachypodium distachyon* cDNA developmental stages 2-5. The gene was cloned with an affinity tag for protein purification using an MBP-tag in the pMAL-c5x vector. This vector was designed with a TEV cleavage recognition site immediately after the MBP-tag and directly upstream of the gene of interest. Bradi1g36980 was inserted in the vector using the InFusion recombination system in the PstI restriction site.

Protein purification with MBPTrap column using FPLC

Affinity chromatography was performed using an FPLC and the affinity column MBPTrap (GE Healthcare) MBP-tag Bradi1g36980. For MBP-tag fusion proteins, buffers used were Binding (20 mM Tris-HCl, pH 7.5, 200 mM NaCl) and Elution (20 mM Tris-HCl, pH 7.5, 200 mM NaCl, 10 mM maltose). After binding to the column, elution of the Bradi1g36980-MBP-fusion protein occurred by an immediate ramp to 100% Elution Buffer for 5 CVs. The peak corresponding to eluted protein was analyzed using SDS-PAGE.

Enzymatic assays

Assays were performed in 50 μ L volumes in 50 mM Tris-HCl, pH 7.5. CoA substrates and acceptor substrates were added to 1 mM final concentrations. Incubations were performed at room temperature at varying lengths of time, but typically for 1 hour (unless stated otherwise). To stop the reactions, 1 μ L of 10 M HCl was added to each tube and 150 μ L of 2:1 methanol:ddH₂O was added to adjust the final concentration of methanol to 50%. The reaction

products were separated from proteins using an Amicon 10K 0.5 mL filter and 10 minutes of centrifugation at 15,000 rpm. Ten microliters of the filter reaction product were injected into the UPLC for analysis.

UPLC RPC analysis

All assays were analyzed using RPC on a Waters UPLC H-series using a BEH C18 column, 1.7 μ M, 2.1 x 100 mm. Flow rate is 0.300 mL/min with Solvent A: ddH₂O with 0.1% TFA and Solvent B: 100% ACN. The linear gradient began with 87% Solvent A, moving to 60% Solvent A from 0 to 7 minutes. The gradient goes to 100% B from 7 to 8 minutes. From 10-11 minutes, it returns to 87% A for column equilibration. Detection was performed by TUV detector at 262 and 340 nm wavelengths.

Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS) with C18 column

Enzymatic assay reaction products were analyzed at the Mass Spectrometry Facility at Michigan State University using a G2-XS QToF instrument. Separation was performed on the a Waters Acquity BEH C18 column, 1.7 μ M, 2.1 x 100 mm, with a slightly modified RPC gradient: Solvent A is ddH₂O with 0.1% formic acid and Solvent B is 100% ACN. The flow rate is 0.300 mL/min, and begins with 87% A. From 0 to 7 minutes, the Solvent A goes from 87% to 60%, and between 7 and 8 minutes ramps up to 100% B (until 10 minutes). At 10.01 minutes, it returns to 87% A and equilibrates until 12 minutes. Mass spectrometry analysis was performed in positive and negative electrospray ionization (ESI) modes searching for a mass range between 50 and 1500 Da, with ramped collision from 20-80 eV to fragment the compounds. Capillary was set at 2.0 kV and lockspray capillary at 2.5 kV. Scan duration was 0.200 seconds, with source and desolvation temperatures at 100 and 350 degrees Celsius, respectively. Cone gas flow and desolvation gas flow rates were at 25.0 and 600 L/Hr, respectively.

Time course enzyme kinetic assay methods and analysis

All enzyme kinetic assays were performed in 50 mM Tris-HCl, pH 8, buffer. The 150 μ l reaction mixture contained 0.001-1 mM CoA thioester donor substrate (feruloyl- or *p*-coumaroyl-CoA) and 0.005-1 mM phenylamine acceptor substrate (tyramine, tryptamine, or serotonin). The reaction was started with the addition of 1.5 μ g purified enzyme (MBP-fusion). At 5 minute (for feruloyl-CoA and tryptamine) or 10 minute (for serotonin) intervals, 20 μ l of reaction mixture was removed and stopped with 1 M HCl. These reactions were brought up to 50% methanol in 100 μ l final volumes and analyzed using the RPC method (as described above). Standard curves of products were performed to calculate the product concentration using peak areas from the RPC. Each reaction time point was analyzed by measuring the product formed over time. Initial velocities were calculated and plotted versus substrate concentration using a Michaelis-Menten approximation model ($y = ax/(1+bx)$) and a 2-parameter asymptotic exponential approximation model ($y = a(1-e^{-bx})$) in RStudio (Crawley, 2007).

Exposure of *Brachypodium* to methyl-jasmonate (MeJA)

Brachypodium seeds were sterilized and plated in magenta boxes (four seeds each) containing about 100 ml of MS + vitamin agar. Boxes were kept in the dark in a cold room for three days stratification, after which were moved to a percival, at 22°C at 16/8 hour light conditions and 130-150 μ E. After 14 days, plants were subjected to MeJA by adding 2.2 μ l MeJA (final concentration 5 μ M) to cotton twine and placed in each box. At various times, seedlings were collected into liquid nitrogen. Controls were treated with a mock twine of 100% methanol, which was added to their boxes, and collected at 0, 8, 24, and 48 hours. Four to 6 replicates were collected for each time point.

Quantitative Real-Time PCR (qRT-PCR) of MeJA-treated plants

Bradi1g36980 expression analysis was performed using qPCR SYBR Green Master Mix with 5 ng/ μ l cDNA, 2 μ M each of forward and reverse primers, and 2.5 μ l ddH₂O for a total 15 μ l reaction volume per well. The following primers were used: forward primer 5'-GGAGGCACCTCAGGTCATTT and reverse primer 5'-ATAGCGGTCATTGTCTTGCG for Bradi4g00660, housekeeping gene UBC18; forward primer 5'-TGGACTACCTCGAGTACCCGC and reverse primer 5'-GAACCCGACCACGAAGCC for Bradi1g36980; forward primer 5'-GAGATCAGCAAGAGCAGGAGG and reverse primer 5'-CATCAGGCTCAGCGTCGT for Bradi1g72490, JAZ1 protein. Samples were analyzed in triplicate with all biological samples.

Tobacco *Agrobacterium*-infiltration of Bradi1g36980

Nicotiana benthamiana plants were grown at 16/8 hour light cycles and leaves were infiltrated with *Agrobacterium* at 4-5 weeks. Bradi1g36980 was cloned with an N-terminal 6xHis-tag and inserted in the vector pORE-O4 (Coutu *et al.*, 2007) using InFusion cloning and transformed into the *Agrobacterium* GV3101 competent cells. *Agrobacterium* was grown overnight at 28C to stationary phase in Luria Broth, then pelleted and rinsed twice with infiltration buffer (10 mM 2-(N-morpholino)-ethanesulphonic acid (MES), pH 5.6, 10 mM MgCl₂, and 100 μ M acetosyringone). *Agrobacterium* cells were resuspended in infiltration buffer to OD₆₀₀ of 1. The constructs p19 empty vector and YFP in pORE-O4 (35S promoter) were also injected in tobacco leaves as negative controls.

Analysis of *Agrobacterium*-infiltrated *N. benthamiana* leaves

Infiltrated tobacco leaves of 36980 overexpression, YFP, and empty vector p19, as well as non-infiltrated leaves, were collected 72 hours after infiltration, weighed, and frozen in liquid

nitrogen. To extract soluble metabolites from tobacco leaves for LC-MS/MS analysis, 1 mL of 70% ethanol was to each sample and milled for 2 minutes at 20 Hz. This was repeated a total of three times. With the resulting alcohol-insoluble residue (AIR), 400 μ L of sodium phosphate buffer pH 7.4 was added and 100 μ L 10 M NaOH to saponify overnight. The next morning the reaction was neutralized with 8.5 M acetic acid and its saponifiable hydroxycinnamates were extracted with ethyl acetate. All extracts were then analyzed using the LC-MS/MS C18 column method described above.

RESULTS

Characterization of Bradi1g36980 with feruloyl-CoA and substrates

Bradi1g36980 was successfully cloned, heterologously expressed in *E. coli*, and purified using an MBPTrap column on the FPLC (**Figure 16**). The purified enzyme was assayed with the artificial acceptor methanol to test for protein activity and determine preference for feruloyl- or *p*-coumaroyl-CoA. The assay showed that Bradi1g36980 has activity with feruloyl-CoA and no activity with *p*-coumaroyl-CoA.

In the paper by Sibout *et al.*, Bradi1g36980 was overexpressed in Arabidopsis under the cinnamate-4-hydroxylase (*AtC4H*) promoter, and plants had an increase in alkali-releasable *p*-coumarate from cell walls. These results lead the researchers to conclude that Bradi1g36980 is a *p*-coumarate monolignol transferase (PMT), and referred to it as PMT2 (Sibout *et al.*, 2016). While our methanol assays with this enzyme showed no activity with *p*-coumaroyl-CoA, we determined its activity with *p*-coumaroyl-CoA and the acceptor substrates sinapyl and coniferyl alcohols. The enzyme was assayed separately with *p*-coumaroyl- and feruloyl-CoA with both coniferyl and sinapyl alcohols. Neither of the CoA-substrates showed any activity with the monolignol acceptor substrates.

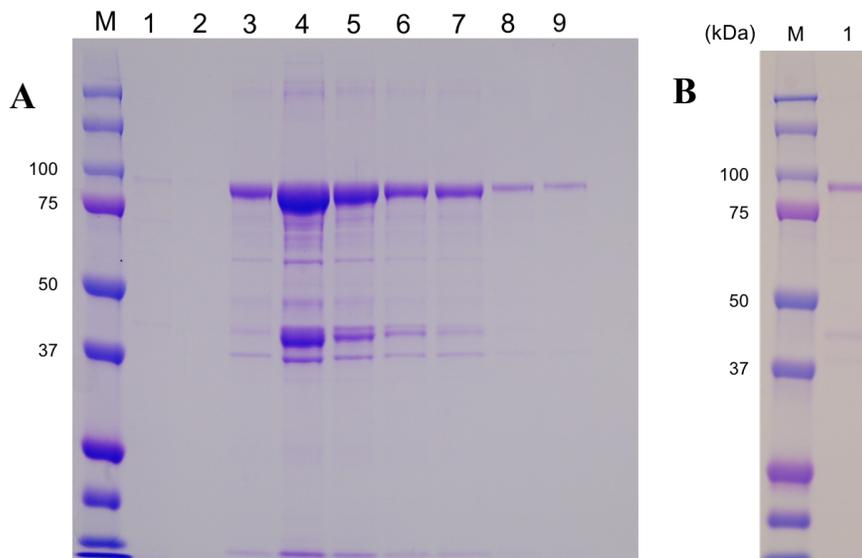


Figure 16: SDS-PAGE gels of Bradi1g36980 MBP-tag affinity purification. In panel A, the purification process of Bradi1g36980 MBP-fusion protein off of the MBPTrap column is shown; all lanes show the elution off the column with maltose. Panel B is 1.5 µg of Bradi1g36980 after a gel filtration column.

The enzyme was then assayed with a series of compounds known to be hydroxycinnamated (HCA'd) in grasses (**Table 3**). When Bradi1g36980 was incubated with feruloyl-CoA and tyramine and separated using RPC, a novel peak occurred at ~4.9 minutes retention time. To determine if Bradi1g36980 was adding ferulate to the hydroxyl or amino group of tyramine, we assayed it with feruloyl-CoA and 2-phenylethylamine (2-PEA). This compound is identical to tyramine, but 2-PEA has no hydroxyl group. We hypothesized that if the enzyme was *N*-acylating tyramine, it would use 2-PEA as an acceptor substrate as well. The result of that assay showed it did react with 2-PEA, giving evidence that our enzyme was *N*-acylating these two phenethylamine compounds (**Figure 17**).

Compound	Type	Functional groups
agmatine	aliphatic polyamine	amino groups
putrescine	aliphatic polyamine	amino groups
spermidine	aliphatic polyamine	amino groups
spermine	aliphatic polyamine	amino groups
serotonin	phenylamine	phenolic, amino groups
tryptamine	phenylamine	phenolic, amino groups
tyramine	phenylamine	phenolic, amino group
2-phenylethylamine	phenylamine	benzene with amino group
dopamine	phenylamine	hydroxyl groups, amino group
(-)-norepinephrine	phenylamine	hydroxyl groups, amino group
octopamine	phenylamine	hydroxyl groups, amino group
hordenine	alkaloid of phenethylamine	phenolic, N,N-dimethyl group
D-(-)-Quinic acid	cyclic carboxylic acid	carboxylic acid, hydroxyl group
16-hydroxyhexadecanoic acid	hydroxylated fatty acid	carboxylic acid, hydroxyl group
sucrose	disaccharide	glucoside, fructoside, hydroxy groups
kuromanin	anthocyanin	glucoside, phenolic, hydroxyl groups
kaempferol	flavonoid	polyphenol hydroxyl groups
quercetin	flavonoid	polyphenol, hydroxyl groups
quercitrin	flavonol glycoside	rhamnoside, polyphenol, hydroxyl groups
rutin	flavonol glycoside	disaccharide rutinose, polyphenol, hydroxyl groups
β -sitostanol	phytosterol	hydroxyl group, sterol
β -sitosterol	phytosterol	hydroxyl group, sterol

Table 3: List of compounds known to be hydroxycinnamated in plants. These compounds were collected to assay our AraFAT candidates in an effort to narrow down our search for AraFAT. Compounds are organized based on type, grouping similar compounds together.

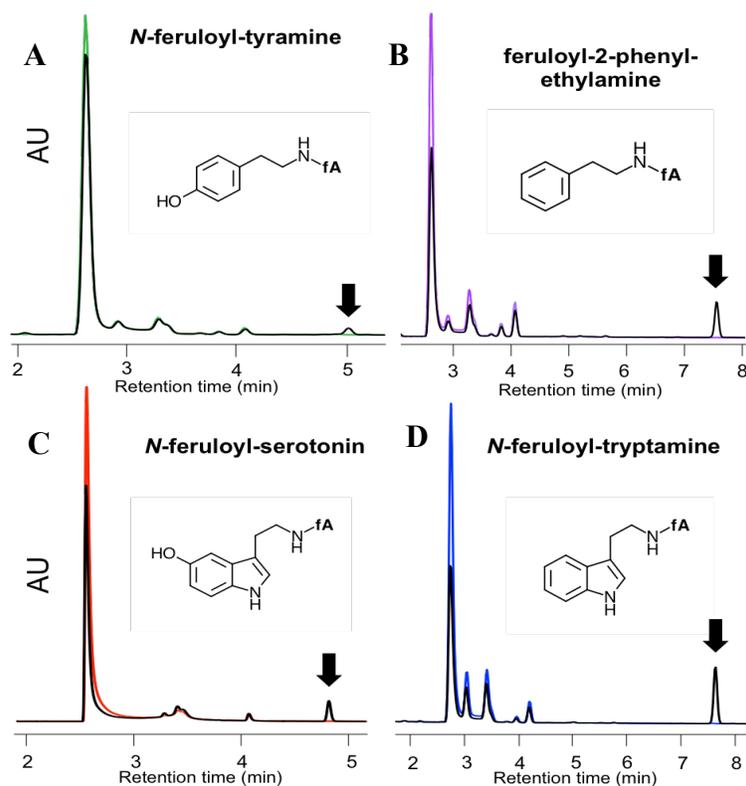


Figure 17: RPC separation of Bradi1g36980 assay products with feruloyl-CoA and phenylamines. Enzyme was active with tyramine (A), 2-phenylethylamine (B), serotonin (C), and tryptamine (D). Black arrows indicate the product peak corresponding to each feruloyl-conjugate. Simplified structures of the products are shown, where **fA** is ferulate.

In an effort to characterize the Bradi1g36980 enzyme and the substrates it uses, we assayed the protein with other phenylamine substrates that are known to be hydroxycinnamated in the plant. Serotonin is a neurotransmitter that naturally occurs in plants, and acts as a growth regulator and in stress responses (Erland and Saxena, 2017). Both tryptamine and serotonin assays produced novel peaks, corresponding to feruloyl-tryptamine and feruloyl-serotonin (**Figure 17**). To our knowledge, this is the first time a BAHD AT has been proven to make feruloyl-conjugates with serotonin (Peng *et al.*, 2016).

There was no activity with feruloyl-CoA and any aliphatic amines tested (agmatine, spermidine, spermine, and putrescine). A few phenylamines that showed no activity with our enzyme and feruloyl-CoA were dopamine, norepinephrine, and octopamine (**Figure 18**). These compounds are structurally similar to the active phenylamines, such as tyramine, but have a few key differences. In the case of dopamine and norepinephrine, both have an additional hydroxyl group on the aromatic ring at the 3-position. This could interfere with substrate binding in the active site of Bradi1g36980. While octopamine does not have this di-(3,4)-hydroxyl on the phenyl group, it does contain a hydroxyl group on the carbon-1 position of ethyl tail. Norepinephrine also has this ethyl-hydroxyl group. Two other substrates that did not have activity with the enzyme were amino acids tyrosine and tryptophan. The lack of activity with these compounds gives us additional insights into the substrate preferences in grasses.

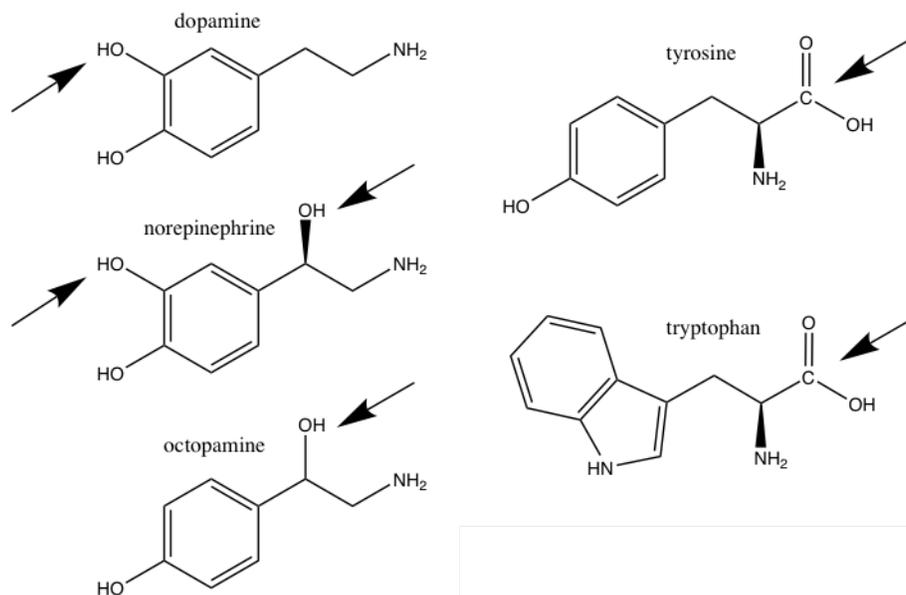


Figure 18: Amino-containing compounds assayed with Bradi1g36980 resulting in no activity. On the left, dopamine, norepinephrine, and octopamine all have hydroxyl group on the ethyl tail and/or on the aromatic ring. On the right, tyrosine and tryptophan both contain carboxylic groups. Arrows indicate hydroxyl or carboxyl group that could potentially explain the lack of activity with the enzyme.

The assay products for feruloyl-tryptamine, feruloyl-serotonin, feruloyl-tyramine, and feruloyl-2-PEA were analyzed using LC-MS/MS to verify *N*-feruloyl-conjugation (**Figure 19**). All products showed a parent ion consistent with expected mass of the conjugate. Ferulates in positive ion mode produce a series of m/z 177, 145, and 117. In negative ion mode, a feruloyl-ester typically fragments at m/z 193, but because our products are feruloyl-amides, we do not see this fragment. Based on the fragmentation pattern of the compounds (**Figure 20**), we are confident that ferulate was added on the amino group of these phenylamines.

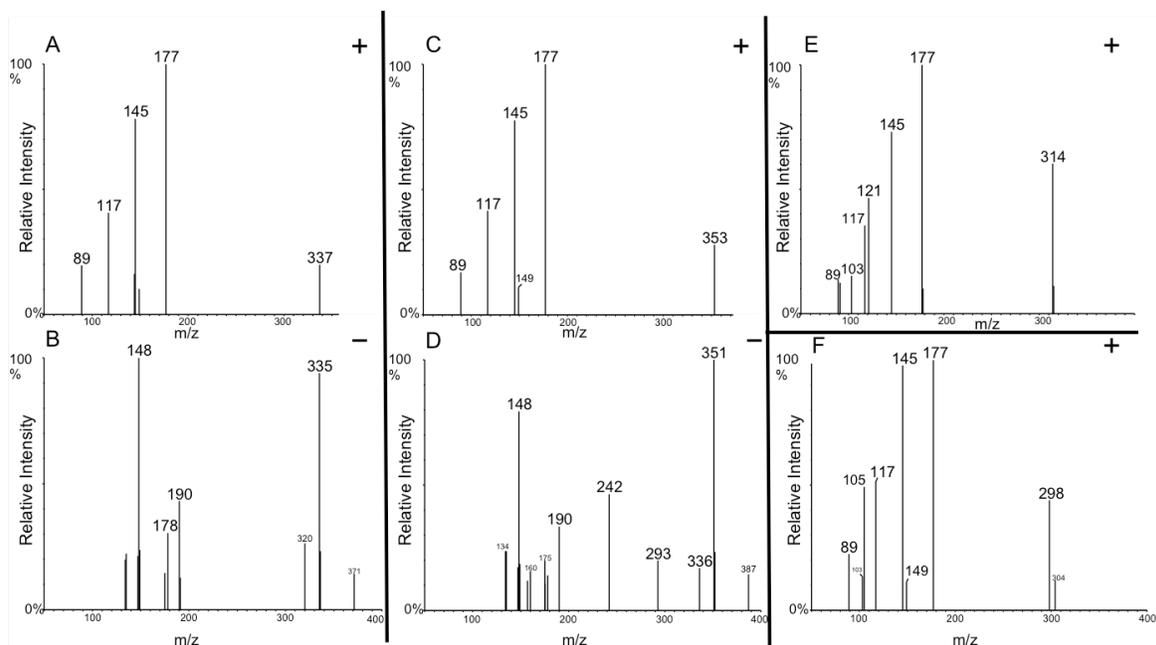


Figure 19: Mass spectra of hydroxycinnamic acid amides made by Bradi1g36980 enzyme. All spectra are shown in ramped collision energy of 20-80 eV. **A** and **B** are *N*-feruloyl-tryptamine (FerTrp), in positive and negative ion modes, respectively. FerTrp has m/z 337 $[M+H]^+$ and m/z 335 $[M-H]^-$; **C** and **D** show *N*-feruloyl-serotonin, in positive and negative ion modes, respectively. **E** is feruloyl-tyramine (FerTyr), and **F** is feruloyl-2-phenylethylamine (Fer2PEA), both in positive ion mode. FerTyr has an m/z 314 corresponds to $[M+H]^+$. Likewise, Fer2PEA $[M+H]^+$ peak is 298. See **Figure 20** for ferulate fragmentation analysis.

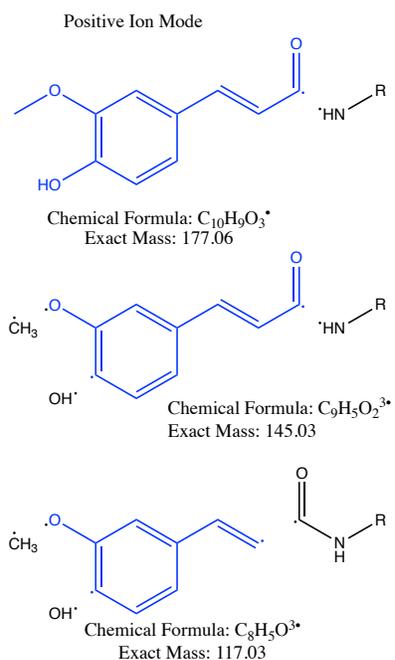


Figure 20: Fragmentation patterns of feruloyl-amides in LC-MS/MS in positive ion mode. Fragments of ferulate compounds in positive ion mode are m/z 177, 145, and 117.

Michaelis-Menten kinetic constants for substrates

We performed time course enzyme kinetic analyses on the Bradi1g36980 enzyme in order to determine the correct substrates, and to evaluate the likelihood that these phenylamines act as acceptor substrates *in vivo* (**Figures 21, 22, and 23**). We analyzed the initial rates versus concentration using the non-linear regression models Michaelis-Menten and 2-parameter asymptotic function. The Michaelis constant (K_M) and the enzyme turnover number (k_{cat}) were calculated for the donor substrate feruloyl-CoA using Michaelis-Menten asymptotic approximation, $y = ax/(1 + bx)$, where a is the regression intercept and b is the slope (Crawley, 2007). We attempted to calculate *p*-coumaroyl-CoA's kinetic parameters, but Bradi1g36980 did not have any activity with it. The enzyme shows high preference for feruloyl-CoA based on the Michaelis-Menten approximation, and its calculated K_M value is 2.5 μM , with a k_{cat} of 4037 min^{-1} (**Figure 24**).

The rate values collected after 10 μM feruloyl-CoA appear to have low variance (with the exception of 40-50 μM feruloyl-CoA). This could be due to several technical reasons, including poor pipetting skills or improper master mix preparation. It could also be an issue of the UPLC, as we recently had transducer errors that required replacement and instrument maintenance. One problem is that these reactions were performed on different days, as we were limited by number of vials run on the UPLC. One set of reactions I performed had a number of outliers, as highlighted in **Figure 25**. Although it is unclear why this day's reactions had high variation compared to other days' runs, it provides a possibility that the kinetic assays were skewed due to human error.

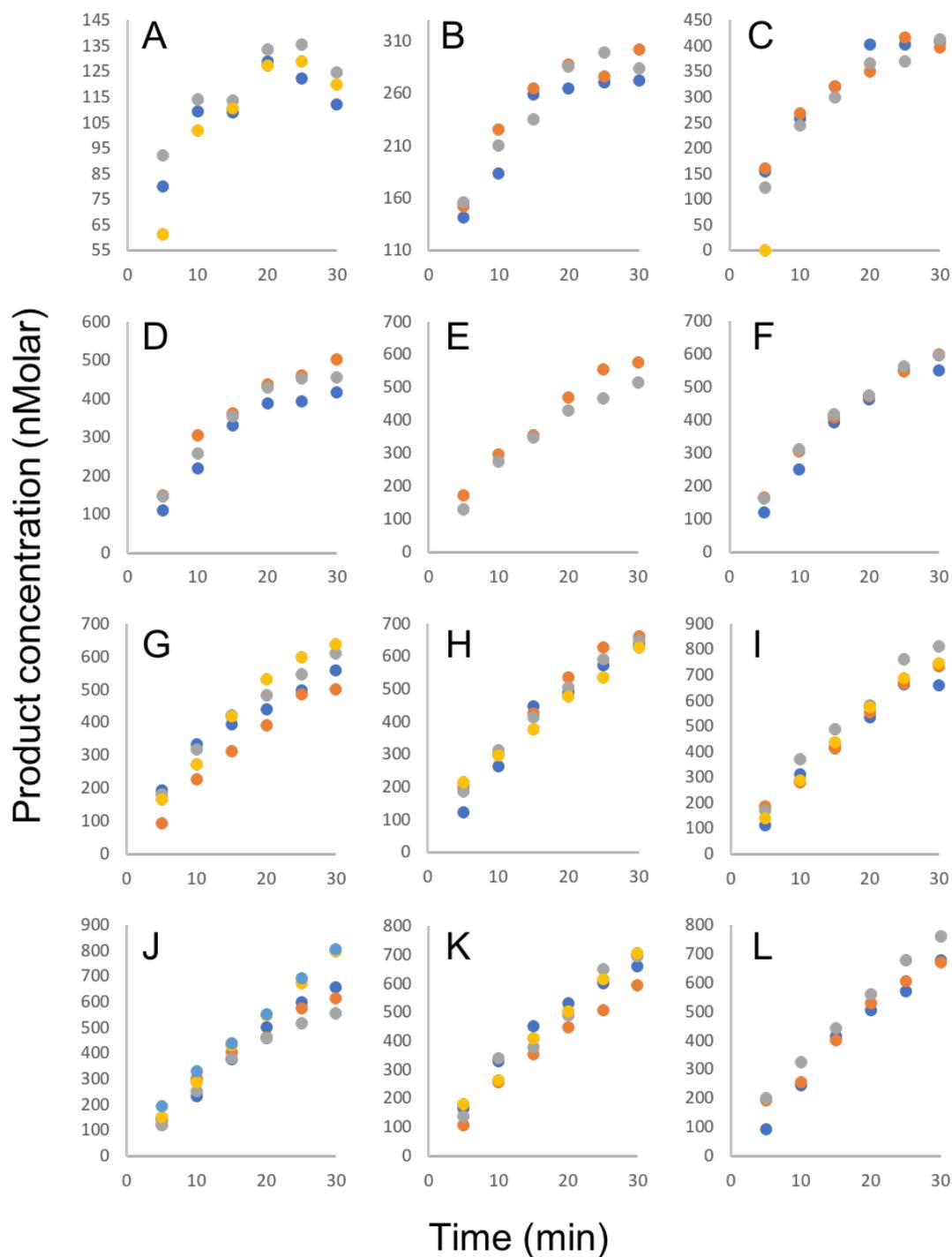


Figure 21: Saturation kinetics of Bradi1g36980 with varied feruloyl-CoA concentrations, measuring feruloyl-tryptamine product formation versus time. The concentrations of feruloyl-CoA used were 1 μM (A), 2 μM (B), 3 μM (C), 4 μM (D), 6 μM (E), 8 μM (F), 10 μM (G), 20 μM (H), 30 μM (I), 40 μM (J), 50 μM (K), and 60 μM (L). Two to five replicates were performed at each concentration.

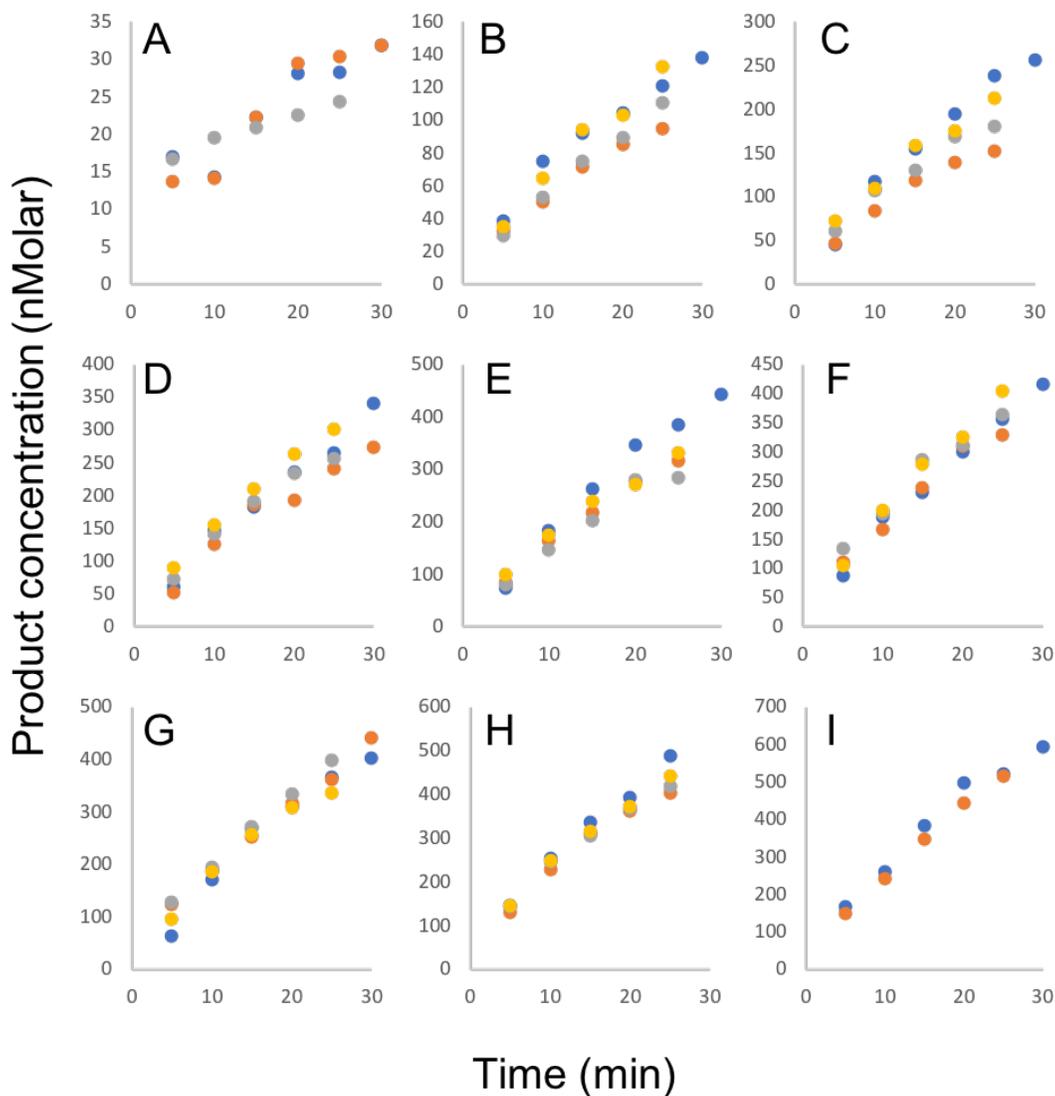


Figure 22: Saturation kinetics of Bradi1g36980 with varied tryptamine concentrations, measuring feruloyl-tryptamine product formation versus time. The concentrations of tryptamine used were 20 μM (A), 100 μM (B), 200 μM (C), 300 μM (D), 400 μM (E), 500 μM (F), 600 μM (G), 700 μM (H), and 800 μM (I). Two to four replicates were performed at each concentration.

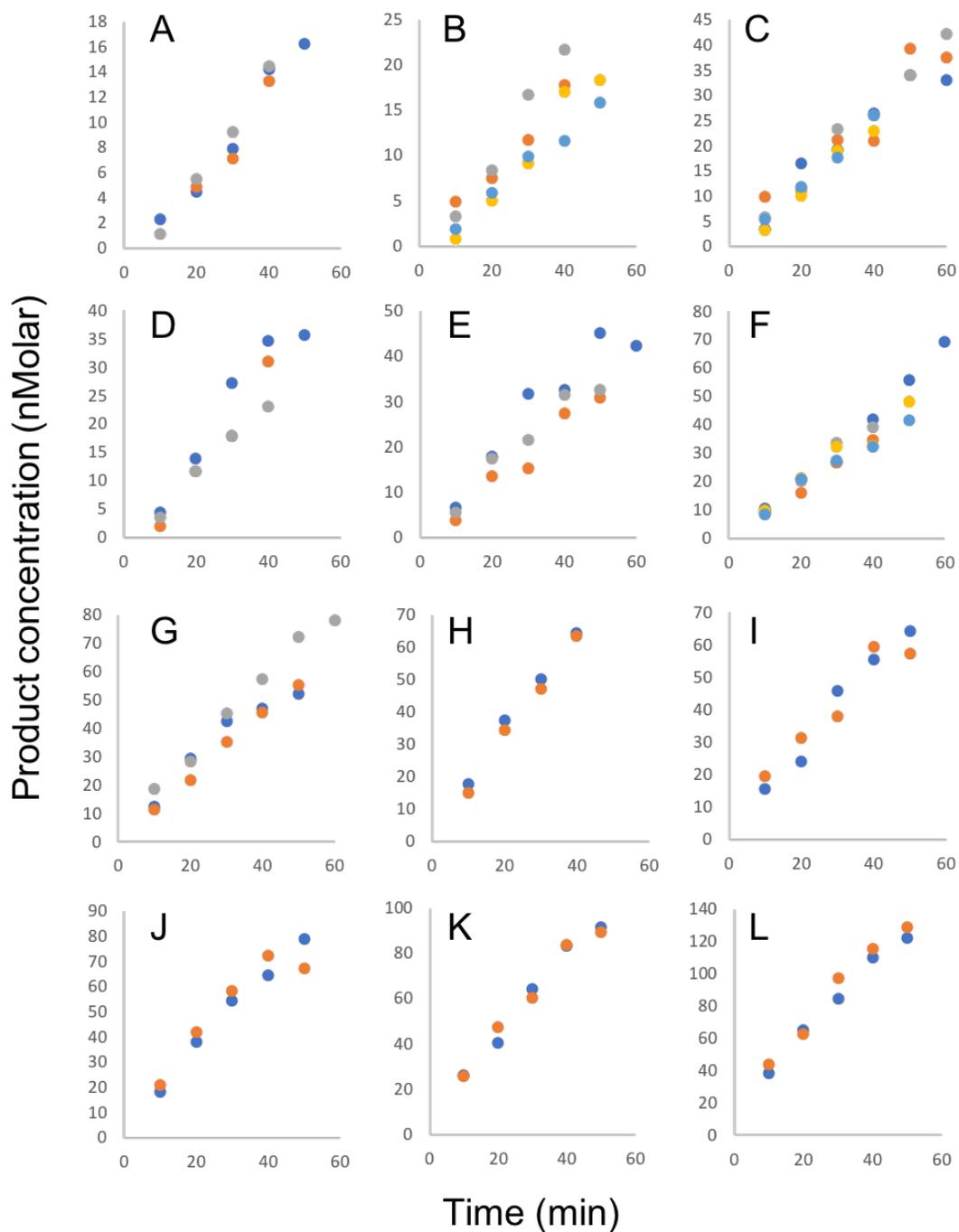


Figure 23: Saturation kinetics of Bradi1g36980 with varied serotonin concentrations, measuring feruloyl-serotonin product formation versus time. The concentrations of serotonin used were 100 μM (A), 125 μM (B), 150 μM (C), 175 μM (D), 200 μM (E), 250 μM (F), 300 μM (G), 400 μM (H), 500 μM (I), 600 μM (J), 700 μM (K), and 800 μM (L). Two to four replicates were performed at each concentration.

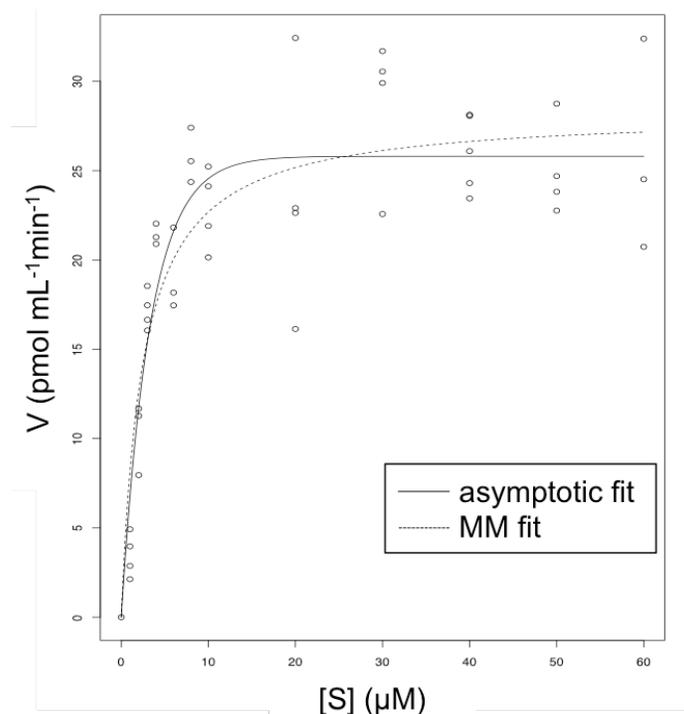


Figure 24: Michaelis-Menten and asymptotic fits of feruloyl-CoA by Bradi1g36980 enzyme. These assays were performed by saturating with 1 mM tryptamine. The two non-linear regression fits are compared, and V_{\max} is calculated to be $25.8 \text{ pmol mL}^{-1}\text{min}^{-1}$ and the K_M is $2.5 \text{ }\mu\text{M}$ based on this Michaelis-Menten approximation. The k_{cat} is calculated to be 4037 min^{-1} . Three to five replicates were performed at each substrate concentration.

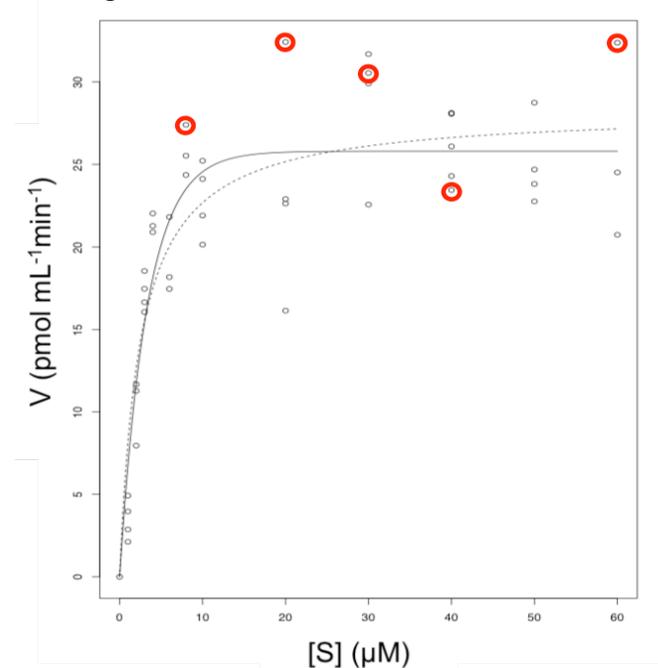


Figure 25: Potential errors in data series of kinetic assays of feruloyl-CoA by Bradi1g36980 enzyme. This series of kinetic data (from **Figure 24**) collected show high variation from the other replicates, showing potential errors in the experiment. The red circles depict one set of experiments collected on the same day that contain a number of outlying data points.

Other reasons could be related to the protein's activity. I did not have time to optimize the protein concentration for these reactions, and began with 1.5 μg in a 150 μl reaction volume. I am performing more kinetic assays to determine the best amount of enzyme needed for this reaction. It is also possible that at these higher concentrations of feruloyl-CoA we are getting protein inactivation or product inhibition.

Preliminary time course kinetic assays were performed on two acceptor substrates, tryptamine and serotonin, and Michaelis-Menten approximations were calculated for these data, but a lack of time and optimization has left us without proper K_M and V_{\max} values. From our preliminary data and calculations, these substrates do not appear to be likely substrates for Bradi1g36980 (**Figure 26**). The asymptote, known as V_{\max} , has not been reached in either substrate series, so we have not reached saturation of the substrates. The estimated K_M values for tryptamine and serotonin with these data points are 0.423 mM and 0.893 mM, respectively, which are relatively high K_M values and suggest this enzyme does not bind very efficiently using these as substrates. More concentrations and replicates would need to be performed to calculate the actual K_M values.

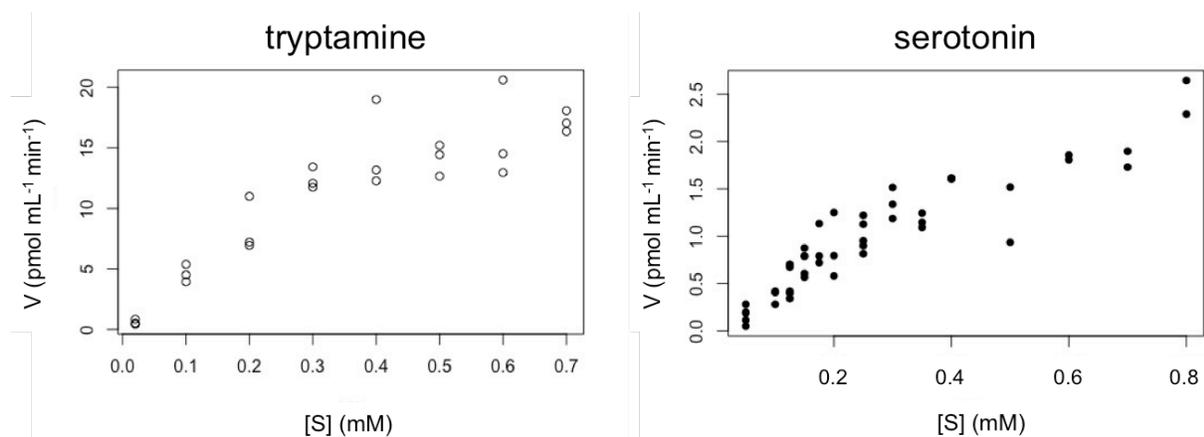


Figure 26: Preliminary data on initial velocity of product formed versus substrate concentration for tryptamine and serotonin substrates by Bradi1g36980. These assays were performed by saturating with 1 mM feruloyl-CoA. Two to four replicates were performed for each concentration.

Brachypodium seedling exposure to methyl-jasmonate

One piece of evidence that Bradi1g36980 may be involved in a wounding, jasmonate response pathway, or ethylene response pathway was analyzing the RNA-seq dataset of Brachypodium developmental stages. The expression pattern of Bradi1g36980 at different developmental stages in Brachypodium showed that of the three biological replicates (performed by a postdoc in our lab, Dr. Jacob Jensen) was significantly more highly expressed than the others (**Figure 27**). Further, when looking at the gene enrichment of the ratio of Replicate A to Replicate B, in the top forty genes (ranked by highest ratio of normalized counts) there were several genes involved in both jasmonate and ethylene response pathways (**Table 4**). Unfortunately, the reason for this induced or heightened expression of Bradi1g36980 in

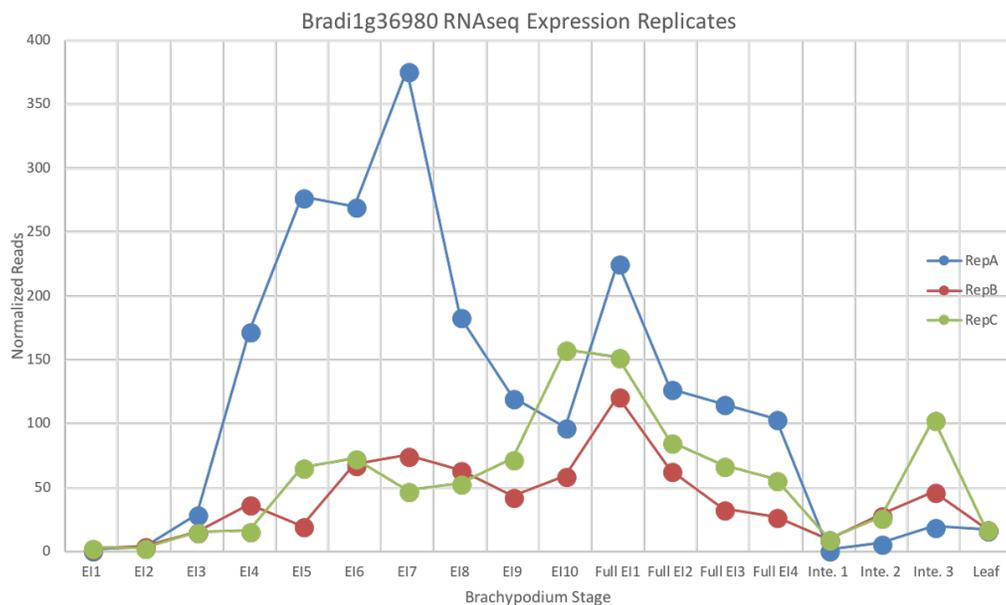


Figure 27: Graph of three biological replicates of RNA-seq data of Bradi1g36980 at different developmental stages of Brachypodium. There is an increased expression in Replicate A in elongating internodes (EI) 4-8, compared to Replicates B and C. EI1-10, elongating internodes 1-10; Full EI1-4, fully elongated internodes 1-4; Inte. 1-3, integuments 1-3; Leaf, leaf.

Name	definition	Set A	Set B	A / B
Bradi3g35920	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein LOC_Os08g32160.1	10.88	0	Inf
Bradi3g30716		10.65	0	Inf
Bradi3g29580	RIPER6 - Ripening-related family protein precursor	11.52	0.2	57.91
Bradi5g23135	LOC_Os04g54210.1 expressed protein	44.09	0.91	48.34
Bradi2g04050	AT2G30230.1 LOC_Os01g07150.1 expressed protein	6.12	0.14	43.09
Bradi3g58630	EXO Phosphate-responsive 1 family protein LOC_Os02g52010.1	5.7	0.17	33.55
Bradi3g52590	LOC_Os04g51320.1 transcription factor TF2 putative expressed	11.79	0.36	32.97
Bradi1g33070	Wound-responsive family protein LOC_Os06g46970.1	50.33	1.66	30.33
Bradi1g56010	AT3G22530.1 LOC_Os07g07670.1 expressed protein	14.28	0.5	28.41
Bradi1g71970		17.34	0.62	28.14
Bradi2g24106		24.54	1.03	23.73
Bradi1g37080	ATMBF1C MBF1C multiprotein bridging factor 1C LOC_Os06g39240.1	10.96	0.49	22.58
Bradi3g23180	JAZ11 TIFY3A jasmonate-zim-domain protein 11 LOC_Os10g25230.1	7.43	0.35	21.43
Bradi2g55920	ATNHL10 NHL10 YLS9 Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family L	133.43	6.44	20.72
Bradi2g61150	Calcium-binding EF-hand family protein LOC_Os01g72530.1	8.5	0.42	20.05
Bradi3g50490	ERF-1 ethylene responsive element binding factor 1 LOC_Os02g43790.1	12.26	0.61	20.05
Bradi2g51716	serine-rich protein-related LOC_Os02g09090.1	29.59	1.53	19.38
Bradi3g45220	MKS1 MAP kinase substrate 1 LOC_Os02g33600.1	7.39	0.39	19.18
Bradi4g35214	LOC_Os08g42800.1 expressed protein	8.57	0.45	19.14
Bradi5g19640	Protein of unknown function (DUF581) LOC_Os04g49670.1	14.82	0.81	18.28
Bradi3g18070	ATERF7 ERF7 ethylene response factor 7 LOC_Os09g26420.4	10.78	0.6	18.11
Bradi3g59690	LOC_Os02g50710.1 expressed protein	8.36	0.46	18.05
Bradi1g50120	LOC_Os06g04230.1 expressed protein	12.86	0.72	17.96
Bradi1g13940	PLA1IIA PLP7 patatin-like protein 6 LOC_Os03g43880.1	22.25	1.25	17.87
Bradi3g58620	EXO Phosphate-responsive 1 family protein LOC_Os02g52040.1	21.87	1.24	17.67
Bradi2g54800	Protein of unknown function (DUF668) LOC_Os01g62670.1	12.98	0.74	17.61
Bradi5g05110	basic helix-loop-helix (bHLH) DNA-binding superfamily protein LOC_Os04g23550.1	39.16	2.28	17.19
Bradi5g21250	ATERF-4 ATERF4 ERF4 RAP2.5 ethylene responsive element binding factor 4 LOC_Os04g52090.1	42.27	2.47	17.09
Bradi1g44170	atnudt21 NUDT21 nudix hydrolase homolog 21 LOC_Os06g14420.1	37.01	2.21	16.71
Bradi3g36475		15.26	0.95	16.12
Bradi2g12030	sequence-specific DNA binding transcription factors LOC_Os12g06640.1	8.34	0.53	15.74
Bradi2g38025		11.24	0.73	15.35
Bradi4g08010	Protein kinase superfamily protein LOC_Os12g23700.1	7.67	0.5	15.2
Bradi1g72590	JAZ1 TIFY10A jasmonate-zim-domain protein 1 LOC_Os03g08330.1	60.49	4.08	14.83
Bradi2g35850	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein LOC_Os05g05680.1	23.01	1.56	14.76
Bradi3g44910	Glycosyl hydrolase superfamily protein LOC_Os02g33000.1	8	0.55	14.54
Bradi5g23158	Wound-responsive family protein LOC_Os04g54240.1	224.84	15.52	14.49
Bradi3g43480	Protein of unknown function (DUF679) LOC_Os02g27800.1	16.09	1.11	14.47
Bradi1g33550	ATERF-4 ATERF4 ERF4 RAP2.5 ethylene responsive element binding factor 4 LOC_Os06g47590.1	39.3	2.74	14.32

Table 4: Gene enrichment table of the ratio of Replicates A to B of Bradi1g36980 RNA-seq data. The top forty enriched genes are shown, ranked by highest A/B ratio (except for the first two, which could not be inferred due to no reads in Rep B). The rows highlighted in red are involved in the jasmonate response pathway, and in green, the ethylene response pathways. Some genes in this table have not been annotated (blank “definition”) or are of unknown function, which means there could be potentially more genes that are involved in one or both of these response pathways.

Replicate A is unknown, but the enrichment in genes involved in wounding response pathways gives us an indication that perhaps Replicate A plants were under some sort of stress, such as insect herbivory, fungal pathogens, or wounded prior to collection.

Taking what we know of Bradi1g36980, that the enzyme is making phenylamide conjugates that are known to occur in wounded plants and that one replicate of RNA-seq data showed an enrichment in genes involved in either jasmonate or ethylene response pathways, we

hypothesized that Bradi1g36980 must be involved in or respond to one or both of these pathways. To test this hypothesis, we first began with the jasmonate response pathway, which has previously been shown to induce phenylamide conjugations in grasses (Bassard *et al.*, 2010; Jang *et al.*, 2004; Lee *et al.*, 1997; Onkokesung *et al.*, 2012). We exposed 2-week old Brachypodium seedlings to 5 μ M MeJA in sealed magenta boxes for various lengths of time, from 30 minutes to 48 hours. While there seems to be a slight increase in expression after 4 hours exposure, we measured no significant increase in Bradi1g36980's gene expression level in response to MeJA at any of these time points (**Figure 28**). We are confident MeJA was detected by the plants using JAZ1 gene as a positive control. Based on our findings, Bradi1g36980 is not transcriptionally regulated by the jasmonate pathway. It is possible that this gene could be involved in the ethylene response pathway, or in a convergence of the two (see **Discussion**).

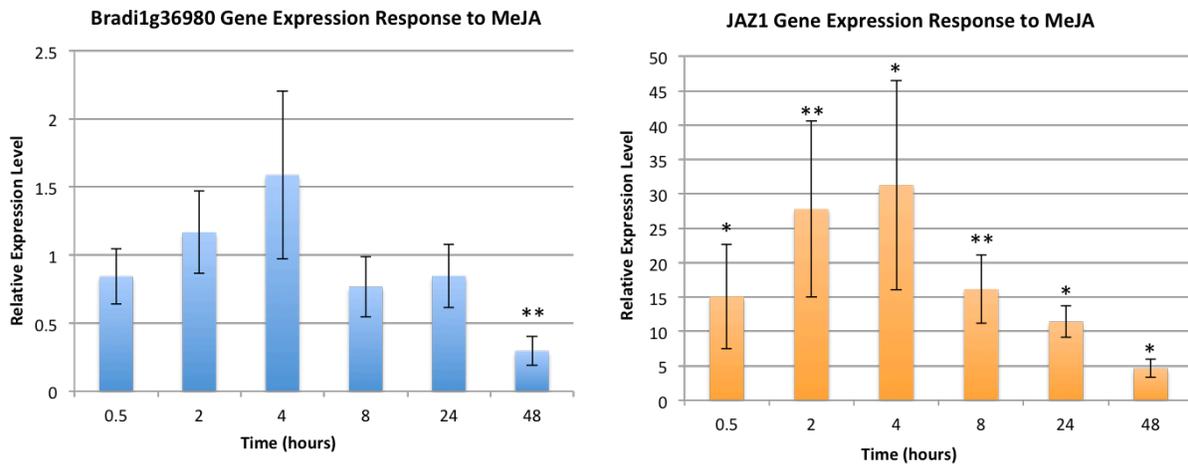


Figure 28: Relative gene expression in response to methyl-jasmonate (MeJA) over time analyzed by qRT-PCR. On the left in blue, the gene of interest Bradi1g36980 shows no significant increase in expression after exposure to MeJA. On the right in orange was our control gene, JAZ1 (Bradi1g72490), to show that added MeJA was detected by the plants to induce the jasmonate response pathway at each time point. * p-value \leq 0.05; ** p-value \leq 0.01

Agrobacterium-infiltrated tobacco plants with overexpressed gene

Another approach to determining substrate use and subsequently enzyme function in the plant is through *Agrobacterium*-mediated transient overexpression of our gene in tobacco leaves. This approach is advantageous for our purposes because it is a relatively quick and simple way to transiently overexpress our gene of interest in a plant and analyze its metabolites for changes compared to wild-type plants. It allows us to genetically study our gene *in planta* without performing a lengthy transformation in *Brachypodium*, a process that often takes 20-31 weeks to get transformed seeds (Bragg *et al.*, 2014).

We infiltrated tobacco leaves with *Agrobacterium* containing Bradi1g36980 in an overexpression vector and collected the leaves 3 days after infiltration. We used 70% ethanol to extract the soluble metabolites in the transiently expressed and untreated leaves for analysis on the LC-MS/MS. This comparison between Bradi1g36980 OX and untreated leaves was performed with the hypothesis that we would discover ferulate conjugates that occurred as a result of overexpression of our gene. When searching the positive mode ion trace of both untreated and OX samples for the characteristic ferulate fragment 177.055, we found a significant increase in feruloylated compounds in the OX sample (**Figure 29**). In an effort to identify some of these compounds using MS, we looked at the fragmentation patterns of a few select peaks in the 2 to 5 minute range, when feruloyl compounds were highest (**Figure 30**). We saw a characteristic pattern of m/z 177, 145, 117, and 89, giving us confidence that these were feruloyl conjugations. The fragment m/z 177 is found in each mass spectra, giving us confidence that there is a ferulate compound in each of these peaks.

Despite having the mass spectra of some of these feruloyl-containing peaks from transient overexpression of Bradi1g36980, it is difficult to determine what the parent ions are

that are being feruloylated in these traces. At larger masses, the number of possible chemical formulas increases, and it becomes hard to identify these compounds. The lack of fragmented masses, aside from the ferulate patterns, does not give us much information on possible structures of feruloyl compounds.

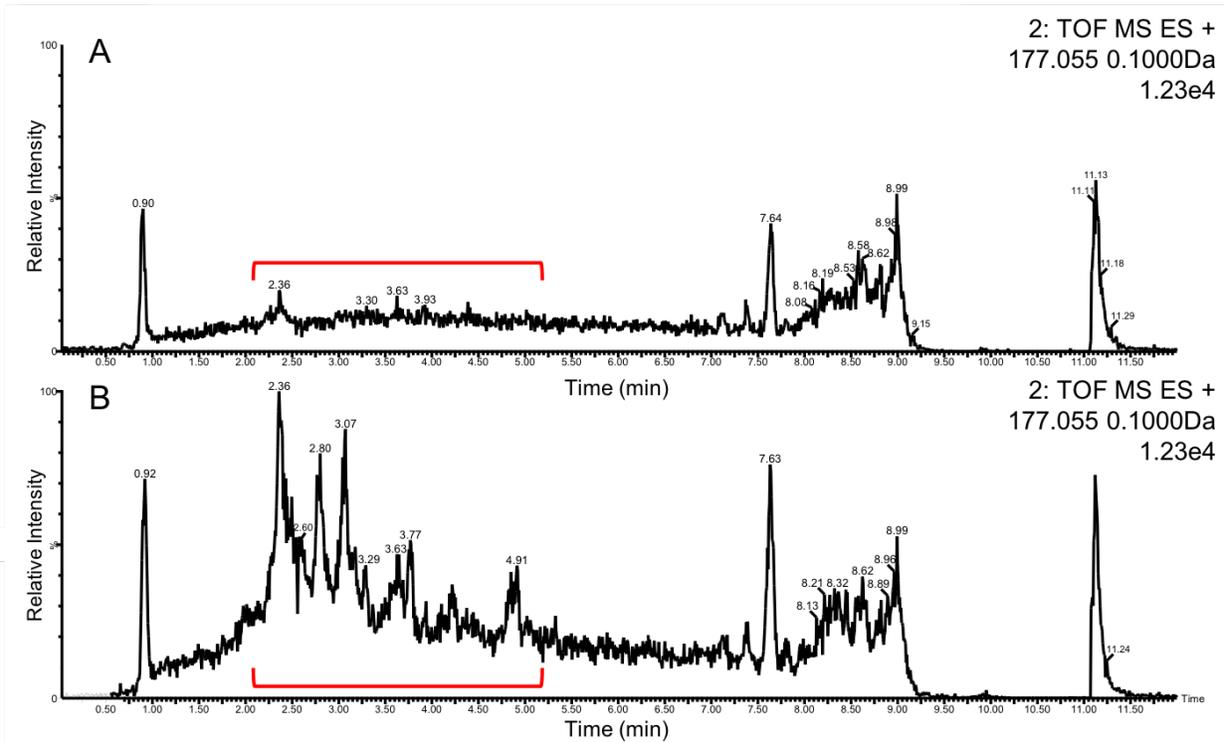


Figure 29: LC-MS/MS data of soluble extracts of transiently overexpressed Bradi1g36980 and untreated leaves in tobacco leaves. When searching for mass 177.055 in positive ion mode, a mass corresponding to a fragmentation pattern of ferulate, a large difference can be seen between ~2-5 minutes (red brackets) between untreated (A) and Bradi1g36980 OX (B). This provides evidence that this gene is responsible for the feruloylation of compounds in the plant. Some of these peaks, at 2.36, 2.60, 2.80, and 3.77 minutes, were further analyzed using MS (Figure 30).

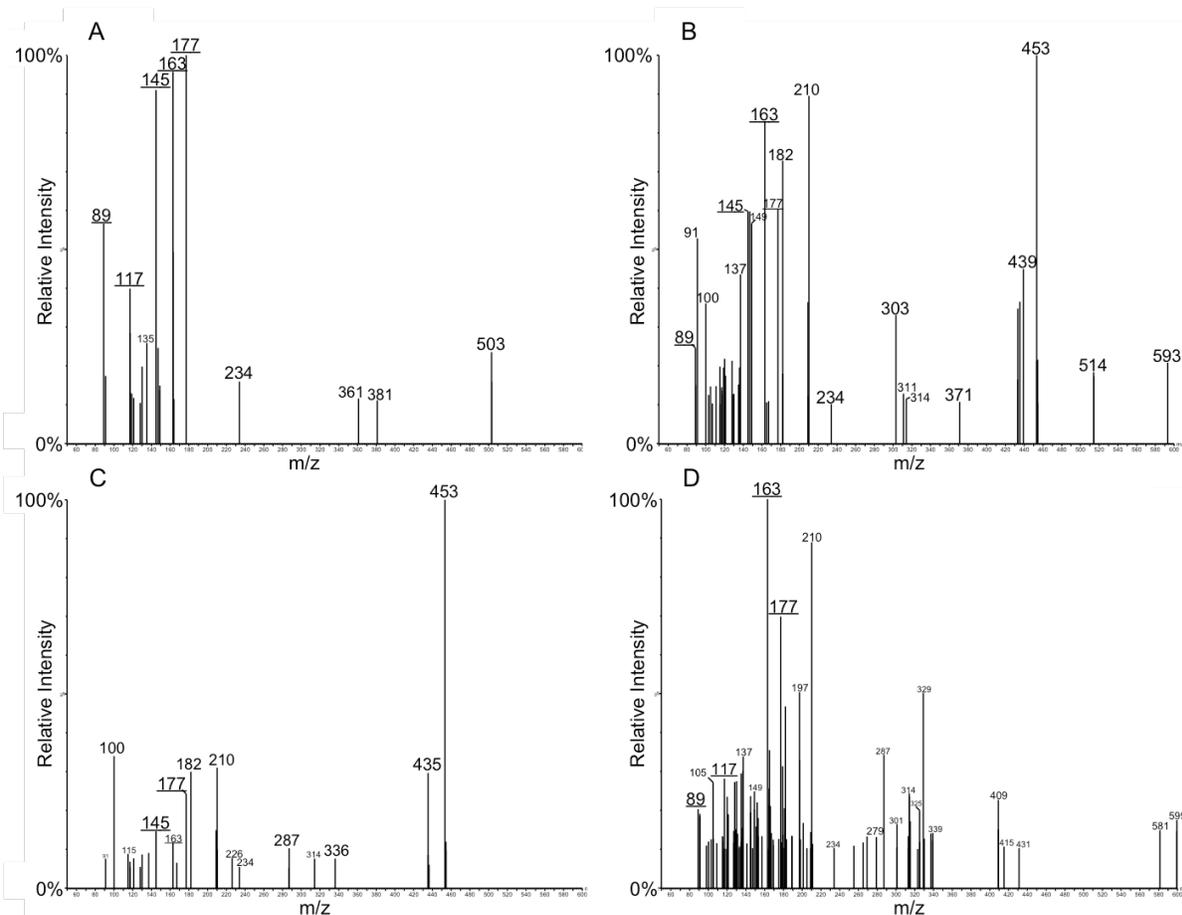


Figure 30: Mass spectra of feruloyl-containing peaks in Bradi1g36980 overexpression in tobacco leaves. These peaks were selected for MS analysis from **Figure 29B**. **A** is the spectra of peak at 2.36 min, **B** is at 2.60 min, **C** is at 2.80 min, and **D** is at 3.77 min. Masses characteristic of ferulate fragmentation patterns (m/z 177, 163, 145, 117, 89) are underlined. All spectra are produced from 20-80 eV ramped collision energy.

DISCUSSION

This chapter presents the work we have completed towards characterizing the gene Bradi1g36980, a grass-specific BAHD acyltransferase that is a member of the PF02458 clade. When performing protein assays with this enzyme, we discovered that it uses feruloyl-CoA and conjugates with phenethylamines to make *N*-acylated conjugations. The substrates that it has activity with are tyramine, 2-phenylethylamine, tryptamine, and serotonin. This enzymatic reaction with tryptamine and serotonin demonstrates the first proven activity of a BAHD AT

with either acceptor substrate. However, our preliminary enzyme kinetics data using Michaelis-Menten approximations suggest that neither acceptor substrate is likely the natural substrate for this enzyme, as their K_M values are much larger than other BAHD acyltransferase acceptor substrates. For comparison, both PMT and FMT had K_M values ranging from 35-200 μM for their acceptor substrates (Wilkerson *et al.*, 2014; Withers *et al.*, 2012). Further investigation with alternative acceptor substrates need to be performed to determine K_M constants to see which might be the natural substrate. It appears that the enzyme is highly specific to feruloyl-CoA based on its calculated K_M value of 2.5 μM . We did not perform kinetic analyses with other CoA-substrates, as Bradi1g36980 did not use *p*-coumaroyl-CoA, but it is possible that the enzyme may use others, such as caffeoyl-CoA. Bradi1g36980 did not have any activity with aliphatic amines, nor with dopamine, norepinephrine, octopamine, tyrosine, or tryptophan, suggesting that the additional hydroxyl or carboxyl groups on these compounds may interfere with access to the substrate-binding pocket. Bradi1g36980 would need to be knocked-out and overexpressed in *B. distachyon* and analyzed for metabolic changes before being able to conclude what natural substrates this enzyme uses in the plant.

These conjugations are known to occur in grasses, and appear to be in response to wounding, exposure to pathogens, or in response to jasmonate. In order to test if the jasmonate response pathway regulates this gene's expression, we performed an experiment with *Brachypodium distachyon* seedlings and MeJA. Seedlings were exposed to MeJA for various time points, between 0.5 to 48 hours, and qRT-PCR was performed to measure relative gene expression levels. At each time point, we did not see any significant increase in Bradi1g36980 gene expression, suggesting that this gene is not transcriptionally regulated by jasmonate. It is possible that it could be involved in another wound response pathway.

In one paper, researchers tried a similar experiment with a serotonin *N*-hydroxycinnamoyltransferase (SHT), a GNAT enzyme that makes *p*-coumaroyl- and feruloyl-serotonin. They tested to see if transgenic rice plants overexpressing SHT produced increased HCA amides in response to jasmonic acid or wounding, but they found these plants did not respond to either treatment (Jang *et al.*, 2004). They concluded that either JA or wounding does not stimulate serotonin synthesis, or that SHT is not involved due to compartmentalization of substrates. There are differences between their experiment and ours, including that these enzymes are in different classes, they were using transgenic plants overexpressing SHT and were measuring only for increase in HCA amides, not gene expression. However, it does add to the story that enzymes producing feruloyl-phenylamides may not be involved in the jasmonate wound response pathway, and that they may play different roles in the plant.

There are alternative hormones and pathways that are involved in plant response to and defense against pests, including ethylene, salicylic acid (SA), and abscisic acid (ABA). Plants that lack ethylene sensitivity are known to be more susceptible to several species of fungi (Knoester *et al.*, 1998; Hoffman *et al.*, 1999; Thomma *et al.*, 1999) and the soft-rot bacterium *E. carotovora* (Norman-Setterblad *et al.*, 2000), demonstrating its importance in plant responses to pathogen attack. The phytohormones ethylene and jasmonate have been shown by research groups to function in a synergistic way to elicit plant defense responses (Xu *et al.*, 1994; O'Donnell *et al.*, 1996). In one study, researchers found that both jasmonate and ethylene were required to activate transcription of the ethylene response factor 1 (ERF1), a transcription factor responsible for regulating genes in defense responsive pathways (Lorenzo *et al.*, 2002).

The gene enrichment table (**Table 4**) of genes highly co-expressed with Bradi1g36980 RNA-seq data in Replicate A compared to Replicate B shows a number of genes, highlighted in

green, that are involved in the ethylene response pathway. One of these is ERF1, the transcription factor mentioned above that responds to both jasmonate and ethylene simultaneously. This enriched gene, along with the other ethylene pathway genes, gives us a good indication that Bradi1g36980 might respond to ethylene or to both ethylene and JA. Our first attempt to test this in *Brachypodium* seedlings used jasmonic acid and 1-aminocyclopropanecarboxylic acid (ACC), a precursor to ethylene, and qRT-PCR to measure changes in relative gene expression. This experiment unfortunately failed, as we did not see a response in the positive controls for JAZ1 and ERF1, indicating that the plants did not recognize JA and ethylene. We plan to test this again, possibly using ethylene gas and methyl-jasmonate instead of ACC and jasmonic acid, respectively, to ensure that plants are exposed to and the volatile phytohormones.

Agrobacterium-infiltration in tobacco leaves of Bradi1g36980 in an overexpression construct showed an increase in feruloyl-conjugations compared to untreated tobacco leaves, suggesting that this gene is involved in feruloylation of compounds in plants. While we were unable to identify these compounds formed, we are confident that they contain ferulate due to the fragmentation patterns. If given more time and possibly by adjusting the MS collision energy to get new compound fragments, it is possible we could more easily identify these feruloyl-conjugates to determine Bradi1g36980's preferred acceptor substrate.

Another experiment that would be important to try but we did not have time to do would be to analyze the *Agrobacterium*-infiltrated tobacco leaves for conjugates that have been covalently bound into the cell wall. Some of these feruloyl-conjugates, including feruloyl-serotonin and feruloyl-tyramine, are able to cross-link to lignin via ether bonds, potentially as a method of repairing damaged cell walls (Edreva *et al.*, 2007; McLusky *et al.*, 1999; Négrel *et al.*,

1996). One way to analyze the ether-linkages of the cell wall is with thioacidolysis, a method of cleaving ether bonds while keeping ester and amide linkages intact (Lapierre *et al.*, 1996). It is possible that some of these feruloyl-conjugates made by overexpressed Bradi1g36980 protein in tobacco leaves are being covalently bound to the cell wall, removing them from our soluble extract samples run on the LC-MS/MS. If the ether-bound compounds can be extracted from the cell wall using thioacidolysis and then analyzed with mass spectrometry, we may be able to identify conjugates and substrates used by Bradi1g36980 to then assay *in vitro*. However, many substrates will not survive the harsh thioacidolysis treatment, so further investigation into finding cell wall-bound conjugates will be required.

Bradi1g36980 was previously studied using genetic approaches by Sibout *et al.* (2016), who overexpressed the gene in Arabidopsis under the control of the Arabidopsis C4H promoter, which targets gene expression to vascular tissue. This overexpression led to a substantial increase in alkali-soluble *p*-coumarates from the cell wall, comparable to the amounts measured in grass lignin levels (8% to 9% by weight). The researchers concluded that this gene is responsible for the addition of *p*-coumarate on monolignols, the second PMT, namely PMT2. As we have shown by our research, this enzyme is not performing the functions of PMT, previously characterized by our lab (Withers *et al.*, 2012). We believe this study is another example of an indirect phenotypic effect resulting from reverse genetics. While the OX lines of Bradi1g36980 did result in measurable changes in alkali-labile *p*-coumarate in lignin, we believe this increase in *p*-coumarate on lignin may be a result of depleted ferulate pools in the cell due to ferulate consumption by Bradi1g36980 protein. Without available feruloyl-CoA for lignin modification, the transgenic plant would likely have free *p*-coumaroyl-CoA to use for lignin acylation. We were interested in seeing the change in feruloyl content of lignin in these Bradi1g36980 OX lines

in Arabidopsis, however these results were not published in this paper. As this study was performed in Arabidopsis, which is a eudicot and therefore lacks arabinoxylan, the change in *p*CA and fA content on AX could not be measured. With additional time, I would be interested in transforming Brachypodium plants with both overexpression and knock-down constructs to study the changes in the cell wall content and hydroxycinnamate modifications, as well as search for any increase in ferulate conjugates.

Bradi1g36980 is a member of the PF02458 clade, a highly differentially expressed clade of BAHD ATs thought to contain the enzyme responsible for addition of ferulate on arabinoxylan. Other members of this clade have been found to be involved in cell wall biosynthesis, including PMT. The discovery that this enzyme is *N*-feruloylating phenylamines breaks from the clade's functions in the cell wall biosynthesis. The work conducted with this grass-specific BAHD acyltransferase proves that some members of BAHD ATs are able to *N*-acylate serotonin and tryptamine, a reaction never before shown. Our initial hypothesis that this gene may be involved in the wounding response pathway has not been proven, as the gene did not respond to exposure to MeJA. While the role of this gene in the plant has yet to be determined, it is a valuable contribution to our understanding of this grass-specific clade as a whole.

REFERENCES

REFERENCES

- Anders, N., Wilkinson, M.D., Lovegrove, A., Freeman, J., Tryfona, T., Pellny, T.K., Weimar, T., Mortimer, J.C., Stott, K., Baker, J.M., Defoin-Platel, M., Shewry, P.R., Dupree, P., and Mitchell, R.A.C. (2012). Glycosyl transferases in family 61 mediate arabinofuranosyl transfer onto xylan in grasses. *Proceedings of the National Academy of Sciences*, 109(3), 989–93.
- Anderson, M.J., Stark, J.C., Hodgman, C.E., and Jewett, M.C. (2015). Energizing eukaryotic cell-free protein synthesis with glucose metabolism. *FEBS Letters*, 589(15), 1723–1727.
- Bartley, L.E., Peck, M.L., Kim, S.-R., Ebert, B., Manisseri, C., Chiniquy, D.M., Sykes, R., Gao, L., Rautengarten, C., Vega-Sánchez, Benke, P.I., Canlas, P.E., Cao, P., Brewer, S., Lin, F., Smith, W.L., Zhang, X., Keasling, J.D., Jentoff, R.E., Foster, S.B., Zhou, J., Ziebell, A., An, G., Scheller, H.V., and Ronald, P.C. (2013). Overexpression of a BAHD acyltransferase, OsAT10, alters rice cell wall hydroxycinnamic acid content and saccharification. *Plant Physiology*, 161(4), 1615–33.
- Bassard, J.E., Ullmann, P., Bernier, F., and Werck-Reichhart, D. (2010). Phenolamides: Bridging polyamines to the phenolic metabolism. *Phytochemistry*, 71(16), 1808–1824.
- Boerjan, W., Ralph, J., and Baucher, M. (2003). Lignin biosynthesis. *Annual Review of Plant Biology*, 54(1), 519–546.
- Bokern, M., Wray, V., and Strack, D. (1991). Accumulation of Phenolic-Acid Conjugates and Betacyanins, and Changes in the Activities of Enzymes Involved in Feruloylglucose Metabolism in Cell-Suspension Cultures of *Chenopodium rubrum* L. *Planta*, 184(2), 261–270.
- Bontpart, T., Cheynier, V., Ageorges, A., and Terrier, N. (2015). BAHD or SCPL acyltransferase? What a dilemma for acylation in the world of plant phenolic compounds. *New Phytologist*, 208(3), 695–707.
- Boudet, A.M. (2007). Evolution and current status of research in phenolic compounds. *Phytochemistry*, 68, 2722–2735.
- Bragg, J.N., Anderton, A., Nieu, R., and Vogel, J.P. (2014). *Brachypodium distachyon*. *Agrobacterium Protocols: Volume 1*, 1223, 17–33.
- Buanafina, M.M. de O. (2009). Feruloylation in grasses: current and future perspectives. *Molecular Plant*, 2(5), 861–72.
- Carpita, N.C. (1996). Structure and Biogenesis of the Cell Walls of Grasses. *Annual Review of Plant Physiology and Plant Molecular Biology*, 47, 445–76.

- Cassab, G.I. (1998). Plant Cell Wall Proteins. *Annual Review of Plant Physiology and Plant Molecular Biology*, 49, 281–309.
- Chiniquy, D., Sharma, V., Schultink, A., Baidoo, E.E., Rautengarten, C., Cheng, K., Carroll, A., Ulvskov, P., Harholt, J., Keasling, J.D., Pauly, M., and Ronald, P.C. (2012). XAX1 from glycosyltransferase family 61 mediates xylosyltransfer to rice xylan. *Proceedings of the National Academy of Sciences*, 109(42), 17117–17122.
- Coutu, C., Brandle, J., Brown, D., Brown, K., Miki, B., Simmonds, J., and Hegedus, D.D. (2007). PORE: A modular binary vector series suited for both monocot and dicot plant transformation. *Transgenic Research*, 16(6), 771–781.
- Crawley, M.J. (2007). *The R Book*. Chichester: Wiley & Sons Ltd.
- D’Auria, J.C. (2006). Acyltransferases in plants: a good time to be BAHD. *Current Opinion in Plant Biology*, 9, 331–40.
- De Pino, V., Borán, M., Norambuena, L., González, M., Reyes, F., Orellana, A., and Moreno, S. (2007). Complex formation regulates the glycosylation of the reversibly glycosylated polypeptide. *Planta*, 226, 335–345.
- de Souza, W.R., Martins, P.K., Freeman, J., Pellny, T.K., Michaelson, L.V., Sampaio, B.L., Vinecky, F., Ribeiro, A.P., da Cunha, B.A.D.B., Kobayashi, A.K., de Oliveira, P.A., Campanha, R.B., Pacheco, T.F., Martarello, D.C.I., Marchiosi, R., Ferrarese-Filho, O., dos Santos, W.D., Tramontina, R., Squina, F.M., Centeno, D.C., Gaspar, M., Braga, M.R., Tiné, M.A.S., Ralph, J., Mitchell, R.A.C., and Molinari, H.B.C. (2018). Suppression of a single BAHD gene in *Setaria viridis* causes large, stable decreases in cell wall feruloylation and increases biomass digestibility. *New Phytologist*, 218(1), 81-93.
- Dixon, R.A. (2001). Natural products and plant disease resistance. *Nature*, 411, 843–847.
- Doblin, M.S., Kurek, I., Jacob-Wilk, D., and Delmer, D.P. (2002). Cellulose biosynthesis in plants: from genes to rosettes. *Plant Cell Physiol.*, 43, 1407-1420.
- Drakakaki, G., Zabolina, O., Delgado, I., Robert, S., Keegstra, K., and Raikhel, N. (2006). Arabidopsis reversibly glycosylated polypeptides 1 and 2 are essential for pollen development. *Plant Physiology*, 142, 1480–1492.
- Edreva, A.M., Velikova, V.B., and Tsonev, T.D. (2007). Phenylamides in Plants. *Russian Journal of Plant Physiology*, 54(3), 287-301.
- Engels, F.M., and Jung, H.G. (1998). Alfalfa Stem Tissues: Cell-wall Development and Lignification. *Annals of Botany*, 82, 561–568.
- Erland, L.A.E. and Saxena, P.K. (2017). Beyond a neurotransmitter: the role of serotonin in plants. *Neurotransmitter*, 4(e1538), 1-12.

- Escamilla-Treviño, L.L., Shen, H., Hernandez, T., Yin, Y., Xu, Y., and Dixon, R.A. (2013). Early lignin pathway enzymes and routes to chlorogenic acid in switchgrass (*Panicum virgatum* L.). *Plant Molecular Biology*, 84(4-5), 565-576.
- Faik, A. (2010). Xylan biosynthesis: news from the grass. *Plant Physiology*, 153, 396–402.
- Foster, C.E., Martin, T.M., and Pauly, M. (2010). Comprehensive compositional analysis of plant cell walls (Lignocellulosic biomass) part I: lignin. *Journal of Visualized Experiments*, 37, 5–8.
- Fry, S.C., and Miller, J.G. (1989). Toward a working model of the growing plant cell wall. *Plant Cell Wall Polymers*, 399, 3–33.
- Gális, I., Simek, P., Narisawa, T., Sasaki, M., Horiguchi, T., Fukuda, H., Matsuoka, K. (2006). A novel R2R3 MYB transcription factor NtMYBJS1 is a methyl jasmonate-dependent regulator of phenylpropanoid-conjugate biosynthesis in tobacco. *The Plant Journal*, 46, 573–592.
- Geisler, D., Sampathkumar, A., Mutwil, M., and Persson, S. (2008). Laying down the bricks: logistic aspects of cell wall biosynthesis. *Current Opinion in Plant Biology*, 11(6), 647–52.
- Geserick, C. and Tenhaken, R. (2013). UDP-sugar pyrophosphorylase is essential for arabinose and xylose recycling, and is required during vegetative and reproductive growth in *Arabidopsis*. *The Plant Journal*, 74(2), 239-247.
- Gibeaut, D.M., and Carpita, N.C. (1994). Biosynthesis of plant cell wall polysaccharides. *Plant Cell Wall Synthesis*, 904–15.
- Grabber, J.H. (2005). How Do Lignin Composition, Structure, and Cross-Linking Affect Degradability? A Review of Cell Wall Model Studies. *Crop Science*, 45(3), 820.
- Grabber, J.H., Hatfield, R.D., Ralph, J., Zoń, J., and Amrhein, N. (1995). Ferulate cross-linking in cell walls isolated from maize cell suspensions. *Phytochemistry*, 40(4), 1077–1082.
- Grabber, J.H., Ralph, J., Lapierre, C., and Barrière, Y. (2004). Genetic and molecular basis of grass cell-wall degradability. I. Lignin–cell wall matrix interactions. *Comptes Rendus Biologies*, 327, 455–465.
- Grisebach, H. (1981) in *The Biochemistry of Plants* (E.K. Stumpf and E.E. Conn, eds.), Vol. 7, p. 457. London, Academic Press.
- Gross, G.G. and Zenk, M.H. (1974). Isolation and Properties of Hydroxycinnamate:CoA Ligase from Lignifying Tissue of *Forsythia*. *European Journal of Biochemistry*, 42, 453-459.

- Hatfield, R.D., Marita, J.M., Frost, K., Grabber, J.H., Ralph, J., Lu, F., and Kim, H. (2009). Grass lignin acylation: p-coumaroyl transferase activity and cell wall characteristics of C3 and C4 grasses. *Planta*, 229, 1253–67.
- Hatfield, R.D., Ralph, J., and Grabber, J.H. (1999a). Cell Wall Structural Foundations: Molecular Basis for Improving Forage Digestibilities. *Crop Science*, 39, 27–37.
- Hatfield, R.D., Wilson, J.R., and Mertens, D.R. (1999b). Composition of cell walls isolated from cell types of grain sorghum stems. *Journal of The Science of Food and Agriculture*, 79, 891-899.
- Hatfield, R.D., Ralph, J., and Grabber, J.H. (1999c). Cell wall cross-linking by ferulates and diferulates in grasses. *Journal of the Science of Food and Agriculture*, 79, 403–407.
- Hoffman, T., Schmidt, J.S., Zheng, X., and Bent, A. (1999). Isolation of ethylene-insensitive soybean mutants that are altered in pathogen susceptibility and gene-for-gene disease resistance. *Plant Physiology*, 119, 935–949.
- Hohlfeld, H., Schurmann, W., Scheel, D., and Strack, D. (1995). Partial Purification and Characterization of Hydroxycinnamoyl-Coenzyme A:Tyramine Hydroxycinnamoyltransferase from Cell Suspension Cultures of *Solanum tuberosum*. *Plant Physiology*, 107(2), 545–552.
- Iiyama, K., Lam, T.B.T., and Stone, B.A. (1990). Phenolic Acid Bridges Between Polysaccharides and Lignin in Wheat Internodes. *Phytochemistry*, 29(3), 733–737.
- Ishihara, A., Kawata, N., Matsukawa, T., and Iwamura, H. (2000). Induction of N-hydroxycinnamoyltyramine synthesis and tyramine N-hydroxycinnamoyltransferase (THT) activity by wounding in maize leaves. *Bioscience Biotechnology and Biochemistry*, 64(5), 1025-1031.
- Ishihara, A., Hashimoto, Y., Miyagawa, H., and Wakasa, K. (2008a). Induction of serotonin accumulation by feeding of rice striped stem borer in rice leaves. *Plant Signaling and Behavior*, 3(9), 714-716.
- Ishihara, A., Hashimoto, Y., Tanaka, C., Dubouzet, J.G., Nakao, T., Matsuda, F., Nishioka, T., Miyagawa, H., and Wakasa, K. (2008b). The tryptophan pathway is involved in the defense responses of rice against pathogenic infection via serotonin production. *The Plant Journal*, 54(3), 481-495.
- Ishii, T. (1991). Isolation and characterization of a diferuloyl arabinoxylan hexasaccharide from bamboo shoot cell-walls. *Carbohydrate Research*, 219, 15–22.
- Jang, S.-M., Ishihara, A., and Back, K. (2004). Production of coumaroylserotonin and feruloylserotonin in transgenic rice expressing pepper hydroxycinnamoyl-coenzyme A:serotonin N-(hydroxycinnamoyl)transferase. *Plant Physiology*, 135(1), 346–56.

- Jung, H.G. (2003). Maize stem tissues: ferulate deposition in developing internode cell walls. *Phytochemistry*, 63(5), 543–549.
- Jung, H.G., and Phillips, R.L. (2010). Putative Seedling Ferulate Ester (sfe) Maize Mutant: Morphology, Biomass Yield, and Stover Cell Wall Composition and Rumen Degradability. *Crop Science*, 50(1), 403–418.
- Kang, K., Jang, S.M., Kang, S., and Back, K. (2005). Enhanced nutraceutical serotonin derivatives of rice seed by hydroxycinnamoyl-CoA:serotonin N-(hydroxycinnamoyl)-transferase. *Plant Science*, 168(3), 783–788.
- Kang, S., Kang, K., Chung, G.C., Choi, D., Ishihara, A., Lee, D.-S., and Back, K. (2006). Functional analysis of the amine substrate specificity domain of pepper tyramine and serotonin N-hydroxycinnamoyltransferases. *Plant Physiology*, 140(2), 704–715.
- Keegstra, K. (2010). Plant cell walls. *Plant Physiology*, 154, 483–6.
- Knoester, M., van Loon, L.C., van den Heuvel, J., Henning, J., Bol, J.F., and Linthorst, H.J.M. (1998). Ethylene-insensitive tobacco lacks nonhost resistance against soil-borne fungi. *Proc. Natl. Acad. Sci. USA*, 95, 1933–1937.
- Konishi, T., Ohnishi-Kameyama, M., Funane, K., Miyazaki, Y., Konishi, T., and Ishii, T. (2010). An arginyl residue in rice UDP-arabinopyranose mutase is required for catalytic activity and autoglycosylation. *Carbohydrate Research*, 345(6), 787–91.
- Konishi, T., Takeda, T., Miyazaki, Y., Ohnishi-Kameyama, M., Hayashi, T., O'Neill, M., and Ishii, T. (2007). A plant mutase that interconverts UDP-arabinofuranose and UDP-arabinopyranose. *Glycobiology*, 17(3), 345–54.
- Koo, A.J.K., and Howe, G.A. (2009). The wound hormone jasmonate. *Phytochemistry*, 70(13–14), 1571–1580.
- Langeveld, S.M.J., Vennik, M., Kottenhagen, M., Van Wijk, R., Buijk, A., Kijne, J.W., and de Pater, S. (2002). Glucosylation activity and complex formation of two classes of reversibly glycosylated polypeptides. *Plant Physiology*, 129(May), 278–289.
- Lapierre, C., Pollet, B., Négrel, J. (1996). The phenolic domain of potato suberin: structural comparisons with lignins. *Phytochemistry*, 42(4), 949–953.
- Lee, J., Vogt, T., Schmidt, J., Parthier, B., and Lobler, M. (1997). Methyl jasmonate-induced accumulation of coumaroyl conjugates in barley leaf segments. *Phytochemistry*, 44(4), 589–592.
- Lehfeldt, C., Shirley, A.M., Meyer, K., Ruegger, M.O., Cusumano, J.C., Viitanen, P.V., Strack, D., and Chapple, C. (2000). Cloning of the SNG1 gene of Arabidopsis reveals a role for a

- serine carboxypeptidase-like protein as an acyltransferase in secondary metabolism. *The Plant Cell*, *12*, 1295–1306.
- Lenucci, M.S., Piro, G., and Dalessandro, G. (2009). In muro feruloylation and oxidative coupling in monocots: A possible role in plant defense against pathogen attacks. *Plant Signaling and Behavior*, *4*(3), 228–230.
- Lerouxel, O., Cavalier, D.M., Liepman, A.H., and Keegstra, K. (2006). Biosynthesis of plant cell wall polysaccharides - a complex process. *Current Opinion in Plant Biology*, *9*(6), 621–30.
- Li, S., Bashline, L., Lei, L., and Gu, Y. (2014). Cellulose Synthesis and Its Regulation. *The Arabidopsis Book*, *12*, e0169.
- Liu, C.-J. (2010). Biosynthesis of hydroxycinnamate conjugates: implications for sustainable biomass and biofuel production. *Biofuels*, *1*(5), 745–761.
- Lorenzo, O., Piqueras, R., Sanchez-Serrano, J.J., and Solano, R. (2003). ETHYLENE RESPONSE FACTOR1 Integrates Signals from Ethylene and Jasmonate Pathways in Plant Defense. *The Plant Cell*, *15*(1), 165–178.
- Lu, C., Yuan, S., and Li, J. (2006). Contribution of both cell elongation and cell division to wheat coleoptile elongation. *Belgian Journal of Botany*, *139*(2), 167–172.
- Lu, L., Berkey, K.A., and Casero, R.A. (1996). RGF GIGS Is an Amino Acid Sequence Required for Acetyl Coenzyme A Binding and Activity of Human Spermidine / Spermine N¹ Acetyltransferase. *Biochemistry*, *271*(31), 18920–18924.
- Ma, X., Koepke, J., Panjekar, S., Fritzsche, G., and Stockigt, J. (2005). Crystal structure of vinorine synthase, the first representative of the BAHD superfamily. *Journal of Biological Chemistry*, *280*, 13576–13583.
- Marti, G., Erb, M., Boccard, J., Glauser, G., Doyen, G.R., Villard, N., Robert, C.A.M., Turlings, T.C.J., Rudaz, S., and Wolfender, J.-L. (2013). Metabolomics reveals herbivore-induced metabolites of resistance and susceptibility in maize leaves and roots. *Plant, Cell and Environment*, *36*(3), 621–639.
- McLusky, S.R., Bennett, M.H., Beale, M.H., Lewis, M.J., Gaskin, P., and Mansfield, J.W. (1999). Cell wall alterations and localized accumulation of feruloyl-3'-methoxytyramine in onion epidermis at sites of attempted penetration by *Botrytis allii* are associated with actin polarisation, peroxidase activity and suppression of flavinoid biosynthesis. *The Plant Journal*, *17*, 523–534.
- Mitchell, R.A.C., Dupree, P., and Shewry, P.R. (2007). A novel bioinformatics approach identifies candidate genes for the synthesis and feruloylation of arabinoxylan. *Plant Physiology*, *144*, 43–53.

- Mohnen, D. (2008). Pectin structure and biosynthesis. *Current Opinion in Plant Biology*, 11(3), 266–277.
- Molinari, H.B.C., Pellny, T.K., Freeman, J., Shewry, P.R., and Mitchell, R.A.C. (2013). Grass cell wall feruloylation: distribution of bound ferulate and candidate gene expression in *Brachypodium distachyon*. *Frontiers in Plant Science*, 4(50), 1-10.
- Morré, D.J., and Mollenhauer, H.H. (1964). Isolation of the Golgi apparatus from plant cells. *The Journal of Cell Biology*, 23(2), 295–305.
- Mottiar, Y., Vanholme, R., Boerjan, W., Ralph, J., and Mansfield, S.D. (2016). Designer lignins: harnessing the plasticity of lignification. *Current Opinion in Biotechnology*, 37, 190-200.
- Mugford, S.T., Qi, X., Bakht, S., Hill, L., Wegel, E., Hughes, R.K., Papadopoulou, K., Melton, R., Philo, M., Sainsbury, F., Lomonosoff, G.P., Roy, A.D., Goss, R.J.M., and Osbourn, A. (2009). A Serine Carboxypeptidase-Like acyltransferase is required for synthesis of antimicrobial compounds and disease resistance in oats. *The Plant Cell*, 21(8), 2473–2484.
- NASEM (2018). The National Academies presents: What You Need to Know About Energy. [online] Retrieved from <http://needtoknow.nas.edu/energy/energy-sources/fossil-fuels/> on July 1, 2018.
- Négrel, J., and Martin, C. (1984). The biosynthesis of feruloyltyramine in *Nicotiana tabacum*. *Phytochemistry*, 23(12), 2797–2801.
- Négrel, J., Pollet, B., and Lapierre, C. (1996). Ether-linked ferulic acid amides in natural and wound periderms of potato tuber. *Phytochemistry*, 43(6), 1195-1199.
- Neutelings, G. (2011). Lignin variability in plant cell walls: contribution of new models. *Plant Science*, 181, 379–86.
- Nikolić, D., Gödecke, T., Chen, S.-N., White, J., Lankin, D. C., Pauli, G.F., and van Breemen, R. B. (2012). Mass spectrometric dereplication of nitrogen-containing constituents of black cohosh (*Cimicifuga racemosa* L.). *Fitoterapia*, 83(3), 441–460.
- Norman-Setterblad, C., Vidal, S., and Palva, E.T. (2000). Interacting signal pathways control defense gene expression in *Arabidopsis* in response to cell wall-degrading enzymes from *Erwinia carotovora*. *Molecular Plant-Microbe Interactions*, 13, 430–438.
- Obel, N., Porchia, A.C., and Scheller, H.V. (2002). Dynamic changes in cell wall polysaccharides during wheat seedling development. *Phytochemistry*, 60(6), 603–610.
- Obel, N., Porchia, A.C., and Scheller, H.V. (2003). Intracellular feruloylation of arabinoxylan in wheat: evidence for feruloyl-glucose as precursor. *Planta*, 216(4), 620–9.

- O'Donnell, P.J., Calvert, C., Atzorn, R., Wasternack, C., Leyser, H.M.O., and Bowles, D.J. (1996). Ethylene as a signal mediating the wound response of tomato plants. *Science*, 274, 1914–1917.
- Onkokesung, N., Gaquerel, E., Kotkar, H., Kaur, H., Baldwin, I.T., and Galis, I. (2012). MYB8 Controls Inducible Phenolamide Levels by Activating Three Novel Hydroxycinnamoyl-Coenzyme A:Polyamine Transferases in *Nicotiana attenuata*. *Plant Physiology*, 158(1), 389–407.
- Pauly, M. and Keegstra, K. (2008). Cell-wall carbohydrates and their modification as a resource for biofuels. *The Plant Journal*, 54(4), 559–568.
- Perlack, R.D., Wright, L.L., Turhollow, A.F., Graham, R.L., Stokes, B.J., and Erbach, D.C. (2005). Biomass as feedstock for a bioenergy and bioproducts industry: the technical feasibility of a billion-ton annual supply. U.S. Department of Energy.
- Petrik, D.L., Karlen, S.D., Cass, C.L., Padmakshan, D., Lu, F., Liu, S., Le Bris, P., Antelme, S., Santoro, N., Wilkerson, C.G., Sibout, R., Lapierre, C., Ralph, J., and Sedbrook, J.C. (2014). *p*-Coumaroyl-CoA:monolignol transferase (PMT) acts specifically in the lignin biosynthetic pathway in *Brachypodium distachyon*. *The Plant Journal*, 77(5), 713–726.
- Piston, F., Uauy, C., Fu, L., Langston, J., Labavitch, J., and Dubcovsky, J. (2010). Down-regulation of four putative arabinoxylan feruloyl transferase genes from family PF02458 reduces ester-linked ferulate content in rice cell walls. *Planta*, 231(3), 677–91.
- Premium PLUS Expression Kit_E_ver.1.2, Feb. 8, 2013. CellFree Sciences Co., Ltd. Matsuyama, Japan. Accessed on July 21, 2018. [online] Available at <http://www.cfsciences.com/eg/119-entry-1/325-premium-plus-expression-kit#manual>
- Ralph, J. (2010). Hydroxycinnamates in lignification. *Phytochemistry Reviews*, 9(1), 65–83.
- Ralph, J., Helm, R.F., Quideau, S., and Hatfield, R.D. (1992). Lignin–feruloyl ester cross-links in grasses. Part I. Incorporation of feruloyl esters into coniferyl alcohol dehydrogenation polymers. *Journal of Chemical Society, Perkin Transactions 1*, (21), 2961–2969.
- Ralph, J., Quideau, S., Grabber, J.H., and Hatfield, R.D. (1994). Identification and Synthesis of New Ferulic Acid Dehydrodimers Present in Grass Cell Walls. *Journal of Chemical Society, Perkin Transactions 1*, (23), 3485–3498
- Rautengarten, C., Birdseye, D., Pattathil, S., McFarlane, H.E., Saez-Aguayo, S., Orellana, A., Persson, S., Hahn, M.G., Scheller, H.V., Heazlewood, J.L., and Ebert, B. (2017). The elaborate route for UDP-arabinose delivery into the Golgi of plants. *Proceedings of the National Academy of Sciences*, 114(16), 4261–4266.
- Rautengarten, C., Ebert, B., Herter, T., Petzold, C.J., Ishii, T., Mukhopadhyay, A., Usadel, B., and Scheller, H.V. (2011). The interconversion of UDP-arabinopyranose and UDP-

- arabinofuranose is indispensable for plant development in Arabidopsis. *The Plant Cell*, 23(4), 1373–90.
- Saulnier, L., Crépeau, M.J., Lahaye, M., Thibault, J.F., Garcia-Conesa, M.T., Kroon, P.A., and Williamson, G. (1999). Isolation and structural determination of two 5,5'-diferuloyl oligosaccharides indicate that maize heteroxylans are covalently cross-linked by oxidatively coupled ferulates. *Carbohydrate Research*, 320(1–2), 82–92.
- Saulnier, L., Sado, P. E., Branlard, G., Charmet, G., and Guillon, F. (2007). Wheat arabinoxylans: Exploiting variation in amount and composition to develop enhanced varieties. *Journal of Cereal Science*, 46(3), 261–281.
- Sawasaki, T., Hasegawa, Y., Tsuchimochi, M., Kamura, N., Ogasawara, T., Kuroita, T., and Endo, Y. (2002). A bilayer cell-free protein synthesis system for high-throughput screening of gene products. *FEBS Letters*, 514, 102–105.
- Scalbert, A., Monties, B., Lallemand, J., Guittet, E., and Rolando, C. (1985). Ether linkage between phenolic acids and lignin fractions from wheat straw. *Phytochemistry*, 24(6), 1359–1362.
- Scheller, H. V., and Ulvskov, P. (2010). Hemicelluloses. *Annual Review of Plant Biology*, 61(1), 263–289.
- Schilmiller, A.L., and Howe, G.A. (2005). Systemic signaling in the wound response. *Current Opinion in Plant Biology*, 8(4), 369–377.
- Sibout, R., Le Bris, P., Legée, F., Cézard, L., Renault, H., and Lapierre, C. (2016). Structural Redesigning Arabidopsis Lignins into Alkali-Soluble Lignins through the Expression of p-Coumaroyl-CoA:Monolignol Transferase PMT. *Plant Physiology*, 170(3), 1358–66.
- Simmons, B. A., Loqué, D., and Ralph, J. (2010). Advances in modifying lignin for enhanced biofuel production. *Current Opinion in Plant Biology*, 13, 313–20.
- Singh, D.G., Lomako, J., Lomako, W.M., Whelan, W.J., Meyer, H.E., Serwe, M., and Metzger, J.W. (1995). β -Glucosylarginine: a new glucose-protein bond in a self-glucosylating protein from sweet corn. *FEBS Letters*, 376, 61-64.
- Smith, M.M., and Hartley, R.D. (1983). Occurrence and Nature of ferulic acid substitution of cell-wall polysaccharides in graminaceous plants. *Carbohydrate Research*, 118, 65–80.
- Somerville, C. (2006). Cellulose synthesis in higher plants. *Annual Review of Cell and Developmental Biology*, 22, 53-78.
- Steffens, J.C. (2000). Acyltransferases in Protease's Clothing. *The Plant Cell*, 12, 1253–1256.

- Stehle, F., Brandt, W., Stubbs, M.T., Milkowski, C., and Strack, D. (2009). Sinapoyltransferases in the light of molecular evolution. *Phytochemistry*, 70(15–16), 1652–1662.
- Strack, D. (1982). Development of 1-O-sinapoyl-Beta-D-glucose: L-malate sinapoyltransferase activity in cotyledons of red radish (*Raphanus sativus* L. var. *sativus*). *Planta*, 155(1), 31–36.
- Strack, D. and Sharma, V. (1985). Vacuolar localization of the enzymatic synthesis of hydroxycinnamic acid esters of malic acid in protoplasts from *Raphanus sativus* leaves. *Physiologia Plantarum*, 65, 45-50.
- Strack, D. (2001). Enzymes involved in hydroxycinnamate metabolism. *Methods in Enzymology*, 335(3), 70–81.
- Takai, K., Sawasaki, T., and Endo, Y. (2010). Practical cell-free protein synthesis system using purified wheat embryos. *Nature Protocols*, 5(2), 227–238.
- Tameshige, T., Hirakawa, Y., Torii, K.U., and Uchida, N. (2015) Cell walls as a stage for intercellular communication regulating shoot meristem development. *Frontiers in Plant Science*, 6, 324.
- Thines, B., Katsir, L., Melotto, M., Niu, Y., Mandaokar, A., Liu, G., Nomura, K., He, S.Y., Howe, G.A., and Browse, J. (2007). JAZ repressor proteins are targets of the SCFCOI1 complex during jasmonate signalling. *Nature*, 448(7154), 661–665.
- Thomma, B.P.H.J., Eggermont, K., Tierens, K.F.M.-J., and Broekaert, W.F. (1999). Requirement of functional Ethylene-Insensitive 2 gene for efficient resistance of Arabidopsis to infection by *Botrytis cinerea*. *Plant Physiology*, 121, 1093–1101.
- Vanholme, R., Demedts, B., Morreel, K., Ralph, J., and Boerjan, W. (2010). Lignin biosynthesis and structure. *Plant Physiology*, 153(3), 895–905.
- Vanholme, R., Morreel, K., Ralph, J., and Boerjan, W. (2008). Lignin engineering. *Current Opinion in Plant Biology*, 11(3), 278–85.
- Vijayan, P., Shockey, J., Levesque, C.A., Cook, R.J., and Browse, J. (1998). A role for jasmonate in pathogen defense of Arabidopsis. *Proc. Natl. Acad. Sci. USA*, 95, 7209–7214.
- Vismeh, R., Lu, F., Chundawat, S.P.S., Humpala, J.F., Azarpira, A., Balan, V., Dale, B.E., Ralph, J., and Jones, A.D. (2013). Profiling of diferulates (plant cell wall cross-linkers) using ultrahigh-performance liquid chromatography-tandem mass spectrometry. *The Analyst*, 138, 6683–92.
- Vogel, J. (2008). Unique aspects of the grass cell wall. *Current Opinion in Plant Biology*, 11(3), 301–7.

- Vogt, T. (2010). Phenylpropanoid biosynthesis. *Molecular Plant*, 3, 2–20.
- Von Röpenack, E., Parr, A., and Schulze-Lefert, P. (1998). Structural analyses and dynamics of soluble and cell wall-bound phenolics in a broad spectrum resistance to the powdery mildew fungus in barley. *Journal of Biological Chemistry*, 273(15), 9013–9022.
- Whetten, R., and Sederoff, R. (1995). Lignin Biosynthesis. *The Plant Cell*, 7, 1001–1013.
- Wilkerson, C.G., Mansfield, S.D., Lu, F., Withers, S., Park, J.Y., Karlen, S.D., Gonzales-Vigil, E., Padmakshan, D., Unda, F., Rencoret, J., and Ralph, J. (2014). Monolignol Ferulate Transferase Introduces Chemically Labile Linkages into the Lignin Backbone. *Science*, 344, 90-93.
- Withers, S., Lu, F., Kim, H., Zhu, Y., Ralph, J., and Wilkerson, C.G. (2012). Identification of grass-specific enzyme that acylates monolignols with p-coumarate. *The Journal of Biological Chemistry*, 287(11), 8347–55.
- Wood, P.J. (2007). Cereal β -glucans in diet and health. *Journal of Cereal Science*, 46(3), 230-238.
- Xu, Y., Chang, P.F., Liu, D., Narasimhan, M.L., Raghothama, K.G., Hasegawa, P.M., and Bressan, R.A. (1994). Plant defense genes are synergically induced by ethylene and methyl jasmonate. *The Plant Cell*, 6, 1077–1085.
- Zhang, L., Zhang, F., Melotto, M., Yao, J., and He, S.Y. (2017). Jasmonate signaling and manipulation by pathogens and insects. *Journal of Experimental Botany*, 68(6), 1371–1385.