Characterization of Two *Physcomitrella patens* Caffeoyl Coenzyme A *O*-Methyltransferase-Like Proteins

By

Sergei Pilipetskii

A Thesis submitted to the Graduate School-New Brunswick

Rutgers, The State University of New Jersey

And The Graduate School of Biomedical Sciences

University of Medicine and Dentistry of New Jersey

in partial fulfillment of the requirements for the degree of Master of Science

Graduate Program in Microbiology and Molecular Genetics

Written under the direction of

Dr. Faith Belanger

and approved by

_____

_____

_____

New Brunswick, New Jersey

January 2013

ABSTRACT OF THE THESIS

Characterization of Two Characterization of Two *Physcomitrella patens* Caffeoyl Coenzyme A *O*-Methyltransferase-Like Proteins

By SERGEI PILIPETSKII

Thesis Director:

Faith Belanger

Lignin is one of the most abundant biopolymers. It accounts for up to 25% of plant biomass in the biosphere. Lignified cell walls represent an important evolutionary step in plant adaptation to land life. This adaptation provided structural support, nutrient transport, protection from UV radiation, desiccation stress and protection from herbivores and pathogens. Due to lignin's long half-life, it has become a major carbon sink, causing depletion of the CO_2 levels and an increase in O_2 levels. Currently lignin research is important for the forage and biofuel industries. It has been seen that a reduction in lignin content leads to higher digestibility and higher release of cellulose from plants.

Caffeoyl coenzyme A *O*-methyltransferases (CCoAOMTS) are a class of enzymes that transfer a methyl group from S-adenosylmethionine to an alcohol group. These enzymes are necessary for lignin production in plants. Previous work has shown that knocking down the function of this enzyme in higher plants caused a decrease in lignin content.

Physcomitrella patens is a moss, which is considered not to synthesize lignin but it contains genes for the full lignin biosynthesis pathway. Currently there is some

evidence of lignin-like compounds occurring in moss, but evidence of a truly lignified cell wall does not exist. This study looks at two specific enzymes from *Physcomitrella patens,* Pp-OMT1 and Pp-OMT2. These enzymes resemble CCoAOMTs. Both enzymes were tested for substrate specificity and compared to a true caffeoyl coenzyme A *O*-methyltransferase from *Arabidopsis thalina,* AtCCoAOMT1. Furthermore kinetics of Pp-OMT2 and AtCCoAOMT1 were compared. Our results showed that Pp-OMT2 is likely to be a true CCoAOMT while Pp-OMT1 is not.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBRIVIATIONS

AtCCoAOMT1- *Arabidopsis thaliana* caffeoyl coenzyme A *O*-methyltransferase, gene bank accession number: At4g34050. Nomenclature as per Raes et al., 2003.

CCoAOMT- Caffeoyl coenzyme A *O*-methyltransferase

COMT-Caffeic acid-O-methyl transferase

C4H- Cinnamate 4-hydroxylase

F5H-ferulate 5-hydroxylase

G-type lignin- Lignin with guaiacyl linkages

HAL-histidine ammonia-lyase

H-type lignin-lignin with hydroxyphenyl linkages

M.y.a- million years ago

OMTs-*O*-methyltransferases

PAL - phenylalanine ammonia lyase

Pp-OMT1- *Physcomitrella patens O*-methyltransferase 1, gene bank accession number Xm_001754959

Pp-OMT2- *Physcomitrella patens O*-methyltransferase 2, gene bank accession number Xm_001776089

SAM-S-adenosylmethionine

S-type lignin- Lignin with syringyl linkages

INTRODUCTION

The term lignin defines a broad range of aromatic polymers found in plants. This compound is the second most abundant organic compound, first being cellulose, and it accounts for about 25% of terrestrial plant biomass (Humphreys et al., 2002). Lignin forms an important part of the cell wall of plants, having roles in mechanical support, water transport and stress responses. The evolution of lignin is linked to the adaptation of plants to life on land (Humphreys et al., 2002). This innovation of lignified cell walls is thought to have occurred 450 million years ago during the Late Ordovician period (Martone et al., 2009). However, plants began to colonize land much earlier, about 475 m.y.a. (Martone et al., 2009). The first plants to colonize land were the bryophytes, non-vascular plants such as liverworts and moss (Lowry et al.,1980). These early plants that were able to colonize land encountered new challenges to their survival. Without water shielding them, they encountered stressors such as UV-B radiation. Aside from protection from radiation, life in water also provided structural support, and without it early land plants were exposed to desiccation stress. Eventually these early land plants would have to endure land herbivores and pathogens (Lowry et al., 1980). These evolutionary pressures led to the development of secondary metabolism in land plants.

One of the early adaptations was the synthesis of phenylpropanoid compounds (Weng et al., 2010). These compounds are able to absorb the damaging UV-B radiation and offer protection to the plants and their haploid spores, which were extremely vulnerable (Lowry et al., 1980). However these early land plants still lacked structural

support, and thus remained small in size (Batemann et al., 1998). The formation of phenylpropanoid polymers (lignin) and its deposition into the cell walls are thought to have come about with the rise of tracheophytes sometime during the Paleozoic era (450 m.y.a.) (Zimmer et al., 2007). Lignin polymers are able to strengthen secondary cell walls in xylem tissues creating crosslinks with cell wall compounds. This formation of lignin was necessary for structural support for upward growth by preventing the collapse of conductive vessels (Weng et al., 2010). Thus this adaptation not only gave rigidity, but also the ability to transport water at longer distances and allowed the tracheophytes to grow much larger in size then the bryophytes. Lignin polymers also offered protection against pathogens and herbivores due to its resistance to degradation (Boerjan et al., 2003).

The importance of lignin has resulted in phenylpropanoid metabolism becoming a huge carbon sink that could represent up to 30% of total biomass in the biosphere (Weng et al., 2010). Due to this adaptation and other factors tracheophytes were able to grow in abundance and diversity. This had far-reaching consequences for all life. For example, due to lignin's long degradation time, it became a major carbon sink, lowering the levels of CO_2 and increasing O_2 levels (Berner, 1993). Furthermore, the fossilized remains of lignin rich biomass have became coal and oil, the fossil fuels responsible for human industrialization (Hayatsu et al., 1979).

Currently lignin has became a focus of study due to its relevance to the forage and biofuel industries. Digestibility of the cell wall is directly correlated to its lignin content, as well as processing efficiency for biofuels. Genetically engineered alfalfa that expressed antisense constructs that down-regulated lignin biosynthesis had a higher

carbohydrate release level (Chen et al., 2007). It is interesting to note that recent findings showed that changing the composition of lignin (rather then total lignin content) could have positive effects on its digestibility and processing. Studies using a mutant *Arabidopsis* that had a high level of S-lignin showed that it had higher digestibility than wild type, which was rich in G-lignin (Li et al., 2010).

The formation of lignin is a complex multistep process involving over ten different enzymes (Fig. 1).

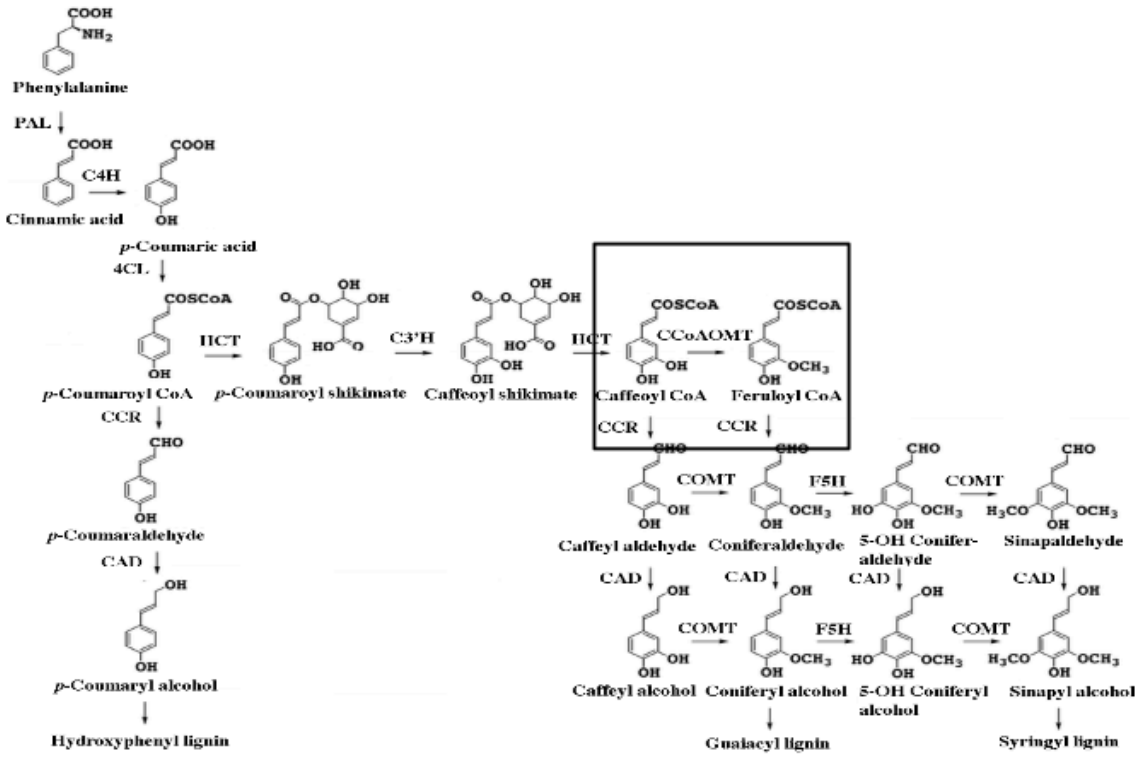


Fig.1. The monolignol biosynthesis pathway (modified from Umezawa et al., 2010). The specific step that caffeoyl CoA OMT (CCoAOMT) enzyme is responsible for is highlighted in the box.

The lignin polymer is made from the hydroxycinnamyl alcohols *p*-coumaryl alcohol, sinapyl alcohol and coniferyl alcohol (Boerjan et al., 2003). The lignin polymers produced from these alcohols are referred to based on their linkages (hydroxyphenyl or H type lignin, guaiacyl or G type lignin, and syringyl or S type). H and G type lignins are found throughout all tracheophytes, while S type is found mostly in angiosperms (Weng et al., 2010). Lignin biosynthesis occurs through two stages; first there is monolignol biosynthesis which is followed by cross-linking of monolignols to form the lignin polymer (Weng et al., 2010). The mechanism of monolignol formation is highly conserved in vascular plants. The pathway begins with the amino acid phenylalanine. Phenylalanine is deaminated by phenyalanine ammonia lyase (PAL) to produce cinnamic acid. Cinnamate 4-hydroxylase (C4H) converts cinnamic acid into *p*-coumaric acid. Following this step, the biosynthesis pathway becomes branched leading to the different lignin subunits (Xu et al., 2009). Some of the enzymes involved in this pathway are multifunctional and are able to act on multiple substrates. For example, the ferulate 5-hydroxylase (F5H) enzyme that controls the ratio of guaiacyl (G) and syringyl (S) monomers, acts on ferulic acid, coniferaldehyde, coniferyl alcohol and other substrates. Caffeic acid-O-methyltransferase (COMT) enzymes are even broader in their substrate preference. *In vitro* they are able to act upon caffeic acid, caffeoyl alcohol, 5-hydroxy-coniferaldehyde, 5-hydroxy-feruloyl CoA and 5-hydroxy coniferyl alcohol (Humphreys et al., 2002). However, the main *in vivo* substrates of COMT enzymes are considered to be caffeoylaldehyde and 5-hydroxy-coniferaldehyde.

The composition of the subunits differs among plant species and could also differ among cell types. It is thought that the pathway evolved through the process of gene duplication. One of the pieces of evidence for this is the enzyme phenylalanine ammonia-lyase (PAL). This enzyme stands at the entry point for the phenylpropanoid pathway. PAL is structurally and functionally similar to an enzyme that is involved in primary metabolism, histidine ammonia-lyase (HAL) (Beadeker et al., 2002). Plants in which the duplication event occurred might have utilized a mutated HAL enzyme to produce phenylpropanoid compounds. The accumulation of phenylpropanoids that could absorb harmful UV radiation provided the basis for selection for this pathway. These compounds would provide a distinct advantage to land-colonizing plants. It has been also suggested that the PAL enzyme could have arisen through horizontal gene transfer with symbiotic fungi or bacteria (Emiliani et al., 2009). The origin of this pathway is still unclear. One thing is certain; whether it came about from duplication events or horizontal gene transfer, the lignin biosynthesis pathway offered a huge advantage to the early land plants. From physical support, to UV absorption and protection, phenylpropanoid molecules were an essential part of land plants adaptations.

Caffeoyl coenzyme A *O*-methyltransferases (CCoAOMTs) are enzymes involved in biosynthesis of lignin. These enzymes are a part of a larger family of enzymes called *O*-methyltransferases (OMTs) (Zhong et al., 2000). The transfer of a methyl group from S-adenosylmethionine (SAM) to a hydroxyl group characterizes their action. This family of enzymes is divided into two classes (Class I and II). Class II OMTs range between 38-40 kDa in size and methylate lignin precursors such as caffeoyl aldehyde and 5-hydroxyconiferaldehyde (Zhong et al., 2000). Class I OMTs range between 26-29 kDa in

size and are represented mostly by caffeoyl coenzyme A OMTs. These enzymes are responsible for the methylation of caffeoyl CoA into feruloyl CoA, and require a Mg++ cation for activity (Fellenberg et al., 2012). It was shown that CCoAOMT is expressed in cells producing lignin. This enzyme catalyzes an essential step for formation of G and S type lignin. In studies in which expression of CCoAOMT was repressed, the lignin content was significantly decreased as well, signifying the essential role of this enzyme. Specifically it was seen that the content of G and S type lignin was decreased (Zhong et al., 2000). This is in contrast to COMT, whose repression leads to only loss of S type lignin (Zhong et al., 2000).

Lignin content is closely correlated to the digestibility of the plant cell wall, and thus its study is important to forage and biofuel industries (Chen et al., 2007). It was shown that decreasing lignin content resulted in biomass that was more readily saccharified, allowing cellulose to be more readily accessible. Furthermore, reducing lignin biosynthesis can lead to higher allocation of resources into production of cellulose and hemicellulose (Zhong et al., 2000). In *Arabidopsis* anthers it was seen that COMT1 and CCoAOMT1 have a functional and spatial separation. CCoAOMT1 was found to be in the tapetum, acting upon caffeoyl CoA, whereas COMT1 was found in epidermal tissues of flower organs and was not detected in the tapetum (Fellenberg et al., 2012).

The true caffeoyl CoA OMT enzymes are highly conserved among plants. However, in addition to the true caffeoyl CoA OMTs there are also caffeoyl CoA OMT-like enzymes. These enzymes share some similarity to true caffeoyl CoA OMTs but are not likely involved in lignin biosynthesis. In *Vanilla planifolia* it was shown that two caffeoyl CoA OMT-like enzymes, Vp-OMT4 and Vp-OMT5, had only low activity with

caffeoyl CoA and lacked similarity in the residues that bind caffeoyl CoA in the true caffeoyl CoA OMTs (Widiez et al., 2011).

Although it is generally thought that lignin only occurs in vascular plants, it has been reported that lignin or lignin-like compounds are found in nonvascular plants such as some mosses and red algae. Studies on a moss, *Sphagnum cuspidatum,* showed that lignin-like epitopes were present in the cell wall, suggesting that the nonvascular plants could have the capability to accumulate and move phenolic compounds into their cell wall (Ligrone et al., 2008). Even more interesting is the finding of lignin in the red alga *Calliarthron cheilosporiodes*. It was shown that this alga possessed H, G, and S lignin units (Martone et al., 2009). The evolutionary split between red algae and vascular plants is thought to have occurred 1.3 billion years ago, prior to the emergence of the terrestrial plants. In the case of red algae, convergent evolution is a likely explanation. Thus far, the origin of evolution of lignin biosynthesis is unknown. The moss *Physcomitrella patens* has the complete pathway for biosynthesis of lignin except for the ferulate 5-hydroxylase (F5H) gene (Xu et al., 2009). With this pathway it would be possible to produce G and H type lignin subunits. The lycophyte *Sellaginella moellendorffii* that does synthesize S-lignin also does not have an F5H gene similar to that of higher plants. The F5H action is carried out by another distinct enzyme (Weng et al., 2010). Currently the existence of lignin in moss is unclear, and the function of this pathway is unknown.

The emergence of land plants is a tremendously important event in earth's history. Bryophytes (non-vascular land plants) are the remnants of the early land plants that first began to appear abut 450 million years ago (Batemann et al., 1998). This makes them ideal for studying of evolutionary changes that took place in the emergence and conquest

of land by the vascular plants. The moss *Physcomitrella patens* was first described in 1849 by Bruch & Schimper (Schaefer et al., 2001). It had its genome sequenced in 2007, first amongst the bryophytes (Shapiro et al., 2008). The life cycle of *Physcomitrella* can be divided into the diploid sporophyte (produces spores) and haploid gametophyte (produces gametes) stages. *Physcomitrella* spores develop into filamentous growth called protonema (Schaefer et al., 2001). Protonema can differentiate into bud growths that eventually develop into gametophores (a more complex leaf-like structure that is 0.5-5 mm in length). Female (archegonia) and male (antheridia) sex organs develop from gametophores. Both of the sex organs can be found on the same plant. *Physcomitrella* sperm are motile, and if water is present, they can travel to fertilize the egg within the archegonium. This produces the diploid sporophyte, which can contain thousands of haploid spores (produced through meiosis) (Schaefer et al., 2001).

The *Physcomitrella* genome contains all but one of the lignin biosynthetic genes, suggesting that the lignin metabolic pathway developed sooner then the rise of tracheophytes (Xu et al., 2009). Currently it is unclear whether mosses do contain lignin. However, they have the necessary machinery for lignin production. The phylogenetic analysis of two *Vanilla* caffeoyl CoA OMT-like enzymes (Widiez et al., 2011), revealed that *Physcomitrella* has genes for two very different caffeoyl CoA OMT-like enzymes. One, designated Pp-OMT2 is at the base of the clade of true caffeoyl CoA OMTs (Fig. 2, clade 1a) and is thus likely to be the progenitor of the higher plant enzymes involved in lignin biosynthesis. The other *Physcomitrella* gene, designated Pp-OMT1 is in a clade of chloroplast localized enzymes (Fig. 2, clade 2).

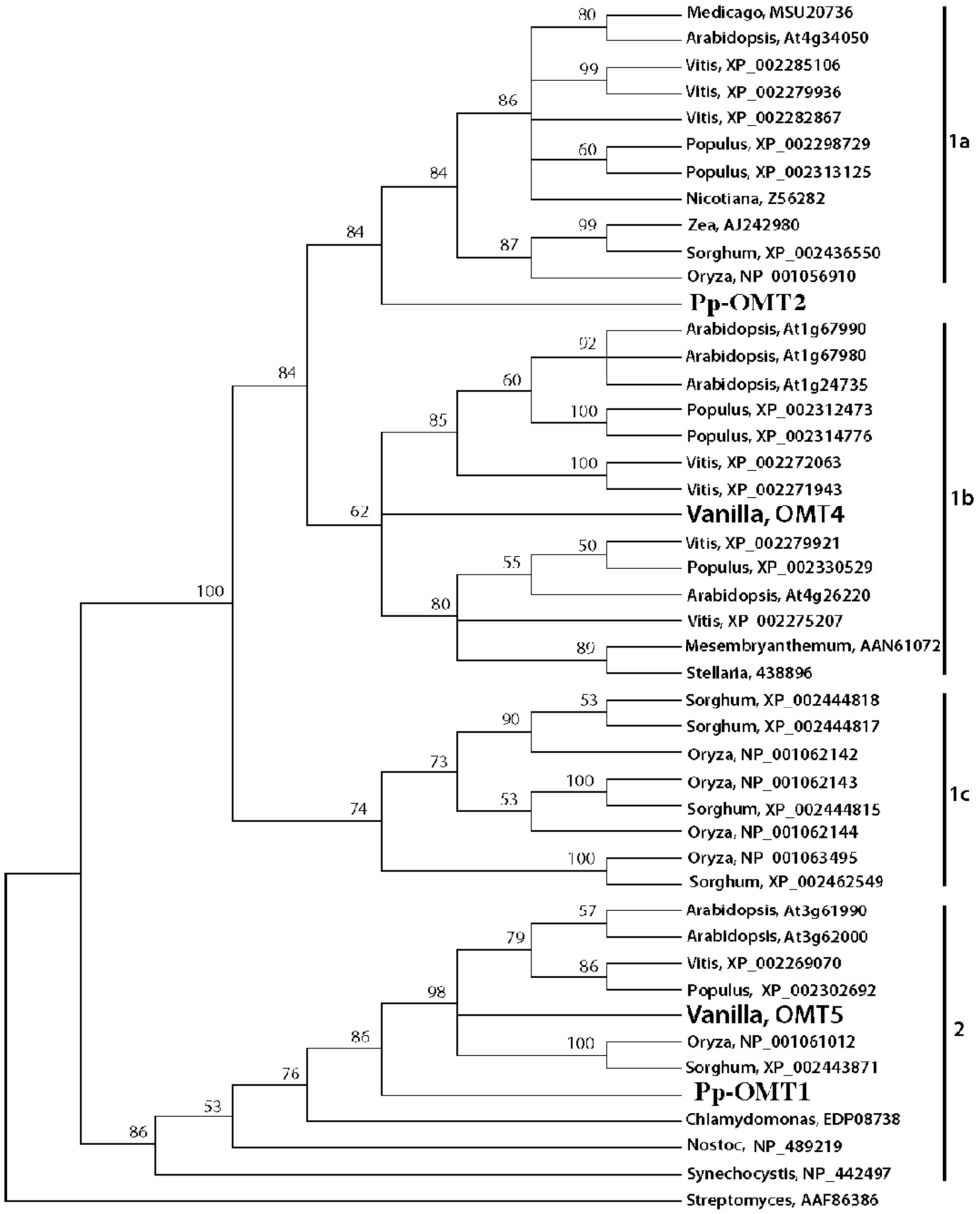


Fig. 2. Rooted 50% majority rule maximum parsimony phylogenetic tree of Class I caffeoyl CoA OMT and caffeoyl CoA OMT-like protein sequences. The *Streptomyces hygroscopicus* sequence was designed as the out-group for rooting the tree. The numbers at the nodes are bootstrap percentages based on 1,000 replications. GenBank accession numbers are given after the genus names (from Widiez et al., 2011).

Here I examined the potential of the *Physcomitrella* lignin biosynthetic machinery at one specific step, specifically I looked at the metabolic activity of the caffeoyl CoA OMT-like enzymes in *Physcomitrella patens,* as well as the *Arabidopsis* true caffeoyl CoA OMT (AtCCoAOMT1, nomenclature according to Raes et al., 2003). Although the *Physcomitrella* caffeoyl CoA OMT protein sequences are both similar to caffeoyl CoA OMTs, they are considerably different from each other (Fig. 3).

Since Pp-OMT2 is at the base of the clade containing the higher plant caffeoyl CoA OMTs it is a possibility that it has a similar activity. A comparison of the protein sequences of the Pp-OMT2 with those of *Arabidopsis* and *Medicago sativa* is shown in Fig. 4. The *M. sativa* protein has been crystalized and its substrate binding residues determined (Ferrer et al., 2005). Pp-OMT2 shares the metal, SAM and substrate binding sites with the other two enzymes. The residue identity and conservation are also largly conserved between all three enzymes. I compared the substrate preferences of the two *Physcomitrella* OMTs and compared the kinetics of the Pp-OMT2 and the *Arabidopsis* enzyme. This will help to elucidate the function of the caffeoyl CoA OMTs in *Physcomitrella*.

```
POMT1 MSNVMLGWPVLTLHKLSTNLKLKCRHVSVLGTRHVQFISYAPNCSRVLCSVDSSATKAVP
POMT2 ----------------------------MVIAGNGEIEKQPFREMGLDLEDVEEVEEVA
                                                  *   * . .: *.

POMT1 SNQLDTLVP-GGKHLPLTDQLYSYILR---HTREPLILQKLREEMAESPGSNMQIPPDQG
POMT2 TQVKGRHVETGHKTLLRSSALYKYILDTSVYPREAGELKELRQLTERHSWNMMATPPDEG
      ::  .  *  * * *  :. **.***    :.**.  *::**:   . . . *  ***:*

POMT1 QFLALLVQLMGARRCIEVGVFHGYSSLAVALVLPEGGKLVACDRDERSLAVARQYYEHAG
POMT2 QFLMLLLRLMNAKRTLEIGVYTGYSLLCTALALPPDGKVIALDISKEWYDIGAPVIQKAG
      *** **::**.*:* :*:**: *** *..**.** .**::* * .:.   :.    ::**

POMT1 VLHKVDIRHGLAADTLNDLLQNGE-AGSYDYAFLDADKMMYREYYELLLQLIKPNGLIVV
POMT2 VAHKIDFRDGPAMDSIEVLLQDDKNHDSFDFIFVDADKDNYLNYHKKLMKLVRVGGLIGY
      * **:*:*.* * *::: ***:.:  .*:*: *:****  * :*:: *::*:: .***

POMT1 DNTLWYG----RTADPLVN--DKRTKFLREFNKFLAEDDRINVSMVPIGDGMTLCRKQ-
POMT2 DNTLWNGALVAGADDPLPKYLHYYKPFILELNSFLAKDPRIQISQVPISDGVTLCRRLF
      ***** *     : *** :  .  . *: *:*.***:* **::* ***.**:****:
```

Fig. 3 Comparison of deduced amino acid sequences of *Physcomitrella* OMT1 (Pp-OMT1), and *Physcomitrella* OMT2 (Pp-OMT2). The cysteine of Pp-OMT1 (in bold) is the predicted start site of the enzyme, with the preceding sequence being the chloroplast targeting sequence. The probable initial methionine of Pp-OMT2 is in bold. An "*" indicates identical residues in all sequences, a ":" indicates strongly conserved residues (score > 0.5) and a "." Indicates weaker conserved residues (score <0.5).

```
Medicago    ------------------------------------------------------MATNEDQK
At4g34050   ---------------------------------------MATTTTEATKTSSTNGEDQK
Pp-OMT2     ----------------------------MVIAGNGEIEKQPFREMGLDLEDVEEVEEVA

                      s                                              s
Medicago    QTESGRHQEVGHKSLLQSDALYQYILETSVFPREHEAMKELREVTAKHPWNIMTTSADEG
At4g34050   QSQNLRHQEVGHKSLLQSDDLYQYILETSVYPREPESMKELREVTAKHPWNIMTTSADEG
Pp-OMT2     TQVKGRHVETGHKTLLRSSALYKYILDTSVYPREAGELKELRQLTERHSWNMMATPPDEG
                 ** *.***:**:*. **:***:***:***   :****::* :*.**:*:*..***

Medicago    QFLSMLLKLINAKNTMEIGVYTGYSLLATALAIPEDGKILAMDINKENYELGLPVIKKAG
At4g34050   QFLNMLIKLVNAKNTMEIGVYTGYSLLATALALPEDGKILAMDVNRENYELGLPIIEKAG
Pp-OMT2     QFLMLLLRLMNAKRTLEIGVYTGYSLLCTALALPPDGKVIALDISKEWYDIGAPVIQKAG
            *** :*::*:***.*:***********.****:* ***::*:*:.:* *::* *:*:***

Medicago    VDHKIDFREGPALPVLDEMIKDEKNHGSYDFIFVDADKDNYLNYHKRLIDLVKVGGVIGY
At4g34050   VAHKIDFREGPALPVLDEIVADEKNHGTYDFIFVDADKDNYINYHKRLIDLVKIGGVIGY
Pp-OMT2     VAHKIDFRDGPAMDSIEVLLQDDKNHDSFDFIFVDADKDNYLNYHKKLMKLVRVGGLIGY
            * ******:***:  :: :: *:***.::************:****:*:.**::**:***

                s            s s   s
Medicago    DNTLWNGSVVAPPDAPLRKYVRYYRDFVLELNKALAVDPRIEICMLPVGDGITICRRIK
At4g34050   DNTLWNGSVVAPPDAPMRKYVRYYRDFVLELNKALAADPRIEICMLPVGDGITICRRIS
Pp-OMT2     DNTLWNGALVAGADDPLPKYLHYYKPFILELNSFLAKDPRIQISQVPISDGVTLCRRLF
            *******::** .* *: **::**: *:****. ** ****:*. :*:.**:*:***:
```

Fig. 4. Comparison of deduced amino acid sequences of *Physcomitrella* OMT2 (Pp-OMT2), *Arabidopsis* AtCCoAOMT1 (At4g34050) and *Medicago sativa* caffeoyl CoA OMT. The second methionine of Pp-OMT2 (in bold) is the likely start site of the enzyme. The binding sites for the substrate are indicated by an "s" for the *M. sativa* sequence. The metal and SAM binding sites are highlighted by bold fonts. An "*" indicates identical residues in all sequences, a ":" indicates strongly conserved residues (score > 0.5) and a "." indicates weaker conserved residues (score <0.5).

METHODS

Compounds used as possible substrates were obtained from Sigma-Aldrich (USA), Indofine Chemical Company Inc. (USA), or were a gift from Dr. R. Dixon (Noble Foundation, USA).

Accession Numbers

Sequence data for AtCCoAOMT1, Pp-OMT1 and Pp-OMT2 can be found in the GenBank databse (http://www.ncbi.nlm.nih.gov) under the accession numbers At4g34050, Xm_001754959, and Xm_001776089, respectively.

RNA isolation and cDNA preparation

RNA was extracted from *Physcomitrella patens* protonema using Tri-Reagent (Sigma-Adrich). Two g of tissue was ground to a fine powder in liquid nitrogen. Twenty ml Tri-Reagent was added and ground in until the liquid was frozen. The thawed sample was then centrifuged at 12,000 rpm for 10 min to remove the debrits. Four mls of chloroform (1/5 volume) was then added to the supernatant and incubated for 15 min. After centrifuging for 10 min, the aqueous phase was removed and chloroform extraction was repeated to remove any residual phenol. The upper aqueous layer was incubated overnight in the freezer with 10 ml isopropanol to precipitate the RNA. The sample was then pelleted by centrifuging for 15 min. The supernatant was discarded, and the RNA pellet was washed with ethanol. The sample was again centrifuged for 10 min and the

supernatant was discarded. The RNA pellet was dissolved in 500 µl of water and the concentration determined. The sample was then purified using the RNeasy cleanup method (Qiagen). Ten µl buffer RDD and 2.5 µl DNase I were added to the sample, which was then incubated at room temperature for 15 min. Following incubation 350 µl of RLT buffer was added. The sample was treated with 250 µl EtOH and transferred to the RNeasy Mini spin column. This was centrifuged at 10,000 rpm for 30 s and the flow-through was discarded. The column was washed with 80% EtOH ,centrifuged and the flow through discarded. The column was dried by centrifuging. The sample was eluted with RNase free water.

cDNA synthesis and Amplification of PpOMT1 and PpOMT2

For 4 µg of RNA, 1 µl of 10 mM dNTPs, 1 µl of RT enzyme, 2 µl of 10 µM oligo dT primer and 32 µl of RT buffer (5X) was added. Water was added to bring the reaction volume up to 11 ul. cDNA product was used in a PCR reaction. The initial denaturation was conducted at 94°C for 2 min, followed by 30 cycles of 30 s denaturation at 94°C, annealaing at 52°C, and 2 min extention at 72°C. An additional final 10 min extention at 72°C was performed. Different forward primers were used in the case of Pp-OMT1 to amplify the full-length coding sequence, and the predicted mature protein sequence after cleavage at the chloroplast targeting sequence. Pp-OMT2 primers reflect the difference in the position of the first Met codon. Forward primers contained an *Nde*I restriction enzyme site, while the reverse primers contained the *Xho*I restiction enzyme site.

The primer sequences used to amplify the genes from the cDNA are given in Table 1.

Table 1. Sequences of oligonucleotide primers used in this study

	Forward Primer 5'-3'	**Reverse Primer 5'-3'**
Expression in *E. coli* Physco OMT-1	GGC AGC CAT ATG CTC GGT TGG CCG GTG	CGG ATC CTC GAG TCA CTG CTT TCG GCA GAG TG
Physco OMT-1 Chloro P	GGC AGC CAT ATG TGT AGC GTC GAC AGC AGT	CGG ATC CTC GAG TCA CTG CTT TCG GCA GAG TG
Physco OMT-2 2nd Met	GGC AGC CAT ATG GGG CTG GAC CTG GAG	CGG ATC CTC GAG TTA AAA CAA CCT TGC GCA TAG
Physco OMT-2 1st Met Extension primer	GGC AGC CAT ATG GTG ATC GCT GGA AAC GGA GAA ATT GAG AAG CAG CCT TTT CGA GAG ATG GGG CTG GAC CTG GAG	CGG ATC CTC GAG TTA AAA CAA CCT TGC GCA TAG
Physco OMT-2 1st Met Expression in *E. coli*	5' GGC AGC CAT ATG AGT AAC GTG ATG CTC GGT TGG CCG GTG	CGG ATC CTC GAG TTA AAA CAA CCT TGC GCA TAG

AmPure (Agencourt) was used to purify the PCR products. AmPure was kept at room temperature for 20 min and then mixed. To the initial PCR products, 70 µl of AmPure was added (0.7X of sample). The solution was mixed by pipetting and left to incubate on a shaker for 10 min at room temperature. Following the incubation the mixture was placed on a magnet for about 2 min (until the mixture was clear). While on the magnet, the supernatant was removed, and 200 µl of 95% EtOH was added without disturbing the beads. The mixture was incubated on a magnet for about 30 seconds, and then the EtOH was removed. The EtOH wash was then repeated. After removing EtOH, the tube was air dried for 20 min off the magnet stand. The DNA bound to the magnetic beads was eluted by adding 40 µl of distilled H_2O and mixed by pipetting. The sample was then placed on a magnet for one minute. The supernatant was then carefully removed to a new tube without disturbing the pellet (done while still placed on the magnet). The absorbance of 2 µl of the sample was measured using NanoDrop (Thermo Scientific). Following the purification, 2 µl of purified product was used in another round of PCR. Five µl of the PCR products were analyzed on a 1% agarose gel. Bands of the correct size were seen for Pp-OMT1 Cloro P, and AtCCoAOMT1. The PCR products were purified using AmPure and their concentrations were measured through Nano Drop. The amplification of Pp-OMT2 from the cDNA was initially unsucessfull. However after perfomring two additional rounds of PCR on the previous PCR product, amplification of Pp-OMT2 was seen.

Transformation, Expression, and Protein Purification

The plasmid pET-15b (Novagen) was used as the expression vector. The plasmid was digested with *Nde*I (New England BioLabs) overnight at 37C. Following the initial digestion, the plasmid was also cut with *Xho*I restriction enzyme (Promega) overnight at 37C. The plasmid was used to transform the XL1 Blue MRF cells by electroporation. Plasmid containing cells were selected on ampicillin-containing media. The insert was verified by sequencing. A plasmid containing the correct insert was purified from a 100 ml LB+ ampicillin culture using QIA miniprep (Qiagen). The purified plasmid was used to transform BL21 Nova blue expression (Stratagene) cells. Protein expression was induced using 0.05 mM concentration of IPTG in 200 ml cultures grown to ~0.5 O.D at 37° C and left to incubate overnight at room temperature. The cultures were then pelleted and lysed using fast break lysis buffer (Promega). Proteins were purified using Ni particles (Promega) and incubation on a magnet. The proteins were concentrated using an Amicon Ultra 10 column (Ultracel). The protein concentration was assayed using a standard curve generated by known concentrations of BSA. Concentration was determined by measuring the wavelength after treating the sample with the Bio-Rad Protein Assay reagent (BioRad).

Enzyme assays were carried out as described by Widiez et al. (2011). The following substrates were used in the substrate preference assay at 5 mM concentration: 5-OH coniferaldehyde, caffeic acid, 5-OH ferulic acid, caffeoyl CoA, 5-OH feruloyl CoA, caffeoyl aldehyde, 5-OH coniferaldehyde, tricetin, myrecitin, quercetin, and 3,4-dihydroxybenaldehyde. The reactions were incubated for 30 min at 30°C with 1 µl of ^{14}C SAM. The reaction was stopped after 30 min with HCl. Ethyl acetate was added and SAM was separated from the methylated products through centrifugation. Radiation

counts were measured for 180 µl of the top layer. With the substrates caffeoyl CoA and 5-OH feruloyl CoA, after the 30° C incubation 6 µl of 5N NaOH was added and the reactions incubated at 40° C for 15 min to hydrolyze the CoA from the methylated products. The reactions were neutralized by adding 44 µl of 1N HCl and then extracted with ethyl acetate as above.

RESULTS

Cloning:

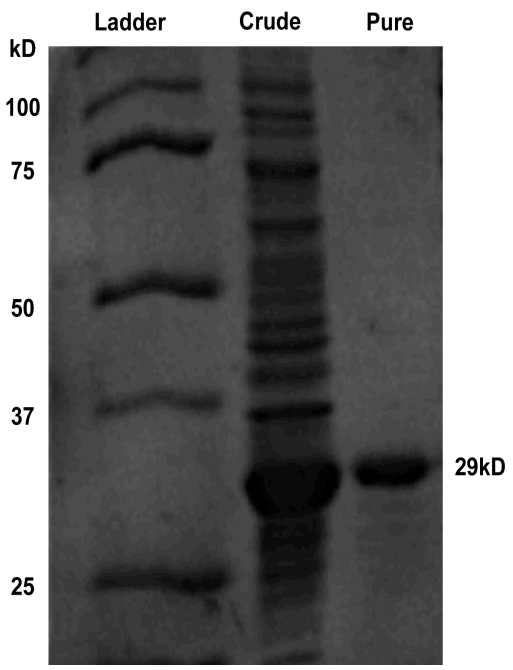
Two caffeoyl CoA OMT-like genes have been proposed to exist in the *Physcomitrella* genome based on the genome sequence (Xu et al., 2009). Based on their position in the phylogenetic tree (Fig. 2), it seemed likely they may have different activities. To compare their activities, I generated expression clones for both and assayed the purified recombinant proteins. Furthermore I looked at two versions of the Pp-OMT2 gene from *Physcomitrella*. During this study the *Physcomitrella* genome was re-annotated in respect to the caffeoyl CoA OMT2 gene (PHYSCO116394), adding 48 bases to the initial start site. I created two constructs, one based on the initial annotation, and one with the 48 base addition to the beginning and tested them both for substrate preference.

To create expression clones for the two genes, cDNA was prepared from the *Physcomitrella* RNA. An *Arabidopsis* cDNA clone from ABRC was used for the true caffeoyl CoA OMT (AtCCoAOMT1) gene isolation. Genes were amplified from *Physcomitrella* cDNA using specific primers unique to the genes. The *Physcomitrella* OMT2 cDNA was initially isolated using a primer that started on the second methionine (refer to Fig. 4), and was elongated using a long forward primer that added the first methionine and the intervening sequence. The genes were inserted into pET15B vector and transformed into XL1Blue MRF cells and BL21(DE) expression cells. In total 13 clones were prepared. Gene sequence validity was checked by sequencing.

Protein Induction/ Purification

BL21DE cells with a pET15b construct were induced using IPTG. The recombinant proteins were purified based on the presence of a His-tag (Fig. 5).

A)



B)

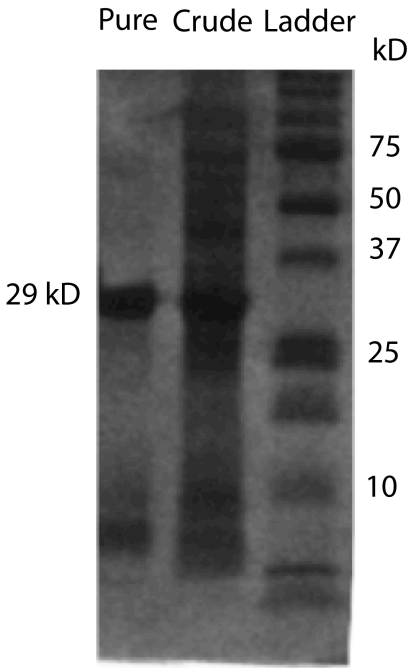


Fig. 5. A) SDS-PAGE gel of purified recombinant AtCCoAOMT1 enzyme. Lane 1 is Precision Plus Protein Dual Xtra Standard ladder. Lane 2 is the crude extract prior to Ni purification and Amicon Ultra 10 concentration. Lane 3 is the purified and concentrated AtCCoAOMT1 enzyme.

B) SDS-PAGE gel of purified recombinant Pp-OMT2 enzyme. Lane 3 is Precision Plus Protein Dual Xtra Standard ladder. Lane 2 is the crude extract prior to Ni purification and Amicon Ultra 10 concentration. Lane 1 is the purified and concentrated AtCCoAOMT1 enzyme.

Substrate assay

The following constructs were tested for substrate preference; Pp-OMT2 (first met), Pp-OMT2 (second met), Pp-OMT1(chloro p) and AtCCoAOMT1. The following substrates were tested; tricetin, myricetin, quercetin, 5-OH coniferaldehyde, 5-OH ferulic acid, caffeoyl aldehyde, 3,4 dihydroxybenzaldehyde, caffeoyl CoA, and caffeic acid. Pp-OMT 2 (first met) had no activity for any substrate. This suggests the re-annotation of the gene was incorrect.

Substrate preference results:

Substrate preferences for Pp-OMT1, Pp-OMT2 and AtCCoAOMT1 were compared (Table 2). Two micrograms of each enzyme was used and all substrates were assayed at 5 mM.

Table 2. Substrate preference data

	Pp-OMT2	AtCCoAOMT1	Pp-OMT1
Substrate	Activity	Activity	Activity
Caffeoyl CoA	26.5±1.9 (100)*	25.2 ± 2.4 (68)	1.3±0.7 (2)
Caffeoylaldehyde	22.7±1.3 (86)	33.7 ± 1.3 (91)	3.7±0.7 (6)
5-OH coniferaldehyde	19.7±2.5 (74)	27.5 ± 0.8 (74)	27.7±2.5 (44)
Caffeic acid	14.6±0.3 (55)	8.9 ± 0.7 (24)	2.8± (0.8)
5-OH ferulic acid	10.9±0.9 (41)	22.1 ± 0.3 (59)	44.8±5.9 (71)
Quercetin	8.9±0.8 (33)	34.7±1.0 (94)	1.8±0.1 (3)
Tricetin	6.2±0.5 (23)	36.9±2.7 (100)	63.2±8.1 (100)
Myricetin	4.6±1.7 (17)	35.6±6.1 (96)	13.2±3.8 (21)
3,4 dihydroxy benzaldehyde	9.2±0.4 (34)	No data	5.3±0.6 (8.4)

* Activities presented are pkat mg^{-1} ± SE (percent relative activity)

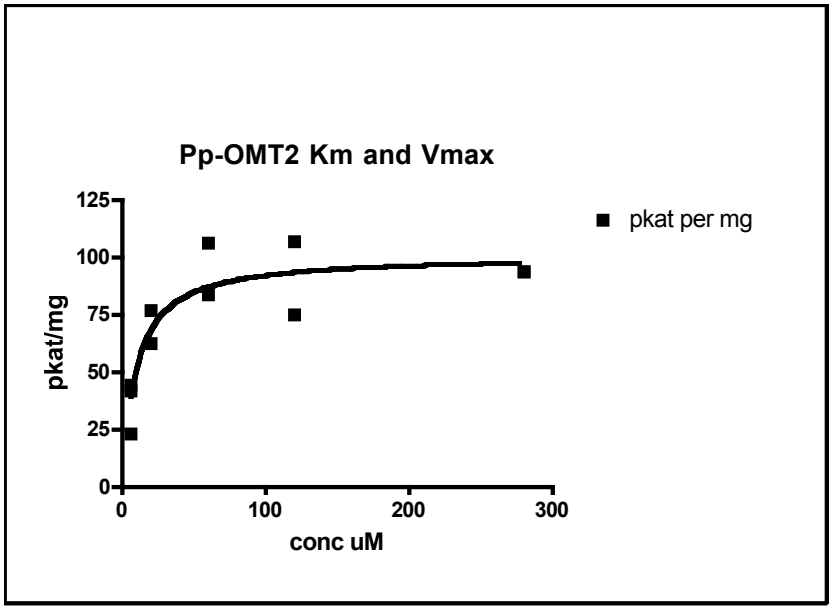
Physcomitrella OMT1 Chloro P had the highest activity with tricetin and the lowest with caffeoyl CoA. *Physcomitrella* OMT2 had the highest activity with caffeoyl CoA. These results confirmed the expectation based on the sequence comparison that the two enzymes were quite different. The *Arabidopsis* caffeoyl CoA OMT had the highest preference for tricetin and had the lowest preference for caffeic acid. AtCCoAOMT1 was confirmed as being involved in lignin biosynthesis by analysis of a T-DNA insertion

mutant (Do et al., 2007). However, its kinetic parameters have not been previously reported. This characterization of AtCCoAOMT1 was the first at the time of this writing.

Km/Vmax Assay

Pp-OMT2 (2^{nd} met) and AtCCoAOMT1 purified proteins were analyzed for the Km and Vmax values with caffeoyl CoA as the substrate. The kcat/Km values were used to extrapolate the efficiencies of both of the enzymes (Fig. 6, Table 3).

A.



B.

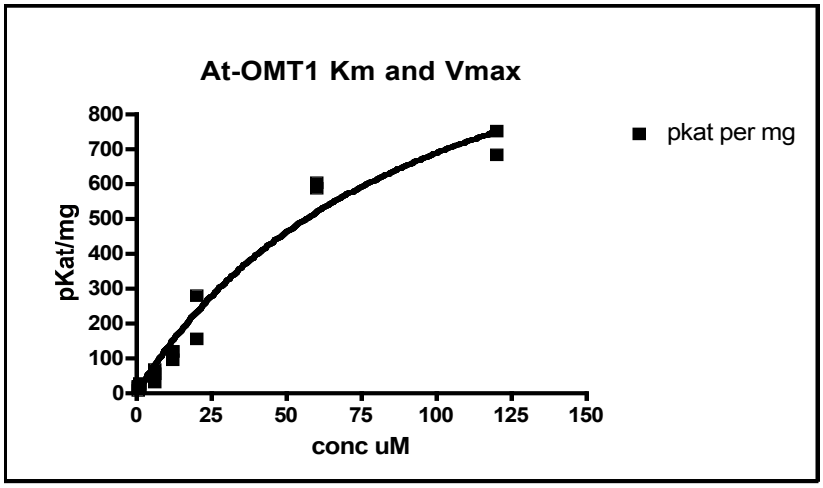


Fig. 6. A. Activity graph of Pp-OMT2 enzyme vs. concentration of caffeoyl CoA substrate. B. Activity graph of AtCCoAOMT1 enzyme vs. concentration of caffeoyl CoA substrate

Table 3. Kinetic parameters of Pp-OMT2 and AtCCoAOMT1 with the substrate caffeoyl CoA

Enzyme	V_{max}, pkatal mg^{-1}	K_M, uM	k_{cat}/K_M, M^{-1} s^{-1}
Pp-OMT2	100.7	9.512	2.25E+02
AtCCoAOMT1	1354	96.8	2.17E+02

DISCUSSION

Here I report the functional characterization of two caffeoyl CoA OMT-like enzymes from the moss *Physcomitrella patens*. Their substrate preferences and kinetic parameters are compared with the *Arabidopsis* OMT1 enzyme. The results suggest that the *Physcomitrella* Pp-OMT1 enzyme does not share a similar function with the *Arabidopsis* AtCCoAOMT1 enzyme. The evidence for this can be seen in the phylogenetic analysis (Fig. 2) in which the two enzymes are distantly related in the tree. Furthermore their substrate preferences are not similar. Specifically the high preference of *Arabidopsis* AtCCoAOMT1 for caffeoyl CoA (the known *in vivo* substrate of this enzyme) (Fellenberg et al., 2012) is in contrast to the extremely low preference of the *Physcomitrella* OMT1 for this substrate. Due to the results of the activity of the Pp-OMT1 protein with the chloroplast transit peptide cleaved off, it is likely that this protein is localized in the chloroplast. Pp-OMT1 preferred the flavone tricetin as a substrate, as did the *Vanilla planifolia* enzyme Vp-OMT5, which was confirmed as being chloroplast localized (Widiez et al.,2011).

Substrate preference analysis of *Physcomitrella* Pp-OMT2 is comparable to the substrate preference of AtCCoAOMT1, in respect to having a high preference for caffeoyl CoA and 5-OH coniferaldehyde. However they differed widely in preference for the flavonoids tricetin, myricitin, and quercetin, with AtCCoAOMT1 having high preference and Pp-OMT2 having a low preference. Comparing the Vmax values it is seen that the AtCCoAOMT1 has a much higher Vmax (1354 pkat/mg vs 100.7 pkat/mg). This implies that AtCCoAOMT1 can convert much more of the substrate (caffeoyl CoA) when it is fully saturated with the substrate. The Km value for the Pp-OMT2 is low when

compared to the AtCCoAOMT1 enzyme. This suggests that Pp-OMT2 needs much lower concentrations of the substrate to be active, while AtCCoAOMT1 needs a higher concentration. This might reflect an environment that is encountered in *Arabidopsis*. Here the substrate caffeoyl CoA might be found in abundance, thus the enzyme will be optimized for high substrate concentration conditions. Interestingly, although the Vmax and the Km of the two enzymes differ, the overall catalytic efficiency is almost identical as indicated by the ratio of kcat/Km.

This study looks at two moss enzymes that resemble the true caffeoyl CoA OMT enzymes that are seen in higher plants. *Physcomitrella* Pp-OMT1, is less likely to share function with the true caffeoyl CoA OMTs. This can be seen through different substrate preferences and evolutionary distance within the caffeoyl CoA OMT phylogenetic tree (Fig. 2). *Physcomitrella* Pp-OMT2 seems to resemble the AtCCoAOMT1 in terms of function. It has similar substrate preference, specifically the preference for caffeoyl CoA. The mosses are the predecessors of all land plants and are thought to represent early plants that colonized land. As of yet they are thought to lack the capability to produce lignin structures. Evolutionarily it was essential for land plants to develop the capability to produce this polymer. The results found here will help elucidate the origin of this capability. It is seen that *Physcomitrella* has the necessary machinery to produce this polymer. Here we can see that *Physcomitrella* Pp-OMT2 is capable of a similar function as AtCCoAOMT1, known to be involved in lignin biosynthesis. Although the Vmax and Km of the two enzymes differ, the catalytic efficiency based on the kcat/Km ratios are nearly identical. What is currently puzzling is the function of this enzyme in moss. If there is no lignin, then what is it doing? With the current findings of lignin and lignin-like

compounds in other primitive plants such as algae it could be possible that there is something resembling lignin also found in moss.

Physcomitrella patens is a useful model system for future determination of the function of Pp-OMT2 because of the ease with which targeted gene knockouts are possible (Schaefer, 2002). So far it is difficult to say where to look for a phenotype change. EST, for Pp-OMT2 have been reported from the sporophyte tissue but none from gametophyte tissue. It is therefore more likely that the gene has a purpose in the sporophyte, and that is where any phenotype change in a knockout mutant maybe observable. The localization of Pp-OMT2 ESTs could indicate that if lignin is being produced, it is being made in the sporophyte, or the spore. There the lignin could protect the spore from radiation or other stress. Also it could be a structural element as the spore is harder then the vegetative tissues of the plant.

WORKS CITED

Baedeker M, Shultz G (2002). Autocatalytic peptide cyclization during chain folding of histidine ammonia-lyase Structure **10**: 61-67

Batemann R, Crane P, DiMichele W, Kendrick P, Rowe N, Speck T, Stein W (1998). Early evolution of land plants: phylogeny, physiology, and ecology of the primary terrestrial radiation. Annual Review of Ecology and Systematics **29**: 263-292

Berner, R (1993). Paleozoic atmoshperic CO_2: importance of solar radiation and planet evolution. Science **261**: 68-70

Boerjan W, Ralph J, Bauchner M (2003). Lignin Biosynthesis. Annual Review of Plant Molecular Biology **54**: 519-546

Chen F, Dixon R (2007). Lignin modification improves fermentable sugar yields for biofuel production. Nature Biotechnology **25**: 759-761

Do C, Pollet B, Thevanin J, Sibout R, DenoueD, Barriere Y, Lapierre C, Jouanin L (2007). Both caffeoyl coenzyme A 3-*O*-methyltransferase 1 and caffeic acid *O*-methyltransferase 1 are involved in redundant functions for lignin, flavonoids and sinapoyl malate biosynthesis in *Arabidopsis*. Planta **226**:1117-1129

Emilani G, Fondi M, Fani R, Gribaldo S(2009). A horizontal gene transfer at the origin of phenypropanoid metabolism: a key adaptation of plants on land. Biology Direct. **4**(7)

Fellenberg C, Ohlen M, Handric V, Vogt T (2012). The role of CCoAOMT1 and COMT1 in *Arabidopsis* anthers. Planta **236**: 51-61

Ferrer J-L, Zubieta C, Dixon R, Noel J (2005). Crystal structures of alfalfa caffeoyl coenzyme A 3-*O*-methyltransferase. Plant Physiol **137**:1009-1016

Hayatsu R, Winans R, McBeth R, Scott R, Moore L, Studier M (1979). Lignin-like polymers in coals. Nature **279**:41-43

Humphreys J, Chapelle C (2002). Rewriting the lignin roadmap. Current Opinion Plant Biology. **5**:224-229

Li X, Ximenses E, Kim Y, Slinger M, Meilan R, Ladish M, Chappel C (2010). Lignin monomer composition affects *Arabidopsis* cell-well degradability after liquid hot water pretreatment. Biotechnology for Biofuels **3** (27)

Ligrone R, Carafa A, Duckett J, Ranzaglia K, Ruel K (2008). Immunocytochemical detection of lignin-related epitopes in cell wall of bryophytes and charalean alga *Nitrella*. Plant Systematics and Evolution. **270:** 257-272

Lowry B, Herbant C, Lee D (1980). The origin of land plants- A new look at an old problem. Taxon **29**: 183-197

Martone P, Estevez J, Lu E, Ruel K, Denny M, Somerville C, Ralph J (2009). Discovery of lignin in seaweed reveals convergent evolution of cell-wall architecture. Current Biology. **19**: 169-175

Raes J, Rohde A, Christensen J, Peer Y, Boerjan W (2003). Genome-wide characterization of the lignification toolbox in *Arabidopsis*. Plant Physiology **133:**1051-1071

Schaefer D (2002). A new moss genetics: Targeted mutagenesis in *Physcomitrella patens.* Annual Review Plant Biology. **53:** 447-501

Schaefer D, Zrÿd J (2001). The moss *Physcomitrella patens,* now and then. Plant Physiology. **127**:1430-1438

Shapiro H, Nishiyama T, Perround P, Lindquist E, Kamisugi Y, Sakakibara K, Fujiata T, Oishi K, Shin-I T, Kuroki Y, Toyoda A, Suzuki Y, Hashimoto S, Yamaguchi K,Sugano S, Kohara Y, Fujiyama A, Anterola A,Aoki S, Ashton N, Barbazuk B, Barker E, Bennetzen J, Blankenship R, Cho S, Dutcher S, Estelle M, Fawcett J, Gundlach H, Hanada K, Heyl A,Hicks K, Hughes J,Lohr M, Mayer K, Melkozernov A, Murata T, Nelson D, Pils B, Prigge M, Reiss B, Renner T, Rombauts S, Rushton P, Sanderfoot A, Schween G, Shiu S , Stueber K, Theodoulou F, Tu H, Van de Peer Y, Verrier P, Waters E, Wood A, Yang L, Cove D, Cuming A, Hasebe M, Lucas S, Mishler B, Reski R, Grigoriev I, Quatrano R, Boore J (2008). The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants. Science **319**: 64-69.

Umezawa T (2010). The cinnamate/monolignol pathway. Phytochem Rev. **9**:1-17

Weng J, Chapelle C (2010). The origin and evolution of lignin. New Phytologist.**187**: 273-285

Weng J, Akiyama T, Bonawitz N, Li X, Ralp J, Chapple C (2010). Convergent evolution of syringyl lignin biosynthesis via distinct pathways in the Lycophyte *Selaginella* and flowering plants. The Plant Cell. **22:**1033-1045

Widiez T, Hartman T, Dubai N, Yan Q, Lawton M, Havkin-Frenkel D, Belanger F (2011). Functional characterization of two new members of the caffeoyl CoA O-methyltransferase-like gene family from *Vanilla planifolia* reveals a new class of plastid-localized O-methyltransferases. Plant Molecular Biology **76**:6 475-488.

Xu Z, Zhand D, Zhou H, Zhou X, Ye X, Rechel K, Stewart N, Syrenne R, Yang X, Gao P, Shi W, Doeppke C, Sykes R, Burris J, Bozell J, Cheng Z, Hayes D, Labbe N, Davis M, Steward N, Yuan J (2009). Comparative genome analysis of lignin biosynthesis gene families across the plant kingdom. Bioinformatics **10** (Suppl 11) S3

Zhong R, Morrison H, Himmelsbach D, Poole F, Ye Z (2000). Essential role of caffeoyl coenzyme A O-methyltransferase in lignin biosynthesis in woody poplar plants. Plant Physiology **124**: 563-577.

Zimmer A, Lang D, Richard S, Frank W, Reski R, Rensing S (2007). Dating the early evolution of plants: detection and molecular clock analyses of orthologs. Mol Genet Genomics **278**:393-402