Title
Biochemical and Functional Characterization of the GH3 Amino Acid-Conjugase PBS3 of Arabidopsis thaliana

Permalink
https://escholarship.org/uc/item/2q972429

Author
Okrent, Rachel Allegra

Publication Date
2010

Peer reviewed|Thesis/dissertation
Biochemical and Functional Characterization of the GH3 Amino Acid-Conjugase PBS3 of *Arabidopsis thaliana*

By

Rachel Allegra Okrent

A dissertation submitted in partial satisfaction of the requirements for the degree of

Doctor of Philosophy

in

Plant Biology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Mary Wildermuth, Chair
Professor Bob Buchanan
Professor Brian Staskawicz
Professor Barry Shane

Spring 2010
Abstract

Biochemical and Functional Characterization of the GH3 Amino Acid-Conjugase PBS3 of *Arabidopsis thaliana*

by

Rachel Allegra Okrent

Doctor of Philosophy in Plant Biology

University of California, Berkeley

Professor Mary Wildermuth, Chair

Plants have evolved sophisticated mechanisms to evade disease caused by pathogens. These resistance responses require intricate signaling networks mediated by phytohormones. Synthesis and accumulation of these molecules are regulated at several levels, including covalent modification through esterification and conjugation to amino acids and sugars. Many enzymes responsible for these modifications have been identified in the model plant *Arabidopsis thaliana*. The PBS3 gene encodes an enzyme belonging to an uncharacterized group of the GH3 family present in many plant species; previously characterized GH3 family members act as phytohormone-amino acid synthetases. The *pbs3* mutants (avrPphB Susceptible 3) exhibit enhanced susceptibility to virulent and avirulent *Pseudomonas syringae* strains, suggesting PBS3 is involved in basal disease resistance.

*PBS3* was found to be crucial for accumulation of induced salicylic acid (SA), a critical mediator of basal resistance, as well as expression of the SA-dependent pathogenesis related marker gene *PRI*. Using a novel high throughput adenylation assay, 4-substituted benzoates 4-hydroxybenzoate (4-HBA) and para-aminobenzoate (pABA) were identified as preferred acyl substrates of PBS3; no observed activity was observed with 2-substituted benzoates like SA. Notably, SA was found to specifically and reversibly inhibit PBS3 activity, perhaps revealing a general mechanism for the rapid regulation of GH3 activity. In addition, *PBS3* influences the elicited accumulation of several compounds, including a novel compound. My study implicates 4-HBA and pABA as potential signaling molecules in plant-pathogen interactions and reveals PBS3 as a mediator of cross-talk between signaling pathways involving chorismate-derived compounds.

Investigation of the evolutionary history of the GH3 family reveals that PBS3 is part of an ancient group of genes. The underexplored GH3 Group III is enriched in genes that have appeared through tandem duplication and insertion, consistent with function in response to environmental stress. However, functional analysis indicates that PBS3 is unique within the family in regards to its critical role in basal and SA-mediated defense responses.
I dedicate this dissertation to the memory of my mother, Pamela.
# TABLE OF CONTENTS

FIGURE LIST

ACKNOWLEDGMENTS

CHAPTER I
Introduction

CHAPTER II
The GH3 Acyl Adenyrase Family Member PBS3 Regulates Salicylic Acid-Dependent Defense Responses

CHAPTER III
PBS3 Conjugates Amino Acids to 4-Substituted Benzoates and is Inhibited by Salicylate

CHAPTER IV
Impact of pbs3 on Metabolism of Small Molecules in *Arabidopsis*

CHAPTER V
Evolutionary History of the GH3 Family in Rosids

CHAPTER VI
Conclusion
FIGURES AND TABLES

CHAPTER I
Figure 1. Structures of phytohormones involved in plant defense
Figure 2. Conjugation of SA as an example of hormone regulation
Figure 3. pbs3 plants are more susceptible to infection with
Pseudomonas syringae

CHAPTER II
Figure 1. The sequence of PBS3 highlighting locations of pbs3-1
and pbs3-2 mutations
Figure 2. Protein sequence alignment of Arabidopsis GH3 family
Figure 3. PBS3 transcript in pbs3-1 and pbs3-2 mutants
Figure 4. Pathogen-induced expression of PRI is delayed and reduced
in the pbs3 mutant
Figure 5. SAG accumulation is dramatically reduced in pbs3 mutants
in response to Pst avrRpt2
Figure 6. Exogenous treatment with SA restores SAG accumulation
Figure 7. Comparative expression of PBS3 and ICS1 in response to bacteria
Figure 8. Proposed biochemical functions of PBS3

Table 1. Salicylic acid levels in pbs3 mutants in pathogen infected leaves

CHAPTER III
Figure 1. Reaction catalyzed by GH3 enzymes
Figure 2. Neighbor-joining phylogenetic tree of the Arabidopsis GH3
family of proteins
Figure 3. PBS3 catalyzes the formation of amino acid conjugates
Figure 4. Predicted secondary structure of PBS3 and alignment of AtGH3 sequences in the corresponding region
Figure 5. Inhibition of PBS3 activity by SA

Table 1. Initial acyl substrate screen
Table 2. Acyl substrate preference and specificity of PBS3
Table 3. Kinetic parameters of PBS3 substrates
Table 4. Fractional inhibition of pABA-AMP formation

CHAPTER IV
Figure 1. The central position of chorismate in aromatic
acid biosynthesis in plants
Figure 2. The flavonoid and general phenylpropanoid
pathways diverge from 4-coumaroyl-CoA
Figure 3. General protocol for metabolite extraction and analysis
Figure 4. Putative indole accumulated 12 h following treatment with UV-C
Figure 5. Fluorescent peak is reproducibly elevated in *pbs3*-2 leaves in response to infection with *Pst avrRpt2* 78

Table 1. Treatment of *pbs3* mutant, wild-type and *PBS3* overexpressing plants for metabolite analysis 74
Table 2. Compounds elevated in response to *Pst avrRpt2* 77

CHAPTER V

Figure 1. Angiosperm phylogeny showing location of paleohexaploidy and tetraploidy events in lineages of interest 87
Figure 2. Synteny between the region of *Arabidopsis* chromosome 5 surrounding *PBS3* and grape and poplar chromosomes 90
Figure 3. Synteny between rice and other species 91
Figure 4. Maximum likelihood phylogenetic tree of rosid GH3 proteins 96
Figure 5. *GH3.7* was inserted within Brassicales 98
Figure 6. Number of GH3 genes in modern species overlayed on angiosperm phylogeny 101

Table 1. Comparison of syntenic genes in *Arabidopsis* chromosomes 5 and 1, grape, polar, *Mimulus* and rice 93
Table 2. GH3 genes in poplar, grape, papaya and *Mimulus* and their relationship to *Arabidopsis* genes 94
Table 3. Comparison of PBS3 protein with other Group IIA members 97
Table 4. Loss and gain of GH3 genes in each syntenic group since divergence from common ancestor 102
Table 5. Expression of AtGH3 genes in response to hormone treatment 103
Table 6. Expression of AtGH3 genes in response to abiotic and biotic stress 104
Table 7. Expression of Group II and Group III AtGH3 genes in various tissue types 106
Table 8. Predicted cis-acting regulatory element binding motifs in AtGH3 promoters 108

CHAPTER VI

Figure 1. Position of PBS3 in pathogen-induced SA metabolism 119
ACKNOWLEDGMENTS

I could not have completed this dissertation and my graduate studies without the guidance, assistance and support of many people. The Wildermuth Laboratory has been a fantastic place for me to grow and develop as a scientist. I would like to thank my fellow graduate students Alexander Jones and Marcus Strawn for sharing this experience with me. Xander frequently inspired me with his enthusiasm for biology and Marc has been a helpful resource for questions of chemistry. Post-doc Divya Chandran has been a great friend and role model for me. Past post-docs Noriko Inada and Sharon Marr were instrumental in my early work in the lab. Several undergraduates have been a great help, especially Aly Wong. Above all, Mary Wildermuth has been an inspirational mentor, continually challenging me to see both the complex details and the big picture of a question and put everything together. I have learned so much from her during my time at Berkeley. I am also grateful for her support in my interest in education and encouraged my teaching endeavors.

Many members of the Plant & Microbial Biology department have helped me both scientifically and personally. Many have been generous in sharing advice, expertise, and equipment during the course of my work, particularly members of my committees. I would also like to thank the PMB graduate students for helping me get through the challenges of graduate school. I am glad to have been part of such an excellent group of graduate students, particularly my cohort, the other PMB Monkeys. I have also been fortunate to have great collaborators in other departments, especially Roger Innes and his laboratory at Indiana University and Richmond Sarpong and his laboratory in Chemistry at UC Berkeley.

As a graduate student, I have had the opportunity to participate in several teaching experiences that have given me an outlet for my enthusiasm regarding science education, including my courses as a GSI, GK-12 Exploring California Diversity program, Berkeley school gardens, and Community Resources for Science. I would particularly like to thank the other members of the GK-12 program for this experience that has so enriched my final year of graduate school.

Finally, I am greatly indebted to my family for their support and encouragement through the rigors of graduate school and my academic career. I would particularly like to thank my parents, Bill Okrent and Connie Wold. My dad has always supported my academic goals and encouraged me to reach higher. Connie has always had faith in me, setting me on a path of science and being there for me as I struggled. Lastly, I could not have finished without the love and support of my husband, Erik Haunreiter. I am so grateful for his love, patience, and thoughtfulness through all of the ups and downs of graduate school.
Chapter I: Introduction

Plants are continually exposed to disease-causing bacterial, fungal, viral, and oomycete pathogens. The economic impact of these pathogens on crop plants is considerable. A conservative estimate is that 14% of crops are lost worldwide due to pathogen-borne diseases, leading to a $220 billion loss a year, with a greater proportion in the developing world (Agrios, 2005). Current agricultural practices typically employ costly and potentially harmful pesticides to combat these pathogens. Another approach is to take advantage of the natural resistance of plants. Research in the field of plant-microbe interactions has revealed much about the mechanism employed by plants to resist pathogen attack. In particular, investigation of the resistance response has been greatly facilitated by use of the model plant Arabidopsis thaliana.

MAMPs and R genes

Plant resistance to pathogens consists of both passive and active defenses. Passive defenses include strong cell walls and waxy cuticles on leaves to prevent access to plant nutrients. Active defenses are induced upon pathogen attack. The co-evolution of resistance in plants and virulence in pathogens is reflected in the three layers of plant-pathogen interaction. In the first layer of the resistance response, plants detect microbe-associated molecular patterns (MAMPs\(^1\)), also called pathogen-associated molecular patterns (PAMPs). MAMPs are polymers common to a wide range of microbes but not the host, and include flg22, a 22 amino acid long segment of flagellin, lipopolysaccharides (LPS) and elongation factor Tu (EF-Tu) from gram-negative bacteria, chitin, ergosterol and \(\beta\)-glucans from fungi, and Nep1-like proteins from oomycetes (Qutob et al., 2006). Pretreatment with MAMPs can enhance resistance to subsequent pathogen attack (Zipfel et al., 2004). The detection of MAMPs by the plant cell through MAP kinase cascades result in a suite of changes, including activation of defense-related genes, production of reactive oxygen species (ROS), and strengthening of the cell wall, and production of antimicrobial proteins and compounds (Nürnberger et al., 2004; Chisholm et al., 2006). This type of resistance response is sometimes called basal resistance, and can limit disease symptoms caused by microbes that are pathogenic on a particular plant.

Microbes, however, have evolved mechanisms to evade detection by plant cells, resulting in the second layer of the plant-microbe interaction. These mechanisms include production of virulence factors, typically released into the plant cell by the type III secretion system in bacterial pathogens. These effectors can limit the plant response to MAMPs. In turn, plants have evolved resistance (R) genes which specifically detect these virulence factors or activity, resulting in the third layer of the interaction (Jones and Dangl, 2006). This type of resistance response, termed effector-triggered immunity, is a strong and rapid response typically characterized by the hypersensitive response (HR), localized cell death restricting pathogen growth (Hammond-Kosack and Jones, 1996). In addition, the responses triggered by recognition of MAMPs are also induced during effector-triggered immunity (Nürnberger et al., 2004). One example of an R gene-effector interaction is the R gene RPS2 from the plant Arabidopsis thaliana and the effector

---

\(^1\) Abbreviations. Compounds: IAA, indole-3-acetic acid; JA, jasmonic acid; SA, salicylic acid; SAG, salicylic acid glucoside. Genes: NPR1, non-expressor of pathogenesis related 1; eds16, enhanced disease susceptibility 16; GH3, Gretchen Hagen 3; ICS1, isochorismate synthase 1; JAR1, jasmonic acid resistant 1; PBS3, avrPphb susceptible 3; PR1, pathogenesis related 1. Organisms: Terms: MAMP, microbe associated molecular pattern; SAR, systemic acquired resistance.
avrRpt2 from the bacterial pathogen *Pseudomonas syringae*. When avrRpt2 is not present, the plant protein RIN4 covalently binds to RPS2. When AvrRpt2, a cysteine protease, is inserted into the plant cell, it cleaves RIN4 and removes the negative regulation. (Mudgett, 2005)

**Systemic Acquired Resistance**

Challenge by pathogens can lead to resistance to subsequent attack by a broad range of pathogens in parts of the plant distal to the site of initial infection. This type of resistance is termed systemic acquired resistance (SAR) (Durrant and Dong, 2004) and is long-lasting. The phytohormone salicylic acid (2-hydroxybenzoic acid, SA) plays a critical role in the SAR response. External application of SA or its analogues, including 2,6-dichloroisonicotinic acid (INA) (Conrath et al., 1995) and Benzo(1,2,3)-thiadiazole-7-carbothioic acid S-methyl ester (BTH) (Lawton et al., 1996), is sufficient to induce SAR. Salicylic acid accumulates both at the site of pathogen infection and in the systemic tissue. While the identity of the systemic signal of SAR has been controversial, (Vernooij et al., 1994; Maldonado et al., 2002), recent evidence strongly suggests that the methyl ester of salicylic acid, methyl salicylate, is a systemic signal. The methyl ester is more soluble in the membrane and less reactive than the free acid, and can be converted to SA by a methyl esterase (SABP2 in tobacco) (Forouhar et al., 2005; Vlot et al., 2008). By grafting wild-type tobacco plants with those unable to produce SABP2, Park et al found that the methyl esterase was required in leaf tissue that receives the SAR signal but not in tissue that generates the signal (2007). This indicated that, during SAR, the methyl esterase hydrolyzes the ester in methyl salicylate, producing free salicylic acid. However, there is evidence that other signals, such as a lipid-based molecule, also contribute to SAR (Chaturvedi et al., 2008).

*NPR1* is a key master regulator of SA-dependent and ethylene/jasmonic acid (JA)-dependent defense responses. NPR1 activity is itself regulated by redox conditions in the cell during the response to pathogens. Under normal cellular conditions, NPR1 is present in an oligomeric form due to intermolecular disulphide bonds (Mou et al., 2003). Upon pathogen infection or SA treatment, induced thioredoxin leads to reduction of NPR1 to a monomer (Tada et al., 2008), which is transported to the nucleus and interacts with TGA transcription factors, leading to induction of SAR-related genes (Dong, 2004). Expression of PR genes, such as *PR1*, is mediated by NPR1 and a hallmark of SAR. NPR1 is involved in a negative feedback loop with SA; *npr1* recessive mutants accumulate elevated levels of SA following pathogen attack (Clarke et al., 1998) as NPR1 negatively regulates SA biosynthesis (Zhang et al., 2010).

**Role of phytohormones in plant defense**

Small molecules mediate networks of signaling pathways in many processes, including plant defense. SA, JA, and ethylene are the major signaling molecules involved in defense, although there is an emerging understanding of the role played by indole-3-acetic acid (IAA) and other hormones thought to be primarily involved in growth processes (Kazan and Manners, 2009); structures of phytohormones are shown in Figure 1.1. SA is primarily involved in resistance to biotrophic pathogens, while JA and ethylene are dominant in resistance to necrotrophic pathogens (Glazebrook, 2005) and response to herbivory and wounding (Farmer et al., 2003). There is extensive crosstalk between the separate signaling pathways (Reski, 2006). In general, the SA pathway is antagonistic with the JA/ethylene pathways, although there are exceptions. *NPR1* is one node of intersection: the gene is required for the antagonistic relationship between SA and JA (Spoel et al., 2003).
Figure 1.1. Structures of selected phytohormones involved in plant defense.
In *Arabidopsis*, the bulk of the pathogen-induced SA is synthesized from chorismate via *Isochorismate Synthase 1 (ICS1)* (Wildermuth et al., 2001; Strawn et al., 2007). Long before scientists discovered the role of SA as a plant hormone, willow bark was used as a pain remedy. A modified form is now sold as the familiar pain-reliever aspirin. SA is produced in response to MAMPs, and many MAMP-dependent responses are also dependent on SA (Tsuda et al., 2008). In addition, SA is also produced in response to virulence factors in R-gene mediated defense as well as virulent pathogens. As mentioned earlier, SA is also required for SAR. JA is one of many fatty acid-derived molecules involved in plant defense, as well as fertility and root growth (Farmer et al., 2003) (Wasternack, 2007). Ethylene is a small gaseous hormone involved in the regulation of many plant processes, including fruit ripening, leaf and flower senescence, as well as the response to abiotic and biotic stress.

**Modification of small molecules**

Phytohormones and other small molecules are frequently conjugated to other small molecules or otherwise modified, forming derivatives with altered chemical properties and modulated activity (Wildermuth, 2006). Common modifications include esterification to form an ester conjugation or conjugation to sugars or amino acids. An example of these modifications is shown for SA in Figure 1.2. Esterification increases membrane solubility and volatility, while conjugation to glucose or other sugars results in enhanced water solubility and often reduced reactivity. Conjugated molecules are often storage forms: most of the pathogen induced SA is conjugated to glucose as the SA-O-β-glucoside (SAG), and has been shown to be transported to the plant vacuole in tobacco (Dean and Mills, 2004; Dean et al., 2005). Similarly, many toxic compounds, or xenobiotics, are conjugated and safely stored in the vacuole. The conjugation of particular amino acids to IAA can lead either to a temporary storage form (alanine and leucine) or target the molecule for degradation (glutamate and aspartate) (Woodward and Bartel, 2005). In contrast, conjugation of isoleucine to JA activates the phytohormone: JA-Ile facilitates the interaction between the F-box protein COI1 and JAZ JA repressor family members (Thines et al., 2007). This interaction results in degradation of the JAZ repressor by the ubiquitin-ligase dependent 26S proteasome and activation of subsets of JA-dependent responses (Chini et al., 2007; Yan et al., 2007; Chung et al., 2008).

**Pathogens modulate hormone and signal transduction pathways**

Some plant pathogens can specifically interfere with plant hormone signaling, such as by synthesizing the hormones or hormone mimics or modifying the compounds through conjugation. SA, IAA, and ethylene are all synthesized by some pathogenic bacteria or fungi (van Loon et al., 2006; Chen et al., 2007; Jones et al., 2007). Some microbial species encode enzymes that catabolize the molecules, such as the salicylate hydroxylase nahG from *Pseudomonas putida* (Yamamoto et al., 1965; Gaffney et al., 1993). Levels of phytohormones can also be manipulated through bacterial effectors, such as the *Pseudomonas syringae* effector AvrRpt2 in the elicitation of elevated levels of IAA in plants. There are several examples of pathogens interfering with phytohormone levels through production of conjugate mimics or conjugation of hormones. Pathovars of *Pseudomonas syringae* synthesize the phytotoxin coronatine, which functions as a JA-Ile analogue (Katsir et al., 2008) and is required for full virulence (Brooks et al., 2004). Other pathovars of *P. syringae* contain an auxin-AA synthetase that is required for virulence, resulting in gall production on oleander plants (Glass and Kosuge, 1988).
Figure 1.2. Covalent modification of salicylic acid as an example of phytohormone regulation. Plant and microbial enzymes catalyzing reactions modifying SA are shown above the appropriate arrow. Not all plant enzymes have been identified in *Arabidopsis*. Abbreviations: MeSA, methyl salicylate; SAGT, SA glycosyltransferase; SABP2, SA Binding Protein 2; SAMT, SA Methyltransferase; nahG, salicylate hydroxylase from *Pseudomonas putida*. 
The GH3 family of phytohormone- amino acid conjugases

Conjugation of amino acids to phytohormones is catalyzed by the GH3 family of enzymes. GH3 (Gretchen Hagen 3) genes were originally identified in Glycine max (soybean) as responsive to the phytohormone auxin (Hagen et al., 1984) have been identified in many plant species (Terol et al., 2006). The family is part of the larger acyl-adenylate/thioester-forming enzyme family, also called the firefly luciferase family (Staswick et al., 2002). This superfamily consists of enzymes that catalyze a variety of reactions with a common first step: the transfer of AMP from ATP to the carboxylic acid group of an acyl substrate, forming an activated acyl-adenylate intermediate. GH3 enzymes are unique within this superfamily in that amino acid conjugation occurs in the absence of a thioester intermediate. Biochemical activity has been demonstrated for some of the 19 Arabidopsis GH3 enzymes. For example, JAR1 (AtGH3.11) catalyzes the formation of JA-Ile (Staswick and Tiryaki, 2004) and several of the enzymes conjugate amino acids to IAA (Staswick et al., 2005). One of the enzymes, AtGH3.12 was not active on any of the several phytohormones tested by Staswick et al. (Staswick et al., 2002).

The AtGH3.12 enzyme is encoded by a gene identified in a screen for mutants with enhanced susceptibility to strains of Pseudomonas syringae pv. tomato (Pst DC3000) carrying the effector avrPphB. The gene was named PBS3 (avrPphB Susceptible 3), one of three genes discovered in the screen. The pbs3 mutant displayed enhanced disease susceptibility to both virulent and avirulent P. syringae DC3000 strains, including Pst avrPphB and Pst avrRpt2 (Warren et al., 1999), as shown in Figure 1.3. This phenotype is similar to that observed for SA deficient Arabidopsis mutants such as eds5/sid1 (Rogers and Ausubel, 1997; Nawrath et al., 2002) and sid2/eds16 (Nawrath and Metraux, 1999; Dewdney et al., 2000; Wildermuth et al., 2001). The Wildermuth laboratory’s preliminary investigation of publicly available gene expression data revealed that the expression of PBS3 is highly correlated to that of the SA biosynthetic gene ICS1 and was induced by biotrophic and hemibiotrophic pathogens (data not shown). These observations led us to hypothesize that PBS3 could be involved in SA-mediated disease resistance, perhaps through conjugation of amino acids to SA.

In this dissertation, I describe my work uncovering the function of PBS3. I found that PBS plays a critical role in SA-mediated resistance to bacterial pathogens, as pbs3 mutants are deficient in total SA accumulation and PR1 gene expression. Based on the SA-related function of PBS3, I screened compounds functionally and structurally related to SA as substrates for the PBS3 enzyme. I found that PBS3 does not conjugate amino acids to SA as hypothesized, but instead to structurally related 4-substituted benzoates. Notably, SA effectively inhibited enzyme activity, a novel method of inhibition for the GH3 family. These results point towards a new function of 4-substituted benzoates in pathogen response, and suggest crosstalk between branches of the chorismate pathway. While metabolic analysis did not uncover large changes in flux through chorismate, it did reveal a fluorescent compound reproducibly and specifically elevated in pbs3 mutants following bacterial infection. Among several possibilities, this metabolite could be a new signaling molecule or the in planta substrate of PBS3. In addition, bioinformatic analysis using comparative genomic data suggests PBS3 is more closely related to genes in distantly related species than previously thought, and is part of a rapidly evolving group in Arabidopsis.
Figure 1.3. *pbs3* plants are more susceptible to infection with *Pseudomonas syringae*. Photograph of wild-type Col-0 and *pbs3-2* mutant plants taken three days post infection with *P. syringae pv. maculicola*. Bacterial suspensions were infiltrated into the abaxial surface of three leaves per plant at an OD$_{600}$ of 0.0001 using a needless syringe.
REFERENCES

Brooks, D.M., Hernandez-Guzman, G., Kloek, A.P., Alarcon-Chaidez, F., Sreedharan, A.,
Identification and characterization of a well-defined series of coronatine biosynthetic
mutants of Pseudomonas syringae pv. tomato DC3000. Mol. Plant-Microbe Interact. 17,
162-174.
Chaturvedi, R., Krothapalli, K., Makandar, R., Nandi, A., Sparks, A.A., Roth, M.R., Welti,
a systemic acquired resistance inducing activity in petiole exudates of Arabidopsis thaliana is independent of jasmonic acid. Plant J. 54, 106-117.
Pseudomonas syringae type III effector AvrRpt2 alters Arabidopsis thaliana auxin
Chini, A., Fonseca, S., Fernandez, G., Adie, B., Chico, J.M., Lorenzo, O., Garcia-Casado,
G., Lopez-Vidriero, I., Lozano, F.M., Ponce, M.R., Micol, J.L., and Solano, R.
(2007). The JAZ family of repressors is the missing link in jasmonate signalling. Nature
448, 666.
(2008). Regulation and function of Arabidopsis jasmonate ZIM-domain genes in
from NPR1 and bacterial resistance: Characterization of the dominant Arabidopsis cpr6-1
mutant. The Plant Cell 10, 557-569.
defense responses, 2,6-dichloroisonicotinic acid and salicylic acid, inhibit catalase
tonoplast vesicles by an ATP-binding cassette transporter-type mechanism. Physiol.
Plant. 120, 603-612.
localization, and tonoplast transport of salicylic acid glucose conjugates in tobacco cell
suspension cultures. Planta 221, 287-296.
Dewdney, J., Reuber, T.L., Wildermuth, M.C., Devoto, A., Cui, J., Stutius, L.M.,
identify eds loci required for limiting growth of a biotrophic fungal pathogen. Plant J. 24,
205-218.
42, 185-209.


responsible for inducing systemic acquired resistance but is required in signal transduction. Plant Cell 6, 959-965.


CHAPTER II: The GH3 Acyl Adenylase Family Member PBS3 Regulates Salicylic Acid–Dependent Defense Responses

PREFACE

A portion of the material presented here was previously published in:


SUMMARY

The pbs3-1 mutant, identified in a screen for Arabidopsis thaliana mutants exhibiting enhanced susceptibility to the avirulent Pseudomonas syringae pathogen DC3000 avrPphB, also exhibits enhanced susceptibility to virulent Pseudomonas syringae strains, suggesting it may impact basal disease resistance. As induced salicylic acid (SA) is a critical mediator of basal resistance responses, free and glucose-conjugated SA levels were measured and expression of the SA-dependent pathogenesis related marker PR1 was assessed. Surprisingly, while accumulation of the salicylic acid glucoside (SAG) and expression of PR1 were dramatically reduced in the pbs3-1 mutant in response to P. syringae infection, free SA was elevated. However, in response to exogenous SA, the conversion of free SA to SAG and the induced expression of PR1 were similar in pbs3-1 and wild-type plants. The pbs3-1 mutant contains two point mutations in the C-terminal region of the protein encoded by At5g13320, resulting in non-conserved amino acid changes in highly conserved residues. Additional analyses with T-DNA insertion (pbs3-2) and transposon insertion (pbs3-3) mutations in At5g13320 confirmed the findings with pbs3-1. PBS3 (also referred to as GH3.12) is a member of the GH3 family of acyl-adenylate/thioester-forming enzymes. Characterized GH3 family members, such as JAR1, act as phytohormone-amino acid synthetases. These results thus suggest that amino acid conjugation plays a critical role in SA metabolism and induced defense responses with PBS3 acting upstream of SA, directly on SA, or on a competitive inhibitor of SA.

---

1 Abbreviations. Compounds: IAA, indole-3-acetic acid; JA, jasmonic acid; SA, salicylic acid; SAG, salicylic acid glucoside. Genes: eds16, enhanced disease susceptibility 16; GH3, Gretchen Hagen 3; ICS1, isochorismate synthase 1; JAR1, jasmonic acid resistant 1; PBS3, avrPphb susceptible 3; PR1, pathogenesis related 1; SAGT, Salicylic acid glycosyltransferase. Organisms: Psm, Pseudomonas syringae pv. maculicola; Psp, Pseudomonas syringae pv. phaseicola; Pst, Pseudomonas syringae pv. tomato. Terms: hpi, hours post infection; hpt, hours post treatment.
INTRODUCTION

Disease resistance in plants is often dependent on recognition of infecting pathogens by specific disease resistance (R) proteins (Jones and Dangl, 2006). Once activated, R proteins trigger a complex cascade of defense responses, including production of activated oxygen species, fortification of cell walls, accumulation of antimicrobial proteins known as pathogenesis-related (PR) proteins, production of antimicrobial secondary metabolites, and localized programmed cell death known as the hypersensitive response (HR) (Hammond-Kosack and Jones, 1996). The specific signal transduction steps leading to these various responses are poorly understood, although forward genetic screens have identified many potential regulators of these responses (Innes, 1998; Glazebrook, 2005).

A key second messenger involved in inducing production of PR proteins and amplifying the oxidative burst is salicylic acid (SA) (Ryals et al., 1996; Shirasu et al., 1997). SA levels in uninfected dicot plants are normally very low, but rapidly increase upon infection with avirulent pathogens (i.e. those that activate an R protein) (Klessig and Malamy, 1994). Virulent pathogens also induce accumulation of SA, but not as rapidly as avirulent strains (Zhou et al., 1998; Shapiro and Gutsche, 2003). In Arabidopsis, the bulk of pathogen-induced SA accumulates as the salicylic acid-\(\beta\)-glucoside (SAG). Mutations that reduce SA production (Nawrath and Metraux, 1999; Dewdney et al., 2000), or impair SA perception (Cao et al., 1994; Delaney et al., 1995; Shah et al., 1997) increase susceptibility to both virulent and avirulent pathogens, indicating that SA contributes to both basal and R protein-mediated resistance.

In order to identify additional components of the R protein signal transduction pathway, Warren et al. (1999) screened for Arabidopsis thaliana mutants that displayed enhanced susceptibility after inoculation with an avirulent strain of Pseudomonas syringae pv. tomato (Pst DC3000 \textit{avrPphB}). Resistance to this strain by Arabidopsis ecotype Col-0 is mediated by the \textit{RPS5} gene (Simonich and Innes, 1995), which encodes a member of the nucleotide binding site-leucine rich repeat (NBS-LRR) family of R proteins (Warren et al., 1999). This mutant screen led to the identification of several susceptible alleles of \textit{RPS5}, plus mutations in three additional genes, which were named \textit{PBS1}, \textit{PBS2} and \textit{PBS3} for \textit{avrPphB} susceptible (Warren et al., 1999). The isolation and characterization of \textit{PBS1} and \textit{PBS2} have been reported previously (Swiderski and Innes, 2001; Tornero et al., 2002). The isolation of \textit{PBS3} is described in (Nobuta et al., 2007).

Warren et al. (1999) reported that the \textit{pbs3} mutant displayed enhanced disease susceptibility to both virulent and avirulent \textit{Pst} DC3000 strains, including \textit{Pst avrPphB} and \textit{Pst avrRpt2} (Warren et al., 1999). This phenotype is similar to that reported for SA deficient \textit{Arabidopsis} mutants such as \textit{eds5/sid1} (Rogers and Ausubel, 1997; Nawrath and Metraux, 1999; Nawrath et al., 2002) and \textit{sid2/eds16} (Nawrath and Metraux, 1999; Dewdney et al., 2000; Wildermuth et al., 2001). The results of positional cloning indicated that \textit{pbs3-1} contains two point mutations in the C-terminal region of the protein encoded by \textit{At5g13320} (Nobuta et al., 2007). Additional analyses with T-DNA insertion (\textit{pbs3-2}) and transposon insertion (\textit{pbs3-3}) mutants in \textit{At5g13320} were used to further confirm the findings with \textit{pbs3-1}. \textit{PBS3} (also referred to as \textit{GH3.12}) is a member of the GH3 protein family of acyl-adenylate/thioester forming enzymes (Staswick et al., 2002). Characterized GH3 family members, such as \textit{JAR1}, act as phytohormone-amino acid synthetases (Staswick et al., 2002; Staswick and Tiryaki, 2004; Staswick et al., 2005). Based on our preliminary bioinformatics analyses, we predicted \textit{PBS3} might play a role in SA-mediated defense signaling.
MATERIALS AND METHODS

Bacteria, Plants, and Growth Conditions

*P. syringae* strains *Psm* ES4326 and *Pst* DC3000 have been described previously (Dong et al., 1991; Whalen et al., 1991). The avirulence gene *avrRpt2* was cloned into broad host-range vector pVSP61 and introduced into these strains by triparental mating as described by (Kunkel et al., 1993). Bacteria were cultured on King’s B medium (protease peptone, 10 mg/ml; glycerol 15 mg/ml; K₂HPO₄, 1.5 mg/ml; MgSO₄, 4 mM pH7.0) supplemented with 100 μg/ml streptomycin (ES4326) or 100 μg/ml rifampicin (DC3000) plus 50 μg/ml kanamycin when *avrRpt2* was present in the strain. *Arabidopsis* varieties Columbia-0 (Col-0) and Nössen-0 (No-0) were grown in Metromix 360 in growth rooms under a 9 hr-light/15 hr-dark cycle at 23°C. Isolation of the *pbs3-1* mutant was described by Warren et al. (Warren et al., 1999). T-DNA insertion lines (Col-0 genetic background) were obtained from the Arabidopsis Biological Resource Center at Ohio State (Columbus, OH) (Alonso et al., 2003). A *Ds* transposon insertion allele of *PBS3* (No-0 genetic background) was obtained from the RIKEN Plant Functional Genomics Research Group (Kanagawa, Japan) (Kuromori et al., 2004).

Measurement of SA and SAG levels in Arabidopsis Leaves

*Arabidopsis* plants for salicylic acid analysis were grown in Scotts Metro-Mix 200 (Scotts, Marysville, OH) with a 12 h photoperiod at a PAR of 100-150 μEm⁻²s⁻¹. The EMS mutant *pbs3-1*, the SALK T-DNA insertion line *pbs3-2*, *eds16-1*, and Col-0 were infected at 4 weeks with OD₆₀₀ 0.0001 *Pst* DC3000 containing the avirulence gene *avrRpt2* on the pVSP61 plasmid (Kunkel et al., 1993). Overnight cultures of *Pst avrRpt2* grown in King’s B with rifampicin to OD₆₀₀ 0.7 were pelleted, resuspended in sterile 10 mM MgSO₄, and diluted to OD₆₀₀ 0.0001. Three leaves per plant were infiltrated with either OD₆₀₀ 0.0001 *Pst* in 10 mM MgSO₄ or 10 mM MgSO₄ as a negative control using a needleless syringe. Mature, fully expanded rosette leaves were collected at 1 dpi, frozen in liquid nitrogen, and stored at -80°C.

The protocol for SA extraction and analysis was adapted from Dewdney et al. (2000). Frozen leaf samples (~0.5 g) were ground to a powder in a prechilled mortar and pestle using liquid nitrogen. The ground leaf material was transferred to a glass tube and suspended in 3 mL of 90% MeOH. 500 ng *o*-anisic acid (Aldrich) in 100% MeOH was added to each sample as an internal standard. Samples were vortexed, sonicated in a water bath sonicator for 20 min, and centrifuged at 5000 rpm for 15 min at 4°C. The supernatant was transferred to a new tube, and the brown pellet was resuspended in 2 mL 90% MeOH with vortexing. This suspension was sonicated for 20 min and centrifuged for 15 min at 4°C. The two supernatants from each sample were combined, vortexed to mix, and divided into two equal portions in new tubes (for free and total SA measurement). The solvent was evaporated using a dry vacuum at ~5 torr.

For total SA, 500 μL of 80U/mL β-glucosidase (Fluka) in 100mM sodium acetate (pH 5.2) was added. The samples were sonicated 5 min, vortexed, and incubated for 90 min at 37°C. For both total and free SA, 2.5 mL 5% trichloroacetic acid (Sigma) was added, and samples were vortexed, sonicated 5 min, and centrifuged at 5000 rpm for 15 min at 4°C. The supernatant was transferred to a new tube, and extracted two times with 2.5 mL of a 1:1 mixture of ethyl acetate and cyclopentane. Organic phases were combined in a new tube, and the solvent was evaporated under vacuum as above. The evaporated samples were stored at -80°C until ready to load on the HPLC. Prior to loading, samples were resuspended in 125 μL 20% MeOH, vortexed, sonicated 5 min, and filtered through 0.45 μm PTFE filter (Millipore).
HPLC separation of leaf extracts was performed on a Shimadzu SCL-10A system with a Shimadzu RF-10A scanning fluorescence detector and Shimadzu SPD-M10A photodiode array detector (Shimadzu, Kyoto, Japan). Samples were separated on a 5 μm, 15 cm x 4.6 mm ID Supelcosil LC-ABZ+Plus column (Supelco) preceded by a LC-ABZ+Plus guard column maintained at 27º C. Prior to loading 50 μL sample, the column was equilibrated with 15% acetonitrile in 25 mM KH₂PO₄ pH 2.5 at a flow rate of 1.0 mL/min. The concentration of acetonitrile was increased linearly to 20% over 15 min, followed by isocratic flow at 20% for 5 min, followed by a linear increase from 20% to 43% over 23 min, a linear increase from 43% to 66% over 2 min, isocratic flow at 66% for 5 min, a linear decrease from 66% to 15% over 5 min, and isocratic flow at 15% for 3 min.

o-Anisic acid and salicylic acid were quantified using a fluorescence detector set at 305 nm excitation/365 nm emission for o-anisic acid and 305/407 for salicylic acid. The calibration curves used were y=4104.6x (r² = 0.9997) for o-anisic acid and y = 3893.8x (r² = 0.9988) for salicylic acid, with x in ng and y in area units. Under these HPLC conditions, salicylic acid eluted at approximately 22 min and o-anisic acid at 10 min. The percent recovery of salicylic acid was estimated from that of o-anisic acid and ranged from 60-70% in the three separate experiments. The detection limit for o-anisic acid and salicylic acid were approximately 0.5 ng.

SAG is calculated for paired samples as total SA minus free SA.

Analysis of PBS3, PR1, and ICS1 mRNA Level

Publicly available Affymetrix Gene Chip data were accessed through the Genevestigator web portal (http://www.genevestigator.ethz.ch/at/) (Zimmermann et al., 2004). Using the “Digital Northern” tool on this site, we determined that PBS3 was expressed at low levels (typically called “absent”) under the vast majority of experimental conditions (data from 2507 whole genome ATH1 chips). Exceptions were from experiments examining biotic stress.

Two particularly informative experiments compared expression in plants infected with different bacterial strains. Expression data from PBS3 (At5g13320) and ICS1 (At1g74710) were downloaded from NASCArrays (Craigon et al., 2004). Expression in response to infection with strains of P. syringae was examined in the dataset AtGenExpress: Response to virulent, avirulent, type III-secretion system deficient and nonhost bacteria (NASCArrays-120) performed by Thorsten Nürnberger. Expression in response to E. coli and P. syringae lacking functional flagellin or type III secretion systems was examined in the dataset: Genome-wide transcriptional analysis of the compatible Arabidopsis thaliana-Pseudomonas syringae pv. tomato DC3000 interaction (NASCArrays-340) performed by William Underwood (Thilmony et al., 2006).

Correlation analyses were performed by downloading the full ATH1 expression dataset from NASCArrays (Craigon et al., 2004) and then limiting the dataset to those slides (GeneChips) using plant stock code N1092 (Col-0), a total of 821 GeneChips. The Pearson correlation (Freedman, 2005) between log₂ expression values of ICS1 (At1g74710) and the remaining >22K genes on the ATH1 GeneChip and between PBS3 (At5g13320) and the remaining genes was computed across these 821 GeneChips.

For RT-PCR, RNA was isolated from infected tissue immediately before inoculation, 12, 24, and 48 hrs post inoculation with Pst avrRpt2 at an OD600 of 0.0001. RNA was purified using the TRIzol method (Invitrogen), and cDNA was generated using SuperScript III (Invitrogen) with random primers. Primers used to amplify UBQ5 (ubiquitin 5; At3g62250) were UBQ5F (5’ GTGGTGCTAAGAAGGAGGAAG 3’) and UBQ5R (5’ TCAAGCTTCAACTCCTTTT 3’) yielding a 250 bp product. Primers used to amplify PR1 (At2g14610) were PR1F (5’
TAGCCCAAAAGATTCTAAGG 3’) and PR1R (5’ CTCGTTCACATAATTCCCAC 3’) generating a 391 bp product. Primers used to amplify PBS3 (At5g13320) were PBS3F (5’ GAAGATGAAAACCTTGGTGCAC 3’) and PBS3R (5’ CCTCCATTACCAAAACAACCG 3’) yielding a 391 bp product. The PBS3 primers flank intron 3, the site of the T-DNA insertion in pbs3-2.

Exogenous SA Treatment

Five week old plants were sprayed with a 2.5 mM SA solution adjusted to pH 7.0. Leaves were collected immediately before spraying and 24 hours after spraying. SA analyses were performed as described above. Q-PCR was performed by my collaborators and is described in (Nobuta et al., 2007).

Phylogenetic Analysis of the Arabidopsis GH3 Family

Protein sequences for the nineteen Arabidopsis GH3 proteins were obtained from TAIR and aligned using Megalign with the ClustalW method (Thompson et al., 1994) from the DNASTAR software package. Identical residues were highlighted in black and conserved residues in gray. The three groups are those determined by Staswick et al. (2002) based on phylogeny (Staswick et al., 2002). The three motifs that form an AMP-binding domain in acyl-adenylases (Chang et al., 1997) are present in all 19 proteins.

RESULTS

PBS3 Belongs to the GH3 Family of Acyl Adenylylases.

At5g13320 is a member of the GH3 multigene family that consists of 19 family members in Arabidopsis variety Col-0 (Staswick et al., 2002). Its sequence is shown in Figure 2.1. The first GH3 gene described was isolated from Glycine max as an early auxin responsive gene (Hagen et al., 1984). Since then, homologs have been identified in many plants, including Arabidopsis thaliana, Oryza sativa, Gossypium spp., Lycopersicon spp., Populus spp., Pinus spp., and the moss Physcomitrella patens based on ESTs or sequence data (Terol et al., 2006), and in some bacteria (Okrent, R. and Wildermuth, M; unpublished data). The GH3 proteins are members of a large enzyme superfamily of acyl-adenylate/thioester-forming enzymes that catalyze a variety of reactions with a common first step: the transfer of AMP from ATP to the carboxylic acid group of an acyl substrate forming an activated acyl-adenylate intermediate (Staswick et al., 2002). Characterized family members have been shown to catalyze the ATP-dependent conjugation of amino acids to the phytohormones jasmonic acid (JA) or auxin (IAA) with mutants exhibiting altered phytohormone phenotypes (Staswick et al., 2002; Staswick and Tiryaki, 2004; Staswick et al., 2005). For example, the jar1 mutant was isolated as a jasmonate insensitive mutant and displays enhanced susceptibility to necrotrophic, but not biotrophic, pathogens (Staswick et al., 1998).

Phylogenetic analysis of the 19 AtGH3 family members identified three sequence homology groups, with the known substrate specificity corresponding to their phylogenetic relationships: Group I members, which include JAR1, a JA-amino acid synthetase, Group II members, which are capable of adenylylating IAA, and Group III members which include PBS3 (GH3.12) and act on unknown substrates (Staswick et al., 2002; Staswick and Tiryaki, 2004; Staswick et al., 2005). Recombinant PBS3 did not show significant activity with any of the tested substrates: JA, JA methylester, IAA, abscisic acid, 1-aminocyclopropane-1-carboxylic acid, gibberellic acid, 2-
MKPIFDINETFQKLDTNSVKIQDNLLEEEITPTNKTEYLQRFLIDRFKDKEFKKVNI
PIVSYEDIKPYLDRVNGESSDVISARTITGFILSSGTSQAAKMQMNKNNYLDNLTFIY
DLRMQVITKHVKVEEGKMMFLFTKQESMTPSGLPARRVATSSYFKSDYFKRNPNSWYYS
YTPDEVILCPNTESLYCHLLCGLVQRDEVRTGSIASFVMVRAIEVLKNSWEELCSNI
RSGHLSNWVTDLGQNSVSLVGLGGPRPELADTTIEEEICNQNSWKGVKRLWPTNTYIETVV
II TGSMSGQYVPLNNYCNDDLPLVSTFYGSSETTFGINLDPKLCPEDVSYTFMNMSYFEFIP
III MDGGDNDDVLEDVKLGCTYEPVTVNTFAGLYRMRVDIVLVTGFYNNAPQFKFVRRENV
VLSIDSDKTEEDLFKAVSQALKLVLESSGLDLKKTSTFSGHYVYVEVDTEKEGE
OKTAQFELDEEALSTCLVNEESLNVKRCFRKDGS SQGASTQYKTPIKSGKALQVLETCVVAKFSS

Figure 2.1. The sequence of PBS3 highlighting locations of \( pbs3\)-1 and \( pbs3\)-2 mutations. Amino acids highlighted in red are mutated in \( pbs3\)-1 (E502K and I519T). The T-DNA insertion site in \( pbs3\)-2 is indicated with a red triangle. The motifs common to acyl-adenylase proteins involved in AMP binding are indicated by blue boxes.
oxophytodienoic acid, linolenic acid, or SA (Staswick et al., 2002) suggesting that the recombinant protein used in these assays was inactive, or that the correct substrate was not provided.

As shown in Figures 2.1 and 2.2, PBS3 has all three AMP-binding motifs that are necessary for adenylation. Residues of import for substrate specificity have not yet been determined but are likely to reside near the AMP-binding motifs as has been shown for other members of the acyl-adenylate/thioester-forming superfamily (Gulick et al., 2004; Nakatsu et al., 2006). The pbs3-1 point mutations result in two amino acid changes (E502K and I519T) that are located in the C-terminal domain of the PBS3 protein, not in or near the AMP-binding motifs. Both residues are highly conserved in Arabidopsis GH3 family members with amino acids containing negatively charged R groups (E or D) at residue 502 and the nonpolar aliphatic R group members I or V at residue 519. The C-terminal region of these GH3 proteins contains no known protein domains or motifs. It is possible that this region facilitates the conjugation of the amino acid to the activated AMP intermediate, though this remains to be determined.

PBS3 transcripts in the pbs3-1 and pbs3-2 mutants were examined. As shown in Figure 2.3, PBS3 transcript was detected in wild-type and pbs3-1 plants in response to Pst avrRpt2 but not in the insertion mutant pbs3-2 using primers that flank intron 3, the location of the T-DNA insertion (see Figure 2.1).

The pbs3-1 Mutation Impairs Accumulation of SAG and PR1 by Pathogens

The enhanced susceptibility of pbs3 mutant plants to both virulent and avirulent pathogens is similar to previously described enhanced disease susceptibility (eds) mutants, many of which are compromised in the accumulation of SA and the expression of the SA-dependent marker PR1 (Glazebrook et al., 1996, 1997; Dewdney et al., 2000). We therefore assayed levels of PR1 transcript in the pbs3 mutant before and after infection with the avirulent P. syringae avrRpt2. As seen in Figure 2.4, the PR1 was strongly induced by 12 hours post inoculation (hpi) in wild-type plants, while PR1 transcript was not detected in pbs3-1 or pbs3-2 plants until 48 hpi. Similar results were also found following inoculation with virulent P. syringae (Nobuta et al., 2007).

The reduced levels of PR1 transcript in pbs3 mutants suggested that production and/or perception of SA might be compromised. We therefore measured both free SA and SA glucoside (SAG) levels in pbs3-1 and pbs3-2 mutant plants before and 24 hpi with Pst avrRpt2. At this time point, wild-type plants exhibit significant expression of PR1, whereas pbs3 mutants do not (Figure 2.4). We found significant accumulation of SAG in wild-type plants, with an average 5-fold reduction in accumulated SAG in the pbs3 mutant plants (Figure 2.5 and Table 2.1). Surprisingly, the pbs3 mutants accumulated approximately 2-fold more free SA than wild-type at this time point (Figure 2.5 and Table 2.1). Because the majority of SA is found as SAG, the total SA present in the pbs3 mutants 24 hpi was significantly lower (2-3 fold) than in wild-type plants. Though mutants with reduced (e.g. pad4 (Zhou et al., 1998), or abrogated (e.g. ics1/sid2-1/eds16-1 (Nawrath and Metraux, 1999; Dewdney et al., 2000; Wildermuth et al., 2001)) induction of SA and SAG in response to pathogen have been reported, this is the first report of a mutant exhibiting elevated free SA and dramatically reduced SAG accumulation in response to pathogen. A direct comparison of PR1 induction and SA and SAG accumulation in the pbs3 mutants with the SA biosynthetic mutant eds16-1, confirmed that whereas pathogen-induced PR1 and SAG accumulation are abrogated in the eds16-1 mutant, they are dramatically reduced in the pbs3 mutants (Figures 2.4 and 2.5).
Figure 2.2. Protein sequence alignment of Arabidopsis GH3 family. The Arabidopsis GH3 protein sequences were aligned using ClustalW. Identical residues are highlighted in black and conserved residues in gray. The three GH3 family subgroups (Staswick et al., 2002) are indicated. Asterisks above the sequences are mutations in pbs3-I (E502K and I519T). AMP binding motifs I – III are indicated.
Figure 2.3. PBS3 transcript in the pbs3-1 and pbs3-2 mutants. RT-PCR was performed using RNA isolated from mature Arabidopsis leaves treated with Pst avrRpt2 at OD_{600}=0.0001 before (0), 12, 24, and 48 hrs post inoculation. M= molecular weight marker. Primers are detailed in Materials and Methods. The PBS3 primers flank intron 3, the site of the T-DNA insertion in pbs3-2.
Figure 2.4. Pathogen-induced expression of *PRI* is delayed and reduced in the *pbs3* mutant. RT-PCR analysis of *PRI* expression in wild type, *pbs3* mutants, and *eds16-1* in response to *Pst avrRpt2* (OD$_{600}$ 0.0001).
Figure 2.5. SAG accumulation is dramatically reduced in pbs3 mutant plants in response to Pst avrRpt2. A, Free SA and SAG levels in wild type Col-0, pbs3-1, and pbs3-2 inoculated with Pst avrRpt2 (OD$_{600}$ 0.0001), or with 10 mM MgSO$_4$ at 24 hrs post inoculation. B, Free SA and SAG levels in wild type, pbs3-1, and pbs3-2 compared with the SA biosynthetic mutant eds16-1 at 24 hrs post inoculation with Pst avrRpt2 (OD$_{600}$ 0.0001). Each bar represents the average of three replicates ± standard deviation. These results suggest that the PBS3 protein contributes to, or regulates, SAG biosynthesis and total SA accumulation, as well as expression of the SA-dependent marker PR1, and resistance to Pseudomonas syringae pathogens. The finding that free SA levels were elevated (not reduced) suggests that PBS3 might act directly on SA and that the product formed by PBS3 impacts accumulation of the SA glucose conjugate and expression of PR1.
Free SA and SAG in response to *Pst avrRpt2*

<table>
<thead>
<tr>
<th></th>
<th>Exp 1</th>
<th>Exp 2</th>
<th>Exp 3</th>
<th>AVG</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Free SA</td>
<td>1.99</td>
<td>1.76</td>
<td>2.55</td>
<td>2.10</td>
<td>0.41</td>
</tr>
<tr>
<td>Free SA</td>
<td>1.42</td>
<td>2.05</td>
<td>2.44</td>
<td>1.96</td>
<td>0.51</td>
</tr>
<tr>
<td>SAG</td>
<td>0.15</td>
<td>0.18</td>
<td>0.35</td>
<td>0.23</td>
<td>0.12</td>
</tr>
<tr>
<td>SAG</td>
<td>0.07</td>
<td>0.10</td>
<td>0.22</td>
<td>0.13</td>
<td>0.07</td>
</tr>
</tbody>
</table>

Free SA and SAG in μg/g fresh weight normalized to Col-0 24 hpi with *Pst avrRpt2* (OD<sub>600</sub> 0.0001) for three independent experiments each done in triplicate. SA and SAG levels were calculated based on the average percent recovery of an internal o-anisic acid standard for each experiment.
**Exogenous SA Restores SAG Accumulation, PR1 Expression and Resistance**

To determine whether pbs3 mutants have a defect in processing free SA, and whether this processing is required for wild-type pathogen-induced accumulation of SAG, we treated wild-type and pbs3-1 mutant plants with exogenous SA. SAG formation was comparable for pbs3-1 and wild-type leaves (Figure 2.6) at 24 hpt with SA. SAG is calculated by subtracting measured free SA (no hydrolysis) from measured total SA (after hydrolysis of SAG to SA) as described in Materials and Methods. Significant SAG (~25 µg/FW) was formed in leaves of wild-type and pbs3-1 mutants 24 hpt with exogenous SA, resulting in reproducible differences in measured total - free SA values. It should be noted, however, that direct measurement of SAG would need to be performed to ascertain subtle changes in SAG in response to exogenous SA. PR1 expression and resistance to Psm was also restored through exogenous SA treatment (data not shown). Given these findings, it appears that the pbs3 mutants are not defective in the perception or processing of SA and that PBS3 function is not required for the conversion of SA to SAG or for the SA-dependent induction of PR1 expression.

**PBS3 is Induced by Pathogens and is Highly Correlated with ICS1**

Analysis of publicly available Arabidopsis ATH1 Affymetrix GeneChip microarray data (Craigon et al., 2004; Zimmermann et al., 2004) revealed that PBS3 (At5g13320) expression is typically quite low, but is strongly induced by Pseudomonas pathogens with some induction by oxidative stresses such as ozone and UV-B. As shown in Figure 2.7A, PBS3 is expressed most rapidly and strongly in response to non-host strains of P. syringae, and to a somewhat lesser degree by avirulent and virulent strains. Furthermore, the type III secretion system (TTSS) is required for maximal induction of PBS3 expression as Pst hrcC mutants, which cannot deliver type III effectors to host cells, do not induce PBS3 as strongly as does Pst (Figure 2.7A). Average ratios for PBS3 gene expression in the Pst hrcC mutant compared with Pst were 0.35, 0.48, and 0.7 at 2, 6, and 24 hpi, respectively. Analysis of the Thilmony et al (2006) microarray dataset (Thilmony et al., 2006) provides even stronger evidence in support of TTSS involvement, as PBS3 induction is dramatically reduced in the Pst hrpS and hrpA mutants, which cannot express or secrete TTSS effector genes and proteins, respectively (Figure 2.7B).

As shown in Figure 2.7, expression of the pathogen-induced SA biosynthetic gene ICS1 (SID2/EDS16) parallels that of PBS3. As PBS3 impacts pathogen-induced free SA and SAG accumulation as well as expression of the SA-dependent marker PR1, the correlation of ICS1 and PBS3 across all experiments was assessed using wild-type Col-0 plants in the NASCArray Database (821 ATH1 GeneChips; Craigon et al., 2004). PBS3 expression is well correlated with ICS1 (At1g74710), with a Pearson correlation coefficient of 0.63 (see Materials and Methods); whereas the Pearson correlation coefficient for PR1 (At2g14610) with ICS1 is 0.42. As shown in Figure 2.7C, when significant expression of ICS1 and PBS3 are observed, their expression is very highly correlated. Indeed, PBS3 is the second most correlated gene with ICS1 of the >22K genes on the ATH1 GeneChip, while PR1 is the 140th most correlated gene. There was no significant correlation in expression observed between PBS3 and any of the other Arabidopsis GH3 family members, suggesting that PBS3 plays a unique role among its family.
Figure 2.6. Exogenous treatment with 2.5 mM SA of pbs3-1 restores SAG accumulation at 24 hr post treatment. SAG levels in pbs3-1 and wild-type plants in response to exogenous SA treatment. Each bar represents the average of three replicates with standard deviation.
Figure 2.7. Comparative expression of PBS3 and ICS1 in response to bacteria. A, Comparative expression of PBS3, ICS1, and PR1 in response to Pseudomonas syringae DC3000 strains of type non-host (Psp), avirulent (Pst avrRpt2), virulent (Pst), a type III secretion deficient mutant (Pst hrcC−), and mock MgCl2 treatment is shown. Expression data was acquired from NASCArrays (Craigon et al., 2004) for the experiment performed by Nürnberger.

B, Ratio of expression induced by type III secretion versus type III secretion wild-type P. syringae strains for the PBS3 and ICS1 genes. The Pst hrpA and hrpS mutants are deficient in type-III secretion, and Pst COR− is deficient in the synthesis of the toxin coronatine. Expression data was acquired from NASCArrays for the experiment performed by Underwood, published in Thilmony, et al. (2006).

C, NASCArray Gene Correlation Plot for log PBS3 (At5g13320) vs. log ICS1 (At1g74710) expression using the full ATH1 Affymetrix wild-type NASCArray dataset. Perfect correlation would be an upward diagonal line. Note the high degree of correlation in slides (GeneChips) with significant induction of both genes.
DISCUSSION

In *Arabidopsis* and many other plants, SA is a key molecule that activates plant defense genes, and its accumulation is known to be necessary for local and systemic acquired resistance (Durrant and Dong, 2004). Although SA is necessary for triggering plant defense pathways, an excess amount of free SA may be phytotoxic (e.g. Lee et al., 1995; Kenton et al., 2000). Plants regulate free SA levels in part by glycosylation forming salicylic acid–2-O-β-glucoside (SAG) as the dominant form of SA detected in plants (Enyedi et al., 1992; Malamy et al., 1992). Infection by pathogens such as TMV (tobacco mosaic virus) and *P. syringae* pathovars rapidly induces the accumulation of free SA and SAG (Chong et al., 2001, Malamy, 1992 #1904; Shapiro and Gutsche, 2003). The conversion of free SA to SAG has been followed using radiolabelled SA (Dean et al., 2005) and appears to be catalyzed by UDP-glucose:salicylic acid glycosyltransferases (SAGT). In tobacco and *Arabidopsis*, putative SAGTs have been identified based on correlation of expression with accumulation of SAG, induction of SAGT expression by free SA, and the ability of these enzymes to catalyze the formation of SAG in vitro (Enyedi and Raskin, 1993; Lee and Raskin, 1999; Lim et al., 2002; Song, 2006).

Although SA appears to be synthesized in the chloroplast (Strawn et al., 2007), SAG is formed in the cytoplasm and then transported into the vacuole (Dean and Mills, 2004; Dean et al., 2005). The vacuolar localization of SAG suggests that it is primarily a storage form of SA. Whether SAG can be exported back out of the vacuole is not known. Because SAG can be rapidly hydrolyzed to form free SA by endogenous hydrolases, SAG is hypothesized to function as an inactive pool for the rapid sustained induction of systemic acquired resistance. In support of this function, exogenous application of SAG, but not the non-hydrolyzable SAG analog thio-SAG, results in the accumulation of free SA and expression of *PR1* (Hennig et al., 1993). The *pbs3* mutants exhibit ~5-fold reduction in pathogen-induced SAG accumulation compared to wild-type *Arabidopsis* plants (Figure 2.5 and Table 2.1). In contrast to SAG levels, the level of free SA is significantly higher in the *pbs3* mutant relative to wild-type at 24 hpi with *Pst avrRpt2*. This unusual phenotype suggests that PBS3 may regulate the conversion of SA to SAG. Total SA levels (SA + SAG) are 2- to 3-fold lower in pathogen-infected *pbs3* mutant plants than in wild-type, thus PBS3 is not only affecting accumulation of SAG, but somehow is impacting the overall metabolism of SA such that net SA biosynthesis is reduced, or SA turnover is increased.

Pathogen-induced expression of the SA-dependent marker *PRI* is severely compromised in the *pbs3* mutants a ~36 hr delay in induction in response to avirulent *Pst avrRpt2* (Figure 2.4). The pathogen-induced expression of *PRI* requires the ankyrin-repeat containing master regulator NPR1 (Cao et al., 1994; Delaney et al., 1995). Expression of the SA biosynthetic gene *ICS1* and *PBS3* are correlated (Figure 2.7C) and *ICS1* and *PBS3* expression precede the dramatic induction of *PRI* expression (Figure 2.7A, most obvious in response to *Ps* and avirulent *Pst*). In the *pbs3* mutants, despite the elevated induced free SA levels at 24 hpt with *Pst avrRpt2*, *PRI* is not induced (Figure 2.4). This result suggests that elevated free SA alone is not sufficient to activate *PRI* expression.

Perhaps, at these time points, total SA (free SA plus SAG) better represents the plant cells’ exposure to free SA during the first 24 hrs after pathogen inoculation. Loss of PBS3 function would then lead to a reduction in *PRI* expression and other defenses because it results in a reduction in total SA levels. The dramatic reduction of *PRI* expression at 24 hpi with *Pst avrRpt2* compared to a 2- to 3-fold reduction in total SA levels in the *pbs3* mutants can be
explained if an SA threshold (best assessed as total SA) is required for induction of PR1 expression. This SA threshold is consistent with the previously proposed SA amplification loop (e.g. Shah, 2003; Jirage et al. 1999).

A review of the relevant literature indicates that when total leaf SA levels are less than 3 µg/g FW, as reported here for the pbs3 mutants, PR1 induction is not typically observed (e.g. Zeier et al., 2004). In addition, when transgenic plant lines overexpressing bacterial SA biosynthetic genes exhibit constitutive expression of total SA of ≤ 3 µg/gFW, PR1 is not significantly induced (Mauch et al., 2001; Verberne et al., 2000). High levels of SA provided by exogenous SA application would then surpass the SA threshold and restore induced PR1 expression in the pbs3-1 mutant as observed. Exogenous SA application also restores PR1 expression in mutants thought to participate in the SA amplification loop (Parker et al., 1996; Glazebrook et al., 1997; Zhou et al., 1998; Falk et al., 1999; Jirage et al., 1999; Feys et al., 2001). An alternate explanation of our results is that SAG is the active form of SA and is required for PR1 expression. However, SAG is unlikely to be active itself as its hydrolysis is required for activation of PR1 induction (Hennig et al., 1993).

In either case, these findings suggest that PBS3 functions upstream of SA synthesis, either in a regulatory capacity (which may include an amplification loop) or in SA biosynthesis. However, this function still requires an explanation of how free SA levels could be elevated in the pbs3 mutants while SAG (and total SA) is dramatically reduced. Another viable hypothesis is that SA needs to be modified by PBS3 to form SAG and to activate PR1. However, in this case exogenous SA application should not restore wild-type SAG accumulation and expression of PR1 in pbs3-1 plants. These confounding findings as well as the uncoupling of free SA levels from induced PR1 are discussed in the context of proposed biochemical activities for PBS3.

**Proposed Biochemical Function of PBS3**

PBS3 is a member of the GH3 family of acyl-adenylate/thioester forming enzymes known to conjugate amino acids (AA) to phytohormones (Staswick and Tiryaki, 2004; Staswick et al., 2005), thus it is tempting to speculate that PBS3 may act directly on SA as an SA- amino acid synthetase to form an SA-AA conjugate. Alternatively, PBS3 could act on a competitive inhibitor of SA, or upstream of SA biosynthesis either in a regulatory capacity or through action on a precursor of SA. As discussed earlier, very little is known about SA metabolism. Though SA synthesis appears to be plastid-localized (Strawn et al. 2007), no predicted chloroplast transit sequence for PBS3 was found, similar to other GH3 proteins. Thus PBS3, like SAGT, is expected to act in the cytosol, though this requires experimental confirmation.

**Could PBS3 act as an SA – AA synthetase?**

As shown in Figure 2.8 (top panel), an SA-AA conjugate may itself be the active form of SA required for PR1 induction. This function of PBS3 would be similar to that of JAR1, which activates JA by forming specific JA-AA conjugates (Staswick and Tiryaki, 2004). Amino acid conjugates of SA have been reported in Arabidopsis (Zhang et al., 2007), grape (Steffan et al., 1988) and bean (Bourne et al., 1991). If the SA-AA conjugate itself is the active form of SA required for PR1 induction, then exogenous SA application should not result in PR1 expression in the pbs3 mutant. However, at 24 hpt with SA, PR1 was expressed similarly in leaves of pbs3 and wild-type plants. This suggests that either SA-AA is not the active form of SA required for PR1 induction or that another GH3 family member, perhaps with lower affinity for SA, may compensate for loss of PBS3 function when supplied with a great excess of the substrate SA.
A. PBS3 acts on SA

\[
\begin{align*}
\text{SA} \xrightarrow{\text{PBS3}} \text{SA-AA} & \quad \bullet \, \text{SAG accumulation} \\
& \quad \bullet \, \text{ Activation of SA-dependent transcriptional responses (e.g. PR1)} \\
\end{align*}
\]

B. PBS3 acts on a competitive inhibitor (X) of SA

\[
\begin{align*}
\text{SA} \xrightarrow{\text{PBS3}} \text{SA-AA} & \quad \bullet \, \text{SAG accumulation} \\
& \quad \bullet \, \text{ Activation of SA-dependent transcriptional responses (e.g. PR1)} \\
\end{align*}
\]

C. PBS3 acts upstream of SA

\[
\begin{align*}
\text{PBS3} \xrightarrow{} \text{SA} & \quad \bullet \, \text{SAG accumulation} \\
& \quad \bullet \, \text{ Activation of SA-dependent transcriptional responses (e.g. PR1)} \\
\end{align*}
\]

**Figure 2.8. Proposed biochemical functions of PBS3.** These models assume PBS3 activity results in conjugation of an amino acid to a small molecule (such as a phytohormone), similar to characterized GH3 family members. *A.* PBS3 acts directly on salicylic acid (SA) to form an SA-amino acid conjugate. In this case, SA-AA, itself, may be the active form of SA required for SAG accumulation and induction of PR1 expression, as in the top panel. Alternatively, the formation of SA-AA may be required for a subsequent conversion to an unknown active form of SA. In the bottom panel, the formation of an SA-AA conjugate is required for the proper spatial localization of SA. The SA-AA conjugate is translocated and then hydrolyzed to release SA, resulting in SAG accumulation and PR1 induction. *B.* PBS3 acts on a competitive inhibitor (X) of SA to inactivate it by forming X-AA, resulting in full SAG accumulation and activation of SA-dependent transcriptional responses. The differential impact of X on SAG accumulation vs. PR1 expression could be explained by different binding affinities of the inhibitor for distinct enzymes controlling these processes. *C.* PBS3 acts upstream of SA either in a regulatory capacity or on a precursor of SA. In this case, net SA biosynthesis is reduced, resulting in reduced SAG accumulation and PR1 expression. Activators of SAG accumulation and PR1 expression are shown in green whereas inhibitors are shown in red.
One candidate for this compensatory role is GH3.5 (At4g27260), the only tested GH3 family member to exhibit adenylase activity with SA as a substrate in vitro (Staswick et al., 2002). This function of PBS3 would explain why PR1 is not induced in the pbs3 mutants despite elevated free SA levels.

Alternatively (Figure 2.8A, bottom), PBS3 could act on SA to form an SA-AA conjugate that is required for the proper spatial localization of the active form of SA, which may be free SA. Amino-acid conjugation of auxin appears to regulate its subcellular and tissue-specific distribution and hydrolysis of phytohormone amino acid conjugates by specific amidohydrolases (e.g. IAA-amidohydrolases) allows for fine regulation of active forms of the phytohormone (Woodward and Bartel, 2005). For example, in response to Pst avrRpt2 free SA would be synthesized in the chloroplast of a cell as an early event associated with the hypersensitive response. It could then be modified to SA-AA by PBS3 in the cytosol of the HR-undergoing cell and then exported to neighboring cells. In the neighboring cells the SA-AA conjugate could be hydrolyzed, releasing free SA that could then induce PR1 expression. SAG formation catalyzed by a cytosolic SAGT could also occur in these neighboring cells. In this case, in the pbs3 mutant, free SA would be elevated as it would not be converted to SA-AA, and SAG formation in neighboring cells would be reduced. Exogenous SA application would negate this requirement for SA transport, thus restoring PR1 expression and SAG formation to wild-type levels in the pbs3 mutant. Consistent with this model, in tobacco free SA accumulation preceded visible HR symptoms by >24 hrs, as observed using an SA reporter strain, with higher levels of free SA detected in neighboring cells surrounding HR lesions as they became apparent (Huang et al., 2006). In addition, induced PR1 accumulation was not observed in cells undergoing an HR (and subsequently dying) but specifically around HR lesions in tobacco and Arabidopsis (e.g. Stone et al., 2000; Dorey et al., 1997).

Could PBS3 act on a competitive inhibitor of SA?

In this scenario, PBS3 would inactivate a competitive inhibitor X by converting it to an inactive form X-AA (Figure 2.8B). Amino acid conjugation can either inactivate phytohormones or target them for degradation pathways (Woodward and Bartel, 2005). In the pbs3 mutants, the competitive inhibitor would compete with SA in binding to the active site of enzymes that use SA as a substrate, such as SAGT. The competitive inhibitor could also bind to enzymes that are regulated by SA-binding. For example, in E. coli, binding of SA to the multiple antibiotic resistance repressor MarR reduces its DNA-binding affinity thus allowing for expression of genes associated with multiple antibiotic resistance (Alekshun and Levy, 1999). In Arabidopsis, SA affects the interaction of the master regulator NPR1 with TGA transcription factors enhancing DNA binding (Despres et al., 2003, Fan and Dong, 2002). If PBS3 acted on a competitive inhibitor of SA, free SA would be elevated in the pbs3 mutants, and SAG and PR1 would be reduced. The differential reduction in SAG accumulation (5-fold) vs. PR1 expression (~50-fold) could be explained by different binding affinities of the inhibitor for distinct enzymes controlling SAG synthesis or PR1 expression. Exogenous SA (2.5 mM) would likely dominate a reversible competitive inhibitor resulting in restored SAG and PR1 formation. The competitive inhibitor would likely be a small molecule similar to SA, such as a substituted benzoic acid. In support of this hypothesis, the Arabidopsis SAGT1 enzyme, in addition to SA, can also use benzoic acid and a subset of other substituted benzoates as substrates (Lim et al., 2002; Song, 2006). Similarly, the DNA-binding affinity of E. coli MarR can be regulated by SA and other small phenolics (Alekshun and Levy, 1999). If PBS3 inactivates an inhibitor, the reduction in
total SA could then be explained by a feedback inhibition on SA biosynthesis by elevated free SA.

**Could PBS3 act upstream of SA biosynthesis?**

PBS3 may act upstream of SA biosynthesis either in a regulatory capacity or through action on a precursor of SA ([Figure 2.8C](#)). For example, adenylation (and amino acid conjugation) of an SA precursor might either be required for its subsequent biochemical conversion or for its proper spatial localization. If PBS3 acted upstream of SA in either a regulatory or biosynthetic capacity, the application of exogenous SA would restore SAG formation and PR1 expression in the *pbs3* mutants as is observed. However, both induced free SA and SAG levels would be reduced in the *pbs3* mutants in this scenario; this was not observed. One possible explanation for the observed increase in free SA with decreased SAG at 24 hpi is that lack of PBS3 function delays SA biosynthesis. If free SA is first made and then subsequently converted to SAG, it is possible that by 24 hpi with *Pst avrRpt2*, the bulk of free SA has been converted to SAG in the wild-type and thus free SA levels have decreased from their maximum, whereas SAG levels are still increasing. This decrease in free SA levels as SAG continues to increase has been observed (e.g. Zhou et al., 1998). In contrast, if SA synthesis is delayed in the *pbs3* mutant, free SA might be at its maximum at 24hpi, with less of it having been converted to SAG thereby resulting in higher free SA and lower SAG levels in the *pbs3* mutant. In this scenario, the induction of PR1 expression requires some threshold of total SA (free + SAG) that was not met at 24 hpi with *Pst avrRpt2* in the *pbs3* mutants.

**CONCLUSION**

These results suggest that PBS3 plays an important role in pathogen-induced SA metabolism and that this function is critical to SAG accumulation, PR1 expression, and both basal and R-protein mediated resistance. Given the function of GH3 family members of acyl-adenylase thioester-forming enzymes, it is likely that PBS3 acts as a small molecule-amino acid synthetase, either upstream of SA biosynthesis, on SA, or on a competitive inhibitor of SA. Further work is clearly needed to unravel the complexity of SA synthesis, activation, processing, and catabolism. In addition, analysis of the *pbs3*-1 mutant, which contains two point mutations resulting in non-conserved amino acid changes in the C-terminus of the PBS3 protein, highlights the importance of this uncharacterized region of GH3 family members and provides a framework for detailed mechanistic analyses of GH3 function.

**REFERENCES**


CHAPTER III: PBS3 Conjugates Amino Acids to 4-Substituted Benzoates and is Inhibited by Salicylate

PREFACE

This research was originally published in:


SUMMARY

Salicylate (SA, 2-hydroxybenzoate) is a phytohormone best known for its role as a critical mediator of local and systemic plant defense responses. In response to pathogens such as Pseudomonas syringae, SA is synthesized and activates widespread gene expression. In gh3.12/pbs3 mutants of Arabidopsis thaliana, induced total SA accumulation is significantly compromised as is SA-dependent gene expression and plant defense. AtGH3 subfamily I and II members have been shown to conjugate phytohormone acyl substrates to amino acids in vitro, with this role supported by in planta analyses. Here we sought to determine the in vitro biochemical activity and kinetic properties of GH3.12/avrPphB susceptible 3 (PBS3), a member of the uncharacterized AtGH3 subfamily III. Using a novel high throughput adenylation assay, I characterized the acyl substrate preference of PBS3. I found PBS3 favors 4-substituted benzoates such as 4-aminobenzoate and 4-hydroxybenzoate, with moderate activity on benzoate and no observed activity with 2-substituted benzoates. Similar to known GH3 enzymes, PBS3 catalyzes the conjugation of specific amino acids (e.g. Glu) to its preferred acyl substrates. Kinetic analyses indicate 4-aminobenzoate and 4-hydroxybenzoate are preferred acyl substrates as PBS3 exhibits both higher affinities (apparent K_m 153 and 459 μM, respectively) and higher catalytic efficiencies (k_cat/K_m 0.0179 and 0.0444 μM⁻¹ min⁻¹, respectively) with these acyl substrates compared with benzoate (apparent K_m 867 μM, k_cat/K_m 0.0046 μM⁻¹ min⁻¹). Notably, SA specifically and reversibly inhibits PBS3 activity with an IC₅₀ of 15 μM. This suggests a general mechanism for the rapid, reversible regulation of GH3 activity and small molecule crosstalk. For PBS3, this may allow for coordination of flux through diverse chorismate-derived pathways.

1 Abbreviations. Compounds: BA, benzoate; 4-HBA, 4-hydroxybenzoate; CA, chorismate; IAA, indole-3-acetate; IC, isochorismate; INA, 2,6-dichloroisonicotinic acid; JA, jasmonate; pABA, para-aminobenzoate; SA, salicylate; MeSA, methyl salicylate; SAG, SA glucoside; VA, vanillate. Enzymes: ICS1 isochorismate synthase 1; IPL, isochorismate pyruvate lyase; PBS3, avrPphB susceptible 3 (GH3.12); SAMeT, SA methyltransferase; SGT1, SA glucosyl transferase. Organisms: Pseudomonas syringae pv. tomato (Pst), Pseudomonas syringae pv. maculicola (Psm).
INTRODUCTION

Mutant screens in Arabidopsis have successfully identified many components of plant disease resistance pathways, with a number of these mutants exhibiting altered phytohormone synthesis, perception, or signaling (Glazebrook, 2001). The pbs3-1 EMS mutant was first isolated in a screen for enhanced susceptibility to avirulent Pseudomonas syringae pv. tomato (Pst) DC3000, carrying the effector avrPphB, resistance to which is mediated by the plant resistance (R) gene RPS5 in Arabidopsis (Warren et al., 1999). Our subsequent analysis with Roger Innes revealed that the pbs3-1 mutant is also more susceptible to virulent P. syringae strains including P. syringae pv. maculicola (Psm) ES4326, suggesting a more general role for PBS3 than in directly mediating effector-R gene interactions (Nobuta et al., 2007). Furthermore, in response to P. syringae pathovars, the pbs3-1 EMS mutant and the pbs3-2 T-DNA insertion line exhibited reduced total SA accumulation and expression of the SA-dependent pathogenesis-related gene PRI (At2g14610) in comparison with wild-type plants (Chapter II). The mutation in pbs3-1 was mapped and cloned, and PBS3 was found to be AtGH3.12 (At5g13320), a member of the AtGH3 multigene family (Nobuta et al., 2007).

The first GH3 gene was characterized by Guilfoyle and Hagen as an early auxin-responsive gene from Glycine max (Hagen et al., 1984). Homologs have since been identified in many other plant species and in bacteria (Terol et al., 2006). The GH3 family is part of the broader acyl-adenylate/thioester-forming enzyme family, also called the firefly luciferase family (Staswick et al., 2002). This superfamily consists of enzymes that catalyze a variety of reactions with a common first step: the transfer of AMP from ATP to the carboxylic acid group of an acyl substrate, forming an activated acyl-adenylate intermediate. GH3 enzymes are unique within this superfamily in that amino acid conjugation occurs in the absence of a thioester intermediate. Figure 3.1 shows the reaction catalyzed by GH3 enzymes, with benzoate as the acyl substrate. To-date, biochemical activity has only been demonstrated for select plant GH3 enzymes. Mechanistic analyses are hampered by the lack of sequence conservation between GH3 and other acyl-adenylase/thioester-forming superfamily members and no structural information for any GH3 family member.

The known substrate specificity of the nineteen Arabidopsis AtGH3 proteins corresponds to their phylogenetic relationship (Staswick et al., 2002) which divides the family into three groups (Figure 3.2). Group I includes JAR1, which has been shown to conjugate the formation of JA-Ile, an active form of JA (Staswick and Tiryaki, 2004). JA-Ile promotes the interaction of the F-box protein COI1 with JAZ JA repressor family members (Thines et al., 2007). This interaction results in degradation of the JAZ repressor by the ubiquitin-ligase dependent 26S proteasome and activation of subsets of JA-dependent responses (Chini et al., 2007; Yan et al., 2007; Chung et al., 2008). Group II is composed of several proteins that conjugate the formation of IAA-amino acid conjugates (Staswick et al., 2005). Conjugation with Asp and Glu have been shown to target IAA for catabolism, while Ala and Leu conjugates seem to serve as inactive storage forms (Woodward and Bartel, 2005). IAA directly binds to TIR1, the F-box protein subunit of SCFTIR1, enabling its interaction with the Aux/IAA family of auxin transcriptional repressors (Dharmasiri et al., 2005; Kepinski and Leyser, 2005). In contrast, IAA-amino acid conjugates are too bulky to fit into the auxin binding pocket of TIR1 and do not promote TIR1-Aux/IAA complex formation (Mark Estelle, personal communication). Group III, which includes PBS3
Figure 3.1. Reaction catalyzed by GH3 enzymes, shown with benzoate as the acyl substrate. Adenylation of the acyl substrate is followed by formation of the amino acid conjugate, shown with a generic amino acid.
Figure 3.2. Neighbor-joining phylogenetic tree of the *Arabidopsis* GH3 family of proteins. Sequences were aligned using ClustalW and constructed using the neighbor-joining algorithm in MEGA 4.0 (Tamura et al., 2007). Groups I, II, and III are described in (Staswick et al., 2002) and correspond to substrate specificity and global sequence similarity.
(AtGH3.12), contains proteins for which the substrate specificity has not yet been reported, as these proteins were inactive on tested substrates (Staswick et al., 2002).

The amino acid conjugation reaction catalyzed by characterized GH3 enzymes plays a critical role in plant hormone homeostasis and plant-microbe interactions (e.g., (Staswick and Tiryaki, 2004; Woodward and Bartel, 2005)). Not only do plants regulate active phytohormone forms through amino acid conjugation, but some plant pathogens specifically manipulate plant conjugation of phytohormones or produce mimics of amino acid-conjugated hormones to interfere with phytohormone-conjugate driven processes. For example, *P. syringae* pv. *savastanoi* encodes an auxin-AA synthetase that is required for virulence, resulting in gall production on oleander plants (Glass and Kosuge, 1988). *Pst* DC3000 and *Psm* ES4326 synthesize the phytotoxin coronatine, which is required for full virulence (e.g. (Brooks et al., 2004)) and functions as a JA-Ile analogue (Katsir et al., 2008). And, notably, the HopW1-1 effector of *Psm* ES4326 targets PBS3, altering host SA signal transduction pathways (Lee et al., 2008).

Because PBS3 is a GH3 protein and *pbs3* mutants exhibit SA-deficient phenotypes, we proposed that PBS3 might act upstream of SA, on SA, or on a compound that could compete with SA (see Chapter II and Figure 2.8). In the last case, such a compound might bind to a metabolic or regulatory enzyme that normally binds SA. Here, I systematically examine the acyl substrate and amino acid preference of PBS3 followed by determination of the kinetic and catalytic properties of PBS3. In addition, the influence of putative inhibitors on enzymatic activity is assessed. The impact of each of the two point mutations in the lack of function *pbs3-1* EMS mutant is also ascertained. Together, these findings provide a framework to explore the role of PBS3 in pathogen-induced SA accumulation and disease resistance. Furthermore, they support an underexplored and significant role for 4-substituted benzoates in plant-pathogen interactions.

**MATERIALS AND METHODS**

**Materials and General Protocols**

All specialty reagents and chemicals were purchased from Sigma-Aldrich unless otherwise specified. HPLC-grade solvents (EMD Biosciences) were employed in the HPLC analyses. TLC plates were silica gel 60 with F254 from EMD Biosciences. Commonly utilized protein and molecular biology reagents and protocols were prepared and used as described in *Current Protocols of Molecular Biology* (Ausubel et al., 2005).

**AtPBS3 Cloning and Purification**

The *AtPBS3* coding sequence was amplified from cDNA isolated from *Arabidopsis thaliana* ecotype Col-0 and inserted into a pET-28a vector (Novagen) as an NdeI/BamHI fragment using forward primer 5’<catgacatatgaagccaatcttcgata and reverse primer 5’<cacgtggtgggatcctaaatactgaagaatt. The resulting construct, pET-His-PBS3, contains an N-terminal histidine tag fused to the coding region of *AtPBS3*. Crude cell extracts were prepared from a 2 L culture of *E. coli* Rosetta2 (DE3) cells transformed with pET-His-PBS3 grown in TB media containing 0.2% glucose, 50 µg/ml kanamycin and 30 µg/ml chloramphenicol. Cultures were grown at 37°C to mid-log phase, shifted to 18°C, and 0.1 mM IPTG was added to induce His-PBS3 synthesis. Cells were harvested after shaking overnight (~25 g wet weight), and stored
at –20°C. His-PBS3 was then purified using Ni-NTA His-Bind Resin (Novagen) according to manufacturer’s directions. His-PBS3 was the only visible protein when 10 µg of this eluant was resolved by SDS-PAGE and visualized using Coomassie Blue. Thrombin (Novagen) was used to cleave the histidine tag. As His-PBS3 and thrombin-cleaved PBS3 displayed similar activity on a set of 5 substrates, His-PBS3 was utilized unless specified. Protein concentrations were determined by a Bradford assay modified for 96-well plate format using Coomassie Blue G-250 (EM Biosciences) with bovine serum albumin as the standard. Aliquots of the purified recombinant PBS3 proteins (5.3 to 9.5 mg/ml in 100 mM Tris, pH 7.7, 10% glycerol, 1 mM DTT) were stored at –80°C. The same batch of enzyme was used for each set of assays reported in a single figure or table, with assays repeated with enzyme from at least on other batch to confirm reproducibility of results. In all assays, an amount of enzyme concentrations was used such that the velocity was linear with increasing enzyme concentration.

**Generation and Purification of PBS3 Mutants**

Amino acid substitutions were introduced into the pET-His-PBS3 plasmid using the QuikChange II Site-Directed Mutagenesis Kit (Stratagene) with the assistance with forward primer 5’caccactcttatctgagaggcggctcgatctctgatcgacatc and reverse primer 5’gatgtcgattcaagcgactccttcattaccaaacaacagttg to make the E502R mutation and forward primer 5’caccactcttatctgagaggcggctcgatctctgatcgacatc and reverse primer 5’gatgtcgattcaagcgactccttcattaccaaacaacagttg to make the I519T mutation. The double mutant was made using the pET-His-I519K plasmid with the E502R forward and reverse primers. The His-tagged single and double mutant proteins were purified as described for wild-type PBS3, above.

**Adenylation Reaction**

His-PBS3 adenylation activity was monitored spectrophotometrically at 340 nm by coupling the production of pyrophosphate to oxidation of NADH using Pyrophosphate reagent (Sigma, O'Brien, 1976). The reagent contained the coupling enzymes fructose-6-phosphate kinase, aldolase, triosephosphate isomerase, glycerophosphate dehydrogenase, and appropriate substrates and cofactors, including NADH. The Pyrophosphate reagent was reconstituted in 4 mL ddH2O and used at a volume of 65µL per 200 µL reaction. For the screen of substrates, reaction mixtures (200 µL) were in 45 mM imidazole (pH 7.4) and contained 5.0 mM MgCl2, 2.5 mM ATP, 1 mM DTT, 0.1-10 mM substrate, and 100 µg His-PBS3 enzyme in addition to the Pyrophosphate reagent. For determination of kinetic parameters for ATP, reactions contained 50 µg/mL His-PBS3 enzyme, Pyrophosphate reagent, 1 mM DTT, and 10 mM acyl substrate. The concentration of ATP was varied while either (a) maintaining a 1.5:1 Mg2+ to ATP ratio or (b) with 10 mM MgCl2. Three replicates were performed for each substrate or concentration value. The reactions were initiated with the addition of substrate to a 96-well plate preheated to 30°C and analyzed using a Spectromax Plus microplate spectrophotometer (Molecular Devices). The change in absorbance at 340 nm was measured every 10-15 sec for at least 20 min and converted to velocity by least squares fitting of each curve using the accompanying program SOFTmax PRO 3.0 with manual assessment/confirmation of the linear range. The velocity of a no His-PBS3 control was subtracted; this velocity was on average 0.16 µM/min. An extinction coefficient of 6.22 µM⁻¹ cm⁻¹ for NADH was used to convert velocity values from mAU/min to µM/min. Kinetic parameters were estimated by fitting initial velocity values to the Hanes equation (Rudolph et al., 1979).
Amino Acid Conjugation Reaction (Full Reaction)

Reaction mixtures (200 µL) were in 50 mM Tris-HCl (pH 8.5) and contained 5.0 mM MgCl₂, 2.5 mM ATP, 1.0 mM DTT, 1.0 mM acyl substrate, 1.0 mM amino acid, and 100 µg/mL His-PBS3 enzyme. Reactions were incubated at 30° C for 2 hr and monitored by TLC and/or HPLC. Reaction mixtures and standards were spotted on silica gel 60 F₂₅₄ plates (EMD Biosciences) and developed in either 30:60:10 dichloromethane-ethyl acetate-formic acid (for most amino acids) or 8:1:1 isopropanol, ammonium hydroxide-water (for basic and hydrophilic amino acids). The method used to detect the conjugates varied by substrate. pABA conjugates were detected by staining TLC plates with vanillin reagent (6% (w/v) vanillic acid and 1% sulfuric acid in ethanol) and by visualizing TLC plates under short wavelength UV. 4-HBA, vanillic acid, and trans-cinnamic acid conjugates were detected by visualizing TLC plates under short wavelength UV. BA conjugates were detected by HPLC using a UV detector at 233 nm. HPLC conditions are the same as those used in Km determination (below). SA was analyzed by TLC visualized under short wavelength UV and by HPLC using a fluorescence detector (ex 305nm/em 407 nm); however, no spots or peaks associated with SA-conjugates were detected.

Kₘ Determinations

Reaction mixtures (200 µL) were in 50 mM Tris-HCl (pH 8.5) and contained 5.0 mM MgCl₂, 5.0 mM ATP, 1 mM DTT, 10 mM glutamate, varying concentrations of benzoate, and 50 µg/mL His-PBS3 (reaction with 4-HBA) or 125 µg/mL His-PBS3 (reactions with pABA and BA). Benzoate substrate concentrations ranged from 50 µM to 5 mM. Reactions were quenched every 2 min by diluting the reaction mixture 10 fold into 2.5% HCl. Samples were filtered through a 0.2 µm Millex-LG syringe filter (Millipore), and a 50 µL aliquot was injected into a Shimadzu SCL-10AVP series HPLC system equipped with a Shimadzu SPD-10AVP photodiode array detector and a Shimadzu RF-10AXL fluorescence detector. A 5 µm, 15 cm x 4.6 mm inner diameter Supelcosil LC-ABZPlus column (Supelco) preceded by a LC-ABZ-Plus guard column was pre-equilibrated in 5% acetonitrile with 25 mM potassium phosphate buffer (initial HPLC buffer), pH 2.5, at a flow rate of 1.0 ml/min. The elution program and detection method varied by substrate. For BA, the initial HPLC buffer was run for 10 min, then changed to 52% acetonitrile and 48% 25 mM potassium phosphate buffer over 12 min. Under these conditions, BA-Glu eluted at 15.5 min and BA at 19 min. Elution was monitored by UV absorbance at 233 nm. Authentic samples were used to generate calibration curves for BA (y = 0.0397x) and BA-Glu (TCI America; y = 0.0407x) with x in absorbance units and y in nM compound. For pABA, pABA eluted at 7.5 min and pABA-Glu eluted at 5.0 min using the BA program. Elution was monitored by fluorescence at ex290/em340 nm. Authentic samples were used to generate calibration curves for pABA (y = 0.2585x) and pABA-Glu (y = 0.1045x), with x in fluorescence units and y in nM compound. For 4-HBA, the initial HPLC buffer was run for 4 min, then changed to 52% acetonitrile and 48% 25 mM potassium phosphate buffer over 7 min. Elution was monitored by UV absorbance at 254 nm. An authentic sample of 4-HBA was used to generate a calibration curve (y = 0.0274x) and synthesized 4-HBA-Glu (see below) was used to generate a calibration curve (y = 0.0318x) with x in UV absorbance area units and y in nM compound. Under these conditions, 4-HBA eluted at 9.7 min and 4-HBA-Glu eluted at 8.3 min. The reaction velocity with each substrate concentration was determined from a linear fit of four time points. Kinetic parameters were estimated by fitting initial velocity values to the Hanes
equation as above, with data from at least 6 substrate concentrations. Nonlinear curve fitting using Synergy KaleidaGraph version 4.0 gave very similar results.

**Mass Spectrometry to Verify 4-HBA-Glu Formation**

A 5 µm, 15 cm x 4.6 mm inner diameter Prevail C18 column (Alltech) preceded by a 7.5 cm x 4.6 mm guard column was pre-equilibrated in 5% acetonitrile with 25 mM formate buffer (initial HPLC buffer), pH 2.8, at a flow rate of 1.0 ml/min. The initial buffer was run for 7 min, then changed to 52% acetonitrile over 15 min. The putative 4-HBA-Glu peak eluting at 11.3-11.9 min was collected and dried down under vacuum. This sample was rerun using our standard HPLC assay for verification. Samples for LC-MS analysis were resuspended in a 50:50 mixture of 25 mM ammonium formate and acetonitrile and analyzed using a Thermo Fisher Surveyor HPLC with a Phenomenex C18, 2 x 150mm column and a Thermo Fisher LCQ Classic Mass Spectrometer at the Vincent Coates Foundation Mass Spectrometry Laboratory (Stanford University, USA). LC-MS with electrospray ionization (ESI) positive-ion analysis employed a C18 column with similar elution scheme. MS-MS fragmentation was performed on the putative dominant ion peak for further confirmation of compound identity. The sample was also analyzed by nanospray infusion on the Q-TOF mass spectrometer (Waters Micromass Q-TOF hybrid quadrupole time of flight) with MS-MS performed on 268.1 m/z.

**Synthesis of 4-HBA-Glu Standard**

As 4-HBA-Glu is not commercially available, it was synthesized in collaboration with the Sarpong lab for calibration of the HPLC. Briefly, 4-acetoxynbenzoic acid was converted to the acyl chloride and condensed with diethyl L-glutamate to form diethyl 4-acetoxybenzoyl-L-glutamate. The protecting groups were removed in base to yield 4-HBA-Glu as a white crystalline solid with 99% purity by HPLC. The structure was confirmed by 1H-NMR and 13C-NMR. See supplementary method S1 in (Okrent et al., 2009) for complete procedures and spectral data.

**Inhibition Studies**

Inhibition of PBS3 amino acid conjugation was assessed with SA using our HPLC assay to monitor pABA-Glu formation. Reaction conditions were the same as for the $K_m$ experiment (above), with the addition of SA (5 to 600 µM). A pABA concentration equivalent to the $K_m$ (150 µM) was used. The adenylation assay was then used to measure the effect of potential inhibitors on adenylation velocity using pABA as the acyl substrate. Reaction mixtures (200 µL) were in 45 mM imidazole (pH 7.4) and contained 5.0 mM MgCl$_2$, 5.0 mM ATP, 1 mM DTT, 150 µM pABA, 100 µg/mL His-PBS3 enzyme, 65 µL Pyrophosphate reagent reconstituted in 4.0 mL ddH$_2$O, and potential inhibitors. SA and MeSA were tested over a range of concentrations from 10 to 600 µM. Other potential inhibitors (INA, 2,4-DHBA, 3-HBA, IAA, and JA) were tested at 30 and 300 µM only. To determine whether SA act as a competitive inhibitor, the effect of using 150 µM or 1.5 mM of pABA with 300 µM SA was tested.

**Potential Inhibition of ICS1-catalyzed SA Biosynthesis**

To evaluate compounds as inhibitors of the AtICS1-catalyzed conversion of chorismate to isochorismate, a coupled spectrophotometric assay was used (Strawn et al., 2007). The conversion of chorismate to isochorismate was coupled to the oxidation of NADH through the coupling enzymes isochorismate pyruvate lyase (recombinant PchB), which converts IC to
pyruvate and SA, and lactic dehydrogenase (Sigma L1254), which converts pyruvate to lactate in a NADH-dependent fashion. The 200 µL reaction contained an effective chorismate concentration of 90 µM (apparent $K_m$ of AtICS1 for chorismate), 0.4 mM NADH, 0.833 µg/mL L-lactic dehydrogenase, 32.0 µg/mL PchB, and 10 µg/mL AtICS1 in 100 mM Tris, pH 7.7 with 10% glycerol, and 10 mM MgCl$_2$. AtICS1 activity was assessed in duplicate for reactions with or without 200 µM 4-HBA, pABA, or pABA-Glu and for the appropriate no enzyme controls. These assays were also performed with 1 mM 4-HBA, pABA, or pABA-Glu. The student’s T-test ($\alpha = 0.1$) was used to evaluate whether any observed small changes in activity were statistically significant.

**Potential Inhibition of Isochorismate Pyruvate Lyase**

To evaluate compounds as inhibitors of the isochorismate pyruvate lyase (IPL) reaction in which isochorismate is converted to SA, a modified version of our coupled ICS spectrophotometric assay was employed, described above. In this case, 50 µg/reaction of recombinant PchB was employed with 60 µM isochorismate, the approximate $K_m$ of our recombinant enzyme for isochorismate under our assay conditions. The reported apparent $K_m$ for purified *P. aeruginosa* PchB for isochorismate is 12.5 µM (Gaille et al., 2002). IPL activity was assessed in duplicate for reactions with or without 100 µM 4-HBA, pABA, or pABA-Glu and for the appropriate no enzyme controls. The student’s T-test ($\alpha = 0.1$) was used to evaluate whether any observed small changes in activity were statistically significant. Isochorismate was prepared and purified as in (Strawn et al., 2007).

**RESULTS**

**Determination of Acyl Substrates of PBS3**

To explore the biochemical activity of PBS3, His-tagged PBS3 was overexpressed and purified recombinantly. A novel high-throughput, 96-well format adenylation assay was also developed for use in determining putative acyl substrate(s) of PBS3. In this assay, the release of pyrophosphate, which occurs when the acyl substrate is adenylated to form the acyl substrate-AMP conjugate (Figure 3.1), is coupled to the oxidation of NADH which can be measured spectrophotometrically (O’Brien, 1976). This allows for rapid kinetic assessment of the GH3 adenylation reaction that is not readily performed using standard radiolabelled PPi exchange assays. The findings are reported using 1 mM acyl substrate to allow for comparison with previously reported GH3 acyl substrate surveys (e.g. (Staswick et al., 2002), (Staswick et al., 2005)). Assessments with 100 µM and/or 10 mM acyl substrates were performed for substrates on which PBS3 displayed moderate and low activity to ensure I had not missed the linear range of activity. No significant differences in the relative activities of PBS3 with these acyl substrates were observed compared with the results using 1 mM substrate.

The reduced accumulation of salicylic acid and expression of *PR1* in response to *P. syringae* in *pbs3* mutants suggested PBS3 might act upstream of SA, on SA, or on a compound that could compete with SA in binding (either as a substrate or modulator of activity) to an SA-binding metabolic or regulatory enzyme, as is discussed in Chapter II. Therefore, the initial screen consisted of three overlapping classes of commercially available compounds: 1) phytohormones including SA and its functional analog INA; 2) compounds that are structurally similar to SA (i.e. benzoates); and 3) compounds in the chorismate pathway. Recombinant PBS3 did not exhibit significant activity on any of the phytohormones tested (Table 3.1), including SA and the
**TABLE 3.1**

Initial acyl substrate screen

<table>
<thead>
<tr>
<th></th>
<th>Activity*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>nmol/min/mg</td>
</tr>
<tr>
<td><strong>Phytohormones</strong></td>
<td></td>
</tr>
<tr>
<td>Indole-3-acetate</td>
<td>1.5 ± 0.1</td>
</tr>
<tr>
<td>Jasmonate</td>
<td>2.9 ± 0.1</td>
</tr>
<tr>
<td>Gibberellate</td>
<td>0.1 ± 0.1</td>
</tr>
<tr>
<td>Abscisate</td>
<td>0.8 ± 0.1</td>
</tr>
<tr>
<td>Salicylate</td>
<td>1.7 ± 0.1</td>
</tr>
<tr>
<td>2,6-dichloroisonicotinate (INA)</td>
<td>0.1 ± 0.2</td>
</tr>
<tr>
<td><strong>Benzoates</strong></td>
<td></td>
</tr>
<tr>
<td>Benzoate</td>
<td>59.6 ± 1.1</td>
</tr>
<tr>
<td>Phenylacetate</td>
<td>2.3 ± 0.1</td>
</tr>
<tr>
<td>Salicylate (2-Hydroxybenzoate)</td>
<td>1.7 ± 0.1</td>
</tr>
<tr>
<td>4-Hydroxybenzoate</td>
<td>135.2 ± 8.6</td>
</tr>
<tr>
<td><strong>Chorismate pathway compounds</strong></td>
<td></td>
</tr>
<tr>
<td>Chorismate†</td>
<td>32.9 ± 2.0</td>
</tr>
<tr>
<td>Prephenate</td>
<td>10.1 ± 0.4</td>
</tr>
<tr>
<td>Anthranilate (2-Aminobenzoate)</td>
<td>0.7 ± 1.1</td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>0.0 ± 0.0</td>
</tr>
<tr>
<td>Tyrosine</td>
<td>0.0 ± 0.0</td>
</tr>
<tr>
<td>Tryptophan</td>
<td>0.0 ± 0.0</td>
</tr>
<tr>
<td>trans-Cinnamate</td>
<td>21.3 ± 1.8</td>
</tr>
</tbody>
</table>

* Activities are the mean of three replicates ± 1 S.D. and are determined for 1 mM acyl substrate using the coupled adenylation assay. Activities below 10 nmol/min/mg are minimal. Additional experiments yielded similar results. † Activity is due to 4-HBA in chorismate preparation.
functional SA analog INA (2,6-dichloroisonicotinic acid). In contrast, recombinant PBS3 was very active on benzoate (BA) and 4-hydroxybenzoate (4-HBA); activities assessed with 1 mM substrate were 59.6 and 135.2 nmol/min/mg for BA and 4-HBA respectively. Moderate activity was observed for the chorismate pathway compounds chorismate (32.9 nmol/min/mg), a precursor of SA, and trans-cinnamate.

Except for chorismate and prephenate, the acyl substrates on which PBS3 showed activity contain aromatic rings. Therefore, we wanted to establish whether PBS3 is truly active on chorismate (and prephenate) or is active on breakdown products of these compounds. Known breakdown products of chorismate are phenyl pyruvate, prephenate, and 4-HBA (Gibson, 1964; Gajewski et al., 1987; Holden et al., 2002). As shown in Table 3.1, PBS3 was highly active on 4-HBA with low activity on the non-aromatic prephenate (10.1 nmol/min/mg), and no significant activity on the aromatic breakdown product phenyl pyruvate. To determine whether the observed activity on chorismate and prephenate was due to contamination with 4-HBA, HPLC analysis was performed on freshly prepared chorismate and prephenate solutions. 4-HBA was present in our chorismate solution at a level sufficient to account for the observed PBS3 activity on chorismate (33.5 µM 4-HBA in 1 mM chorismate). On the other hand, no 4-HBA was detected in the 1 mM prephenate solution, even after incubation for 1 hr under assay conditions (30°C, pH 7.4). This suggests that the low adenylation activity of PBS3 with 1 mM prephenate is valid.

To further examine the acyl substrate preference and specificity of PBS3, compounds structurally related to the active acyl substrates above were examined (Table 3.2). As PBS3 exhibited moderate activity on trans-cinnamate, an additional series of cinnamic acids were tested. PBS3 exhibited higher activity on the parent compound trans-cinnamate than on any of the tested derivatives. As PBS3 exhibited significantly greater adenylation activity with 4-HBA and BA than trans-cinnamate, I further investigated its substrate preference on substituted benzoates.

**PBS3 Strongly Prefers para-substituted Benzoates as the Acyl Substrate**

The adenylation activity of PBS3 on mono-substituted benzoates substituted with hydroxyl (-OH) or amino (-NH₂) groups at either the 2-, 3- or 4- position was examined (Table 3.2). The compounds on which PBS3 exhibited the highest activity, 4-HBA (135.2 nmol/min/mg) and pABA (91.1 nmol/min/mg), are substituted at the para-position and were favored over BA. In contrast, substitution at the meta position with either –OH or –NH₂ resulted in significantly lowered PBS3 adenylation activity compared with BA, with adenylation activities of 13.3 and 11.5 nmol/min/mg for 3-HBA and 3-ABA, respectively. Finally, substitution at the ortho (2-) position was strongly disfavored with minimal activities using SA (1.7 nmol/min/mg) or 2-ABA (0.6 nmol/min/mg).

The impact of additional hydroxyl or methoxy (-MeO) substitutions on PBS3 adenylation activity (Table 3.2) was next examined. Again para-hydroxyl substituted benzoates were favored with ortho-hydroxyl substituted benzoates strongly disfavored as follows: 4-HBA > BA > 3,4-DHBA > 3-HBA >> [SA; 3,4,5-THBA; 2,4-DHBA; and 2,3-DHBA]. At the meta-position, Me substitutions were preferred over additional hydroxyl substitutions. For both vanillate (4-hydroxy-3-methoxy benzoate) versus 3,4-DHBA and syringate (4-hydroxy-3,5-dimethoxy benzoate) versus gallate (3,4,5-trihydroxybenzoate), PBS3 exhibited at least 4-fold higher adenylation activity with MeO compared with the hydroxyl substitutions. Additional modifications (hydroxyl substituents) to pABA also reduced PBS3 adenylation activity, with loss
**TABLE 3.2**  
Acyl substrate preference and specificity of PBS3.

<table>
<thead>
<tr>
<th>Cinnamates</th>
<th>R&lt;sub&gt;2&lt;/sub&gt;</th>
<th>R&lt;sub&gt;3&lt;/sub&gt;</th>
<th>R&lt;sub&gt;4&lt;/sub&gt;</th>
<th>R&lt;sub&gt;5&lt;/sub&gt;</th>
<th>Activity* (nmol/min/mg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>trans-Cinnamate</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>21.3 ± 1.8</td>
</tr>
<tr>
<td>ortho-Coumarate</td>
<td>OH</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>12.1 ± 0.4</td>
</tr>
<tr>
<td>meta-Coumarate</td>
<td>H</td>
<td>OH</td>
<td>H</td>
<td>H</td>
<td>12.4 ± 0.7</td>
</tr>
<tr>
<td>para-Coumarate</td>
<td>H</td>
<td>H</td>
<td>OH</td>
<td>H</td>
<td>5.8 ± 0.7</td>
</tr>
<tr>
<td>Caffeate</td>
<td>H</td>
<td>OH</td>
<td>OH</td>
<td>H</td>
<td>6.9 ± 3.9</td>
</tr>
<tr>
<td>Ferulate</td>
<td>H</td>
<td>OCH&lt;sub&gt;3&lt;/sub&gt;</td>
<td>OH</td>
<td>H</td>
<td>5.4 ± 0.5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Benzoates</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Benzoate</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>59.6 ± 1.1</td>
</tr>
<tr>
<td>2-Hydroxybenzoate (SA)</td>
<td>OH</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>1.7 ± 0.1</td>
</tr>
<tr>
<td>2-Aminobenzoate (anthranilate)</td>
<td>NH&lt;sub&gt;2&lt;/sub&gt;</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>0.7 ± 1.1</td>
</tr>
<tr>
<td>3-Hydroxybenzoate</td>
<td>H</td>
<td>OH</td>
<td>H</td>
<td>H</td>
<td>13.3 ± 1.2</td>
</tr>
<tr>
<td>3-Aminobenzoate</td>
<td>H</td>
<td>NH&lt;sub&gt;2&lt;/sub&gt;</td>
<td>H</td>
<td>H</td>
<td>11.5 ± 0.2</td>
</tr>
<tr>
<td>4-Hydroxybenzoate</td>
<td>H</td>
<td>H</td>
<td>OH</td>
<td>H</td>
<td>135.2 ± 8.6</td>
</tr>
<tr>
<td>4-Aminobenzoate (pABA)</td>
<td>H</td>
<td>H</td>
<td>NH&lt;sub&gt;2&lt;/sub&gt;</td>
<td>H</td>
<td>91.1 ± 1.4</td>
</tr>
<tr>
<td>2,4-Dihydroxybenzoate</td>
<td>OH</td>
<td>H</td>
<td>OH</td>
<td>H</td>
<td>0.5 ± 0.9</td>
</tr>
<tr>
<td>3,4-Dihydroxybenzoate</td>
<td>H</td>
<td>OH</td>
<td>OH</td>
<td>H</td>
<td>23.2 ± 0.4</td>
</tr>
<tr>
<td>2,3-Dihydroxybenzoate</td>
<td>OH</td>
<td>OH</td>
<td>H</td>
<td>H</td>
<td>0.7 ± 0.2</td>
</tr>
<tr>
<td>4-Amino-2-hydroxybenzoate</td>
<td>OH</td>
<td>H</td>
<td>NH&lt;sub&gt;2&lt;/sub&gt;</td>
<td>H</td>
<td>1.1 ± 0.3</td>
</tr>
<tr>
<td>4-Amino-3-hydroxybenzoate</td>
<td>H</td>
<td>OH</td>
<td>NH&lt;sub&gt;2&lt;/sub&gt;</td>
<td>H</td>
<td>30.4 ± 7.2</td>
</tr>
<tr>
<td>Vanillate</td>
<td>H</td>
<td>OCH&lt;sub&gt;3&lt;/sub&gt;</td>
<td>OH</td>
<td>H</td>
<td>95.2 ± 2.4</td>
</tr>
<tr>
<td>Syringate</td>
<td>H</td>
<td>OCH&lt;sub&gt;3&lt;/sub&gt;</td>
<td>OH</td>
<td>OCH&lt;sub&gt;3&lt;/sub&gt;</td>
<td>10.6 ± 0.7</td>
</tr>
<tr>
<td>Gallate</td>
<td>H</td>
<td>OH</td>
<td>OH</td>
<td>OH</td>
<td>1.6 ± 0.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>4-mono substituted benzoates</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>4-Hydroxybenzoate</td>
<td>H</td>
<td>H</td>
<td>OH</td>
<td>H</td>
<td>135.2 ± 8.6</td>
</tr>
<tr>
<td>4-Aminobenzoate (pABA)</td>
<td>H</td>
<td>H</td>
<td>NH&lt;sub&gt;2&lt;/sub&gt;</td>
<td>H</td>
<td>91.1 ± 1.4</td>
</tr>
<tr>
<td>4-Fluorobenzoate</td>
<td>H</td>
<td>H</td>
<td>F</td>
<td>H</td>
<td>84.5 ± 0.9</td>
</tr>
<tr>
<td>4-(Aminomethyl) benzoate</td>
<td>H</td>
<td>H</td>
<td>CH&lt;sub&gt;2&lt;/sub&gt;NH&lt;sub&gt;2&lt;/sub&gt;</td>
<td>H</td>
<td>2.8 ± 0.2</td>
</tr>
<tr>
<td>4-(Aminoethyl) benzoate</td>
<td>H</td>
<td>H</td>
<td>CH&lt;sub&gt;3&lt;/sub&gt;CH&lt;sub&gt;2&lt;/sub&gt;NH&lt;sub&gt;2&lt;/sub&gt;</td>
<td>H</td>
<td>10.4 ± 0.3</td>
</tr>
</tbody>
</table>

*Activities are the mean of three replicates ± 1 S.D. and are determined for 1 mM acyl substrate using the coupled adenylation assay. Activities below 10 nmol/min/mg are minimal. Some compounds are repeated from Table 3.1 for ease of comparison. Additional experiments yielded similar results.
of activity associated with substitution at the 2-position. PBS3 activity was 91.1 nmol/min/mg with pABA (4-ABA) compared with 30.4 nmol/min/mg with 4-A-3HBA and 1.11 nmol/min/mg with 4-A-2HBA. To further examine the effect of different substituents at the para position of benzoate, the adenylation activities of PBS3 on 4-HBA, 4-fluorobenzoate, pABA (4-ABA), 4-aminomethylbenzoate (4-AMeBA), and 4-aminoethylbenzoate (4-AEtBA) were compared (Table 3.2). The highest activity was observed on 4-HBA, with activity on pABA and 4-fluorobenzoate somewhat lower but still high, followed by a dramatic reduction in activity with the bulkier acyl substrates 4-AMeBA and 4-AEtBA.

**PBS3 Catalyzes the Formation of Amino Acid Conjugates**

Once potential acyl substrates were identified, a subset were tested for conjugation to amino acids by recombinant PBS3. Because PBS3 displayed the highest velocity on 4-HBA in the adenylation assay (Tables 3.1 and 3.2), reactions of 4-HBA with 20 amino acids were monitored by TLC. TLC spots were reproducibly detected for reactions of 4-HBA with Glu, His, and Lys with a faint spot visible for Met (Figure 3.3A). Very faint spots indicative of the amino acid conjugate were sometimes visible for reactions with Ile, Leu, Val, Thr, Ser, and Trp. A subset of these reaction mixtures were analyzed by HPLC to verify the presence of peaks indicative of the 4-HBA-amino acid conjugates when incubated with PBS3. Novel peaks indicative of 4-HBA-AA conjugate formation were observed by HPLC for reactions with Glu, His, Met, Thr, and Ser. Figure 3.3B shows the chromatogram for the reaction mixtures with 4-HBA and Glu. No 4-HBA-Glu was formed when Mg$^{2+}$ and ATP were excluded from the reaction mixture (not shown). To verify the identity of 4-HBA-Glu, which is not commercially available, the fraction corresponding to the putative 4-HBA-Glu conjugate was collected and analyzed by LC-MS with electrospray ionization (ESI) positive-ion analysis and by nanospray infusion Q-TOF MS with MS/MS. The mass spectrum of the putative 4-HBA-Glu conjugate fraction is consistent with it being the protonated 4-HBA-Glu conjugate (C$_{12}$H$_{14}$NO$_6$) with m/z=268.1. MS/MS analyses of the dominant ion peak (M+H) at 268.1 m/z result in major ion intensities at m/z= 251.0 consistent with loss of –OH and fission products m/z= 130 and 121 likely associated with cleavage of the amide bond (Figure 3.3B inset). Our mass spectra data are similar to those reported for other benzoyl-amino acid conjugates in which fragments consistent with cleavage of the amide bond were also observed (Suzuki et al., 1988; Bourne et al., 1991). In addition, 4-HBA-Glu was chemically synthesized and analyzed by $^1$H and $^{13}$C NMR in collaboration with the Sarpong lab at UC Berkeley (see (Okrent et al., 2009)). The chemically synthesized 4-HBA-Glu displayed identical HPLC retention time and UV spectral properties to the 4-HBA-Glu from enzymatic reactions catalyzed by PBS3. This further confirms that PBS3 catalyzes the formation of 4HBA-Glu in vitro.
Figure 3.3. PBS3 catalyzes the formation of amino acid conjugates with preferred acyl substrates. 
A, TLC analysis of 4-HBA amino acid conjugate formation by PBS3. 
B, HPLC analysis of 4-HBA-Glu formation by PBS3, inset: Verification of putative 4-HBA-Glu peak by LC-MS. Shown is Q-TOF MS/MS on m/z= 268.1 with 4-HBA-Glu structure. Likely fission site is indicated. 
C, TLC analysis of pABA-Glu formation by PBS3: pABA standard (left), reaction mixture (middle), and pABA-Glu standard (right). 
D, TLC analysis of the PBS3 reactions with vanillic acid (VA), 4-HBA, or trans-cinnamic acid (t-CA) as the acyl substrate and Glu. The arrow indicates a faint spot corresponding to t-CA-Glu. 
E, TLC of the PBS3 reaction with 4-HBA and Glu derivatives. AAD= L-2-amino adipic acid, Gla= γ-carboxy L-glutamic acid, MeGlu= L-gutamic acid γ-methyl ester.
Conjugation of pABA was also tested with the 20 amino acids and monitored by TLC. Unlike the results with 4-HBA, only a spot for pABA-Glu was detected either by UV (similarly to 4-HBA) or vanillin staining (Figure 3.3C). The Rf value of this putative pABA-Glu conjugate was similar to that of the commercially available standard. pABA-Glu formation by PBS3 was also followed by HPLC. The conversion of pABA to pABA-Glu occurred only with PBS3 in the reaction mixture. As PBS3 also formed additional amino acid conjugates with 4-HBA, I confirmed that similar amino acids conjugates were not formed with pABA by analyzing a subset of these reactions by HPLC. Putative pABA-His, pABA-Met, or pABA-Ser peaks were not detected by HPLC (data not shown).

Other compounds with high velocity (BA and vanillate), moderate velocity (trans-cinnamate) and minimal velocity (SA) were also tested as acyl substrates in the Glu conjugation reaction. TLC analyses showed that recombinant PBS3 catalyzed the conjugation of Glu to vanillate and trans-cinnamate (Figure 3.3D) but not to SA. As none of a variety of stains tested worked satisfactorily for BA, the BA reaction mixture was assessed by HPLC. The formation of the BA-Glu conjugate from BA by PBS3 was detected, with commercially available BA and BA-Glu used as standards; formation of BA-Glu required PBS3 (not shown). To further confirm that PBS3 was unable to form the SA-Glu conjugate, the SA reaction mixture was assessed by HPLC. Though SA was readily detected by fluorescence, no novel peaks indicative of SA-Glu were apparent. These results validate the use of the adenylation assay to study substrate specificity by confirming that those acyl substrates with high velocity in the adenylation assay are also capable of forming amino acid conjugates whereas those substrates with minimal activity using the adenylation assay (e.g. SA) are not.

4-HBA and pABA conjugates were formed with Glu (L-glutamic acid) but not Asp. In addition, with pABA as the acyl substrate, pABA-Glu was the only amino acid conjugate detected. Therefore, to further examine the specificity of PBS3 for Glu, the related compounds D-glutamic acid, L-2-amino-adipic acid, L-glutamic acid γ-methyl ester, and γ-carboxy L-glutamic acid were tested. L-2-amino-adipic acid contains an additional carbon in between the α-carbon and the γ-carboxy group, L-glutamic acid γ-methyl ester lacks a second carboxylic acid group, and γ-carboxy L-glutamic acid contains a third carboxylic acid group. PBS3 did not form 4-HBA (Figure 3.3E) or pABA conjugates (not shown) with these glutamic acid derivatives.

Point Mutations in the C-terminus of PBS3 Present in the pbs3-1 Mutant Result in Loss of Biochemical Activity

The pbs3-1 EMS mutant contains two amino acid substitutions in highly conserved amino acids at the C-terminus (E502K and I519T) (Nobuta et al., 2007). This portion of the PBS3 protein has no known motifs and is contained within the region of PBS3 (amino acids 335-575) bound by the P. syringae HopW1-1 effector (Lee et al., 2008). Glu-502 is predicted to be within an α-helix, whereas residue Ile-519 is predicted to reside in a loop region following the aforementioned α-helix (Figure 3.4). The high degree of conservation of Ile-519 in all Arabidopsis GH3 proteins (Ile in all but two GH3s which have a Val at this position) suggests it does not confer substrate specificity; however, it could potentially participate in catalysis. In order to ascertain the impact of these mutations on PBS3 activity, with the assistance of rotation student Mathew Brooks, I expressed and purified recombinant enzyme with either or both of the mutations and compared its activity with the wild type recombinant enzyme. The single and double mutations resulted in a complete loss of activity (< 0.5 nmol/min/mg) as measured by the
The substituted amino acids in the pbs3-1 mutant protein, E502K and I519T, are indicated with arrows above the secondary structure of PBS3 as predicted by PSIPRED (McGuffin et al., 2000). In an alignment of the corresponding residue performed using Megalign with the ClustalW method from the DNASTAR software package, identical residue are highlight in black and conserved residues in gray. The *Pseudomonas syringae* effector HopW1-1 binds to residues 335-575 of PBS3 (Lee et al., 2008). Abbreviations are Conf = confidence, Pred = prediction, C = coiled coil, H= alpha helix.
adenylation assay with 4-HBA, pABA, or BA as acyl substrates. Wild-type recombinant PBS3 exhibited activity similar to that found with previous enzyme preparations (see Table 3.2). The loss of biochemical activity associated with these mutations is consistent with the similar phenotypes observed for the recessive pbs3-1 EMS mutant and the null pbs3-2 T-DNA insertion mutant as in Chapter II and (Nobuta et al., 2007).

**Kinetic and Catalytic Properties of PBS3**

Similar to other characterized GH3 enzymes, the exclusion of Mg\(^{2+}\) or ATP from the reaction mixture resulted in loss of PBS3 activity (data not shown). To determine the pH requirements of the reaction, TLC analyses were performed on the reaction mixtures using either pABA or vanillate as the acyl substrate with Glu as the amino acid at pH 7.5, 8.0, and 8.5. Clearly visible products were observed for the reactions at pH 8.0 and 8.5, but not at pH 7.5 (data not shown).

To further characterize the PBS3 benzoyl amino acid synthetase reaction, its kinetic properties were determined using the three biologically relevant acyl substrates with the highest adenylation activity: 4-HBA, pABA, and BA. For these analyses, the apparent \(K_m\) for ATP were first determined using 4-HBA and pABA as acyl substrates to ensure sufficient ATP in the kinetic assays. With the coupled adenylation assay, PBS3 exhibited standard Michaelis-Menton kinetics with an apparent \(K_m\) for ATP of 791 \(\mu\)M (with 4-HBA) and 636 \(\mu\)M (with pABA). The apparent \(K_m\) of PBS3 was then determined for the acyl substrates 4-HBA, pABA, and BA at pH 8.5 with saturating Mg\(^{2+}\), ATP, and Glu. The concentration of the Glu conjugate produced in the reactions was quantified by HPLC over time, with the results summarized in Table 3.3. With all three acyl substrates, we found PBS3 exhibited standard Michaelis-Menten kinetics. The corresponding Hanes plots indicated that PBS3 had the highest affinity for pABA (\(K_m = 153\) \(\mu\)M), followed by 4-HBA (\(K_m = 459\) \(\mu\)M) and BA (\(K_m = 867\) \(\mu\)M). The \(k_{cat}\) values for pABA and BA were similar (2.74 min\(^{-1}\) and 3.96 min\(^{-1}\)), whereas that for 4-HBA was about 5-fold higher (20.36 min\(^{-1}\)). PBS3 exhibited similar catalytic efficiencies for 4-HBA and pABA (0.0444 and 0.0179 \(\mu\)M\(^{-1}\) min\(^{-1}\) respectively) with reduced catalytic efficiency with BA as the acyl substrate (0.0046 \(\mu\)M\(^{-1}\) min\(^{-1}\)). Nonlinear curve fitting of the kinetic data gave similar results with apparent \(K_m\) values of 460, 120, and 931 \(\mu\)M and calculated \(k_{cat}\) values of 20.83, 2.70, and 4.06 min\(^{-1}\) for 4-HBA, pABA, and BA respectively.

**Does PBS3 Impact SA Metabolism by Altering the Enzymatic Activity of SA Biosynthetic Enzymes?**

One potential manner by which PBS3 could impact SA metabolism is by altering the enzymatic activity of SA biosynthetic enzymes. As the pbs3 mutants appear to be compromised in SA biosynthesis, if our hypothesis is correct, we would expect that 1) PBS3 substrate(s) inhibit SA biosynthetic activity and/or 2) PBS3 product(s) activate SA biosynthetic enzymes. In *A. thaliana*, pathogen-induced SA biosynthesis requires the inducible isochorismate synthase AtICS1 (Wildermuth et al., 2001). AtICS1 catalyzes the reversible conversion of chorismate to isochorismate (Strawn et al., 2007). Isochorismate is likely then converted to SA by an isochorismate pyruvate lyase (IPL) with the concomitant release of pyruvate, as it is in bacteria that synthesize SA (e.g. (Serino et al., 1995; Gaille et al., 2002)). Therefore, we wanted to determine whether the PBS3 substrates 4-HBA or pABA or the PBS3 product pABA-Glu impacts the activity of known SA biosynthetic enzymes (AtICS1 and the *P. aeruginosa* isochorismate pyruvate lyase PchB). We used the *P. aeruginosa* PchB enzyme as a surrogate for a putative, yet-to-be identified, *Arabidopsis* IPL. All assessments were performed with substrate...
<table>
<thead>
<tr>
<th>Acyl Substrate</th>
<th>$k_{cat}$ (min$^{-1}$)</th>
<th>$K_m$ (µM)</th>
<th>$k_{cat}/K_m$</th>
</tr>
</thead>
<tbody>
<tr>
<td>4-HBA</td>
<td>20.36</td>
<td>459</td>
<td>0.0444</td>
</tr>
<tr>
<td>pABA</td>
<td>2.74</td>
<td>153</td>
<td>0.0179</td>
</tr>
<tr>
<td>BA</td>
<td>3.96</td>
<td>867</td>
<td>0.0046</td>
</tr>
</tbody>
</table>

Kinetic constants were determined by following acyl substrate-Glu formation by HPLC. Similar results were obtained for experiments with an independent enzyme preparation.
concentrations at the $K_m$ for the respective enzyme. A coupled spectrophotometric assay was employed to assess the conversion of chorismate to isochorismate by AtICS1 (Strawn et al., 2007) and found no statistically significant difference in activity with the addition of 200 µM 4-HBA, pABA, or pABA-Glu. This was also true when assessed with 1 mM 4-HBA, pABA, or pABA-Glu. Furthermore, the conversion of isochorismate to SA by PchB was also unaffected by these compounds.

**PBS3 Activity is Inhibited by SA, but Not by Other Phytohormones**

As PBS3 impacts SA metabolism and the most active acyl substrates of PBS3 are structurally similar to SA, we sought to determine whether SA could inhibit PBS3 activity. To test this, the effect of increasing concentrations of SA on PBS3-catalyzed pABA-Glu formation was assayed. As shown in Figure 3.5, SA dramatically reduced the amino acid conjugation activity of PBS3 with an estimated IC$_{50}$ of 15 µM. To determine whether the inhibition by SA was competitive, excess (saturating) pABA was added to the reaction mixture with 15 µM SA (the IC$_{50}$). Full activity with maximal velocity was restored (data not shown). This suggests that SA acts as a competitive inhibitor of PBS3 activity.

I then sought to determine whether 3-HBA or phytohormones, including MeSA (the mobile ester of SA) and the functional analog INA, could modulate PBS3 activity. The high throughput adenylation assay was used for this. With this assay, 30 µM SA inhibited PBS3 activity by 33% and 300 µM SA inhibited it by 75%. Although this observed inhibition was of reduced magnitude compared with that found using the full HPLC assay, we verified that the inhibition was due to the impact of SA on PBS3 activity by confirming that SA had no effect on activity of the coupling enzymes with PP$i$ as the substrate. Of the six compounds tested, only SA and the SA functional analog INA dramatically inhibited PBS3 activity (Table 3.4). IAA, JA, MeSA, and 3-HBA did not substantially inhibit PBS3 activity.

**DISCUSSION**

**Insights into PBS3 Active Site and Reaction Mechanism**

The GH3 enzymes differ from other members of the acyl-adenylate/thioester-forming superfamily in that amino acid conjugation occurs directly to the acyl-adenylate without the requirement for a thiol ester intermediate. Adenylation reactions, such as those catalyzed by this superfamily, require Mg$^{2+}$ and ATP to complex with the phosphoester oxygens and enhance the electrophilicity of the phosphorus atoms. As expected, we found PBS3 activity requires Mg$^{2+}$ and ATP. We also found alkaline pH is necessary for amino acid conjugation, but not the adenylation reaction, probably to deprotonate the attacking nucleophile, the $\alpha$-NH$_3^+$ group of the amino acid substrate (pKa~9.5). This is consistent with the pH requirements of other characterized GH3 enzymes (e.g. (Staswick et al., 2005)), but is not a hallmark of all superfamily members. The reactions catalyzed by other superfamily family members, such as CoA ligases, do not use an amino acid as the nucleophile, and thus an alkaline pH would not be required as it is for PBS3.

We are limited in our analysis of PBS3 by the lack of a crystal structure for a GH3 enzyme and by the limited sequence conservation with other acyl-adenylate/thioester-forming superfamily members. However, the three AMP-binding motifs of PBS3 are evident, and the importance of these motifs has been clearly demonstrated for GH3 enzymes. For example, in JAR1, mutation of a conserved residue of AMP-binding motif I results in a complete loss of
Figure 3.5. Inhibition of PBS3 activity by SA. SA inhibits the formation of pABA-Glu catalyzed by PBS3. The reaction velocities were determined by monitoring the formation of pABA-Glu by HPLC of reactions with 150 μM pABA and 10 mM Glu. The experiment was repeated with similar results.
<table>
<thead>
<tr>
<th>Compound</th>
<th>Fractional inhibition* (30 μM)</th>
<th>Fractional inhibition (300 μM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SA</td>
<td>0.33</td>
<td>0.75</td>
</tr>
<tr>
<td>MeSA</td>
<td>0.02</td>
<td>0.00</td>
</tr>
<tr>
<td>INA</td>
<td>0.09</td>
<td>0.70</td>
</tr>
<tr>
<td>IAA</td>
<td>0.04</td>
<td>0.19</td>
</tr>
<tr>
<td>JA</td>
<td>0.02</td>
<td>0.10</td>
</tr>
<tr>
<td>3-HBA</td>
<td>0.04</td>
<td>0.22</td>
</tr>
</tbody>
</table>

* Fractional inhibition compares activity with additional compound to control reaction. Inhibition assessed using the coupled adenylase assay with 150 μM pABA as the acyl substrate and 30 or 300 μM of the compound being tested. Repeat experiments with a separate enzyme batch gave similar results.
activity (Staswick and Tiryaki, 2004). To date, no other known motifs have been identified in PBS3 or other GH3 enzymes. However, our study of the impact of the point mutations contained in pbs3-1 on PBS3 activity, with loss of adenylation activity for both E502K and I519T PBS3 mutants, suggests that the highly conserved C-terminal region of PBS3 including loop residue Ile-519 (Figure 3.4), may participate in substrate binding or catalysis.

In terms of acyl substrate specificity, PBS3 had the highest velocity with 4-HBA and pABA (4-ABA) of the compounds tested (Table 3.2). However, bulkier groups at the para-position, such as 4-AMeBA or 4-AEtBA, were not well tolerated. Substitution at the meta- and ortho-positions led to substantially lower reaction velocity with ortho substitutions resulting in lack of activity, perhaps due to steric hindrance. Meta substitutions were tolerated and more electronegative substituents such as MeO provided a 4-fold enhancement of activity compared with –OH substitutions. In addition, our results comparing PBS3 activity with benzoate versus phenylacetate suggest that it is important for the carboxylate group to be attached directly to the aromatic ring, likely for resonance stabilization. The fact that PBS3 exhibited moderate activity with trans-cinnamate, which has a double bond between the carboxylic acid and aromatic ring, supports this conclusion. However, another explanation for the lack of activity with phenylacetate is that although the cinnamates and benzoates display a planar geometry, phenylacetate does not, and thus may be perceived quite differently by the active site.

The observed substrate preference of PBS3 for subsets of benzoates is consistent with glycosyltransferases and CoA ligases active on benzoate substrates (Lim et al., 2002). Despite the clear preference for benzoates, PBS3 was also moderately active on trans-cinnamate (Table 3.2). This is unusual, as glycosyltransferases and CoA ligases active on benzoates have not been reported to have significant activity on cinnamates (Beuerle and Pichersky, 2002; Song, 2006). Similar to PBS3, the few characterized GH3 proteins have exhibited clear substrate preferences while remaining active on a fairly broad range of substrates. For example, Arabidopsis Group II GH3 proteins are capable not only of adenylylating indole-3-acetate (IAA) in vitro, but also adenylylating other compounds with auxin activity, including indoles such as indole-3-pyruvate and indole-3-butyrate and the structurally distinct phenylacetate and naphthaleneacetate (Staswick et al., 2005). We also observed that SA inhibits both the adenylation and full (adenylation and amino acid conjugation) PBS3 reactions. Furthermore, the restoration of maximal PBS3 velocity when excess pABA is added to the reaction mixture with 15 µM SA suggests SA acts as a competitive inhibitor of PBS3. In this capacity, SA most likely competes with the acyl substrate (e.g. pABA) for the active site in a mutually exclusive fashion. Although less frequently observed, it is also possible that SA could act as an allosteric competitive inhibitor, binding a distinct site on PBS3.

Although PBS3 tolerates a range of benzoates as substrates, there is a strict requirement for Glu as the amino acid substrate with pABA. No activity was observed with the other 19 amino acids, including Asp which differs from Glu by only one methylene group. In addition, modified forms of Glu were also inactive. This suggests a precise configuration of the active site for L-glutamic acid. The strict requirement for Glu is unusual among the few characterized GH3 enzymes, which are capable of conjugating the addition of multiple amino acids to the acyl substrate in vitro. For example, JAR1 catalyzes the addition of jasmonate to at least 11 amino acids and the ethylene precursor 1-aminocyclopropane-1-carboxylate (Staswick and Tiryaki, 2004). JAR1 does not, however, appear to form Glu or Asp conjugates. The Arabidopsis Group II GH3 proteins also form multiple IAA-amino acid conjugates including Glu and Asp conjugates in vitro (Staswick et al., 2005). Some of these IAA-conjugating proteins, such as
AtGH3.17, do exhibit a strong preference for either Glu or Asp (Staswick et al., 2005). Surprisingly, with 4-HBA as the acyl substrate, PBS3 can utilize a number of amino acids including Glu but not Asp (Figure 3.3). What accounts for this difference in amino acid specificity? PBS3 exhibits ~3-fold higher affinity for pABA than for 4-HBA (Table 3.3). Perhaps the pABA-AMP intermediate more precisely fits the active site, constraining acceptable amino acids for use in the conjugation reaction. In terms of in planta activity, this would depend upon the relative availability of acyl and amino acid substrates.

**Kinetic Parameters of PBS3 and Physiological Relevance of Substrates**

We found that PBS3 had the highest affinity for pABA ($K_m=153 \mu M$), compared with 4-HBA (459 $\mu M$) and BA (867 $\mu M$) and a 10-fold higher catalytic efficiency with 4-HBA compared with BA (Table 3.3). One $K_m$ value has been previously reported for a GH3 family member: that of JAR1 for (-)-jasmonic acid (Suza and Staswick, 2008). This value, 190 $\mu M$, is similar to that found here for PBS3. However, in order to assess the relevance of our findings, the physiological concentrations of the substrates of PBS3 must be considered. How do the $K_m$ values we have determined for PBS3 compare to physiological concentrations of these substrates?

In Arabidopsis, pathogen-induced BA levels have been rarely reported, although BA levels of ~4 nmol/g FW assessed using GC-MS, have been reported 2 days post-inoculation with a virulent P. syringae strain (Schmelz et al., 2003). As pathogen-induced BA levels have not been measured with sufficient temporal or spatial resolution to be confident in physiological concentrations, we also reviewed reports of the affinity of pathogen-inducible Arabidopsis enzymes for BA to provide insight into relevant physiological concentrations. The Arabidopsis glucosyltransferases with the highest activity and affinity for BA have reported apparent $K_m$ values for BA of 260 $\mu M$ (Lim et al., 2002) and 540 $\mu M$ (Song, 2006) for UGT74F1, and 80 $\mu M$ (Lim et al., 2002) for UGT74F2. The apparent $K_m$ of the Arabidopsis thaliana SA/BA methyltransferase AtBSMT1 for BA is reported to be 65 $\mu M$ (Chen et al., 2003). These affinities suggest that physiological concentrations of BA are insufficient to support efficient catalysis by PBS3 using BA as the acyl substrate. Therefore, given the low affinity ($K_m= 867 \mu M$) and catalytic efficiency of PBS3 with BA, it appears that BA may not be the favored substrate of PBS3 in planta.

pABA-Glu has been detected in Arabidopsis leaves at 0.3 nmol/g FW and is thought to be present as a breakdown product of folates (Orsomando et al., 2006). Folates consist of three subunits: a pterin moiety, pABA, and varying numbers of glutamic acids. They are essential cofactors in one-carbon transfer reactions, which are central to plant metabolism, and are involved in the synthesis of methionine, pantothenate, purines, pyrimidines, and thymidylate (Hanson and Roje, 2001). It may be that pABA-Glu is not just a breakdown product of folate, but is actually synthesized de novo in response to biotic stress. While an increase in pABA or pABA-Glu has not been detected in response to pathogen attack in Arabidopsis, pABA-Glu would not be present in samples from analytical procedures typically used to measure metabolites like SA (Zhang et al., 2005; Quinlivan et al., 2006). The apparent $K_m$ of the pABA-glucosyltransferase for pABA was found to be 120 $\mu M$ (Eudes et al., 2008), which is quite similar to the value we found for PBS3. Perhaps the formation of pABA-Glu conjugates is a way for plants to restrict access of the pathogen to free pABA. Bacteria can commonly import pABA but few have been found that import pABA-Glu (e.g. (Hussein et al., 1998)). Furthermore, there is precedent for pABA limitations on the growth of phytopathogens, as
observed for growth of a folate auxotroph of *Ralstonia solanacearum* on tobacco (Shinohara et al., 2005). Therefore, pABA remains an attractive candidate as a potential substrate for PBS3 in planta.

4-HBA is typically known as an intermediate in the synthesis of ubiquinone, a lipid-soluble electron carrier of the mitochondrial electron transport chain (Swiezewska, 2004). Therefore, all plants normally produce 4-HBA in small quantities. However, some plants use 4-HBA in the synthesis of specialized metabolites such as shikonin (Yazaki et al., 2002), and 4-HBA is often associated with the plant cell wall, typically as a component of lignin (Lu et al., 2004). Pathogen or elicitor-induced 4-HBA and incorporation into the plant cell wall has been reported for a number of plants, including *Arabidopsis* (Tan et al., 2004) and carrot (Veit et al., 2001). For example, in *Arabidopsis* in response to virulent and avirulent *P. syringae pv. tomato* DC3000, 4-HBA accumulated to 0.2 nmol/mg cell wall by 72 hours post-infection (Tan et al., 2004) or ~5 nmol/g FW. This concentration was calculated using a conversion factor for milligrams of cell wall to g FW I obtained through cell wall extractions performed using this published cell wall preparation protocol. Enzymes that use 4-HBA as a substrate include five putative *Arabidopsis* 4-HBA glucosyltransferases, which exhibit $K_m$ values of 210-290 µM for 4-HBA (Lim et al., 2002) and *Lithospermum erythrorhizon* geranyl-diphosphate, 4-HBA geranyltransferases involved in shikonin synthesis, with reported $K_m$ values for 4-HBA ranging from 18.4-45.9 µM (Yazaki et al., 2002). Therefore, 4-HBA remains an attractive candidate for the *in planta* substrate for PBS3, with 4-HBA or 4-HBA-Glu perhaps playing a role in pathogen-induced cell wall remodeling.

As the analytical methods required to extract, enrich, and detect these putative conjugates are highly specific, our *in vitro* analysis now allows us to employ methods specific to these compounds and to the analysis of compounds incorporated into the cell wall. To-date, *in planta* activities of GH3 enzymes appear consistent with *in vitro* findings. For example, IAA-Asp is one of the dominant IAA-amino acid conjugates formed by GH3.6 as evaluated by TLC, and overexpression of GH3.6 results in a 4.5-fold increase in IAA-Asp in *Arabidopsis* leaves (Staswick et al., 2005). Similarly, *jar1* mutants exhibit >7-fold reduction in JA-Ile levels and JA-Ile is a dominant JA-amino acid conjugate observed by TLC (Staswick and Tiryaki, 2004). Consistent with the preference of PBS3 for Glu, the only endogenous benzoyl-amino acid conjugates detected *in planta* have been either Glu or Asp conjugates (e.g. pABA-Glu (Orsomando et al., 2006), 4-HBA-Glu (Trennheuser et al., 1994), SA-Asp (Steffan et al., 1988), and BA-Asp (Suzuki et al., 1988). As PBS3 is not active with Asp, it will be interesting to determine whether another *Arabidopsis* GH3 family member exhibits overlapping acyl substrate specificity but altered AA preference, as is the case for the *Arabidopsis* GH3 family members acting on IAA (Staswick et al., 2005). This possibility may also require the use of a PBS3 overexpresser to detect significantly altered levels of putative Glu conjugates.

**How Does PBS3 Function Impact SA Metabolism and SA-dependent Gene Expression?**

Previous analyses of *pbs3/gdg1/win3* mutants by us and others showed that the lack of function of AtGH3.12 (PBS3) reduces total pathogen-induced SA accumulation, activation of SA-dependent gene expression, and disease resistance (Chapter II and (Jagadeeswaran et al., 2007; Lee et al., 2007; Nobuta et al., 2007)). Furthermore, exogenous application of SA to *gh3.12* mutants induces the expression of SA-dependent genes such as *PRI1* and rescues the enhanced disease susceptibility phenotype (Jagadeeswaran et al., 2007; Nobuta et al., 2007). Therefore, the most likely explanation for the impact of PBS3 on SA accumulation and SA-
dependent responses is that SA synthesis is delayed (and reduced) in response to *P. syringae* pathogens. The action of PBS3 upstream of SA rather than on SA itself is supported by the observed substrate specificity of PBS3, with PBS3 inactive on SA, and the inhibition of PBS3 activity by SA (IC$_{50}$ = 15 µM SA). Of interest, AtBSMT1, which catalyzes the formation of MeSA from SA, functions at these low concentrations of SA (K$_{m}$ = 16 µM for SA) (Chen et al., 2003). In contrast, the K$_{m}$ value of the pathogen-induced *Arabidopsis* SA glycosyltransferase 1 (UGT74F1; At2g4380), which converts free SA to SAG, is 90-230 µM for SA (Lim et al., 2002), (Song, 2006), suggesting higher concentrations of SA are required for SAG formation. 

In situ quantification of free SA has not been reported in *Arabidopsis*. In tobacco, tobacco mosaic virus inoculation resulted in free SA accumulation of 6.5 µM at 16 hpi with maximal detected local concentrations of >200 µM at 40 hpi (Huang et al., 2006). Therefore, it is likely that PBS3 functions upstream of SA, early in the defense response when SA levels are low. In this scenario, PBS3 activity leads to conjugate formation (e.g. 4HBA-Glu), which signals or primes SA biosynthesis. PBS3 activity is then no longer required once SA synthesis has been sufficiently initiated. The inhibition of PBS3 by SA serves as a mechanism to rapidly reduce synthesis of the conjugate when it is no longer necessary for SA priming. In this scenario, we would expect a reduced conjugate (e.g. 4-HBA-Glu) concentration as the defense response progresses. A role for PBS3 upstream of SA in a regulatory capacity is also supported by recent expression profiling comparing expression of wild type and a variety of *Arabidopsis* mutants in defense signaling at 24 hpi with *P. syringae pv. maculicola* ES4326 using a custom mini array (Wang et al., 2008). In this study, PBS3 was placed upstream of SA and NPR1, as the pbs3 mutation impacted the expression of many more genes than the standard SA biosynthesis and signaling mutants ics1 (sid2), eds5, and npr1 (Wang et al., 2008). Of interest, the expression of chalcone synthase, a key enzyme directing flux to flavonoid biosynthesis from 4-coumaroyl-CoA, was strongly down-regulated at 24 hpi with *P. syringae pv. maculicola* in wild type and ics1 but was up-regulated in the pbs3 mutant (Wang et al., 2008). In other systems such as elicited carrot cell cultures, up-regulation of 4-HBA synthesis and enhanced flux through the general phenylpropanoid pathway was associated with down-regulation of flavonoid biosynthesis from 4-coumaroyl-CoA (Gleitz et al., 1991). This suggests that flux from 4-coumaroyl-CoA to phenylpropanoid cell wall precursors may be up-regulated to facilitate the observed cell wall remodeling. Therefore, it is possible that PBS3 coordinately modulates pathogen-induced SA, 4-HBA, and their associated responses with 4-HBA playing a role in cell wall remodeling and the direction of chorismate flux (from 4-coumaroyl-CoA) for this purpose. Future work will address this possibility and potential mechanisms of action. For example, given the known function of phytohormone-AA conjugates in promoting or inhibiting the ubiquitin ligase-dependent proteasomic degradation of the associated transcriptional repressor, it is tempting to speculate that 4-HBA-Glu may act in a similar manner. With knowledge of the *in vitro* specificity of PBS3 gained through this study, potential PBS3 products could be used in assays with candidate F-box proteins and repressors to identify F-box-repressor interactions mediating SA and/or 4-HBA-associated signaling (similar to (Thines et al., 2007) for JA). This is critical as these interactions require binding of the appropriate molecule, e.g. JA-Ile for JA signaling.
CONCLUSION

This study further supports the critical role of amino acid conjugation of small molecules in phytohormone regulation and homeostasis. To date, this conjugation is catalyzed by members of the GH3 enzyme family. Here I characterize the in vitro activity of the first GH3 subfamily III member to be analyzed, PBS3 (AtGH3.12). In the process, I developed a high throughput adenylation assay to screen for preferred acyl substrates of GH3 enzymes. This high throughput assay can be applied to the identification of substrates for other GH3 enzymes for which no acyl substrates have been identified or inferred through analysis of mutant phenotypes. I also determine kinetic parameters for recombinant PBS3 and demonstrate that PBS3 is specifically inhibited by low concentrations of SA. This is important as it presents a novel mechanism for rapidly and reversibly fine-tuning the activity of small molecules, including phytohormones. It could also be a mechanism for rapid and reversible crosstalk between phytohormones or other small molecules, such as SA and 4-HBA. These in vitro analyses of PBS3 suggest a potential, underexplored role for 4-HBA and/or pABA and their respective amino acid conjugates in plant/microbe interactions and provide a framework to explore the mechanism by which PBS3 impacts pathogen-induced SA accumulation and disease resistance. In addition, using knowledge of the substrate preferences and biochemical and kinetic properties of PBS3, one could test and optimize PBS3 for use in biotechnological applications. For example, PBS3 might be used to remove and purify 4-HBA, an inhibitory by-product of cellulosic ethanol production (Palmqvist et al., 1999; Palmqvist and Hahn-Hagerdal, 2000), for resale to the crystal polymer industry (Figuly, 1996).

REFERENCES


major regulators and the phytotoxin coronatine. Mol Plant Microbe Interact 21, 1408-1420.


Chapter IV: Impact of pbs3 on Metabolism of Small Molecules in Arabidopsis

SUMMARY

Past biochemical experiments revealed that 4-substituted benzoates serve as substrates of PBS3\(^1\) *in vitro*. Therefore, a series of experiment were undertaken to determine whether these PBS3 products could be detected *in planta*. Although these products were not detected, there was evidence of the impact of PBS3 in specialized metabolism and flux through the chorismate pathway. Furthermore, a novel compound was detected that was elevated in the pbs3 mutant in response to pathogen infection, which could be the *in planta* substrate of PBS3 or a novel signaling molecule involved in small molecule crosstalk.

---

\(^1\) Abbreviations. Compounds: BA, benzoate; 4-HBA, 4-hydroxybenzoate; CA, chorismate; IAA, indole-3-acetate; IC, isochorismate; JA, jasmonate; pABA, *para*-aminobenzoate; SA, salicylate; MeSA, methyl salicylate; SAG, SA glucoside; VA, vanillate. Enzymes: ICS1 isochorismate synthase 1; IPL, isochorismate pyruvate lyase; PBS3, avrPphB susceptible 3 (GH3.12). Organisms: *Pseudomonas syringae* pv. *tomato* (*Pst*), *Pseudomonas syringae* pv. *maculicola* (*Psm*). Terms: hpi, hours post infection; SPE, solid phase extraction.
INTRODUCTION

_Arabidopsis_ plants with mutations in the PBS3 gene demonstrate an enhanced disease susceptibility phenotype. Like many mutants with this phenotype (Glazebrook et al., 1996; Rogers and Ausubel, 1997; Dewdney et al., 2000), they are compromised in the accumulation of pathogen-induced salicylic acid-glycoside (SAG), considered to be an inactive storage form of SA. SAG accumulates in the vacuole (Dean and Mills, 2004; Dean et al., 2005) and is hypothesized to function as an inactive pool for rapid and sustained induction of systemic acquired resistance through hydrolysis to the active free SA. Intriguingly, levels of free SA are not necessarily reduced in the _pbs3_ mutants. Free SA is sometimes elevated in the mutants compared to wild-type following infection with _Pseudomonas syringae_ (Ps), as I found 24 hours post infection (hpi) with _P. syringae pv. tomato avrRpt2_ and other researchers found 24 hpi with _P. syringae pv. maculicola avrRpm1_ (Lee et al., 2007). However, elevated free SA is not always observed following infection with _P. syringae_. Free SA is reduced compared to wild-type following infection with other combinations of strains and time points (Jagadeeswaran et al., 2007; Lee et al., 2007). Even when SA levels are elevated in _pbs3_-2, the SA-dependent marker gene _PRI_ is not induced in _pbs3_ mutants in response to _Pst avrRpt2_ until 48 h, an approximate 36 h delay compared with wild-type. It appears that the presence of free SA is not sufficient to result in _PRI_ induction, but instead that total levels must reach a certain threshold. When free SA levels are elevated, it is not enough to compensate for reduced SAG. In _pbs3_ mutants, total levels of SA are reduced 2 to 5-fold in _pbs3_ mutants infected by pathogens compared to wild-type plants. When total leaf SA levels are less than 3 μg/g fresh weight, as reported in Chapter II for the _pbs3_ mutants, _PRI_ induction is not typically observed (Strawn et al., 2007). In addition, when transgenic plant lines overexpressing bacterial SA biosynthetic genes exhibit constitutive expression of total SA of ≤3 μg/g fresh weight, _PRI_ is not significantly induced (Verberne et al., 2000; Mauch et al., 2001). Both SAG accumulation and _PRI_ expression can be restored to normal levels in the mutant by application of SA. Presumably, these high levels of SA surpass the SA threshold to restore induced _PRI_ expression (Chapter II). This suggests that the phenotype in _PBS3_ is due to altered accumulation of SA rather than processing of the signal. My results from the biochemical characterization of the PBS3 enzyme support placement of PBS3 upstream of SA biosynthesis. SA inhibits activity of PBS3 (discussed in Chapter III), indicating that PBS3 could participate in a feedback loop in SA biosynthesis. Evidence for a role of PBS3 upstream of SA in a regulatory capacity is also supported by expression profiling of wild-type _Arabidopsis_ and a variety of mutants in defense signaling following infection with _Psm_ ES4326 (Wang et al., 2008). The _pbs3_ mutation influenced the expression of more genes than the known SA biosynthetic and signaling mutants. That study thus supports placement of PBS3 upstream of SA biosynthesis.

We had originally hypothesized that PBS3 might conjugate amino acids to SA directly in a step necessary for SAG accumulation. However, biochemical assays discussed in Chapter III indicated that SA is not a substrate of PBS3 _in vitro_. Instead, PBS3 catalyzes the conjugation of amino acids, particularly glutamate, to 4-substituted benzoates. Two of these 4-substituted benzoates, pABA and 4-HBA, are known to be present in _Arabidopsis_. All plants contain these compounds in small quantities, as they are the precursors of essential metabolites; pABA is a precursor of the co-factor folate (Ravanel et al., 2001) while 4-HBA is a precursor of the electron transport chain carrier ubiquinone (Okada et al., 2004) (Figure 4.1). Kinetic characterization is consistent with identification of these compounds as potential substrates of the reaction catalyzed...
by PBS3. The amino acid conjugates that are products of the reaction are anticipated to be present transiently and in low amounts. While pABA-Glu has been previously found in *Arabidopsis* (Orsomando et al., 2006), it is thought to be a breakdown product of folate rather than synthesized *de novo* in the plant. pABA-Glu has not been found to be induced by infection with pathogens or treatment with elicitors. 4HBA-Glu has not been identified in *Arabidopsis*, although 4HBA has been shown to accumulate in the cell wall in response to pathogens (Hagemeier et al., 2001; Tan et al., 2004) and is known to covalently modify lignins (Lu et al., 2004).

As PBS3 influences accumulation of SA, it could also potentially alter the accumulation of other metabolites through flux through a common pathway. SA, like many other primary and secondary metabolites, is synthesized through the chorismate pathway. (Figure 4.1) Other compounds deriving from the chorismate pathway include compounds identified as substrates of the PBS3 enzyme *in vitro*, including the benzoates BA, 4-HBA, and pABA (see Chapter III). Aromatic amino acids, indoles, lignin, flavonoids and anthocyanins are also synthesized through this pathway. Expression data in *pbs3* and wild-type *Arabidopsis* in response to *P. syringae* pv *maculicola* ES4326 suggest possible alteration in flux through the chorismate pathway. For example, chalcone synthase, a key enzyme in flavonoid biosynthesis, was down-regulated 24 hpi in WT and *ics1*, but up-regulated in *pbs3* (Wang et al., 2008). A tension between the general phenylpropanoid pathway, including 4-HBA, and the flavonoid pathway, has previously been noted (Figure 4.2). For example, up-regulation of 4-HBA was found to correlate with down-regulation of flavonoid biosynthesis through 4-coumaroyl-CoA in elicited carrot cell cultures (Gleitz et al., 1991). If PBS3 is active on 4-HBA *in planta*, PBS3 could coordinately modulate pathogen-induced SA, 4-HBA, and their associated responses through flux through the chorismate pathway. If PBS3 were required for transport of 4-HBA to the cell wall as a 4-HBA-Glu conjugate, the accumulation of 4-HBA should be reduced in *pbs3* cell walls.

In order to investigate the impact of PBS3 on small molecules metabolism, extracts of soluble and cell wall-bound metabolites from elicited leaf samples were analyzed. The *pbs3* mutants and wild-type plants were treated with elicitors such as *Pseudomonas syringae* pv. *tomato* and UV-C radiation combined with application of potential benzoate substrates.

**MATERIALS AND METHODS**

**General Materials and Methods**

All chemicals are from Sigma-Aldrich unless otherwise indicated. Solvents are of HPLC purity from EMD and filtered prior to use.

**Growth Conditions of Bacteria and Plants**

*Pseudomonas syringae* strains *Pst* DC3000 and *Pst* carrying the avirulence gene *avrRpt2* were cultured in King’s B medium (10 mg/mL protease peptone; 15 mg/mL glycerol; 1.5 mg/mL K$_2$HPO$_4$; 4 mM MgSO$_4$, pH 7.0) supplemented with 25 µM kanamycin and 50 µM rifampicin (*Pst avrRpt2*) or kanamycin alone (virulent *Pst*) and grown at 28°C. Overnight cultures were grown to OD$_{600}$ 0.5-0.9. *Arabidopsis thaliana* ecotype Col-0 was grown in Scott’s Metro-Mix 200 with a 12-h photoperiod at a photosynthetically activated radiation of ~100 m$^{-2}$ s$^{-1}$ in growth rooms under a 12 h light/dark cycle at 23°C. Isolation and characterization of the *pbs3* mutants are described in (Nobuta et al., 2007). The PBS3 overexpressor lines (35S::PBS3) consist of
Figure 4.1. The central position of chorismate in aromatic acid biosynthesis in plants. Boxes indicate important metabolites synthesized from chorismate. Atoms derived from the reactions of chorismate are labeled in red. Abbreviations are: CM = Chorismate mutase; CPL = Chorismate pyruvate lyase; ADC = 4-Amino-4-deoxychorismate; AS = Anthranilate synthase; 4-HBA = 4-Hydroxybenzoic acid; pABA = para-Aminobenzoic acid. Adapted from (Strawn, 2010) Four enzymes have been shown to convert CM in plants. 4-HBA is synthesized via Phe in 7 steps rather than directly from chorismate by a CPL.
Figure 4.2. The flavonoid and general phenylpropanoid pathways diverge from 4-coumaroyl-CoA. Up-regulation of the general phenylpropanoid pathway can be accompanied by down-regulation of the flavonoid pathway upon treatment with pathogen elicitors (Gleitz et al., 1991). Shown are representative compounds that have been identified in *Arabidopsis*. R₁ and R₂ in kaempferol and cyanidin represent sugar moieties.
PBS3 under the control of the 35S promoter and were provided by Dr. Ramesh Raina of Syracuse University.

Infection of Plants with Pseudomonas syringae

Leaves of the Arabidopsis wild-type and mutant lines were infected at 4.5 weeks with OD$_{600}$ 0.0001 virulent Pst DC3000 or Pst DC3000 containing the avirulence gene $avrRpt2$ for syringe inoculation or OD$_{600}$ 0.002 for spray infection. Overnight bacterial cultures were pelleted, resuspended in sterile 10 mM MgSO$_4$ for syringe infiltration or dH$_2$O for spray infection and diluted to the appropriate concentration. For the syringe infiltration, three mature, fully expanded rosette leaves per plant were infiltrated with either OD$_{600}$ 0.0001 Pst in 10 mM MgSO$_4$ or 10 mM MgSO$_4$ as a negative control using a needleless syringe. Leaves were collected at 1 dpi, frozen in liquid nitrogen, and stored at -80°C. For the spray infiltration, Silwet-60 was added to a concentration of 0.02% and rosette leaves were sprayed uniformly to confluence with a manual pump spray bottle containing either the bacterial suspension or 0.02% Silwet-60 in dH$_2$O as a negative control. Leaves were collected at 3 dpi, frozen in liquid nitrogen, and stored at -80°C.

Treatment of Plants with UV-C and Benzoates

Plants were sprayed with 1 mM pABA or 4-HBA in 0.02% Silwet using a standard spray bottle and irradiated with UV-C (254 nm) light at 5 kJ/m$^2$ in a Stratagene UV Stratalinker 1800 thirty minutes after spraying. Leaves were collected after 3 or 12 h, frozen in liquid nitrogen, and stored at -80°C.

Measurement of Cell Wall-Bound Metabolites in Arabidopsis Leaves

Extraction of cell wall-bound metabolites was performed as described in (Tan et al., 2004) with a few modifications. Frozen tissue (2.5 g) was homogenized in 90% MeOH (20 mL), shaken at RT for 15 min and centrifuged 15 min at 4000 g. The supernatant was removed, and the residue was washed sequentially with: MeOH, water, 0.5% SDS, 1 M NaCl, water, MeOH, acetone and n-hexane (2 x 8 mL, shaken, and centrifuged). The white crystalline residues were evaporated using a dry vacuum at approximately 5 Torr and stored at -20°C. Cell wall samples (~25mg/sample) were transferred to serum bottles, suspended under N$_2$ in 3 ml 1 M NaOH, and incubated in darkness for 24 h at 80°C. Mixtures were acidified to ~ pH 3 with concentrated HCl and extracted 2 times with ethyl acetate. Organic layers were combined and dried as above. Residues were redissolved in 160 μL 20% MeOH for HPLC analysis.

Measurement of Soluble Metabolites in Arabidopsis Leaves

Leaves were treated with Pst, UV-C and/or benzoates as described above. Frozen leaf samples (approximately 0.5 g) were ground to a powder in a prechilled mortar and pestle using liquid nitrogen. The ground leaf material was transferred to a glass tube and suspended in 3 mL of 90% MeOH. Samples were vortexed, sonicated in a water bath sonicator for 20 min, and centrifuged at 5,000 rpm for 15 min at 4°C. The supernatant was transferred to a new tube and the brown pellet was resuspended in 2 mL 90% MeOH with vortexing. This suspension was sonicated for 20 min and centrifuged for 15 min at 4°C. The two supernatants from each sample were combined, vortexed to mix and transferred to a new tube. The solvent was evaporated using a dry vacuum at approximately 5 Torr.
The extracts from the UV-C experiments and the trial Pst avrRpt2 experiments were partially purified by passing suspensions through a solid phase extraction (SPE) column (Biotage Isolute 101, 200 mg bed, 6 mL capacity). Suspensions in 2 mL 25 mM potassium phosphate (KH$_2$PO$_4$) buffer, pH 2.5, were loaded onto a column, washed with 3 mL buffer plus 1% methanol, and eluted sequentially with 2 x 0.5 mL and 2 x 1mL 1% acetic acid in 80% methanol. The eluates were dried down and analyzed by HPLC as described below.

The extracts were resuspended in 160 – 180 µL 20% MeOH or starting HPLC buffer, vortexed, sonicated for 20 min, and filtered through a 0.22 µm syringe filter (Millipore). HPLC separation of leaf extracts was performed on a Shimadzu SCL-10A system with a Shimadzu RF-10A scanning fluorescence detector and a Shimadzu SPD-M10A photodiode array detector. Samples were separated on a 5-µm, 15 cm x 4.6-mm i.d. Supelcosil LC-ABZ Plus column (Supelco) preceded by a LC-ABZ Plus guard column maintained at 27°C. Prior to loading the 50-µL sample, the column was equilibrated with 5% acetonitrile in 25 mM KH$_2$PO$_4$, pH 2.5, at a flow rate of 1.0 mL/min. After 10 min, the concentration of acetonitrile was increased linearly to 43% over 25 min, followed by isocratic flow at 43% for 5 min, followed by a linear decrease from 43% to 5% over 3 min, and isocratic flow at 5% for 5 min.

SA-Asp was synthesized by adding SA to a mixture of 20 µg His-GH3.5 enzyme, 1 mM aspartic acid, 2.5 mM ATP, 5 mM Mg$^{2+}$, 1 mM DTT in 50 mM Tris buffer, pH 8.5, as described for PBS3 amino acid conjugation reactions in Chapter III.

**Bulk Collection of Peaks of Interest**

Metabolite extraction was similar to that described above, except the starting samples were 1 – 2 g. Samples are extracted into 2 x 5 mL 90% MeOH, which was evaporated under vacuum. Extracts were prepared for injection onto the HPLC column as above. For the first round of purification, samples were separated on a 5-µm, 15 cm x 4.6-mm i.d. Supelcosil LC-ABZ Plus column (Supelco) preceded by a LC-ABZ Plus guard column maintained at 27°C. Prior to loading the 50 µL sample, the column was equilibrated with 5% acetonitrile in 25 mM ammonium formate, pH 2.8, at a flow rate of 1.0 mL/min. After 7 min, the concentration of acetonitrile was increased linearly to 48% over 23 min, followed by isocratic flow at 43% for 5 min, a linear decrease from 43% to 5% over 3 min, and isocratic flow at 5% for 5 min. For the second round of purification, samples were separated on a 5-µm, 15 cm x 4.6-mm i.d. Altech C18 column with the same buffer and method as in the first round. Fractions corresponding to the peaks of interest were collected using an automatic fraction collector. Peaks were detected using the fluorescence detector set to ex 305/ em 407 nm.

**RESULTS**

Strategies to identify products of the reaction catalyzed by PBS3 in planta include treatment of leaves with elicitors such as UV-C radiation, spraying with potential benzoate precursors, infection with *Pseudomonas syringae* pv. *tomato*, and use of plant lines overexpressing PBS3. The extraction and analysis protocol is shown in Figure 4.3 and treatment regimes summarized in Table 4.1. In addition, extracts of soluble and cell-wall bound metabolites were analyzed for evidence of altered flux through the chorismate pathway by identifying classes of compounds synthesized from chorismate. One compound, reproducibly elevated in the pbs3 mutant in response to *Pst*, could be a novel signaling molecule or the in planta substrate of PBS3. Further
Figure 4.3. General protocol for metabolite extraction and analysis. Treated tissues from *Pst* and *Psm* infections and UV-C irradiation experiments are extracted, purified and analyzed. Cell wall extraction and analysis was performed on *Pst* and *Psm*-treated samples. Selected samples were also isolated by HPLC and analyzed for mass spectroscopy analysis. SPE = solid phase extraction, used in a crude purification of UV-C treated extracts prior to HPLC analysis.
The experimental treatments of plants for extraction of soluble and cell-wall bound metabolites are shown.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Lines</th>
<th>Time collected post-treatment</th>
<th>Soluble Fraction</th>
<th>Cell wall Fraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>UV-C + 4-HBA, pABA</td>
<td>Col-0, pbs3-2, 35S::PBS3</td>
<td>3, 12 h</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td><em>Pst avrRpt2</em></td>
<td>Col-0, pbs3-2, eds16-1</td>
<td>24 h</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td><em>Psm</em></td>
<td>Col-0, pbs3-1, pbs3-2, 35S::PBS3</td>
<td>72 h</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>
characterization through MS/MS and NMR analysis will elucidate the structure of the compound.

**Accumulation of Metabolites in Arabidopsis Cell Walls were Unaltered in Response to Infection with P. syringae**

Although elevated 4-HBA in Arabidopsis cell walls in response to infection with Pst has been previously reported (Tan et al., 2004), I did not observe this alteration. No compounds were found to be significantly and consistently altered in wild-type cell walls in response to avirulent or virulent strains of *Pseudomonas syringae* (Pst) (data not shown). A few minor differences were found between wild-type and pbs3 samples, but none were significant. Highly elevated salicylic acid was observed in several samples but this result was not reproducible. Because of the lack of significant and reproducible results, soluble fractions were the focus of the remainder of this study.

**Treatment with UV-C Radiation and Potential Substrates of PBS3 Did Not Lead to Conjugate Formation**

In order to encourage formation of potential PBS3-catalyzed benzoyl-amino acid conjugates, Arabidopsis leaves were treated with pABA or 4-HBA and irradiated with UV-C. Leaf samples were collected 3 and 12 hours later, corresponding to time-points in which PBS3 transcript level is elevated (data not shown). Approximately 45 – 70 nmol/g fresh weight 4-HBA and 8 nmol g/fresh weight pABA was quantified in the samples (data not shown). As equal solutions of equal molarities were sprayed on the plants, this suggests that 4-HBA is absorbed more effectively by the leaves or that pABA is more rapidly metabolized and converted to other forms. I looked for peak areas that differed between treatments or lines, particularly those corresponding to the retention time, fluorescence properties, and UV-profile of known benzoyl-amino acid conjugates. No putative pABA-Glu or 4HBA-Glu was observed. However, there were several compounds that were elevated in response to UV-C treatment and slightly reduced in pbs3-2. For example, a putative indole that eluted at 7 min followed this trend in both the 3 h and 12 h samples. It was detected using fluorescence at ex 290/em 340 nm. While accumulation was slightly lower in pABA-treated samples than in 4-HBA-treated samples, accumulation did not require benzoate treatment (Figure 4.4). As UV-C was not effective in inducing conjugate formation, treatment with Psm and Pst was tried instead with the idea that factors associated with bacterial infection could be required for formation of conjugates.

**Analysis of Soluble Metabolites from Pst avrRpt2-treated Leaf Tissue**

*Arabidopsis* leaves were infected with virulent *Pst* or avirulent *Pst avrRpt2*. No HPLC peaks corresponding to potential benzoyl-amino acid conjugates were observed in the extracts. However, two peaks were consistently and significantly elevated 24 hpi with infiltrated *Pst avrRpt2* in soluble leaf extracts from pbs3 mutants compared to wild-type (compounds 1 and 3 in Table 4.2). One of the peaks, corresponding to compound 1, was found to be reproducibly elevated in several experiments, shown in Figure 4.5. Peak area was quite low in untreated samples. This compound fluoresces at ex 305/em 407 nm but is only slightly UV absorbent (data not shown). This peak was not found in leaf samples treated with UV-C nor when *Pst avrRpt2* is sprayed rather than infiltrated into leaves (data not shown).

Isolation of the compound of interest was hampered by the presence of co-eluting compounds (two putative anthocyanin glycosides and sinapoyl malate) (Hagemeier et al., 2001).
Figure 4.4. Putative indole accumulated 12 h following treatment with UV-C. The peak area with corresponding to compound that eluted at ~ 7.2 min was measured using a fluorescence detector at 290/340 nm ex/em. Peak areas shown correspond to a single replicate were divided by 1000. Similar results were found 3 h post-treatment. Inset The UV profile corresponding to the peak is similar to that of known indoles.
### TABLE 4.2
Metabolite analysis from leaves treated with *Pst avrRpt2*

<table>
<thead>
<tr>
<th>Compound</th>
<th>Class</th>
<th>RT (min)</th>
<th>Identity</th>
<th>UV maxima</th>
<th>Fluorescence</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Unknown</td>
<td>22.2</td>
<td>Unknown</td>
<td>-</td>
<td>305/407</td>
<td>Elevated in <em>pbs3-2</em> with <em>Pst avrRpt2</em></td>
</tr>
<tr>
<td>2</td>
<td>Unknown</td>
<td>25.4</td>
<td>Unknown</td>
<td>310</td>
<td>305/407</td>
<td>Elevated in <em>pbs3-2</em> with <em>Pst avrRpt2</em> in Set 1</td>
</tr>
<tr>
<td>3</td>
<td>Unknown</td>
<td>27.8</td>
<td>Unknown</td>
<td>325</td>
<td>305/407</td>
<td>Elevated in <em>pbs3-2</em> with <em>Pst avrRpt2</em></td>
</tr>
<tr>
<td>4</td>
<td>Flavonoid</td>
<td>21.4</td>
<td>Kaempferol Glycoside</td>
<td>255, 370</td>
<td>-</td>
<td>Elevated in Col-0</td>
</tr>
<tr>
<td>5</td>
<td>Flavonoid</td>
<td>25.4</td>
<td>Kaempferol Glycoside</td>
<td>255, 370</td>
<td>-</td>
<td>Elevated in Col-0</td>
</tr>
<tr>
<td>6</td>
<td>Benzoate</td>
<td>11.6</td>
<td>Salicylic acid glucoside</td>
<td>210, 310</td>
<td>305/407</td>
<td>Elevated in Col-0 with <em>Pst avrRpt2</em></td>
</tr>
</tbody>
</table>

**Set 1**

<table>
<thead>
<tr>
<th>Compound</th>
<th>Col-0 Avg Peak Area*</th>
<th>Stdev†</th>
<th><em>pbs3-2</em> Avg Peak Area</th>
<th>stdev</th>
<th>Col-0 <em>Pst avrRpt2</em> Avg Peak Area</th>
<th>stdev</th>
<th><em>pbs3-2 Pst avrRpt2</em> Avg Peak Area</th>
<th>stdev</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8.9E+05</td>
<td>1.3E+04</td>
<td>8.1E+05</td>
<td>6.9E+04</td>
<td>2.5E+05</td>
<td>3.9E+04</td>
<td>6.2E+06</td>
<td>1.9E+06</td>
</tr>
<tr>
<td>2</td>
<td>4.7E+05</td>
<td>2.0E+05</td>
<td>4.5E+05</td>
<td>9.1E+04</td>
<td>3.8E+05</td>
<td>3.0E+05</td>
<td>5.1E+05</td>
<td>7.0E+04</td>
</tr>
<tr>
<td>3</td>
<td>3.1E+05</td>
<td>7.3E+04</td>
<td>2.1E+04</td>
<td>1.1E+04</td>
<td>3.3E+05</td>
<td>2.8E+05</td>
<td>1.1E+06</td>
<td>2.4E+05</td>
</tr>
<tr>
<td>4</td>
<td>3.5E+07</td>
<td>4.4E+06</td>
<td>2.7E+07</td>
<td>2.1E+06</td>
<td>3.6E+07</td>
<td>7.6E+06</td>
<td>2.9E+07</td>
<td>4.4E+06</td>
</tr>
<tr>
<td>5</td>
<td>4.9E+07</td>
<td>6.3E+06</td>
<td>2.3E+07</td>
<td>2.1E+07</td>
<td>4.7E+07</td>
<td>6.7E+06</td>
<td>4.0E+07</td>
<td>2.0E+06</td>
</tr>
<tr>
<td>6</td>
<td>1.1E+07</td>
<td>8.5E+06</td>
<td>1.3E+06</td>
<td>7.8E+05</td>
<td>1.6E+07</td>
<td>2.0E+06</td>
<td>1.5E+06</td>
<td>1.1E+06</td>
</tr>
</tbody>
</table>

**Set 2**

<table>
<thead>
<tr>
<th>Compound</th>
<th>Col-0 Avg Peak Area</th>
<th>Stdev</th>
<th><em>pbs3-2</em> Avg Peak Area</th>
<th>stdev</th>
<th>Col-0 <em>Pst avrRpt2</em> Avg Peak Area</th>
<th>stdev</th>
<th><em>pbs3-2 Pst avrRpt2</em> Avg Peak Area</th>
<th>stdev</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.8E+05</td>
<td>2.5E+05</td>
<td>5.1E+05</td>
<td>1.3E+05</td>
<td>8.8E+05</td>
<td>1.0E+06</td>
<td>8.2E+06</td>
<td>2.9E+06</td>
</tr>
<tr>
<td>2</td>
<td>2.4E+05</td>
<td>2.2E+05</td>
<td>1.9E+05</td>
<td>1.7E+05</td>
<td>4.5E+05</td>
<td>6.4E+05</td>
<td>3.7E+06</td>
<td>1.6E+06</td>
</tr>
<tr>
<td>3</td>
<td>2.3E+05</td>
<td>2.8E+05</td>
<td>2.1E+05</td>
<td>1.4E+05</td>
<td>3.0E+05</td>
<td>4.3E+05</td>
<td>2.6E+06</td>
<td>1.4E+06</td>
</tr>
<tr>
<td>4</td>
<td>2.3E+07</td>
<td>5.0E+06</td>
<td>1.7E+07</td>
<td>3.0E+06</td>
<td>2.8E+07</td>
<td>9.5E+06</td>
<td>1.2E+07</td>
<td>7.8E+06</td>
</tr>
<tr>
<td>5</td>
<td>3.8E+07</td>
<td>1.2E+07</td>
<td>2.4E+07</td>
<td>6.6E+06</td>
<td>4.2E+07</td>
<td>1.2E+07</td>
<td>2.6E+07</td>
<td>8.8E+06</td>
</tr>
<tr>
<td>6</td>
<td>1.5E+07</td>
<td>2.5E+06</td>
<td>2.1E+06</td>
<td>2.5E+05</td>
<td>1.2E+07</td>
<td>7.0E+06</td>
<td>2.2E+06</td>
<td>1.1E+06</td>
</tr>
</tbody>
</table>

Results of HPLC analysis of soluble extracts from ~0.5 g leaf samples treated with *Pst avrRpt2* or MgSO4 alone as negative control

* Avg Peak Area is the average peak area from HPLC analysis divided by grams fresh weight of three duplicate samples
† Stdev is the standard deviations of peak areas
Figure 4.5. Fluorescent peak is reproducibly elevated in pbs3-2 leaves in response to infection with *Pseudomonas syringae* pv. *tomato* *avrRpt2*, but not in wild-type Col-0. Leaves were infiltrated with bacterial suspensions at a dose of OD$_{600}$ 0.0001 and collected 24 hpi. Soluble compounds were extracted with methanol and analyzed by HPLC. The same pattern of accumulation has been observed in three separate experiments.
Two different HPLC columns (Supelco ABZ+ and Altech Prevail C18) with a variety of solvent combinations (phosphate/acetonitrile and formate/acetonitrile) and elution gradients were tested. No single HPLC method was found that would separate the peak of interest from the two interfering compounds. Instead, an isolation protocol was devised using two rounds of HPLC: separation and elution on the ABZ+ column followed by separation and elution on the C18 column. Soluble metabolites have been extracted from ~10 g each of Col-0 and pbs3-2 and pooled for HPLC separation using this protocol, and purification is in progress.

**Examination of Altered Flux Through the Chorismate Pathway**

Several classes of compounds derived from the chorismate pathway were identified in soluble leaf extracts based on fluorescence properties and UV absorption profiles. These include indoles, flavonoids, anthocyanins and benzoates. As shown in Table 4.2, slight reduction in the flavonoids kaempferol glycosides (Veit and Pauli, 1999; Tan et al., 2004) were found in pbs3-2 extracts compared to wild-type. While this difference was not statistically significant, it was consistent between experimental repeats and for the two kaempferol glycosides identified in the samples. There was no change in kaempferol glycosides between uninfected and infected samples 24 hours post infection with *Pst avrRpt2*. Levels of anthocyanins (Bloor and Abrahams, 2002) and indole compounds (Hagemeier et al., 2001) were consistent between samples in *Pst* experiments. However, in response to UV-C with exogenous addition of pABA or 4-HBA, a putative indolic compound is elevated (Figure 4.4). This could be further explored and may represent crosstalk between benzoates and indoles through the chorismate pathway. Therefore, although my findings do not indicate that PBS3 has a dramatic influence on flux through the general phenylpropanoid/flavonoid pathways, there is potential for flux to specific compounds to be altered and this could be further investigated.

**DISCUSSION**

**Cell Wall Extractions**

While other researchers found elevated levels of 4-HBA in *Arabidopsis* leaves in response to *Pst* (Tan et al., 2004), I did not detect 4-HBA in leaf samples (data not shown). It may be that levels of these metabolites from the cell wall are sensitive to specific experimental conditions such as light levels, soil lot, relative humidity, or specific pathogen strain which may vary between laboratories. The precise timing of the experiment could also influence metabolite accumulation. I have noticed that levels of some metabolites vary widely. For example, I found that SA levels were quite high in samples extracted one day from one set of plants but not on a different day from the same experiment (data not shown). While growth conditions and experimental protocols are tightly controlled, there is always some variability.

**Search for Benzoyl-Amino Acid Conjugates**

In order to encourage formation of potential PBS3-catalyzed benzoyl-amino acid conjugates, *Arabidopsis* leaves (wild-type Col-0, pbs3 mutants, and PBS3 overexpressors) were infected with *Pst* or *Psm* or treated with pABA or 4-HBA and irradiated with UV-C. None of these treatments led to the successful identification of a benzoyl-amino acid conjugate. This could be due to a number of reasons: the compound could be made transiently and under specific conditions or the compound is made in such small amounts that it is undetectable with the analytical methods employed. The use of isotopically labeled precursors could be used to
facilitate detection of the product. I did not take this approach as the isotopically-labeled benzoates are expensive and require dedicated equipment and we had not established that these benzoates were the actual \textit{in planta} substrate. In addition, as is the case with any application of starting material, the conditions are artificial and may not occur under normal conditions. Another possible approach is to use isolated protoplasts instead of whole leaves. This would ensure a uniform population of cells and could enhance the signal. Use of plant lines expressing bacterial effectors, such as \textit{avrRpt2}, could also enhance resolution with a greater number of responding cells and a synchronized response.

\textbf{Isolation of Compounds Elevated in \textit{pbs3} Mutants}

While prospective benzoyl-amino acid conjugates were not identified in soluble extracts, there were several compounds that different between samples. In the UV-C treated samples at both 3 and 12 h, a putative indole was elevated in samples from UV-C treated leaves (Figure 4.4). There was less accumulation in \textit{pbs3}-2 plants than in wild-type, although insufficient replicates were performed to be confident in this difference. Repeating the experiment with more replicates would help determine if PBS3 is involved in the accumulation of this compound in response to UV-C. If this were the case, it could help further define the function of PBS3. One peak is readily apparent in only the \textit{pbs3}-2 samples 24 hpi with \textit{Pst avrRpt2} (Figure 4.5). It is very low or absent in the uninfected samples and in Col-0 with or without treatment. While this compound is not the conjugate I was looking for, it is of great interest. I have not previously detected it in extractions using the standard SA protocol, described in Chapter II. In that method, the initial methanol extract is separated into two portions. One is treated with \(\beta\)-glucosidase, and then both portions are extracted with an organic mixture of ethyl acetate/cyclopentanes to extract SA from the aqueous layer, which is discarded. Non-hydrolyzed SAG remains in the aqueous layer, along with many of the other compounds detected when the initial methanol extract is analyzed directly on the HPLC without the subsequent steps. This suggests that the compound of interest, similar to SAG, is water-soluble. It may also be conjugated to a sugar, as are the cyanidin glycosides and sinapoyl malate with similar retention times (data not shown). It also has weak UV absorbance and fluoresces at 305/407 nm, the same excitation and emission wavelengths as SA. The peak of interest was not detected in samples from leaves infected by spray infection with OD\textsubscript{600} 0.01 \textit{Pst avrRpt2} or \textit{Psm} but only in syringe-infiltrated leaves. Infiltration, either using the syringe or vacuum method, is the preferred method of inoculation for many \textit{Arabidopsis} researchers. Spray infection, even with the addition of Silwet surfactant, requires uptake from the leaf surface, while infiltration is a more direct method of introducing the bacteria through the stomata.

There are several possible explanations for why the compound of interest is elevated in \textit{pbs3}-2 but not wild-type in response to infection with \textit{Pst avrRpt2}. First, this compound could be a substrate of PBS3 that is synthesized in response to bacterial infection and is typically converted into an amino acid conjugate when PBS3 is active. Without PBS3 activity, the compound is not converted and accumulates instead. The problem with this idea is that the conjugate should be detectable in Col-0 following infection; however it could be rapidly converted into another product. Second, accumulation of the compound could be a more indirect result of reduced SA accumulation, rather than directly related to PBS3 conjugation activity. In support of this idea is that the peak was also very low in the SA biosynthetic mutant \textit{eds16-1}. The peak has a similar retention time to SA-Asp (data not shown), although the expected UV-profile is absent. I have found that SA-Asp can be synthesized by \textit{AtGH3} (data not shown), and other researchers have
observed that it is elevated in plants overexpressing AtGH3.5 (Zhang et al., 2007). Third, the compound could be a metabolite downstream of the substrate. For example, if the in planta substrate were 4-HBA, free 4-HBA levels would accumulate in the pbs3 mutant. This could lead to increased flux through the ubiquinone pathway and build up of intermediates. Mass spectroscopy analysis on the purified compound will identify the compound and determine which of these scenarios, if any, occurs in the plant. Initial analyses of samples purified from small-scale experiments were not successful at identifying the compound; a scaled-up experiment with more plant material is in progress.

CONCLUSION

General flux to phenylpropanoids or flavonoids is not dramatically impacted in pbs3 mutants, however there is evidence of altered indole accumulation in response to UV-C with application of 4-HBA or pABA. Although the in planta product of the PBS3-catalyzed reaction has not yet been isolated, metabolite analysis revealed several compounds that are altered in pbs3 mutants following abiotic or biotic stress. One of these compounds was found to be elevated in pbs3-2 plants, but not Col-0, 24 hpi with Pst avrRpt2. This compound is quite water-soluble, highly fluorescent, weakly UV-absorbent, and of similarity polarity to glycosylated phenylpropanoids. This compound could function as a novel signaling molecule, a substrate of PBS3, or a product derived from the PBS3 substrate. The isolation of this compound has been hampered by interference from other similar compounds; a two-stage HPLC method has been developed to facilitate its isolation. Once purified in sufficient quantities, the compound can be analyzed by MS/MS and NMR in an effort to determine its identity. Identification of compounds altered in pbs3 would provide important insights into PBS3 function in basal resistance to pathogens and small molecule crosstalk.

REFERENCES


Chapter V: Evolutionary History of the GH3 Family in Rosids

SUMMARY

GH3\(^1\) amino acid conjugases have been identified in many plant and bacterial species. Here I focus on the evolution of GH3 genes in plant species using the sequenced rosids papaya, poplar, and grape and the asterid *Mimulus*. Comparison of co-linear genes in regions surrounding GH3 genes between species helps reconstruct the evolutionary history of the family. Combining analysis of synteny with phylogenetics and gene expression data leads to a new understanding of the Group III GH3 genes, to which *PBS3* belongs. I found this group to be rich with tandem duplication and transposition, reflective of a role in response to environmental stress. Contrary to previous reports that restrict *PBS3* to *Arabidopsis* and its close relatives, I found evidence that *PBS3* has descended from the same ancestral gene as *AtGH3.17*, suggesting that PBS3 has evolved from the Group II IAA-conjugating enzymes. Based on these analyses, I have reclassified *AtGH3.17* and the similar *AtGH3.9* into an expanded Group III.

---

\(^1\) Abbreviations. Compounds, BTH, 1,2,3-benzothiadiazole-7-carbothioic acid S-methyl ester; IAA, indole-3-acetic acid; JA, jasmonic acid; SA, salicylic acid. Genes: bZIP, basic-domain leucine-zipper; *ERF*, ethylene response factor; *GDG1*, GH3-like defense gene 1; *GH3*, Gretchen Hagen 3; *ICS1*, isochorismate synthase 1; *JAR1*, jasmonic acid resistant 1; *PBS3*, *avrPphB* susceptible 3, *AtGH3.12*; *WIN3*, *HopW1*-1-interacting 3. Organisms: At, *Arabidopsis thaliana*; Cp, *Carica papaya* (papaya); Mg, *Mimulus guttatus* (*Mimulus*); Os, *Oryza sativa* (rice); Pt, *Populus trichocarpa* (poplar); Vv, *Vitis vinifera* (grape). Terms: ML, maximum likelihood; MP, maximum parsimony; NJ, neighbor joining; PID, percent identity; WGD, whole genome duplication.
INTRODUCTION

GH3 (Gretchen Hagen 3) genes were originally identified in *Glycine max* (soybean) as responsive to the phytohormone auxin (Hagen et al., 1984), and have since been identified in many plant species (Terol et al., 2006). Several *Arabidopsis* genes were identified in genetic screens for altered phytohormone-mediated responses to auxin (DFL1) (Nakazawa et al., 2001), or jasmonic acid (JAR1) (Staswick et al., 1992). However, the molecular function of these genes remained unknown until Staswick *et al* identified similarity between the Arabidopsis GH3 and the firefly luciferase-like superfamily of proteins based on analysis of JAR1 with the 3D-PSSM platform (Staswick et al., 2002). The firefly luciferase-like superfamily, also called the adenylate-forming superfamily, is a diverse group of enzymes that catalyzes the addition of AMP to carboxyl groups on a wide variety of substrates. This family includes nonribosomal peptide synthetases, 4-coumarate-CoA ligases, acyl-CoA ligases, and oxidoreductases (Conti et al., 1996). They typically contain three conserved motifs that form a binding pocket for AMP and the substrate (Chang et al., 1997). Staswick *et al* identified the three conserved motifs in the Arabidopsis GH3 proteins. They performed *in vitro* assays and found that one, JAR1, catalyzes the addition of amino acids to the plant hormone jasmonic acid (Staswick et al., 2002) and that several others catalyze the addition of amino acids to auxin (Staswick et al., 2005). PBS3, along with several others, was not active on any of the phytohormone substrates investigated in that study. However, my results discussed in Chapter III indicate that 4-substituted benzoates serve as substrates of PBS3 *in vitro*. There are no known substrates for the other proteins in Group III.

Published phylogenetic trees constructed using distance methods divide plant GH3 proteins into three major clades, identified as Groups I, II, and III (Staswick et al., 2002) (Terol et al., 2006) (Felten et al., 2009). In these studies, Groups I and II contain genes from many more species than does Group III, which contains genes from only 3 species, *Arabidopsis thaliana*, *Brassica napus* (rapeseed), and *Gossypium hirsutum* (cotton), all in the Eurosids II subclade of the Rosids superfamily of dicotyledonous plants (Terol et al., 2006). The placement of the *Gossypium* sequence into Group III is unclear due to low bootstrap support. Substrate specificity of assayed proteins tends to correspond to these phylogenetic relationships (Staswick et al., 2002), (Staswick et al., 2005). The JAR1 enzyme active on JA is in Group I (2 genes), all GH3 enzymes that are active on IAA are in Group II (8 genes), and enzymes active on neither of these compounds, including PBS3, are in Group III (9 genes).

Group III appears to be highly expanded in *A. thaliana*. Past analyses showed that the clade containing PBS3 is present only in species closely related to *A. thaliana*. This suggests that PBS3-like genes are recently derived. However, past analyses depended on searches of EST databases without regard to chromosomal position and phylogenetic trees constructed exclusively using distance methods. The recent sequencing of multiple plant genomes coupled with new computational tools can leverage information from whole genomes to infer descent from a common ancestral gene by analyzing co-linearity of neighboring genes. The CoGe suite of comparative genomics programs (Lyons and Freeling, 2008; Lyons et al., 2008) is ideal for this sort of analysis among rosids. The rosids *Arabidopsis thaliana* and *Arabidopsis lyrata*, order Brassicales; *Carica papaya* (papaya), order Brassicales, both in the eurosid I clade; *Populus trichocarpa* (poplar), order Malpighiales in the eurosid II clade; and *Vitis vinifera* (grape), order Vitales, were used for this analysis. The asterid *Mimulus guttatus*, order Lamiales and the monocot grass *Oryza sativa* (rice) were used as outgroups for comparison. Note that there
remains some uncertainty regarding the classification of grape. However, sequence comparisons of chloroplast genomes suggest that grape is likely a basal rosid (Jansen et al., 2006). What is clear is that the ancestor of grape diverged from the rosid lineage following the pre-rosid hexaploidy event (Jaillon et al., 2007). The sequencing and annotation efforts on these genomes have progressed to varying degrees. At this time, A. thaliana, P. trichocarpa, V. vinifera, and O. sativa genes have been assigned to chromosomes, while C. papaya and M. guttatus have not yet been assigned (Phytozome 5.0).

Analysis of synteny is complicated by gene duplication, either due to whole genome duplication (WGD) or local duplication, gene loss or insertion (Lyons et al., 2008). Plant genomes are heavily duplicated, with WGD events quite common in the evolutionary history of many plant lineages. Following episodes of genome duplication, selective gene loss, called fractionation, is typically observed. Alternatively, some genes may evolve to serve different functions (sub or neo-functionalization) (Semon and Wolfe, 2007). All genes are not lost at an equal rate; some functional classes of genes, typically those encoding transcription factors, protein kinases, and ribosomal proteins are retained disproportionately, while the additional copies of others that are involved with more ancient biological functions are often lost more frequently (Freeling, 2009). Although A. thaliana has the smallest genome of any sequenced plant, two rounds of WGD have occurred since its divergence from the primary Rosid lineage (Blanc et al., 2003). A simplified phylogenetic tree showing the occurrence of WGD events is shown in Figure 5.1. One round of WGD has also occurred in the P. trichocarpa lineage. This can lead to as many as four copies of A. thaliana and two of P. trichocarpa for each V. vinifera or C. papaya gene. However, many of these have been lost due to genome fractionation. In addition, all of these species show evidence of a pre-rosid hexaploidy event with subsequent fractionation. Whether the lineage leading to M. guttatus has undergone a WGD event is not yet clear. M. guttatus is frequently used as a model system to study ecological processes. Several members of the Mimulus genus have been shown to have evidence of polyploidy (Wu et al., 2007), although the genome annotation has not yet progressed enough to discern polyploidy in the sequenced species.

If the Group III GH3 genes were only present in a small group of species, it would suggest the encoded enzymes have evolved a new function. Possibly, this function, e.g. acting on a unique substrate, could be specifically required by these species. An in depth exploration of the evolutionary history of the GH3 family, focusing particularly on group III, will help address this question. This analysis is coupled with investigation of expression patterns of the GH3 genes, particularly those for Group III genes, to gain insight on potential gene function. The results discussed in this chapter suggest a reclassification of several GH3 genes based on analysis of synteny. PBS3, four nearby locally duplicated genes, and AtGH3.17, typically classified as Group II, were found to be syntenic to genes in poplar, grape, and rice, suggesting descent from a common ancestral gene. Based on this analysis, I have reclassified AtGH3.17 and the similar AtGH3.9 into an expanded Group III. However, functionally PBS3 is quite distinct from these other two enzymes, and presumably from the grape, poplar and rice orthologs as well. My analysis suggests that the evolution of the distinct function of PBS3 has occurred following the duplication event giving rise to AtGH3.17 and PBS3.
Figure 5.1. Angiosperm phylogeny showing location of paleohexaploidy and paleotetraploidy events in lineages of interest. Figure adapted from (Freeling, 2009) with phylogenetic information from the Missouri Botanical Garden’s Angiosperm Phylogeny Project.
MATERIALS AND METHODS

**Identification of Syntenic Regions Flanking GH3 genes**

The CoGe platform for comparison of genome sequences (http://synteny.cnr.berkeley.edu /CoGe/) (Lyons and Freeling, 2008) was used to identify regions of potential synteny between *Arabidopsis thaliana* (v9) *Arabidopsis lyrata* (v1) and other sequenced rosids, including *Carica papaya* (v0.4, ASPGB draft genome (Ming et al., 2008)), *Vitis vinifera* (v1, French-Italian Public Consortium for Grapevine Genome Characterization (Jaillon et al., 2007)) and *Populus trichocarpa* (JGI, v2 (Tuskan et al., 2006)) as well as the Asterid *Mimulus guttatus* (JGI, v1). The regions of potential synteny compiled in (Lyons et al., 2008) were used as a starting point, modified with new genome versions, evaluated for accuracy, and expanded with additional analyses. The *V. vinifera* genome, less subject to rearrangements, duplications and gene loss than other genomes (Jaillon et al., 2007) (Semon and Wolfe, 2007), was used as a bridge between *A. thaliana* and the other genomes to identify regions of co-linearity and possible synteny.

Three *Arabidopsis* GH3 genes were used as seeds to identify GH3 genes in the other plant genomes. The CDS sequences of PBS3 (*AtGH3.12, At5g13320*), *AtGH3.6* (*At5g54510*), JAR1 (*AtGH3.11, At2g46370*), and *AtGH3.10* (*At4g03400*) were used as BLAST seeds against papaya, grape, poplar, and *Mimulus* genomes in CoGeBlast. Additional BLASTN searches were performed with the CDS sequence from each species against the genome of origin to find any other possible matches. Sequence length and gene models were retrieved from the Phytozome v5.0 website (http://www.phytozome.net/) from the Department of Energy's Joint Genome Institute and the Center for Integrative Genomics. The CDS sequences were translated to protein sequences using the online interface of Virtual Ribosome v1.1 (http://www.cbs.dtu.dk/services/VirtualRibosome/) from the Center for Biological Sequence Analysis at the Technical University of Denmark (Wernersson, 2006). *Arabidopsis* GH3 peptide sequences were retrieved directly from TAIR (v9) and rice sequences from MSU Rice Genome Annotation project (v6).

**Sequence Alignment and Phylogenetic Tree Construction**

CDS and peptide sequences were aligned with MUSCLE (Edgar, 2004) via the online phylogeny.fr platform (http://www.phylogeny.fr) (Dereeper et al., 2008) using default parameters. Distance (BioNJ), maximum likelihood (PhyML) and maximum parsimony (TNT) tree construction methods were used and MEGA 4.0 (Tamura et al., 2007) used to visualize and annotate the phylogenetic trees. DNA regions containing sequences related phylogenetically were tested for synteny using the GEvo tool of the CoGe browser, described above. Pair-wise identity scores were converted from p-distance values calculated using MEGA4 (Tamura et al., 2007) from an alignment of IIA protein sequences constructed with the ClustalW algorithm under default conditions.

**Analysis of Expression Data**

The expression patterns of *AtGH3* genes were explored using tools from Genevestigator (Zimmermann et al., 2004), NascArrays (Craigon et al., 2004), the eFP browser (Toufighi et al., 2005; Winter et al., 2007) (http://www.bar.utoronto.ca/efp/cgi-bin/efpWeb.cgi) and analysis of the relevant literature. Experiments in which GH3 genes were expressed in response to abiotic or biotic stress or hormone treatments were identified in Genevestigator and the eFP browser,
and data downloaded. Most experiments analyzed were from the AtGenExpress series. The Nascarrays experiment reference numbers are shown in the corresponding data tables in the Results section. Values greater than 2-fold increase or decrease relative to control experiments were reported and significance tested using student’s t-tests at $\alpha = 0.05$ for experiments with three replicates. Experiments with two replicates were manually examined for reproducibility of experimental and control samples. *Pt*GH3 and *Vv*GH3 expression patterns were analyzed using the Plant Expression Database (Plexdb, http://plexdb.org (Wise et al., 2007)).

**Analysis of Transcription Factor Binding Motifs**

Potential transcription factor binding sites 1 kb upstream of AtGH3 start sites were identified using the Arabidopsis *cis*-regulatory element database (AtcisDB, http://arabidopsis.med.ohio-state.edu/AtcisDB/) (Molina and Grotewold, 2005) of the Arabidopsis Gene Regulatory Information Server from Ohio State University. Motifs associated with transcription factors known to be involved in response to biotic stress were tallied and summarized. Motifs are defined as follows: WRKYs recognize the W-box: ttgact/c; the MYB and MYB-like recognize aaccaac (MYB), taactaac (MYB3), aacaaac (MYB4), and aaatct (MYB-related CCA1); bZip recognize actcat (ATB2/AtbZip53) and acacttg (DPBF1&2); ARF1 recognizes tgtctc; and DRE motif containing proteins recognize tgcgacaa, gaccgacct, and aacgacca.

**RESULTS**

**Identification of Syntenic Regions Flanking GH3 genes**

Using the GEvo tool of the CoGe genome comparison browser (Lyons and Freeling, 2008), ten genomic regions of *Arabidopsis* flanking GH3 genes were found to demonstrate patterns of co-linearity of genes with the other Rosids, indicative of synteny and probable descent from a common ancestral chromosomal region. These include several areas of *Arabidopsis* chromosomes 1, 2, 4, and 5 and account for 15 of the 19 *Arabidopsis* GH3 (AtGH3) genes. An example of this synteny is shown in Figure 5.2, which compares a section of approximately 40 kB of *Arabidopsis* chromosome 5 with 100 kB of *Vitis vinifera* chromosome 1, 90 kB *Populus trichocarpa* chromosome 1, and 60 kB of *P. trichocarpa* chromosome 3. This *Arabidopsis* chromosome 5 region contains PBS3 (AtGH3.12, At5g13320) and four other AtGH3 (AtGH3.13-16) genes, presumably the result of local duplication. In addition, 70 kB of *Arabidopsis* chromosome 1 containing AtGH3.17, a region of *Oryza sativa* chromosome 11 containing OsGH3.13 (Os11g32520) (Terol et al., 2006), and three regions of *Mimulus guttatus* also demonstrated evidence of synteny (see Figure 5.3 for comparison of rice region with *Arabidopsis*, grape, and one of the *Mimulus* regions). No corresponding syntenic region was identified in papaya. The five AtGH3 genes in the chromosome 5 region, including PBS3, correspond to a single gene in each of the other genomes compared. This suggests that the five GH3 genes in this region of *Arabidopsis* result from a local duplication. Two additional genes (At5g13330 and At5g13340, annotated as an ERF/AP2 transcription factor family and of unknown function, respectively) are located between PBS3 and the other four AtGH3 genes. As expected due to the fractionation of *Arabidopsis* genes following the two episodes of WGD, the other genomes share more syntenic genes with each other than with *Arabidopsis*.

Besides the GH3 genes, six *Arabidopsis* genes in the chromosome 5 region and 4 in the chromosome 1 region display synteny to genetic regions of other species. These genes and their
Figure 5.2. Synteny between the region of Arabidopsis chromosome 5 surrounding PBS3 and grape and poplar chromosomes. A screenshot of the BLASTz output from the CoGe browser is shown. Each large horizontal bar represents one genomic region, with the dashed line dividing the top (5’ on left) and bottom (5’ on right) strand. The genome origin is indicated on the right (At = Arabidopsis thaliana, Vv = Vitis vinifera, Pt = Populus trichocarpa). The colored arrows represent gene models: green are CDS, blue are RNA, and gray are introns. The areas of similarity from BLASTz between genomic regions are shown as colored blocks above or below the gene models. Each pair-wise comparison is shown in a different color. Brown and pink lines are drawn between similar regions of At and the other species. Lines connecting other pair-wise comparisons were omitted for clarity. The analysis can be regenerated at http://tinyurl.com/y8p9kdp.
Figure 5.3. Synteny between the region of rice chromosome 11 surrounding OsGH3.13 and Arabidopsis, grape, and Mimulus chromosomes. A screenshot of the BLASTz output from the CoGe browser is shown, with features as Figure 5.1. The genome origin is indicated on the right (At = Arabidopsis thaliana, Vv = Vitis vinifera, Mg = Mimulus guttatus, Os = Oryza sativa). Lines only connect Os genes to their syntelogs for clarity.
annotations are summarized in Table 5.1. Although three Mimulus regions demonstrate co-linearity, none of the regions contain a sequence corresponding to a GH3 gene. One of the regions is truncated at the location of the GH3 gene in the other species (Mimulus scaffold 39), suggesting that there could be a GH3 gene that was missed in sequencing; the section of Table 5.1 with missing gene information represents this region. Genes in the region encode a variety of different kinds of proteins; however, many of them are annotated as being associated with the mitochondrion or chloroplast. The other syntenic sets are described below.

**Identification of GH3 Genes**

Potential orthologs of the AtGH3 genes were identified by BLASTN search using CoGe Blast with AtGH3 genes as seeds. Six papaya, thirteen poplar, eight grape, and ten mimulus orthologs were identified. The sequences were aligned using MUSCLE (Edgar, 2004), visualized in JalView and evaluated for global alignment and presence of AMP-binding motifs (Chang et al., 1997). In addition, gene structure models were compared to BLASTX results in the Phytozome Genoscope v1 Vitis vinifera genome browser (http://www.phytozome.net/cgi-bin/gbrowse/grape/). The gene names, locus, chromosome, contig, or scaffold number, predicted protein length, number of exons, and Arabidopsis syntelog identified from the synteny analysis are shown in Table 5.2 for poplar, grape, papaya and Mimulus. Poplar names are as in (Felten et al., 2009) with the addition of PtGH3.15, grape, papaya and Mimulus sequences were named approximately based on their Arabidopsis syntelogs.

Several sequences were found to be incomplete. GSVIV00026990001 (VvGH3.7) was missing motif II and the protein sequence was shorter than expected by approximately 200 amino acids based on comparison with other GH3 proteins, typically around 600 amino acids long. Comparison to other peptides in the BLASTX results of the Phytozome genome browser suggested that extra introns incorrectly interrupt exon 3. A BLASTN search of the Genbank EST database with the CDS sequence as the query verified that the missing region is expressed. Sequences were found (gi 110368758 and gi 110698541) that overlapped the annotated and missing regions. In addition, a search of grape sequences in Genbank with BLASTX identified a peptide containing the missing region and otherwise identical to GSVIV00026990001 (gi 225454466). The corrected sequence length and exon number are shown in Table 5.2. Two papaya sequences, EVM prediction supercontig_1065.2 (CpGH3.3) and EVM prediction supercontig_9.204 (CpGH3.4) were both truncated due to missing sequence data. Some additional base pairs in the C-terminus of the p1065.2 gDNA sequence were identified as potentially coding, and added to the CDS sequence. However, this correction did not account for all of the missing sequence at the C-terminus. No ESTs were identified that contain the missing regions for either of the papaya sequences, so they remain incomplete.

**Classification of GH3 proteins into Groups Based on Synteny and Phylogenetic Relationship**

The corrected protein sequences from Arabidopsis, papaya, grape, and poplar were aligned using MUSCLE (Edgar, 2004), and curated by removing gaps or using Gblocks (Castesana, 2000), a more sophisticated curation algorithm, through the phylogeny.fr server (Dereeper et al., 2008). Alignments were constructed with or without the two incomplete papaya sequences and clustered into phylogenetic trees using maximum parsimony (TNT), maximum likelihood (PhyML), and neighbor joining (BioNJ) algorithms. The tree found to best correspond to the syntenic relationships identified above was constructed using a multiple sequence alignment.
Regions are syntenic to At chromosome 5 due to WGD within the lineage (At1 - At2) or speciation (At-Vv, At-Pt, At-Mg, At-Os).

<table>
<thead>
<tr>
<th>At locus 1</th>
<th>At locus 2</th>
<th>Vv Locus 1</th>
<th>Pt locus 1</th>
<th>Mg locus 1</th>
<th>Mg locus 2</th>
<th>Mg locus 3</th>
<th>Os locus</th>
<th>Annotation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ch 5</td>
<td>Ch 1</td>
<td>Ch 1</td>
<td>Ch 1</td>
<td>Scaffold 39</td>
<td>Scaffold 7</td>
<td>Scaffold 54</td>
<td>Chrom 11</td>
<td></td>
</tr>
<tr>
<td>none</td>
<td>none</td>
<td>GSVV00006222001</td>
<td>POPTR_0001s12830</td>
<td>POPTR_0003s15950</td>
<td>none</td>
<td>none</td>
<td>Os11g32750</td>
<td>NUDIX family Hydrolase</td>
</tr>
<tr>
<td>At5g13250</td>
<td>At1g28030</td>
<td>none</td>
<td>none</td>
<td>none</td>
<td>mgf018531m</td>
<td>mgf004629m</td>
<td>none</td>
<td>Unknown</td>
</tr>
<tr>
<td>none</td>
<td>none</td>
<td>GSVV00006233001</td>
<td>none</td>
<td>none</td>
<td>mgf001431m</td>
<td>mgf004629m</td>
<td>none</td>
<td>Unknown</td>
</tr>
<tr>
<td>At5g13320</td>
<td>At1g28130</td>
<td>GSVV00006222001</td>
<td>POPTR_0001s12850</td>
<td>POPTR_0003s15970</td>
<td>none</td>
<td>none</td>
<td>Os11g32520</td>
<td>PBS3, GH3 family protein</td>
</tr>
<tr>
<td>At5g13330</td>
<td>At1g28160</td>
<td>none</td>
<td>POPTR_0001s12820</td>
<td>POPTR_0003s15940</td>
<td>none</td>
<td>none</td>
<td>none</td>
<td>ERF/AP2 transcription factor family</td>
</tr>
<tr>
<td>At5g13340</td>
<td>none</td>
<td>none</td>
<td>none</td>
<td>none</td>
<td>none</td>
<td>none</td>
<td>none</td>
<td>Unknown</td>
</tr>
<tr>
<td>At5g13350</td>
<td>At1g28130</td>
<td>none</td>
<td>none</td>
<td>none</td>
<td>mgf005654m</td>
<td>mgf005654m</td>
<td>Os11g32500</td>
<td>Rubredoxin Family protein</td>
</tr>
<tr>
<td>At5g13360</td>
<td>At1g28130</td>
<td>none</td>
<td>none</td>
<td>none</td>
<td>mgf012815m</td>
<td>none</td>
<td>Os11g32470</td>
<td>Chloroplast NEF1 pollen function</td>
</tr>
<tr>
<td>At5g13370</td>
<td>At1g28130</td>
<td>none</td>
<td>none</td>
<td>none</td>
<td>mgf016173m</td>
<td>none</td>
<td>None</td>
<td>Proton-dependent oligopeptide transport</td>
</tr>
<tr>
<td>At5g13380</td>
<td>At1g28130</td>
<td>none</td>
<td>none</td>
<td>none</td>
<td>mgf016173m</td>
<td>none</td>
<td>None</td>
<td>Cytochrome P450</td>
</tr>
<tr>
<td>At5g13390</td>
<td>none</td>
<td>GSVV00006219001</td>
<td>POPTR_0001s12860</td>
<td>POPTR_0003s15980</td>
<td>mgf012815m</td>
<td>mgf012815m</td>
<td>Os11g32240</td>
<td>Peptidyl-prolyl cis-trans isomerase</td>
</tr>
<tr>
<td>At5g13400</td>
<td>none</td>
<td>GSVV00006214001</td>
<td>POPTR_0001s12890</td>
<td>none</td>
<td>mgf012815m</td>
<td>mgf012815m</td>
<td>Os11g32240</td>
<td>Peptidyl-prolyl cis-trans isomerase</td>
</tr>
<tr>
<td>At5g13410</td>
<td>none</td>
<td>GSVV00006210001</td>
<td>POPTR_0001s12920</td>
<td>none</td>
<td>mgf012815m</td>
<td>mgf012815m</td>
<td>Os11g32240</td>
<td>Peptidyl-prolyl cis-trans isomerase</td>
</tr>
<tr>
<td>At5g13420</td>
<td>none</td>
<td>GSVV00006209001</td>
<td>POPTR_0001s12930</td>
<td>none</td>
<td>mgf012815m</td>
<td>mgf012815m</td>
<td>Os11g32240</td>
<td>Peptidyl-prolyl cis-trans isomerase</td>
</tr>
<tr>
<td>At5g13430</td>
<td>none</td>
<td>GSVV00006206001</td>
<td>POPTR_0001s12960</td>
<td>none</td>
<td>mgf015946m</td>
<td>mgf015946m</td>
<td>Os11g32240</td>
<td>Ubiquinol-cytochrome C reductase</td>
</tr>
<tr>
<td>At5g13440</td>
<td>none</td>
<td>GSVV00006205001</td>
<td>POPTR_0001s12960</td>
<td>POPTR_0003s16060</td>
<td>mgf004895m</td>
<td>mgf004895m</td>
<td>Os11g32240</td>
<td>Ubiquinol-cytochrome C reductase</td>
</tr>
<tr>
<td>At5g13450</td>
<td>none</td>
<td>GSVV00006202001</td>
<td>POPTR_0001s12960</td>
<td>POPTR_0003s16060</td>
<td>mgf004895m</td>
<td>mgf004895m</td>
<td>Os11g32240</td>
<td>Ubiquinol-cytochrome C reductase</td>
</tr>
</tbody>
</table>

TABLE 5.1
Comparison of syntenic genes in *Arabidopsis* chromosomes 5 and 1, grape, poplar, *Mimulus* and rice in region around PBS3 (AtGH3.12, At5g13320)
### TABLE 5.2

GH3 genes in poplar, grape, papaya and *Mimulus* and their relationship to *Arabidopsis* genes through common position on an ancestral chromosome

<table>
<thead>
<tr>
<th>Name*</th>
<th>Locus</th>
<th>Pos.</th>
<th>Protein Length</th>
<th>Exons</th>
<th>Group</th>
<th>At syntelog(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PtGH3.1</td>
<td>POPTR_0007s10350</td>
<td>7</td>
<td>597</td>
<td>3</td>
<td>II</td>
<td>AtGH3.2, AtGH3.3</td>
</tr>
<tr>
<td>PtGH3.2</td>
<td>POPTR_0009s09590</td>
<td>9</td>
<td>690</td>
<td>3</td>
<td>II</td>
<td>AtGH3.1</td>
</tr>
<tr>
<td>PtGH3.3</td>
<td>POPTR_0001s30560</td>
<td>1</td>
<td>596</td>
<td>3</td>
<td>II</td>
<td>AtGH3.1</td>
</tr>
<tr>
<td>PtGH3.4</td>
<td>POPTR_0013s14740</td>
<td>13</td>
<td>608</td>
<td>3</td>
<td>II</td>
<td>-</td>
</tr>
<tr>
<td>PtGH3.5</td>
<td>POPTR_0011s13330</td>
<td>11</td>
<td>611</td>
<td>3</td>
<td>II</td>
<td>AtGH3.5, AtGH3.6</td>
</tr>
<tr>
<td>PtGH3.6</td>
<td>POPTR_0001s43990</td>
<td>1</td>
<td>611</td>
<td>3</td>
<td>II</td>
<td>AtGH3.5, AtGH3.6</td>
</tr>
<tr>
<td>PtGH3.7</td>
<td>POPTR_0001s12850</td>
<td>4</td>
<td>606</td>
<td>4</td>
<td>III</td>
<td>AtGH3.12, AtGH3.17</td>
</tr>
<tr>
<td>PtGH3.8</td>
<td>POPTR_0003s15970</td>
<td>3</td>
<td>594</td>
<td>5</td>
<td>III</td>
<td>AtGH3.12, AtGH3.17</td>
</tr>
<tr>
<td>PtGH3.9</td>
<td>POPTR_0002s32070</td>
<td>2</td>
<td>568</td>
<td>4</td>
<td>II</td>
<td>AtGH3.9</td>
</tr>
<tr>
<td>PtGH3.10</td>
<td>POPTR_0013s14050</td>
<td>13</td>
<td>595</td>
<td>4</td>
<td>II</td>
<td>AtGH3.10</td>
</tr>
<tr>
<td>PtGH3.11</td>
<td>POPTR_0018s17450</td>
<td>19</td>
<td>564</td>
<td>4</td>
<td>II</td>
<td>AtGH3.10</td>
</tr>
<tr>
<td>PtGH3.14</td>
<td>POPTR_0014s09120</td>
<td>14</td>
<td>576</td>
<td>4</td>
<td>II</td>
<td>AtGH3.11</td>
</tr>
<tr>
<td>PtGH3.15</td>
<td>POPTR_0002s16960</td>
<td>2</td>
<td>576</td>
<td>4</td>
<td>I</td>
<td>AtGH3.11</td>
</tr>
<tr>
<td>VvGH3.1</td>
<td>GSVIV0007718001</td>
<td>3</td>
<td>598</td>
<td>4</td>
<td>II</td>
<td>AtGH3.1</td>
</tr>
<tr>
<td>VvGH3.2</td>
<td>GSVIV0001961001</td>
<td>7</td>
<td>600</td>
<td>3</td>
<td>II</td>
<td>AtGH3.2, AtGH3.3</td>
</tr>
<tr>
<td>VvGH3.3</td>
<td>GSVIV00027472001</td>
<td>19</td>
<td>605</td>
<td>4</td>
<td>II</td>
<td>AtGH3.5, AtGH3.6</td>
</tr>
<tr>
<td>VvGH3.4</td>
<td>GSVIV0002612001</td>
<td>12</td>
<td>588</td>
<td>5</td>
<td>II</td>
<td>AtGH3.9</td>
</tr>
<tr>
<td>VvGH3.5</td>
<td>GSVIV00027564001</td>
<td>7</td>
<td>596</td>
<td>4</td>
<td>II</td>
<td>AtGH3.9</td>
</tr>
<tr>
<td>VvGH3.6</td>
<td>GSVIV00026000001</td>
<td>12</td>
<td>592</td>
<td>4</td>
<td>II</td>
<td>AtGH3.10</td>
</tr>
<tr>
<td>VvGH3.7</td>
<td>GSVIV00026990001</td>
<td>15</td>
<td>596</td>
<td>4</td>
<td>I</td>
<td>AtGH3.11</td>
</tr>
<tr>
<td>VvGH3.8</td>
<td>GSVIV00066220001</td>
<td>1</td>
<td>593</td>
<td>5</td>
<td>III</td>
<td>AtGH3.12, AtGH3.17</td>
</tr>
<tr>
<td>CpGH3.1</td>
<td>EVM supercontig_292.1</td>
<td>292</td>
<td>599</td>
<td>3</td>
<td>II</td>
<td>AtGH3.1</td>
</tr>
<tr>
<td>CpGH3.2</td>
<td>EVM supercontig_6.74</td>
<td>6</td>
<td>599</td>
<td>3</td>
<td>II</td>
<td>AtGH3.2, AtGH3.3</td>
</tr>
<tr>
<td>CpGH3.3</td>
<td>EVM supercontig_1065.2</td>
<td>1065</td>
<td>4551</td>
<td>3</td>
<td>II</td>
<td>AtGH3.5, AtGH3.6</td>
</tr>
<tr>
<td>CpGH3.4</td>
<td>EVM supercontig_9.204</td>
<td>9</td>
<td>491</td>
<td>3</td>
<td>II</td>
<td>AtGH3.9</td>
</tr>
<tr>
<td>CpGH3.5</td>
<td>EVM supercontig_34,122</td>
<td>34</td>
<td>607</td>
<td>4</td>
<td>II</td>
<td>AtGH3.10</td>
</tr>
<tr>
<td>CpGH3.6</td>
<td>EVM supercontig_1483.1</td>
<td>1483</td>
<td>591</td>
<td>4</td>
<td>I</td>
<td>AtGH3.11</td>
</tr>
<tr>
<td>MgGH3.1</td>
<td>mgf012326m</td>
<td>128</td>
<td>588</td>
<td>3</td>
<td>II</td>
<td>AtGH3.1 AtGH3.2</td>
</tr>
<tr>
<td>MgGH3.2</td>
<td>mgf023419m</td>
<td>505</td>
<td>354</td>
<td>3</td>
<td>II</td>
<td>AtGH3.1 AtGH3.2</td>
</tr>
<tr>
<td>MgGH3.3</td>
<td>mgf002259m</td>
<td>123</td>
<td>498</td>
<td>2</td>
<td>I</td>
<td>AtGH3.1</td>
</tr>
<tr>
<td>MgGH3.4</td>
<td>mgf018205m</td>
<td>220</td>
<td>580</td>
<td>3</td>
<td>II</td>
<td>AtGH3.2 AtGH3.3</td>
</tr>
<tr>
<td>MgGH3.5</td>
<td>mgf006862m</td>
<td>18</td>
<td>597</td>
<td>3</td>
<td>II</td>
<td>AtGH3.1 AtGH3.2</td>
</tr>
<tr>
<td>MgGH3.6</td>
<td>mgf010505m</td>
<td>24</td>
<td>618</td>
<td>3</td>
<td>II</td>
<td>AtGH3.5 AtGH3.6</td>
</tr>
<tr>
<td>MgGH3.7</td>
<td>mgf005874m</td>
<td>83</td>
<td>612</td>
<td>3</td>
<td>II</td>
<td>AtGH3.5 AtGH3.6</td>
</tr>
<tr>
<td>MgGH3.8</td>
<td>mgf016598m</td>
<td>1</td>
<td>608</td>
<td>4</td>
<td>III</td>
<td>AtGH3.9</td>
</tr>
<tr>
<td>MgGH3.9</td>
<td>mgf002659m</td>
<td>52</td>
<td>606</td>
<td>4</td>
<td>I</td>
<td>AtGH3.10</td>
</tr>
<tr>
<td>MgGH3.10</td>
<td>mgf008182m</td>
<td>27</td>
<td>575</td>
<td>4</td>
<td>I</td>
<td>AtGH3.11</td>
</tr>
</tbody>
</table>

Syntelogs of PBS3 (AtGH3.12) are in bold. *Poplar names are from Felten et al. (2009), PtGH3.15, grape, papaya and *Mimulus* names first described here. Locus names, protein length and exon number from Phytozome v5.0. Position is the chromosome for Pt and Vv, and supercontig for Cp and Mg. Group and At syntelog from synteny analysis described in Materials and Methods. † Gene model in Phytozome v5.0 corrected based on global alignment and EST data. ‡ Sequencing error results in missing sequence.
curated using Gblocks with relaxed settings and PhyML, and is shown in Figure 5.4. Little difference was found including or omitting the two incomplete papaya sections; the only substantive difference is that AtGH3.17 clades better without including the sequences. When the truncated papaya sequences are not included in the alignment, AtGH3.17 was found to form a monophyletic clade with VvGH3.8, PtGH3.7, and PtGH3.8.

The seven sets of syntenic genes are spread between Groups I, II, and III, and the encoded proteins separate into clades in the phylogenetic tree shown in Figure 5.4. There are two syntenic sets in Group I (IA and IB), four in Group II (IIA1, IIA2, IIB1, and IIB2), and two in Group III (IIIA and IIIB). The syntenic set discussed above which includes PBS3 on Arabidopsis chromosome 5 and AtGH3.17 on chromosome 1, is classified as set IIIA in the phylogenetic tree. Although PBS3 is typically classified as Group III and AtGH3.17 as Group II (Staswick et al., 2005), the analysis of co-linear genes suggests they are descended from the same ancestral chromosomal position, as evidenced by their shared synteny with grape and poplar regions. As the genes have different expression profile and the proteins have different enzymatic activities (Staswick et al., 2005; Okrent et al., 2009), they have clearly diverged in function. AtGH3.17, as well as AtGH3.9, typically clade separately from the other Group II genes. Unlike AtGH3.1-6, AtGH3.17 and AtGH3.9 are not induced by auxin (Hagen and Guilfoyle, 2002; Staswick et al., 2005), although they both are capable of adenylating it (Staswick et al., 2002).

Some of the proteins that clade with PBS3 and the other Group IIIA proteins are not syntenic. This clade is quite interesting in that it shows evidence of local duplication, insertion and gene loss. As discussed above, PBS3 and AtGH3.13, AtGH3.14, AtGH3.15, AtGH3.16 are present in the same region of chromosome 5, with two genes in between PBS3 and the other four. However, the four adjacent genes do not have the highest percent identity (PID) to PBS3. Four AtGH3 proteins have higher PID to PBS3 than do these. (See Table 5.3 for scores for Group IIIA proteins). In fact, the orthologs from other species (PtGH3.7, PtGH3.8, VvGH3.8) also have greater PID to PBS3 than do these tandemly duplicated genes. The gene corresponding to the AtGH3 protein with the highest percent identity to PBS3, AtGH3.7 (At1g23160) appears to have inserted into an ancestor of A. thaliana before the divergence from A. lyrata but after papaya. DNA sequence flanking AtGH3.7 was used as a seed for a BLAST search of the grape, poplar, and papaya genomes. One region of grape and papaya and two from poplar were identified as co-linear and presumably syntenic, but do not contain a GH3 gene (Figure 5.5). However, the corresponding region from A. lyrata, identified by a BLAST search with the sequence from the A. thaliana region, does contain a GH3 gene (data not shown).

Similarly, regions of grape chromosome 16 and poplar chromosomes 12 and 15 display co-linear genes with the region around AtGH3.8 (At5g51470), although no grape GH3 gene is present. This region contains many stress response genes, including PBS2, also known as RAR1, which was picked up in the same screen as PBS3 (Warren et al., 1999) (data not shown, can be recapitulated http://genomevolution.org/r/37d). The other two Group IIIA AtGH3 genes, AtGH3.18 (At1g48670) and AtGH3.19 (At1g48660), have presumably been inserted as well, as the flanking genes are not co-linear to the genes flanking PBS3 or its orthologs in other species (data not shown). An additional truncated GH3 gene (At1g48690) is adjacent, with a predicted peptide of only 191 amino acids, a third of the length of the others in the family. There are several retrotransposons nearby, including At1g48680, which is directly in between AtGH3.19 and the truncated gene; and At1g48710; just downstream.
Figure 5.4. Maximum likelihood phylogenetic tree of GH3 proteins from rosids. The tree was constructed using PhyML on multiple sequence alignments from MUSCLE curated using Gblocks with reliability of internal branch length tested using the aLRT method. Branches with the same symbol represent a set with evidence of synteny. Sequences are from the Rosid species At Arabidopsis thaliana, Cp Carica papaya, Pt Populus trichocarpa Vv Vitis vinifera.
Percent identities scores were converted from p-distance values calculated using MEGA4 (Tamura et al., 2007) from a ClustalW alignment of Group IIIA members.

<table>
<thead>
<tr>
<th>Name</th>
<th>Locus</th>
<th>Percent ID with PBS3</th>
</tr>
</thead>
<tbody>
<tr>
<td>AtGH3.7</td>
<td>At1g23160</td>
<td>74.9</td>
</tr>
<tr>
<td>AtGH3.8</td>
<td>At5g51470</td>
<td>62.3</td>
</tr>
<tr>
<td>AtGH3.18</td>
<td>At1g48670</td>
<td>60.1</td>
</tr>
<tr>
<td>VvGH3.8</td>
<td>GSVIV00006220001</td>
<td>59.9</td>
</tr>
<tr>
<td>PtGH3.8</td>
<td>POPTR_0000s15970</td>
<td>59.5</td>
</tr>
<tr>
<td>PtGH3.7</td>
<td>POPTR_0001s12850</td>
<td>58.5</td>
</tr>
<tr>
<td>AtGH3.19</td>
<td>At1g48660</td>
<td>57.3</td>
</tr>
<tr>
<td>AtGH3.17</td>
<td>At1g28130</td>
<td>54.1</td>
</tr>
<tr>
<td>AtGH3.14</td>
<td>At5g13360</td>
<td>53.3</td>
</tr>
<tr>
<td>AtGH3.15</td>
<td>At5g13370</td>
<td>51.9</td>
</tr>
<tr>
<td>AtGH3.16</td>
<td>At5g13380</td>
<td>51.7</td>
</tr>
<tr>
<td>AtGH3.13</td>
<td>At5g13350</td>
<td>50.5</td>
</tr>
</tbody>
</table>
Figure 5.5. *GH3.7* was inserted within Brassicales. A screenshot of the BLASTz output from the CoGe browser is shown, with features as in Figure 5.1. The genome origin is indicated on the right (At = Arabidopsis thaliana, Vv = Vitis vinifera, Cp = Carica papaya, Pt = Populus trichocarpa). Only *Arabidopsis* contains the GH3 gene.
Group I comprises two syntenic sets, IA and IB, each containing one AtGH3 protein (AtGH3.11 and AtGH3.10). Group II is more complex, comprising 4 syntenic sets. Sets IIA1 and IIA2 group into one clade, with the subclades approximately reflecting syntenic groups. The two syntenic sets are difficult to parse out, as there is some co-linearity of genes retained across the two sets, though weaker than within each set (data not shown, to recapitulate the analysis, see http://genomevolution.org/r/36p, showing synteny in IIA1, http://genomevolution.org/r/36r, showing synteny in IIA2, and http://genomevolution.org/r/180, showing synteny between IIA1 and IIA2.) AtGH3.2 and AtGH3.3 are in set IIA1 and AtGH3.1 and perhaps AtGH3.4 are in set IIA2. The co-linearity of genes in the region surrounding AtGH3.4 is quite weak, so this might not actually be syntenic but rather derive from an insertion. Several of the chromosomes in set IIA2 show a high frequency of transposons (13 in 200kB of At chromosome 2, 4 in 400 kB of Vv chromosome 3, and 2 in 200 kB of At chromosome 1). The similarities between the three *Mimulus* are interesting: one of the regions has similar genes with the two others, but those two share no similarities to each other. This is reflected in the identification of syntelogs in Table 5.2: notice that MgGH3.1 is identified as a syntelog of AtGH3.1, AtGH3.2, and AtGH3.3, while VvGH3.2 is only identified as a syntelog of AtGH3.2 and AtGH3.3 and VvGH3.1 is only identified as a syntelog of AtGH3.1.

Sets IIB1 and IIB2 clade together separately from IIA1 and IIA2. Set IIB1 has two AtGH3 proteins (AtGH3.5 and AtGH3.6), while set IIB2 only has one grape and one poplar protein with none from *Arabidopsis*. WES1 (AtGH3.5, At4g27260) is an interesting enzyme in that it has been shown to use both IAA and SA as a substrate (Staswick et al., 2002; Park et al., 2007; Zhang et al., 2007). The six Arabidopsis genes in sets IIA and IIB are those strongly induced by IAA. There is an additional syntenic set with members typically classified as Group II (AtGH3.9, PtGH3.4, PtGH3.9) here moved in to Group III. The set IIB proteins clade together separately from set IIIA, which contains PBS3 and AtGH3.17. Alternatively, Group III could be abolished and PBS3 and its syntelogs and paralogs could be classified as Group II. However, this approach removes the distinction of a separate group for PBS3, which is clearly functionally distinct from JAR1 and WES1 and the others. *As PBS3 and AtGH3.17 are syntelogs with each other and the same genes in grape and poplar, AtGH3.17 and AtGH3.9 are similar in sequence and clade together, and in light of expression analysis discussed below, it is most appropriate to retain Group III but place AtGH3.17 and AtGH3.9 in it.*

**Loss and Gain of GH3 Genes**

Overall, ten genomic regions of *Arabidopsis* flanking GH3 genes were identified as syntenic to chromosomal regions of other rosids. These include several areas of *Arabidopsis* chromosomes 1, 2, 4, and 5 and account for 15 of the 19 *Arabidopsis* GH3 (AtGH3) genes. GH3 genes from the rosids grape and poplar and the asterid *Mimulus guttatus* are described in Table 5.2 and their *Arabidopsis* syntelogs are indicated. Eight GH3 genes were identified in grape, thirteen in poplar (one newly described gene plus the twelve described in (Felten et al., 2009), six in papaya, and 10 in *Mimulus*. As discussed earlier, there may be additional *Mimulus* GH3 and papaya genes, as the sequencing and annotation of those genomes are still incomplete. The sequence of one *Mimulus* scaffold truncates immediately upstream of the position of the GH3 syntenic set IIIA gene in the other genomes (scaffold 39 in Table 5.1). Based on the number of syntenic sets identified here, it would seem that the common ancestor of all of these species had eight GH3 genes. Logically, this suggests that the common ancestor prior to the pre-rosid hexaploidy WGD event had three GH3 genes, leading to nine after the WGD event, one of which...
was lost in the lineage prior to the divergence of grape (Figure 5.6 and Table 5.4). Grape has retained all eight GH3 genes, but the other rosids analyzed have lost multiple genes. Uniquely, Arabidopsis has lost 21 copies but gained 7, due to whole genome duplication, fractionation, local duplication and transposition. The genes gained are all in syntenic set IIIA.

**Analysis of Group II and III Expression and Promoter Sequences**

Examination of expression data and transcription factor binding motifs in the GH3 genes can help provide insight into gene function. The first GH3 genes were identified as inducible by auxin (Hagen et al., 1984), years before their function was known. Subsequent analysis revealed that some GH3 enzymes catalyze the conjugation of amino acids to the auxin IAA (Staswick et al., 2005), but that other GH3 enzymes use other substrates. JAR1 is active on jasmonic acid (Staswick and Tiryaki, 2004) and PBS3 is active on 4-substituted benzoates (Chapter III) and inhibited by salicylic acid. On the whole, expression patterns of PBS3 are quite different than the other AtGH3 genes. Expression levels of many of the GH3 genes tend to be quite low overall. The expression pattern of other Group III genes is quite poorly correlated with that of PBS3, as determined using the two-gene scatter-plot function of NascArrays (data not shown). The expression of Group III genes was compared to that of Group II in AtGenExpress microarray experiments examining gene expression following application of 1 µM IAA, 10 µM JA, and 10 µM SA (Goda et al., 2008). Data were downloaded via the eFP browser (Toufighi et al., 2005). As the amount of SA used in this experiment was quite low, an additional experiment using the SA analogue 1,2,3-benzothiadiazole-7-carbothioic acid S-methyl ester (BTH) (Wang et al., 2006) was also included in the analysis. Genes showing greater than two-fold expression changes compared to control samples are indicated in Table 5.5. All of the genes characterized here as Group II on the microarray chip are induced at least 2-fold in response to treatment with IAA. This is in contrast to the two genes shown here in Group III but typically classified as Group II, AtGH3.9 and AtGH3.17. Neither of these is expressed in response to IAA, which has also been shown in other gene expression experiments (Staswick et al., 2005; Khan and Stone, 2007) (Hagen and Guilfoyle, 2002). JA does not seem to induce transcription of JAR1 in the same way that IAA induced transcription of the IAA-AA conjugases, as JAR1 was not observed to be induced by JA (data not shown). PBS3 in Group III and AtGH3.3 in Group II are the only two genes in these two groups induced in response to SA or BTH (Table 5.5). Expression of PBS3 in response to a higher level of SA (1 mM) has also been shown by RT-PCR (Jagadeeswaran et al., 2007).

PBS3 has been identified in several screens for genes involved in response to pathogens. Warren et al first identified the gene in a screen for mutants with enhanced susceptibility to P. syringae pathogens containing the avirulence gene avrPphB (Warren et al., 1999), naming it PBS3. We prioritized it based on correlation with expression of the SA-biosynthetic gene ICS1 and pathogen induction for publically available microarray databases. Jagadeeswaran found PBS3 when looking for genes upregulated in response to infection with P. syringae (2007), naming it GDG1. Lee et al identified PBS3 in a screen for proteins that bind to the P. syringae pv. maculicola avirulence factor HopW1-1 (2008), naming it WIN3. The other Group III genes have not been similarly identified in such screens. Are they involved in the defense response against pathogens? Analysis of gene expression data and potential transcription factor binding sites in promoters were used to evaluate the possible role of Group III genes in defense.
Figure 5.6. Number of GH3 genes in modern species overlayed on angiosperm phylogeny. Number of genes is shown in plant species and relevant ancestors. Maize is shown for completeness only as it was not analyzed.
<table>
<thead>
<tr>
<th>Syntenic group</th>
<th>At</th>
<th>Cp</th>
<th>Pt</th>
<th>Vv</th>
</tr>
</thead>
<tbody>
<tr>
<td>IA</td>
<td>-3</td>
<td>NC</td>
<td>-1</td>
<td>NC</td>
</tr>
<tr>
<td>IB</td>
<td>-3</td>
<td>NC</td>
<td>NC</td>
<td>NC</td>
</tr>
<tr>
<td>IIA1</td>
<td>-2</td>
<td>NC</td>
<td>-1</td>
<td>NC</td>
</tr>
<tr>
<td>IIA2</td>
<td>-2</td>
<td>NC</td>
<td>NC</td>
<td>NC</td>
</tr>
<tr>
<td>IIB1</td>
<td>-2</td>
<td>NC</td>
<td>NC</td>
<td>NC</td>
</tr>
<tr>
<td>IIB2</td>
<td>-4</td>
<td>-1</td>
<td>-1</td>
<td>NC</td>
</tr>
<tr>
<td>IIIA*</td>
<td>-2/+7</td>
<td>-1</td>
<td>-1</td>
<td>NC</td>
</tr>
<tr>
<td>IIIB</td>
<td>-3</td>
<td>-1</td>
<td>NC</td>
<td>NC</td>
</tr>
<tr>
<td><strong>Total:</strong></td>
<td><strong>19</strong></td>
<td><strong>5</strong></td>
<td><strong>12</strong></td>
<td><strong>8</strong></td>
</tr>
</tbody>
</table>

NC = no change, At = Arabidopsis thaliana, Cp = Carica payapa, Pt = Populus trichocarpa, Vv = Vitis vinifera. * Set IIIA contains PBS3.
TABLE 5.5
Expression of AtGH3 genes in response to hormone treatment

<table>
<thead>
<tr>
<th></th>
<th>AtGH3.12</th>
<th>AtGH3.7</th>
<th>AtGH3.8</th>
<th>AtGH3.13</th>
<th>AtGH3.14</th>
<th>AtGH3.15</th>
<th>AtGH3.16</th>
<th>AtGH3.18</th>
<th>AtGH3.19</th>
<th>AtGH3.9</th>
<th>AtGH3.17</th>
<th>NASCArrays Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hormone: Treatment of seedlings with 1 µM IAA</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NA</td>
<td>NA</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>175</td>
</tr>
<tr>
<td>Hormone: Treatment of seedlings with 10 µM MeJA</td>
<td>NS</td>
<td>NS</td>
<td>+</td>
<td>NS</td>
<td>NA</td>
<td>NA</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>174</td>
</tr>
<tr>
<td>Hormone: Treatment of seedlings with 10 µM SA</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>192</td>
</tr>
<tr>
<td>Hormone analog: Treatment of adult leaves with 60 µM BTH</td>
<td>+</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NA</td>
<td>NA</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>392</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>AtGH3.6</th>
<th>AtGH3.5</th>
<th>AtGH3.3</th>
<th>AtGH3.2</th>
<th>AtGH3.4</th>
<th>AtGH3.1</th>
<th>NASCArrays Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group II</td>
<td>DFL1</td>
<td>WES1</td>
<td>YDK1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hormone: Treatment of seedlings with 1 µM IAA</td>
<td>+</td>
<td>+</td>
<td>++</td>
<td>NA</td>
<td>NA</td>
<td>++</td>
<td>175</td>
</tr>
<tr>
<td>Hormone: Treatment of seedlings with 10 µM MeJA</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>NA</td>
<td>NA</td>
<td>NS</td>
<td>174</td>
</tr>
<tr>
<td>Hormone: Treatment of seedlings with 10 µM SA</td>
<td>NS</td>
<td>NS</td>
<td>+</td>
<td>NA</td>
<td>NA</td>
<td>NS</td>
<td>192</td>
</tr>
<tr>
<td>Hormone analog: Treatment of adult leaves with 60 µM BTH</td>
<td>NS</td>
<td>NS</td>
<td>+</td>
<td>NA</td>
<td>NA</td>
<td>NS</td>
<td>392</td>
</tr>
</tbody>
</table>

Values shown are significant at $\alpha = 0.05$ or consistent between replicates; NS is not significant.

+ is $\geq 2$-fold increase in expression compared to control treatment, ++ is $\geq$ ten fold increase in expression compared to control treatment.

- is $\geq 2$-fold decrease in expression compared to control experiment; NA is not on array.
<table>
<thead>
<tr>
<th></th>
<th>AtGH3.12</th>
<th>AtGH3.7</th>
<th>AtGH3.8</th>
<th>AtGH3.18</th>
<th>AtGH3.19</th>
<th>AtGH3.13</th>
<th>AtGH3.14</th>
<th>AtGH3.15</th>
<th>AtGH3.16</th>
<th>AtGH3.17</th>
<th>AtGH3.19</th>
<th>NASC-Arrays</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pathogen: Infection of adult leaves with virulent <em>Pst</em></td>
<td>++</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NA</td>
<td>NA</td>
<td>NS</td>
<td>+</td>
<td>NS</td>
<td>120</td>
<td></td>
</tr>
<tr>
<td>Pathogen: Infection of adult leaves with avirulent <em>Pst</em></td>
<td>++</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NA</td>
<td>NA</td>
<td>NS</td>
<td>++</td>
<td>NS</td>
<td>120</td>
<td></td>
</tr>
<tr>
<td>Pathogen: Infection of adult leaves with nonhost <em>Psp</em></td>
<td>++</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NA</td>
<td>NA</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>120</td>
<td></td>
</tr>
<tr>
<td>Pathogen: RT-PCR: Infection of adult leaves with <em>Pst avrRpm1</em></td>
<td>+</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NA</td>
<td>NA</td>
<td>NS</td>
<td>NT</td>
<td>NT</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>Pathogen: Infection of adult leaves with <em>Botrytis cinerea</em></td>
<td>NS</td>
<td>+</td>
<td>NS</td>
<td>+</td>
<td>NS</td>
<td>NA</td>
<td>NA</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>167</td>
<td></td>
</tr>
<tr>
<td>Pathogen: Infection of adult leaves with <em>Phytophthora infestans</em></td>
<td>+</td>
<td>NS</td>
<td>-</td>
<td>NS</td>
<td>NS</td>
<td>NA</td>
<td>NA</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>123</td>
<td></td>
</tr>
<tr>
<td>Abiotic stress: Response of seedlings to cold/osmotic/salt/drought</td>
<td>+</td>
<td>+</td>
<td>NS</td>
<td>+</td>
<td>+</td>
<td>NS</td>
<td>NA</td>
<td>NA</td>
<td>NS</td>
<td>NS</td>
<td>138-141</td>
<td></td>
</tr>
</tbody>
</table>

Values shown are significant at $\alpha = 0.05$; NS is not significant.

+ is $\geq$ 2-fold increase in expression compared to control treatment, ++ is $\geq$ ten fold increase in expression compared to control treatment.

- is $\leq$ 2-fold decrease in expression compared to control experiment; NA is not on array.

* is from data published in (Jagadeeswaran et al., 2007)
The eFB browser (Toufighi et al., 2005; Winter et al., 2007) and Genevestigator (Zimmermann et al., 2004) were used to determine expression in response to abiotic and biotic stress in the AtGenExpress microarray datasets. The results of these analyses are summarized in Table 5.6. Expression is said to be elevated in response to a treatment when the expression level is at least two-fold higher than a negative control, and the expression levels between replicates are consistent and statistically significant at \( \alpha = 0.05 \) when there are three replicates.

While PBS3 is strongly induced following infection of leaf tissue with virulent, avirulent, and non-host strains of the bacterium *Pseudomonas syringae*, the only other Group III gene induced following infection with *P. syringae* is AtGH3.17. This is a trait shared with a subset of AtGH3 genes encoding enzymes that use IAA as a substrate: AtGH3.3, AtGH3.5 and AtGH3.6 are also induced following *P. syringae* infection (data not shown). The results of RT-PCR experiments testing for expression of Group III AtGH3 genes following infection with *P. syringae* support the analysis of microarray data; only PBS3 of the genes tested was expressed. The expression of AtGH3.17 and AtGH3.9 were not tested by the investigators as they were previously classified as Group II GH3 genes. (Jagadeeswaran et al., 2007) As shown in Table 5.6, a few genes were expressed in response to infection with the necrotrophic fungi *Bortyris cinerea*, including AtGH3.7, which encodes the protein with the highest percent identity to PBS3 (Table 5.3), and AtGH3.18. In addition, several of the AtGH3 genes are induced in response to abiotic stress, such as the related cold/osmotic/salt/drought stressors (Table 5.6). Several of the genes in Groups I and II are also induced under these conditions (data not shown).

One common explanation for retention of genes following duplication events is due to subfunctionalization resulting in distinct tissue specificity (Force et al., 1999). In order to investigate whether duplicated Group III GH3 genes have evolved unique tissue specificity, the expression patterns were compared in microarray data from different tissues and developmental stages: cotyledons, rosette leaves, roots, flowers, and seeds (Schmid et al., 2005). Senescing leaves were included as well although they are not a developmental stage per se; those genes expressed in senescing tissue are often associated with response to abiotic and biotic stress (Lin and Wu, 2004). As shown in Table 5.7, most of the genes are expressed in several types of tissue, although a few are expressed in only one type. AtGH3.1 is expressed in the root while AtGH3.14 is expressed only in the seed. Of the tissues shown, PBS3 is only expressed in senescing leaves and flowers. In general, PBS3 is weakly expressed except under conditions of abiotic and biotic stress (Chapter II). While AtGH3.9 has been described as having a role in the root (Khan and Stone, 2007), its expression was identified in the seed in these experiments. This may be due to expression differences in plants of specific ages or under specific experimental conditions.

Resources for analyzing microarray datasets are currently less developed for other plant species compared to *Arabidopsis*. In a published microarray experiment of poplar (*Populus tremula* x *Populus alba*) tree roots interacting with the ectomycorrhizal fungus *Laccaria bicolor*, the PBS3 syntelogs *PtGH3.7*, and *Pt.GH3.8* as well as Group II *PtGH3.1, PtGH3.2* were all induced (Felten et al., 2009). Other auxin response genes were also induced, and the authors noted that the auxin gradient in the root was altered by interaction with the fungus. The Plant Expression Database (Plexdb, http://plexdb.org (Wise et al., 2007)) does contain data from a limited number of poplar and grape experiments. Expression of the poplar Group III genes *PtGH3.7* (probeset PtAffx.140928.1.A1_at) and *PtGH3.8* (probeset PtAffx.210014.1.S1_at) were not significantly elevated in response to infection with the *Melampsora* rust fungus. No probesets were identified for the grape Group III gene *VvGH3.8* or for the *AtGH3.9* syntelogs.
Genes expressed in senescing leaves are typically also associated with response to abiotic and biotic stress. Values shown are significant at $\alpha = 0.05$.

+ is $\geq 2$-fold increase in expression compared to control treatment, ++ is $\geq$ ten fold increase in expression compared to control treatment.

NASCArrays reference and slide numbers are shown for each experiment.

**TABLE 5.7**
Expression of Group II and III AtGH3 genes in various tissue types

<table>
<thead>
<tr>
<th>Group III</th>
<th>Cotyledon</th>
<th>Rosette</th>
<th>Senescing leaf*</th>
<th>Root</th>
<th>Flower</th>
<th>Seed</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AtGH3.12</strong></td>
<td>++</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>AtGH3.7</strong></td>
<td>+</td>
<td>++</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>AtGH3.8</strong></td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>AtGH3.18</strong></td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>AtGH3.19</strong></td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td>+</td>
<td></td>
</tr>
<tr>
<td><strong>AtGH3.16</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>AtGH3.14</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+</td>
</tr>
<tr>
<td><strong>AtGH3.13</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>++</td>
<td></td>
</tr>
<tr>
<td><strong>AtGH3.17</strong></td>
<td>+</td>
<td>+</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td></td>
</tr>
<tr>
<td><strong>AtGH3.9</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>++</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group II</th>
<th>Cotyledon</th>
<th>Rosette</th>
<th>Senescing leaf</th>
<th>Root</th>
<th>Flower</th>
<th>Seed</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AtGH3.6</strong></td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td>+</td>
<td></td>
</tr>
<tr>
<td><strong>AtGH3.5</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+</td>
</tr>
<tr>
<td><strong>AtGH3.3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+</td>
<td></td>
</tr>
<tr>
<td><strong>AtGH3.2</strong></td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>AtGH3.1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+</td>
</tr>
</tbody>
</table>

NASCArrays: 150-1 150-89 150-25 151-9 152-39 154-78
VvGH3.5. There was data for VvGH3.2 (probeset: 1610880 s_at) from Group II, which was elevated in response to infection with Bois Noir phytoplasma (Albertazzi et al., 2009). In summary, a subset of Group III and II genes across species, including PBS3 syntelogs, were found to be induced in response to biotic stress.

Plant transcription factors associated with response to biotic stress include WRKY proteins, ethylene responsive element binding factors (ERF), basic-domain leucine-zipper (bZIP), MYB proteins, and DRE-binding transcription factors (Singh et al., 2002). Auxin Response Factors (ARFs) are known to be involved in auxin response (Guilfoyle and Hagen, 2007). These transcription factors recognize known DNA motifs that are catalogued in databases, including the Arabidopsis cis-regulatory element database (AtcisDB) (Molina and Grotewold, 2005) for Arabidopsis promoters. Motifs potentially recognized by the transcription factors listed above were identified in the first 1 kB upstream of the transcriptional start sites of AtGH3 genes (Table 5.8). Several of the tandemly duplicated genes have fewer than 1 kB between the next gene: AtGH3.15 (288 bp), AtGH3.16 (711 bp) and AtGH3.18 (989 bp). The DRE-motif only occurs in Group III genes and motifs recognized by ARFs in Group II and Group III. There are no other statistical patterns of motifs enriched in the different GH3 groups. Defense-associated WRKYS occur in all Group I, almost all Group III, and several Group II promoters. Comparing duplicated genes resulting from both tandem and segmental duplication events shows evidence for promoter evolution, as frequency of motifs varies between the duplicated genes.

DISCUSSION

Past analyses of the GH3 gene family relying on global sequence similarities (Staswick et al., 2005) (Felten et al., 2009) and search of EST sequences (Terol et al., 2006) suggest that Group III genes, including PBS3, are only present in a small number of species compared to those in Group I and Group II. In this case, when did PBS3 (AtGH3.12, At5g13320) arise? The results discussed in this chapter show that the genomic region surrounding PBS3 and four nearby locally duplicated genes on At chromosome 5 and AtGH3.17 (At1g28130) on At chromosome 1, contains genes arranged in a co-linear order to regions in poplar, grape, and rice, suggesting descent from a common ancestral chromosome dating to before the dicot/monocot split (see Figures 5.2 and 5.3 and Table 5.1). PBS3 and its syntelogs cluster into two sub-clades in phylogenetic trees constructed using the maximum likelihood (ML) method (Figure 5.4). My analysis of synteny and phylogeny indicate that PBS3 is closely related to genes disturbed widely in plants. However, it is not clear that syntelogs of PBS3 share its function; they may be more functionally similar to AtGH3.17, which encodes an IAA-conjugating enzyme (Staswick et al., 2005). The combined analyses suggest that PBS3 evolved from IAA-conjugating enzymes. Analysis of expression data for Group II and III genes supports this hypothesis. Group III has expanded in Arabidopsis, which contains several Group III genes that arose through local duplication and transposition that are not present in other sequenced rosids. The expansion through tandem duplication and insertion is characteristic of a group of genes that respond to environmental stress.

Analysis of Phylogenetic Trees of the GH3 family

A careful analysis of sequence data reconciled with syntenic analysis indicates that GH3 phylogeny is not as straightforward as typically presented. Previous analyses of the GH3 gene family have relied on global sequence similarities (Staswick et al., 2005) (Felten et al., 2009)
Predicted cis-acting regulatory element binding motifs in AtGH3 promoters (1kB) as identified using AtCisDB. Sequences of motifs are described in Materials and Methods.

* Promoters significantly < 1 kB.

<table>
<thead>
<tr>
<th>Name</th>
<th>Group</th>
<th>W-Box</th>
<th>MYB</th>
<th>bZIP</th>
<th>DRE</th>
<th>ARF</th>
</tr>
</thead>
<tbody>
<tr>
<td>AtGH3.11</td>
<td>I</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>AtGH3.10</td>
<td>I</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>AtGH3.1</td>
<td>II</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>AtGH3.2</td>
<td>II</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>AtGH3.3</td>
<td>II</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>AtGH3.4</td>
<td>II</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>AtGH3.5</td>
<td>II</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>AtGH3.6</td>
<td>II</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>AtGH3.9</td>
<td>III</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>AtGH3.17</td>
<td>III</td>
<td>3</td>
<td>6</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>AtGH3.12</td>
<td>III</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>AtGH3.7</td>
<td>III</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>AtGH3.16*</td>
<td>III</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>AtGH3.14</td>
<td>III</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>AtGH3.15*</td>
<td>III</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>AtGH3.8</td>
<td>III</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>AtGH3.18</td>
<td>III</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>AtGH3.19</td>
<td>III</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Representation of the data from Table 5.8, showing average motif frequencies for each GH3 Group. Error bars are ± 1 standard deviation.
(Terol et al., 2006), which are not necessarily indicative of true evolutionary descent. The increasing number of sequenced genomes and new comparative genomic tools makes it possible to use syntenity to evaluate phylogenetic trees. For example, Jun et al. recently evaluated the use of local syntenity for identifying orthologous genes in mammals, and found it quite effective, particularly in cases of local duplication and transposition (Jun et al., 2009). I used the syntenic data to guide the choice of phylogenetic methods in order to best understand how PBS3 is related to the other GH3 genes. I constructed phylogenetic trees using neighbor-joining (NJ), maximum parsimony (MP), and ML methods with various alignment curation methods. The tree constructed via ML best fit the data from syntenic analysis (Figure 5.4). A comparison of tree topology using the same curated multiple sequence alignment with three different tree construction methods is shown in Figure 5.7. The overall topology with Group I in one clade and Group II and III in another is common to all trees constructed using different algorithms. Where the trees differ is the relationship of AtGH3.9, AtGH3.17, PBS3, and their syntenogs in poplar and grape. The major difference between the topology of the NJ tree compared to the ML tree is that syntenic set IIIA is divided into two subclades, both basal to IIIB, IIB, and IIA, in the NJ tree. AtGH3.17, VvGH3.8, PtGH3.7, and PtGH3.8 group together in one subclade (1), while PBS3 and the other members of syntenic set IIIA, all Arabidopsis proteins, group in a second subclade (2) basal to the first and the Group II proteins. While this grouping reflects overall sequence similarity, it does not correspond to the syntenic data. This NJ tree would suggest that the genes in syntenic sets IIA, IIB, IIIB, IIIA (subclade 1) are descended from the same ancestral gene, but not IIIA (subclade 2). However, the synteny data suggests that IIIA subclades 1 and 2 are descended from the same ancestral gene, but IIB and IIA and IIB are not. The MP method fares worse than the NJ method. The MP tree separates Group III sequences into 4 clades, two with only one sequence each, and groups AtGH3.17 and the other PBS3 syntenogs as basal to both Group I and Group II (data not shown). A MP tree with a manually curated alignment handles Group III better, but has very low statistical support (data not shown). Based on this analysis, the tree constructed using the ML method was chosen as the one best fitting all of the available data.

The ML method is a discrete data method that begins with a model of rates of evolutionary change and alters the model until it fits the observed data (Mount, 2004). This is in contrast to NJ and UPGMA, which are distance methods. In distance methods, the percentage of aligned positions that differ between two sequences is computed pair-wise for all sequences in an alignment. These values are arranged such that the most the sequences with lower difference scores are closer together on the tree (Mount, 2004). Distances methods tend to be favored by molecular biologists as they are straightforward and computationally efficient. The drawback of distance methods is that they can be misleading when using an incorrect evolutionary model (Huelsenbeck and Hillis, 1993). While ML methods are traditionally computationally intensive, new algorithms, such as that employed by PhyML (Guindon and Gascuel, 2003) reduce calculation time sufficiently to allow for routine use. There is considerable argument in the literature over the “best” method for phylogenetic analysis, often defined as how much a method can tolerate violations of its assumptions (Huelsenbeck and Hillis, 1993). For example, there is a contentious debate over the relative merits of ML and maximum particularly in the case where different positions in a sequence evolve over different rates (Steel, 2005). Many experts suggest constructing phylogenetic trees using multiple methods and evaluate them carefully.
Figure 5.7. Comparison of topology of condensed phylogenetic trees using protein sequences from rosids constructed using different methods. A, Maximum likelihood tree  B, Neighbor-joining tree  C, Maximum parsimony tree. The maximum likelihood is most consistent with the synteny analysis. Branch labels and colors represent syntenic groups, as defined in Figure 5.4. The clades notated with * contain PBS3. The number of sequences corresponding to each condensed branch is shown to the right of the tree.
When available, addition of data from synteneic analysis can help determine the choice of phylogenetic method. I have found that, in this case, the ML method corresponds best with the analysis of synteny, clustering PBS3 with a subset of the proteins previously classified as Group II.

**Classification of GH3 genes into Groups II and III**

As PBS3 of Group III and AtGH3.17, typically classified as Group II are syntelogs, this raises the question of whether the classification is accurate. How should PBS3, AtGH3.17, their syntelogs in other species, and the related AtGH3.9 be classified in light of the analyses described in this chapter? Functional data can also be used to aid in classification. While these two are typically classified as Group II, they demonstrate differences from other Group II GH3 genes. Unlike the others, they are not generally expressed in response to auxin (Hagen and Guilfoyle, 2002) (Goda et al., 2008). In one report, expression of AtGH3.9 is instead repressed by low levels of IAA (Khan and Stone, 2007). In published neighbor-joining phylogenetic trees, the two tend to clade separately from the other Group II proteins (Staswick et al., 2005), as do their rice (OsGH3.13 and OsGH3.11) (Terol et al., 2006) and poplar (PtGH3.7 - 9) orthologs (Felten et al., 2009). This is also true for the one example in the literature using a consensus tree using both the distance and MP methods (Liu et al., 2005). On the other hand, enzymatic assays support a role for AtGH3.17 in adenylation (Staswick et al., 2002) and conjugation (Staswick et al., 2005) of IAA and AtGH3.9 at least for adenylation activity. Also, plants with loss-of-function mutations in the genes, like plants with mutations in the other Group II genes, are more sensitive to auxin-mediated root inhibition than are wild-type plants (Staswick et al., 2005; Khan and Stone, 2007).

If classification is done only by grouping enzymes with similar enzymatic substrates, that does not group sequences satisfactorily either. For example, it does not account for Group I, in which JAR1 (AtGH3.11) is active on JA but DFL2 (AtGH3.10), is not. The synteny, phylogenetic tree, and functional data need to be reconciled in some way. Returning to the rosid GH3 ML phylogenetic tree in Figure 5.4, there are two clades with high statistical support beginning at the root/base of the tree. One of these is Group I, which in turn forms two subclades: one with JAR1 and its syntelogs (IA) and one with DFL2 and its syntelogs (IB). The second major clade contains the rest of the sequences, which also divides into two subclades: one containing the majority of the traditional Group II sequences (set IIA1, IIA2, IIB1, and IIB2), and one containing PBS3 and its syntelogs, including AtGH3.17 (IIIA); as well as AtGH3.9 and its syntelogs (IIIB). The most parsimonious explanation for the presence of IAA-conjugating enzymes in both the II and IIIA groups is that this ancestral gene encoded an enzyme that used IAA as a substrate, which has since evolved in the Arabidopsis lineage. This would be further supported if VvGH3.8, PtGH3.7, and PtGH3.8 were also active on IAA. It is unclear whether PBS3 or AtGH3.17 are more closely related to their syntelogs in grape and poplar. Based simply on overall sequence similarity as calculated by PID, AtGH3.17 is more similar to the grape and poplar syntelogs than is PBS3. However, in trees constructed using the ML but with different alignments, AtGH3.17 is sometimes monophyletic with VvGH3.8, PtGH3.7, and PtGH3.8 and sometimes is polyphyletic, as in Figure 5.4. This leaves room for the possibility that the orthologs in other species could be more like PBS3 than AtGH3.17. This suggests that PBS3 has evolved a unique role in Arabidopsis or its near relatives. Additional functional analyses with the uncharacterized genes, particularly biochemical assays and pathogen susceptibility assays, would help address this issue.
**Group III is Enriched in Local and Segmental Duplications**

Based on the eight syntenic sets identified for GH3 genes in rosids, it suggests that the common ancestor of all of these species had eight GH3 genes. This can be understood if common ancestor prior to the pre-rosid hexaploidy WGD event had three GH3 genes, leading to nine after the WGD event, one of which was lost in the lineage prior to the divergence of grape. Grape has retained all eight GH3 genes, but the other rosids analyzed have lost multiple genes (Figure 5.6 and Table 5.4). Uniquely, Arabidopsis has lost 21 copies but gained 7, due to whole genome duplication, fractionation, local duplication and transposition. The genes gained in Arabidopsis all clade with PBS3 in syntenic set IIIA (Figure 5.4). AtGH3.7 and AtGH3.8 were inserted into their location in the genome, as evidenced by corresponding regions in papaya, grape, and poplar (shown in Figure 5.5 for AtGH3.7). AtGH3.13 - .16 were locally duplicated, likely from the ancestral gene of PBS3. The sequences of the proteins encoded by the four duplicated genes are more similar to each other than they are to PBS3 (Table 5.3) and clade separately to PBS3 in all phylogenetic trees. The genes AtGH3.18 and AtGH3.19 represent a tandem duplication, and likely inserted into the chromosome as well, although the corresponding region in the other genomes has yet not been identified. All of these rearrangements and insertions suggest a rapidly evolving group of genes. Researchers have shown that genes that respond to environmental stress are overrepresented in arrays of tandemly duplicated genes (Rizzon et al., 2006; Hanada et al., 2008). In addition, the same genes that tend to be locally duplicated also tend to transpose as well within chromosomal regions with high recombination rates. By one estimate, more than half of all genes in the Arabidopsis genome have transposed since the origin of the brassicales. Different types of genes appear to be retained following duplication by local or whole genome duplication. Several researchers have shown that those Arabidopsis genes that tend to be retained following WGD tend to encode regulatory proteins such as transcription factors, protein kinases and ribosomal proteins, while those not retained tend to include basal functions like DNA repair. These results typically agree in Arabidopsis and poplar. The genes over-retained following WGD tend to be under-retained following tandem duplication. (Reviewed in (Freeling, 2009)).

Much of the work studying the divergence in expression of duplicated genes has been done in yeast. Some of the work in plants is inconsistent, perhaps due in part to combining data from genes duplicated in different ways (Li et al., 2005). However, expression patterns of tandemly and segmentally duplicated genes in Arabidopsis have been found to be similar, with a weak but significant correlation between expression pattern and promoter similarity (Haberer et al., 2004). The expression of the tandemly duplicated sets of Arabidopsis genes in the GH3 family, AtGH3.12 and AtGH3.13 – 16; and AtGH3. 18 and AtGH3.19, do not retain correlated expression (data not shown). The expression of the genes in different types of tissues suggests possible specialization (Table 5.7). In addition, they differ in frequency of putative transcription factor binding motifs (Table 5.8), suggesting evidence for promoter evolution.

Analysis of expression data supports evidence from genomic analysis suggesting uncharacterized Group III genes play a role in response to environmental stress. Expression of AtGH3.7 and AtGH3.18 are elevated in response to Botrytis cinerea and several genes are upregulated in response to abiotic stress (Table 5.6) and in senescing leaves (Table 5.7). However, the evidence does suggest that PBS3 is unique in the extent of its involvement in SA-dependent responses. The expression of PBS3 is highly correlated to that of the SA biosynthetic gene ICS1 and is elevated in response to a variety of pathogens and their effectors, particularly Pseudomonas syringae (Chapter II and Table 5.6) as well as SA and its analogs (Table 5.5).
The rare effector HopW1-1 from *Psm* has been shown to bind to PBS3 protein (Lee et al., 2008), indicating unique targeting of PBS3.

CONCLUSION

GH3 genes are widespread in plants. Retention of genes following whole genome duplication, tandem duplication, and transposition has led to many copies in modern plants. Combining analyses of comparative genomics, phylogeny, and gene expression data provided new insight into the evolutionary history of the GH3 family of amino acid conjugases. In particular, my analysis suggests that PBS3 is more closely related to genes in distantly-related species than previously thought. Analysis of co-linear genes can suggest a relationship between genes obscured by sequence similarity alone, as in the syntelogs PBS3 and AtGH3.17. The expansion of Group III in *Arabidopsis*, in which PBS3 belongs, has occurred through both tandem duplication and transposition, reflective of a role in response to environmental stress. Analysis of functional and expression data suggest a role for uncharacterized Group III genes in responding to abiotic and biotic stress, and provide new directions for future work.

REFERENCES


Chapter VI: Conclusion

SUMMARY

The PBS3\(^1\) gene of Arabidopsis thaliana was first identified in a screen for mutants with enhanced susceptibility to the bacterial pathogen Pseudomonas syringae carrying the effector avrPphB (Warren et al., 1999). My early work showed that PBS3 plays an important role in pathogen-induced SA metabolism and that this function is critical to SAG accumulation, PR1 expression, and both basal and R-protein mediated resistance (Chapter II). Given this role of PBS3 in SA metabolism and the function of other characterized GH3 family members as small molecule-amino acid synthetases (Staswick and Tiryaki, 2004; Staswick et al., 2005), it seemed likely that PBS3 would act either upstream of SA biosynthesis, on SA, or on a competitive inhibitor of SA. In vitro assays were undertaken to test these ideas and identify the function of the PBS3 enzyme (Chapter III). PBS3 is the first group III GH3 enzyme to be so characterized. Using a high throughput assay, PBS3 was found to be highly active with 4-HBA and pABA, two small molecules known to occur in Arabidopsis. The reaction kinetics of PBS3 confirmed that these two substrates could serve as in planta substrates. Both of these molecules are intermediates in the biosynthesis of important molecules. pABA is a component of folic acid, an essential co-factor in one carbon metabolism. 4-HBA is an intermediate of ubiquinone biosynthesis and is also elevated in plant cell walls in response to pathogen infection (Hagemeier et al., 2001). The results of the in vitro analyses of PBS3 suggest a potential, under-explored role for these molecules and their respective amino acid conjugates in plant-microbe interactions. While PBS3 was not found to use SA as a substrate, SA and its analog INA specifically inhibit PBS3 activity at low concentrations. This is important as it presents a novel mechanism for rapidly and reversibly fine-tuning the activity of small molecules, including phytohormones. It could also be a mechanism for rapid and reversible cross-talk between phytohormones or other small molecules, such as SA and 4-HBA.

A model of our current understanding of PBS3 function in SA metabolism and pathogen response is shown in Figure 6.1. In response to pathogens such as P. syringae, free SA is synthesized by ICS1 and accumulates. At low levels of SA (~20 μM), SA may be converted to the more volatile MeSA, consistent with the K\(\text{m}\) of the Arabidopsis SAMeT (At3g11480) for SA, which is 16 μM (Chen et al., 2003). Furthermore, the reverse reaction, catalyzed by a MeSA esterase, is dramatically inhibited by low concentrations of free SA (95% inhibition by 2 μM SA for the characterized tobacco enzyme (Forouhar et al., 2005)). Once sufficient free SA accumulates (~200 μM), SAG is formed, SA-dependent alteration of gene expression occurs (shown in the model as the activation of PR1 expression), and negative feedbacks on SA biosynthesis are observed. The K\(\text{m}\) of the putative SA glucosyltransferase SGT1 (UGT 74F1; At2g43840) for SA (190-230 μM) is consistent with this model (Lim et al., 2002; Song, 2006). In what appears to be a negative feedback loop, SA inhibits the activity of PBS3 (Figure 3.5),

---

\(^1\) Abbreviations. Compounds: IAA, indole-3-acetic acid; JA, jasmonic acid; SA, salicylic acid; SAG, salicylic acid glucoside. Genes: NPR1, non-expressor of pathogenesis related 1; eds16, enhanced disease susceptibility 16; GH3, Gretchen Hagen 3; ICS1, isochorismate synthase 1; JAR1, jasmonic acid resistant 1; MeSAE= methyl salicylate esterase; PBS3, avrPphb susceptible 3; PR1, pathogenesis related 1; SAGT, Salicylic acid glucosyltransferase. Organisms: Psm, Pseudomonas syringae pv. maculicola; Pst, Pseudomonas syringae pv. tomato.
with 91% inhibition by 150 µM SA. Another example of negative feed-back on SA accumulation occurs through the SA-response master regulator NPR1 (non-expressor of PR genes) (Cao et al., 1997) as evidenced by elevated SA in npr1 mutants (Wildermuth et al., 2001). Positive feedbacks on SA synthesis (e.g. via PAD4 (Jirage et al., 1999)) are not shown in the model, as the level of SA required for these responses is unclear.

As pbs3 mutants are compromised in accumulation of combined free and glycosylated SA, it is clear that PBS3 plays a role upstream of SA biosynthesis and accumulation. Mechanistically, how might PBS3 impact SA metabolism? The most likely acyl substrates of PBS3 are both structurally similar to SA and derived from chorismate. Therefore, PBS3 function could impact SA biosynthesis by altering the enzymatic activity of SA biosynthentic enzymes, substrate availability, or the amounts of SA biosynthetic proteins. My results described in Chapter III suggest that PBS3 does not influence SA accumulation by affecting expression of ICS1 or by altering activity of the enzyme it encodes. However, the effect of PBS3 on ICS1 protein levels has not yet been tested. This hypothesis is discussed in more detail below.

Alternatively, PBS3 function could alter chorismate partitioning, increasing the flux of available chorismate for SA biosynthesis. The expression and activity of chorismate-utilizing enzymes such as chorismate mutase, anthranilate synthase, aminodeoxychorismate synthase, and chorismate pyruvate lyase tend to be highly regulated (Herrmann and Weaver, 1999). Indeed, in the case of aromatic amino acid biosynthesis, products of one pathway can allosterically modulate the activity of the other chorismate-utilizing enzyme. For example, tryptophan can enhance chorismate mutase (CM) activity required for phenylalanine and tyrosine biosynthesis; whereas, phenylalanine and tyrosine inhibit CM activity (Mobley et al., 1999). pABA, 4-HBA, and SA are structurally similar and all derive from chorismate; therefore, it is possible, that they could act as competing substrates or to modulate the activity of each other’s biosynthetic enzymes. As discussed above, there is precedence for these compounds to act as competing substrates for enzymes such as glucosyltransferases and CoA ligases. On the other hand, though product inhibition has been assessed (e.g., (Stadthagen et al., 2005; Sahr et al., 2006)), there has been little work examining potential cross-regulation of these biosynthetic enzymes by these compounds. My preliminary evidence suggests PBS3 is capable of influencing accumulation of some indoles (Figure 4.4), suggesting crosstalk through the chorismate pathway. However, I did not detect major changes in general phenylpropanoids or flavonoids.

While identification of in vitro substrates of PBS3 permitted a search for the reaction product in planta, this product has not yet been detected (Chapter IV). It is likely produced transiently and at very low concentrations. However, one compound was found which is specifically elevated in leaf tissue of pbs3 mutants in response to Pst avrRpt2. This compound is quite water-soluble, highly fluorescent, weakly UV-absorbent, and of similarity polarity to glycosylated phenylpropanoids. This compound could function as a novel signaling molecule, a substrate of PBS3, or a compound from a competing pathway that uses the substrate of PBS3. The isolation of this compound has been hampered by interference from other similar compounds; a two-stage HPLC method has been developed to facilitate its isolation. Once purified in sufficient quantities, the compound can be analyzed by MS/MS and NMR in an effort to determine its identity. Identification of this compound would provide important insights into PBS3 function in basal resistance to pathogens.

From the start of this project, comparison of PBS3 with its homologs in the GH3 family has provided clues into its function. Broadening this study to include GH3 family members in other species in a comparative genomics framework further enhances understanding of PBS3.
Figure 6.1. **Position of PBS3 in pathogen-induced SA metabolism.** SA biosynthesis is bracketed. Free SA is synthesized via ICS1. At low levels of SA (~20 µM), SA may be converted to the more volatile MeSA. The conversion of MeSA to SA occurs only in the absence of significant SA, as MeSAE activity is inhibited 95% by 2 mM SA. Once sufficient free SA accumulates, shown as the larger SA (~200 mM), SA is converted to SAG (the dominant accumulated form) and SA-dependent changes in gene expression occur. Negative feedback regulation of SA biosynthesis via PBS3 and an NPR1-dependent pathway also occurs. Positive amplification of SA biosynthesis is not shown due to lack of understanding of the relative SA levels required. Blue lines indicate regulation. The size of SA, MeSA, and SAG indicates their relative abundance. SA= salicylic acid, MeSA= methyl salicylate, SAG= SA glucoside, pABA= para-aminobenzoic acid, 4-HBA= 4-hydroxybenzoic acid. ICS1= isochorismate synthase, IPL= isochorismate pyruvate lyase, MeSAE= methyl salicylate esterase, PBS3= avrPphB susceptible 3 (GH3.12), SAMeT= SA methyltransferase, SGT1= SA glucosyl transferase. NPR1= non-expressor of PR genes is an ankyrin repeat master regulator of SA-dependent defense responses.
Analysis of co-linear genes can suggest a relationship between genes obscured by sequence similarity alone, as in the syntelogs PBS3 and AtGH3.17 (Chapter V). Analysis of synteny between rosids suggests that PBS3 has evolved from the same common ancestral gene as AtGH3.17, two genes in poplar and one gene in grape. This suggests that PBS3, and by extension other Group III genes, are more closely related to genes in distantly related species than previously thought. In addition, combining comparative genomics analyses with gene expression and phylogenetic inference expands knowledge of the evolutionary history of the GH3 family. Within the family, retention of genes following whole genome duplication, tandem duplication, and transposition has led to many copies in modern plants. The role of local duplication and transposition is particularly evident for the Group III GH3 genes, of which PBS3 is a member. This is likely reflective of the role of these genes in stress response. Based on analysis of functional and expression data, uncharacterized Group III GH3 genes may play a role in response to abiotic and biotic stress, as does PBS3.

FUTURE DIRECTIONS

While my work has answered many questions about the function of PBS3, it has also led to many new questions. Further investigation of the compound discussed in Chapter IV that is elevated in pbs3-2 in response to infection with Pst avrRpt2 is the most intriguing. As yet, the compound has only been detected in response to infection with P. syringae avrRpt2. Other researchers have shown that this effector may play a unique role in the interaction with PBS3 (Lee et al., 2007). Is the compound elevated in response to pathogens with other effectors and is the effector alone sufficient for synthesis of the compound? To test this, one could analyze the compound in pbs3 and wild-type plants infected with various pathogens and in pbs3 mutants crossed into Arabidopsis lines expressing avrRpt2. Knowing the identity of the compound would also lead to many additional questions. A procedure has been developed to isolate this compound in quantities that should be sufficient for analysis by MS/MS or NMR. These analyses should provide important information about the identity of the compound. Once it is identified, the role of the compound in pathogen response and SA metabolism could be explored.

One question that remains is how PBS3 affects SA accumulation. My research indicates that this does not occur by influencing expression of the SA-biosynthetic gene ICS1 expression or altering activity of the ICS1 enzyme. Another alternative is influence on ICS1 protein levels. PBS3 could affect ICS1 protein levels through altered protein translation or degradation rates. Proteasome-mediated degradation of phytohormone-responsive regulators impacting both phytohormone-dependent gene expression and phytohormone biosynthesis and signaling is a common theme (Dreher and Callis, 2007). JA-Ile has been found to facilitate the physical interaction of the JAZ1 repressor with the F-box protein COI1 in vitro (Thines et al., 2007). In planta, treatment with JA results in degradation of the JAZ1 repressor, which liberates the transcriptional activator MYC2, and results in the subsequent expression of a subset of JA-dependent genes (Chini et al., 2007; Thines et al., 2007). The absence of JA-dependent gene expression and associated functional responses is characteristic of the jar1/gh3.11 mutant (Staswick et al., 1992). Perhaps PBS3 functions similarly to JAR1, but in the regulation of SA-dependent responses. Whether PBS3 affects ICS1 levels could be studied by measuring protein levels during a detailed time-course. This experiment also addresses another enduring question: whether SA accumulation is reduced or delayed in pbs3 mutants. In research performed by myself and others, SA was only measured at one or few time points following pathogen infection.
at which it was reduced (Chapter II, Jagadeeswaran et al., 2007; Lee et al., 2007). Analyzing SA in the time-course in parallel to protein levels would address this issue.

A mechanistic understanding of PBS3 and other GH3 enzymes has been hampered by lack of a solved structure. The only known motifs are the three motifs that form the AMP-binding domain. Even the amino acids that bind to the acyl substrate and determine substrate specificity are not yet identified. Comparison with members of the broader acyl-adenylase superfamily with known structures is not optimal given the divergence of the GH3 proteins from these proteins. There is evidence for an important role of the C-terminal region, although understanding is still incomplete. The pbs3-1 mutant contains two point mutations, resulting in non-conserved amino acid changes in the C-terminus of the PBS3 protein and elimination of activity. The HopW1-1 effector from Psm binds to this C-terminal region of PBS3 (Lee et al., 2008).

The combined analyses using comparative genomics, expression data, and phylogenetic prediction revealed new insights in the evolutionary history of the GH3 family. The group III was found to have expanded in the Arabidopsis lineage as a result of both segmental and local duplication and transposition. Such a rapidly evolving group is indicative of a group of genes that respond to stress. Of these, functions have only been described for PBS3 of the traditionally classified group III AtGH3s, and AtGH3.17 and AtGH3.9 of the expanded group III. Expression data suggest that other group III genes may play a role in response to specific pathogens, such as AtGH3.7 and AtGH3.18 with Botrytis cinerea. To confirm this hypothesis, mutants in these genes could be tested for enhanced susceptibility with those pathogens that induce gene expression. In addition, the high-throughput substrate assay I developed for PBS3 could be used to identify substrates for the group III proteins, in Arabidopsis and other species, with unknown specificity. This investigation would help determine whether the other group III genes function in disease response in a similar way to PBS3, and expand our knowledge of the function of this important gene family.

There are clearly many experiments to investigate the function of PBS3 in pathogen-related SA metabolism. However, there are applied aspects of this research that could also be pursued. Using knowledge of the biochemical and kinetic properties of PBS3, one could test and optimize PBS3 for use in biotechnological applications. For example, PBS3 might be used to remove and purify 4-HBA, an inhibitory by-product of cellulosic ethanol production (Palmqvist et al., 1999; Palmqvist and Hahn-Hagerdal, 2000), for resale to the crystal polymer industry (Figuly, 1996). Isolation of value-added products is a common approach to increase the feasibility of biotechnological processes.

REFERENCES


