

**Examination of 2-Oxoglutarate Dependant Dioxygenases Leading to the Production
of Flavonols in *Arabidopsis thaliana***

Daniel Kenneth Owens

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State
University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

In

Biological Sciences

Brenda S.J. Winkel, Chair

Jill C. Sible

Richard A. Walker

Charles L. Rutherford

John M. McDowell

September 22, 2005

Blacksburg, Virginia

Keywords: 2-oxoglutarate dependant dioxygenases, flavonoid biosynthesis, flavanone
3 β -hydroxylase, flavonol synthase, Michaelis-Menten kinetics

Copyright 2005, Daniel K. Owens

**Examination of 2-Oxoglutarate Dependant Dioxygenases Leading to the Production of
Flavonols in *Arabidopsis thaliana***

Daniel Kenneth Owens

ABSTRACT

The flavonols are a varied and abundant sub-class of flavonoids that are associated with a number of essential physiological functions in plants and pharmacological activities in animals. The 2-oxoglutarate-dependant dioxygenases (2-ODDs), flavonol synthase (FLS) and flavanone 3- hydroxylase (F3H), are essential for flavonol synthesis. The primary goal of this study has been to gain a deeper understanding of the biochemistry of these enzymes in *Arabidopsis*.

To accomplish this goal, an activity assay employing recombinant protein expression and HPLC as a detection system was developed for F3H and adapted for use with FLS. The assay was employed to establish the biochemical parameters of F3H from *Arabidopsis*, and to further characterize the F3H mutant allele, *tt6*(87). Enzymatic activity was demonstrated for F3H enzymes from *Ipomoea alba* (moonflower), *Ipomoea purpurea* (common morning glory), *Citrus sinensis* (sweet orange), and *Malus X domestica* (newton apple), each of which had previously been identified solely based on sequence homology.

Arabidopsis contains six genes with high similarity to *FLS* from other plant species; however, all other central flavonoid pathway enzymes in *Arabidopsis* are encoded by single genes. The hypothesis that differential expression of FLS isozymes with varying substrate specificities is responsible for observed tissue-specific differences in flavonol accumulation was tested. Sequence analysis revealed that *AtFLS2*, *4* and *6* contain premature stop codons that eliminate residues essential for enzyme activity. *AtFLS1* was found to have a strong preference for dihydrokaempferol as a substrate. However, no enzyme activity was observed for *AtFLS3* or

AtFLS5 with a number of different substrates under a variety of reaction conditions. To identify structural elements that may contribute to the observed differences in biochemical activity, homology models for each of the isoforms were generated utilizing Arabidopsis anthocyanin synthase (ANS) as a template. A domain at the N-terminus of AtFLS1 that is missing in the other isozymes was insufficient to convey activity to an AtFLS1/5 chimera. These findings suggest a single catalytically-active form of FLS exists in Arabidopsis. The possibility that the apparently expressed but non-catalytic proteins, AtFLS2, 3, and 5, serve noncatalytic roles in flavonol production were explored by yeast 2-hybrid analysis.

ACKNOWLEDGEMENTS

I thank Dr. Brenda Winkel for the opportunity to be a member of her lab, continuous assistance and support in the lab and beyond, and exceptional skills in her role as mentor. I also thank my committee members Dr. Jill Sible, Dr. Charles Rutherford, Dr. John McDowell, and Dr. Richard Walker for guidance and support. This work would not have been possible without the love, encouragement, support, and sacrifices made by my wife, Melissa Owens, and my daughter, Kayla Owens. I am forever indebted to Dr. Cecelia McIntosh who has given me many opportunities and continues to be an extraordinary mentor. I am grateful to the present and past members of the Winkel lab including: Dr. Mike Santos, Dr. Dave Saslowsky, Elizabeth Tucker, Lori Hill, Michelle Barthet, Dr. Ujwala Warek, Nileshwari Vaghela, Dr. Ana Frazzon, Dr. Anne Alerding, Dr. Roberto Martins, and the too numerous undergraduates to list for camaraderie, advice, and the privilege to work with them. I wish to extend special thanks to Winkel lab members Dr. Chris Dana, for being a good friend as well as training me in the art of protein modeling, and Melissa Ramirez, for being supportive during difficult times. I am forever thankful to my parents, Roger and Margaret Owens, for their love and understanding as well as for teaching me in words and by example that any goal can be obtained if you are willing to work hard enough.

LIST OF ABBREVIATIONS

| | |
|--------|--|
| 2-ODD | - 2-Oxoacid Dependant Dioxygenase |
| 4CL | - 4-Coumaroyl Lyase |
| ANS | - Anthocyanin Synthase |
| BSA | - Bovine Serum Albumin |
| C4H | - Coumaroyl-4-Hydroxylase |
| CHI | - Chalcone Isomerase |
| CHS | - Chalcone Synthase |
| DFR | - Dihydroflavonol Reductase |
| DHK | - Dihydrokaempferol |
| DHM | - Dihydromyricetin |
| DHQ | - Dihydroquercetrin |
| F3H | - Flavanone 3 β -Hydroxylase |
| F3'H | - Flavanone 3' Hydroxylase |
| F3'5'H | - Flavanone 3' 5' Hydroxylase |
| FLS | - Flavonol Synthase |
| FNSI | - Flavanone Synthase I |
| IPTG | - Isopropyl-thio- β -D-Galactoside |
| PAL | - Phenylalanine Ammonia Lyase |
| TRX | - Thioredoxin |

TABLE OF CONTENTS

| | |
|---|-----------|
| CHAPTER 1. LITERATURE REVIEW..... | 1 |
| The Importance of Flavonoids | 2 |
| The Flavonoid Biosynthetic Pathway | 3 |
| The Flavonols..... | 3 |
| Enzyme Complexes | 6 |
| 2-Oxoacid-Dependant Dioxygenases..... | 8 |
| Flavanone 3 β -Hydroxylase | 10 |
| Flavonol Synthase | 11 |
| | |
| CHAPTER 2. Biochemical Characterization of Flavanone 3β-Hydroxylase from <i>Arabidopsis thaliana</i> and Verification of Activity in <i>Citrus sinensis</i>, <i>Malus X domestica</i>, <i>Ipomoea alba</i>, and <i>Ipomoea purpurea</i> | 17 |
| Introduction..... | 18 |
| Materials and Methods..... | 22 |
| Materials | 22 |
| F3H Cloning..... | 22 |
| Site-Directed Mutagenesis of F3H..... | 23 |
| Recombinant F3H Expression | 24 |
| Densitometry..... | 24 |
| F3H Activity Assay..... | 25 |
| HPLC Analysis | 26 |

| | |
|---|-----------|
| Results..... | 26 |
| Discussion..... | 34 |
| | |
| CHAPTER 3. Examination of the Flavonol Synthase Isoforms in <i>Arabidopsis</i> | |
| <i>thaliana</i>..... | 38 |
| Introduction..... | 39 |
| Methods..... | 42 |
| AtFLS Cloning and Enzyme Assay | 42 |
| Homology Modeling..... | 43 |
| Yeast Two-Hybrid Analysis..... | 45 |
| Results..... | 46 |
| Discussion..... | 55 |
| | |
| CHAPTER 4. SUMMARY..... | 61 |
| | |
| REFERENCES..... | 69 |
| | |
| APPENDIX. ENZYME ASSAYS..... | 81 |
| | |
| CURRICULUM VITAE..... | 85 |

LIST OF FIGURES

CHAPTER 1.

| | |
|--|----|
| Figure 1.1. Schematic of the flavonoid biosynthetic pathway leading to flavonol production. | 4 |
| Figure 1.2. Map of FLS gene cluster on Arabidopsis chromosome 5. | 14 |

CHAPTER 2.

| | |
|--|----|
| Figure 2.1. Analysis of pH and temperature effects on activity of recombinant AtF3H. | 28 |
| Figure 2.2. Kinetic analysis of recombinant AtF3H. | 30 |
| Figure 2.3. Representative HPLC chromatogram demonstrating lack of <i>TT6(87)</i> biochemical activity. | 32 |
| Figure 2.4. Assessment of biochemical activity for recombinant F3H from <i>I. alba</i> , <i>I. purpurea</i> , <i>C. sinensi</i> , and <i>M. domestica</i> | 33 |

CHAPTER 3.

| | |
|---|----|
| Figure 3.1. Phylogenetic analysis of the AtFLS isoforms. | 47 |
| Figure 3.2. Biochemical analysis of AtFLS1, AtFLS3, and AtFLS5. | 48 |
| Figure 3.3. Structural comparison of AtFLS1, AtFLS3, and AtFLS5 homology models. | 51 |
| Figure 3.4. Analysis of AtFLS1/AtFLS5 chimera. | 54 |

LIST OF TABLES

CHAPTER 3.

| | |
|--|----|
| Table 3.1. Yeast 2-hybrid interactions between AtFLS1, AtFLS3, AtFLS5 and CHS, CHI, DFR. | 53 |
|--|----|

APPENDIX.

| | |
|---|----|
| Table A.1. Enzyme assay reaction components. | 84 |
| Table A.2. Commercially available flavonoid standards sources. | 84 |

CHAPTER 1

LITERATURE REVIEW

The Importance of Flavonoids

Flavonoids, a diverse group of specialized metabolites found in all species of higher plants, perform a variety of functions essential for plant survival. A number of these roles are universal and therefore apparently arose early in the evolution of the flavonoid biosynthetic pathway (reviewed in Stafford, 1991; Winkel-Shirley, 2001). For example, it has been demonstrated that flavonoids absorb light over a broad spectrum, and provide protection from damage by UV irradiation (Li et al., 1993). There is also evidence that flavonoid compounds act as negative regulators of the plant hormone auxin (Jacobs and Rubery, 1988; Brown et al., 2001). Auxin is involved in various plant activities including development of the embryo, leaf formation, phototropism, gravitropism, apical dominance, fruit development, abscission, and root initiation. Flavonoids also perform functions that are restricted and appear to have evolved independently in various species. Certain flavonoids are required for male fertility in some, but not all plants (Mo et al., 1992; Burbulis et al., 1996). Flavonoid compounds also play a crucial role in nitrogen fixation in legumes by acting as signaling molecules that attract symbiotic bacteria (reviewed in Stafford, 1990).

Besides the important functions that flavonoids serve in plants, they can be beneficial to humans as well. These compounds have been well-studied as the major red, blue, and purple pigments in flowers and fruits and are, therefore, of great interest in the development of ornamental plants (Stafford, 1990). Flavonoids have been shown to contribute to the bitterness or sweetness of fruits and are of commercial interest in agriculture (Jourdan et al., 1985). Flavonoids are also being investigated as phytonutrients (Rice-Evans, 2004), having recognized nutritional and medicinal benefits when consumed by animals. Some flavonoids have identified antioxidant and anti-inflammatory properties and have been shown to be associated with a

decreased incidence of coronary artery disease, stroke, lung and colorectal cancers, and rheumatoid arthritis in cohort studies (reviewed in Arts and Hollman, 2005). The isoflavones are a subgroup of the flavonoids that have been studied extensively as phytoestrogens and are being used in the treatment of hormone-dependent breast and prostate cancers as well as osteoporosis (reviewed in Dixon, 2004).

The Flavonoid Biosynthetic Pathway

Flavonoids consist of a core C-15 ring system that is modified by acylation, methylation, and other means to produce the more than 6000 compounds that have been identified to date in nature (Harborne and Williams, 2000). The initial step in flavonoid biosynthesis is the condensation of three molecules of malonyl-CoA with 4-coumaroyl-CoA by chalcone synthase (CHS) to form the C-15 flavonoid backbone (Figure 1.1). Malonyl-CoA is derived from acetyl-CoA, originating from fatty acid biosynthesis, and CO₂ by the enzyme acetyl-CoA carboxylase. Phenylalanine from the shikimate pathway is modified by a series of steps known as the general phenylpropanoid pathway to produce 4-coumaroyl-CoA. The subsequent steps in the biosynthesis of flavonoid compounds are complicated by the fact that many of the enzymes in the pathway have significant activity with more than one substrate. As a result, the pathway has a complex web-like structure with many of the individual flavonoid products having the ability to be formed by a number of different routes (Figure 1.1).

The Flavonols

Due to the large number of flavonoids that have been identified, the compounds have been divided into several subclasses. The flavonols are the most widespread class of flavonoid

compounds found in land plants (Stafford, 1990). These compounds are represented by three major aglycone structures: kaempferol, quercetin, and myricetin (Figure 1.1). As with the other classes of flavonoids, the basic flavonol compounds are modified by methylation, glycosylation, acylation, and other means to produce the enormous variety of flavonols found in plants. The physiological functions of these diverse compounds are not all known, but flavonols have been specifically indicated as the active agents in a number of key physiological functions associated with flavonoids.

Flavonols are believed to be responsible for much of the UV protective effect of flavonoids. The compounds have absorption maxima in the UV-B range, and are found at high levels in the epidermal layer of leaves and in pollen (reviewed in Shirley, 1996). It has also been shown that UV-B light induces flavonoid gene expression and flavonoid accumulation in many plant species (e.g., Schulze-Lefert et al., 1989; Wingender et al., 1990; Fritze et al., 1991; Liu et al., 1995; Mazza et al., 2000). Direct evidence for flavonol participation in UV photoprotection has come from studies of flavonoid-deficient mutants in *Arabidopsis*, which were shown to be hypersensitive to UV-B radiation due to a lack of kaempferol glycosides (Li et al., 1993; Ormrod et al., 1995). The sinapate esters have also been shown to be critical for protection from UV-B and may work in concert with flavonols and other compounds (Landry et al., 1995).

Flavonols are also involved in the identified flavonoid role of regulating the activity of the plant hormone, auxin. This function was first suggested by the fact that certain flavonoid compounds can function as cofactors for indole acetic acid oxidase (IAAO), an enzyme that actively degrades auxin (Furuya et al., 1962; Galson, 1969). Subsequent studies indicated that flavonoid compounds could inhibit polar auxin transport by binding to the naphthylphthalamic acid (NPA) receptor (Jacobs and Rubery, 1988; Faulkner and Rubery, 1992). *Arabidopsis* flavonoid-

deficient mutants exhibit altered growth patterns consistent with altered auxin transport (Murphy et al., 2000; Brown et al., 2001). Increased rates of transport were observed in these mutants based on assays employing labeled indole acetic acid (IAA). These studies also showed that intermediates early in the flavonoid pathway could inhibit auxin transport.

In addition, flavonols act as both pathogen deterrents and symbiont attractors. Flavonol compounds have been shown to be involved in infection of petunia by *Agrobacterium tumefaciens*, stimulated feeding of flea beetle and boll weevil on horseradish and cotton, and infection by the parasitic plant, Dodder (Harborne and Williams, 1988; Zerback et al., 1989; Kelly, 1990). It is also known that flavonoid gene expression and flavonoid biosynthesis are induced by wounding, jasmonic acid, and fungal elicitors, suggesting a role in plant defense.

Finally, flavonols serve a role in pigmentation. These compounds are typically colorless to the human eye, although glycosylated and methylated forms sometimes appear yellow (Harborne and Williams, 2000). However, flavonols are detectable in the UV range and can provide patterns that attract insect pollinators. They may also serve as co-pigments, which can shift color hues by intermolecular interactions with anthocyanin pigments (reviewed in Harborne and Williams, 2000).

Enzyme Complexes

Metabolic channeling and the formation of enzyme complexes within the cell have been predicted as existing for many years (reviewed in Srere, 2000). A classic series of experiments in the 1960's indicated that the cytosol of the cell contains few if any proteins free in solution (Zalokar, 1960; Kempner and Miller, 1968). An enzyme complex is a group of two or more enzymes that interact together in related metabolic reactions and are localized to a specific

subcellular region (Srere, 1987). This may be advantageous for the cell in a number of ways such as providing a means for maintaining high local substrate concentrations, separating anabolic and catabolic processes, maintaining stereospecificity, and sequestering highly toxic, reactive, or unstable pathway intermediates. Well-established enzyme complexes include the pyruvate dehydrogenase complex, the TCA and Calvin cycles, the enzymes of glycolysis and fatty acid oxidation, and the proteasome (reviewed in Winkel, 2004). However, little is known about less stable and dynamic systems, which are likely to include the majority of metabolic systems.

The absence of identifiable free proteins in the cell, taken in combination with the instability, toxicity, and low solubility of the intermediates in the flavonoid biosynthetic pathway, led Helen Stafford to propose that the pathway may exist as an enzyme complex that may also help to shuttle common pathway intermediates into distinct end products (Stafford, 1974). Initial experimental evidence for a flavonoid enzyme complex was provided by analysis of enzyme activities in permeabilized cells of *Haplopappus gracillis* (Fritsch and Grisebach, 1975). This data suggested that flavonoid biosynthesis occurs partially or completely on microsomal or endoplasmic reticulum (ER) membranes. Subsequent immunocytochemical studies indicated that CHS is located on the cytoplasmic face of the rough ER, and that enzymes of the general phenylpropanoid and flavonoid pathways are anchored to the ER by a membrane-bound protein (Hrazdina et al., 1987; Hrazdina and Jensen, 1992). Co-localization of CHS and chalcone isomerase (CHI) to the ER and tonoplasts of *Arabidopsis* roots has also been observed (Saslowky and Winkel-Shirley, 2001). Recent evidence shows that CHS and CHI can also localize to the nucleus of *Arabidopsis* where the enzymes could either directly or indirectly affect signaling and gene transcription (Saslowky et al., 2005).

As further evidence of the existence of a flavonoid complex, interactions among members of the flavonoid metabolic pathways have been investigated. Channeling, gel filtration, and cell fractionation studies suggested that the phenylpropanoid and flavonoid enzymes, phenylalanine ammonia lyase (PAL), cinnamate-4-hydroxylase (C4H), CHS, and UDP-glucose flavonoid glucosyltransferase (UGFGT) function together in one or more membrane-associated enzyme complexes (Hrazdina and Wagner, 1985). Orientation dependant interactions among CHS, CHI, and dihydroflavonol 4-reductase (DFR) from *Arabidopsis* have been identified by use of yeast two-hybrid assays (Burbulis and Winkel-Shirley, 1999). Affinity chromatography and immunoprecipitation assays have further confirmed interactions between CHS, CHI, and flavanone 3-hydroxylase (F3H) in this system.

2-Oxoacid-Dependant Dioxygenases

The 2-oxoacid-dependant dioxygenases (2-ODDs) represent the largest family of nonheme oxidizing enzymes. These enzymes are ubiquitous in living cells with examples having been reported in plants, animals, microorganisms, and even viruses (Schofield and Zhang, 1999). The 2-ODDs are involved in the biosynthesis of a number of important biological molecules including penicillin, cephalosporin, cephamycin, clavam, ethylene; carnitine, and collagen (reviewed in Prescott and Lloyd, 2000). The number of known and predicted 2-ODD enzymes is increasing rapidly as the amount of sequence data for various genomes becomes available. The completed *Arabidopsis* genome contains 64 predicted open reading frames for potential Fe(II)- and α -ketoglutarate dependent dioxygenases (Ryle and Hausinger, 2002).

The 2-ODDs are soluble enzymes that incorporate both atoms from O_2 into the reaction products. A proposed mechanism involves harnessing the catalytic capability of Fe^{2+} to perform

the reaction of interest coupled with the conversion of 2-oxoglutarate into succinate and CO₂ (Schofield and Zhang, 1999). The reaction must also occur in such a manner that oxygen transfer occurs efficiently, but with minimization of side reactions that could damage the cell, as the reaction mechanism likely involves a free radical intermediate. This mechanism is supported by a measurable increase of enzyme activity for 2-ODDs in the presence of catalase, an enzyme that is known to scavenge hydrogen peroxide. Hydrogen peroxide is a cell-damaging compound that would likely result from reactions between free radical oxygen and water within the cell.

The first published structure of a 2-ODD was for the fungal enzyme, isopenicillin N synthase (IPNS) (Roach et al., 1995). However, this was not an ideal template for plant 2-ODD structural studies as IPNS represents a unique member of the family that does not utilize a 2-oxoacid co-substrate and has low overall identity to plant enzymes (reviewed in Prescott and John, 1996). Subsequently, other 2-ODD structures have been solved including clavamate synthase, deacetoxycephalosporin C synthase, 4-hydroxyphenylpyruvate dioxygenase, and anthocyanin synthase (ANS) from *Arabidopsis* (reviewed in Schofield and Zhang, 1999; and Ryle and Hausinger, 2002; Wilmouth et al., 2002). All 2-ODD structures to date contain a jellyroll motif that is relatively rare in other enzymes, and within which the active site is buried. This arrangement is speculated to aid in sequestering the putative Fe-O reaction intermediate that would be highly reactive if exposed to the aqueous environment of the cell. The active site is predicted to be surrounded by a distorted octahedral geometry with ligands provided by two water molecules, two histidine side chains, an aspartate, and a glutamine close to the carboxy terminus.

The 2-ODDs are characterized by the presence of a conserved amino acid sequence near the carboxy terminus of the protein. However, the overall sequences of these enzymes are not

highly conserved. Interestingly, a degree of conservation in the positioning of introns within 2-ODDs has been observed with up to three introns at four locations having been identified in a variety of dioxygenase genes (reviewed in Prescott and John, 1996). This may indicate that the enzymes arose only once during the course of evolution from a common ancestor.

Flavanone 3B-Hydroxylase

F3H catalyzes the stereospecific hydroxylation of (2S)-naringenin to (2R, 3R)-dihydrokaempferol (DHK) (Figure 1.1), and was the first flavonoid 2-ODD enzyme to be identified. The enzyme was initially characterized in *Petroselinum crispum* and *Matthiola incana* (Britsch et al., 1981). The majority of the subsequent biochemical characterization was carried out with the *Petunia hybrida* protein (Britsch and Grisebach, 1986). A major challenge in this work was the fact that the enzyme undergoes rapid proteolysis in crude plant extracts as well as when expressed as a recombinant protein in *E.coli* (Britsch and Grisebach, 1986; Britsch et al., 1992; Lukacin et al., 2000). It has been established that F3H is active as a monomer, and much work has been dedicated to identifying residues within the enzyme important for activity (Britsch et al., 1993; Lukacin and Britsch, 1997; Lukacin et al., 2000; Lukacin et al., 2000; Wellmann et al., 2004). F3H genes have now been identified and cloned in a number of different plant species based on sequence homology, including *A. thaliana*, *Ammi majus*, *Anethum graveolens*, *Pimpinella anisum*, *P. hybrida*, *Camellia sinensis*, *Citrus paridisii*, *Forsythia X intermedia* and *Zea mays*. However, little biochemical characterization has been reported for any of the corresponding proteins.

Flavonol Synthase

FLS catalyzes the committed step in the production of flavonols, converting the dihydroflavonols, DHK or dihydroquercetin (DHQ), to the corresponding flavonols, kaempferol or quercetin (Figure 1.1). Neither dihydromyricetin (DHM), nor the flavanone 3' 5' hydroxylase enzyme activity also required to produce it, have been identified in *Arabidopsis*. FLS has been characterized as a 2-ODD enzyme based on its requirement for 2-oxoglutarate, ascorbate, and Fe^{2+} to achieve optimal enzyme activity.

FLS was initially identified in enzyme preparations from illuminated *P. crispum* cell suspension cultures that were being tested for F3H activity (Britsch et al., 1981). Spribille and Forkman (1984) subsequently identified and characterized FLS activity in the flower buds of *M. incana*. It was demonstrated that, although DHK and DHQ were both effective substrates *in vitro*, the only product present in measurable quantities *in vivo* was kaempferol. This was attributed in part to the fact that DHK was demonstrated to be a better substrate for FLS in *in vitro* activity assays. Also, the highest levels of FLS activity were shown to occur in very young flower buds. At this stage of development, more DHK than DHQ is present, apparently because flavanone 3'-hydroxylase activity, which forms DHQ, is very low (Figure 1.1).

These studies further suggested that there was coordinated regulation between flavonol and anthocyanidin synthesis. FLS is located at an early branch point in flavonoid biosynthesis, where several enzymes may compete for the same substrate (Figure 1.1). It was demonstrated that FLS activity was highest at early stages of flower development, and decreased rapidly as anthocyanin biosynthesis increased in *M. incana* (Spribille and Forkmann, 1984). Similar experiments were performed with *Dianthus caryophyllus* L. (carnation) buds and flowers (Stich, 1992). As in *M. incana*, flavonol synthesis predominated early and anthocyanin synthesis later

in the development of flowers. There was little if any preference for DHK or DHQ as a FLS substrate in *D. caryophyllus*. However, DFR, an enzyme that competes with FLS and leads to anthocyanin biosynthesis, had a four-fold higher rate using DHQ as a substrate.

The first *FLS* gene was isolated by differential screening of a *P. hybrida* cDNA library (Holton et al., 1993). The identity of this gene was confirmed by performing enzyme activity assays with recombinant protein produced in *Saccharomyces cerevisiae*. Low-stringency DNA blot analysis indicated that a single gene in petunia encodes FLS. A number of *FLS* genes have since been cloned from various species including: *M. incana*, *A. thaliana*, *Citrus unshiu*, *Solanum tuberosum*, *Oryza sativa*, *Vitis vinifera*, *Eustoma grandiflorum*, and *Malus X domestica*.

Much recent work has focused on biochemical analysis of FLS from *C. unshiu*. Initially, a cDNA was isolated based on sequence homology to an Arabidopsis *FLS* EST (153O10T7), and used as a probe to demonstrate that the gene was regulated in a tissue and developmental specific manner (Moriguchi et al., 2002). Southern analysis with a *C. unshiu FLS* probe suggested that several *FLS* genes or related sequences may be present in this species. The optimum pH, optimum temperature, and K_m values for activity with DHK and DHQ were determined for recombinant protein produced in *E. coli* (Wellmann et al., 2002). The enzyme had a six fold higher affinity for DHK than DHQ. Site directed mutagenesis of four strictly conserved 2-ODD residues that had no previously identified functionality and subsequent enzyme assay were performed with FLS from *C. unshiu*. It was found that two of these residues, Gly68 and Gly261, were important for enzyme activity, likely being involved in proper folding of the enzyme, while the other two residues had little or no effect upon activity. The *C. unshiu* enzyme

has also been shown to have promiscuous enzyme activity with naringenin, as originally identified in Arabidopsis (Prescott et al., 2002; Lukacin et al., 2003).

The first *FLS* clone was identified in Arabidopsis as an EST (Accession no. 153O10T7) with high sequence homology to *P. hybrida* and *S. tuberosum* FLS (Pelletier et al., 1997). As the corresponding cDNA clone was not full-length, it was used to screen an Arabidopsis (Landsberg ecotype) genomic DNA library. The identified full-length clone was sub-cloned and used to determine the sequence of the gene. Comparison of genomic and cDNA sequences identified two introns in this gene located at two of the four possible positions that are conserved among all plant 2-ODDs characterized to date. The gene encodes a protein that is 336 amino acids in length with an approximate molecular mass of 38.3 kD. The deduced protein was shown to have substantial homology to both the *P. hybrida* (61%) and *S. tuberosum* (59%) FLS enzymes.

Southern blot analysis was initially used to determine the copy number of *FLS* in Arabidopsis. A single band was observed under high stringency (65°C) in genomic DNA digested with an enzyme that did not cut within the probe region. However when the same blot was probed at low stringency (55°C), several additional bands were observed. This suggested that more than one sequence with homology to *FLS* exist in the Arabidopsis genome.

A total of six genes with *FLS* sequence similarity were subsequently identified as a result of the Arabidopsis genome sequencing project (Arabidopsis Genome Initiative, 2000; A. Bandara and B. Winkel, unpublished data). This makes FLS the only known enzyme of the central flavonoid biosynthetic pathway to be encoded by an apparent gene family in Arabidopsis. These genes were designated as *AtFLS1-6* in order of their identification. All genes are located on chromosome 5, and *AtFLS2-5* are arranged in a tandem array (Figure 1.2). *AtFLS1* expressed in *E.coli* has demonstrated enzyme activity with DHK and DHQ

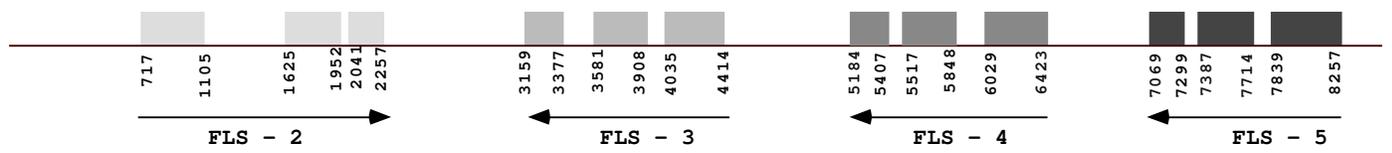


Figure 1.2. Map of FLS gene cluster on Arabidopsis chromosome 5. Gray boxes represent the exons and arrows the direction of transcription for each gene.

(Prescott et al., 2002; Turnbull et al., 2004). AtFLS1 has also been shown to have promiscuous enzyme activity with naringenin (Prescott et al., 2002). *AtFLS2*, *4*, and *6* all contain a premature stop codon that eliminates the known conserved Fe²⁺ coordinating residues and therefore likely represent nonfunctional genes. There have been no ESTs or MPSS signatures identified for AtFLS6, further affirming its classification as a pseudogene.

Although Arabidopsis has a relatively small genome, approximately 140 Mb, multi-enzyme families are relatively common in this organism (Blanc et al., 2000; Arabidopsis Genome Initiative, 2000). One explanation is that large-scale duplications have occurred in the Arabidopsis genome. Blanc et al. (2000) utilized computer-aided sequence analysis to show that over 60% percent of the Arabidopsis genome appears to exist as duplications. Therefore although unique to FLS in the flavonoid biosynthetic pathway, gene families are widespread in Arabidopsis.

Although the physiological significance of the multiple FLS-like sequences in Arabidopsis is unknown, it has been demonstrated that different flavonols are predominant in different tissues of this species. Quercetin is the predominant flavonol in seeds, while in contrast kaempferol predominates in flowers (Shirley et al., 1995). On the other hand, similar levels of the two flavonols were identified in stamens (Burbulis et al., 1996). *AtFLS1* mutants generated with the maize transposable element *En-1* were shown to be deficient in kaempferol accumulation, but to contain wild type levels of quercetin in leaves (Wisman et al., 1998). This suggests a possible mechanism where by different FLS isozymes with varying substrate specificities control the amount and type of flavonols in a particular tissue.

A major objective of this project was the development and utilization of an HPLC based enzyme assay for recombinant flavonoid 2-ODD enzymes. This assay was used to

biochemically characterize Arabidopsis F3H and to investigate the intermediate seed coat color, as compared to wild type and null mutants, observed in the F3H mutant *tt6(87)*. The assay was also used to assess the activities of putative F3H enzymes from *Ipomoea purpurea*, *Ipomoea alba*, *Citrus sinensis*, and *M. domestica*, which had all previously been identified only by sequence homology. In a second study, each of the proposed functional *FLS* genes was expressed as recombinant proteins, in *E.coli*, and assayed for substrate specificity. The findings from these assays were further extended by homology-based structural analysis and yeast 2-hybrid interaction studies of the *FLS* gene family. This work represents one of the few comprehensive analyses of flavonoid gene families and the potential roles of different isoforms in determining the complement of endproducts synthesized in respect to internal and environmental cues.

CHAPTER 2

Biochemical Characterization of Flavanone 3 β -Hydroxylase from *Arabidopsis thaliana* and Verification of Activity in *Citrus sinensis*, *Malus X domestica*, *Ipomoea alba*, and *Ipomoea purpurea*

Daniel K. Owens and Brenda S.J. Winkel

Department of Biological Sciences and Fralin Center for Biotechnology, Virginia Tech,

Blacksburg, VA 24061-0346.

Manuscript to be submitted to Plant Molecular Biology

INTRODUCTION

The 2-oxoglutarate dependant dioxygenases (2-ODDs) are an important class of soluble nonheme iron-containing enzymes that have been identified in both prokaryotes and eukaryotes. In plants, 2-ODD enzymes function in phytohormone biosynthesis and in the production of secondary metabolites such as alkaloids and flavonoids that have key activities in plant growth and development as well as defense against pathogens (Prescott and Lloyd, 2000; Winkel-Shirley, 2001; Swain and Singh, 2005). These metabolites are also of considerable interest from the perspective of human health and nutrition. Mutations in members of this class of enzymes have been implicated as causative agents in diseases such as alcohol liver cirrhosis, Ehlers-Danlos syndrome type VI, tyrosinaemia type III, hawkinsinuria, Refsum's disease, encephalopathy, cardiomyopathy, scurvy, and male pattern baldness (reviewed in Prescott and John, 1996). The enzymes have been extensively studied in fungus due to their involvement in the production of the β -lactam antibiotics. Bacterial dioxygenases have been shown to be involved in antibiotic resistance and sulfate metabolism.

The 2-ODD enzymes are known to perform a variety of reactions including hydroxylation, epoxidation, desaturation, expansion, and ring formation in various metabolic pathways. The enzymes do not share substantial overall sequence identity, but tend to have conservation of three amino acids that represent the Fe^{2+} coordinating residues in the enzyme active site (Hegg and Que, 1997). Due to diversity of action and low sequence homology, enzymes are characterized as 2-ODDs based on biochemical activity and increasingly by tertiary structural elements as crystal structures become available. All 2-ODDs have a requirement for Fe^{2+} as a cofactor for activity, and most employ α -ketoglutarate as a co-substrate. Ascorbate increases activity levels presumably by creating a reducing environment that helps to maintain

the enzyme in the appropriate oxidation state (Englard and Seifter, 1986). In all 2-ODD structures solved to date the active site residues are buried in a jelly-roll core motif, a relatively rare element in enzymes. A crystal structure for the Arabidopsis 2-ODD anthocyanidin synthase (ANS) has been solved, and the Fe²⁺ coordinating and α -ketoglutarate binding residues identified (Wilmouth et al., 2002). Although overall sequence identity between AtF3H and ANS are low, these active site residues are absolutely conserved.

In plants, 2-ODD enzymes that function in flavonoid biosynthesis have been particularly well studied. Flavonoids participate in such processes as flower coloration, UV photoprotection, symbiont attraction, pathogen deterrence, phototropism, and gravitropism (Winkel-Shirley, 2001). Flavonoids also have established nutritional roles with much attention being focused on their proposed anti-cancer and anti-heart disease activities (Yao et al., 2004). It has been suggested that the 2-ODDs of the central flavonoid biosynthetic pathway can be grouped according to their substrate and stereospecificities (Turnbull et al., 2004). Flavonol synthase (FLS) and ANS have broad substrate specificities, catalyzing reactions other than what has been identified as their *in vivo* activity. FLS has been shown *in vitro* to have the ability to convert naringenin to dihydrokaempferol (DHK), the reaction established as being performed by flavanone 3 β -hydroxylase (F3H) *in vivo* (Prescott et al., 2002). Similarly, ANS, *in vitro*, can perform the reactions attributed to F3H and FLS. F3H and flavanone synthase I (FNSI) have very narrow specificities, only having activity with a particular stereoisomer of their corresponding substrates.

F3H catalyzes the stereospecific hydroxylation of (2S)-naringenin to (2R, 3R)-dihydrokaempferol and was the first flavonoid dioxygenase enzyme to be characterized. F3H was initially identified in illuminated parsley cells and *Matthiola incana*, but the majority of

biochemical characterization has been performed on the *Petunia hybrida* protein (Britsch et al., 1981; Britsch and Grisebach, 1986). Initial work focused on determination of biochemical parameters (K_m , V_{max} , optimum pH) for the enzyme purified from Petunia flower petals (Britsch and Grisebach, 1986). A major obstacle in this work was proteolysis of the protein to an inactive form during isolation. The ability to produce F3H in *E. coli* has greatly increased the availability of enzyme for study, but unfortunately the recombinant protein also undergoes proteolysis (Lukacin et al., 2000). Techniques have been developed to address this problem, however, it is still not possible to entirely prevent F3H proteolysis.

Comparison of the *Petunia hybrida* F3H (PhF3H) amino acid sequence with other 2-ODD proteins has led to the identification of eight conserved amino acid residues that could be critical for enzyme function (Lukacin and Britsch, 1997). Site directed mutagenesis and functional analysis demonstrated that His220, His278, and Asp222 are essential for catalysis and likely represent the Fe^{2+} coordinating residues. Similar experiments have confirmed the role of equivalent residues in other 2-ODD enzymes such as isopenicillin N synthase, prolyl hydroxylase, lysyl hydroxylase, aspartyl B-hydroxylase, 1-aminocyclopropane-1-carboxylate oxidase, and hyoscyamine-6B-hydroxylase (reviewed in Prescott and John, 1996). Arg288 and Ser290 have been identified as significant for 2-oxoglutarate binding in PhF3H (Lukacin et al., 2000). PhF3H is active as a monomer (Lukacin et al., 2000). However a tendency to aggregate in solution has been noted, which may imply the enzyme can interact with other proteins *in vivo*. Domain swapping experiments have shown that, unlike the situation in some other 2-ODDs, the C-terminal amino acid residues do not convey substrate specificity in PhF3H (Wellmann et al., 2004).

In *Arabidopsis thaliana*, F3H (At3g55120) is encoded by a single gene, similar to most other members of the central flavonoid pathway in this species (Pelletier and Shirley, 1996; Arabidopsis Genome Initiative, 2000). The full-length AtF3H coding region was previously expressed in *E. coli* for the purpose of antibody production (Pelletier et al., 1999). These antibodies were subsequently used to characterize the expression pattern of AtF3H during seedling development at the protein level. AtF3H was shown to be an “early gene” with mRNA and protein present transiently during the first several days of seedling development.

The first AtF3H allele, *tt6(87)*, was identified in a screen of ethylmethane sulfonate-mutagenized plants by Koornneef (1981) and assigned a function in flavonoid biosynthesis based on its pale-brown seed phenotype. The lack of detectable flavonols and the apparent presence of anthocyanidins in this mutant suggested that the *tt6(87)* locus specified FLS activity (Shirley et al., 1995). However, the sequence of FLS1 in *tt6(87)* plants was subsequently shown to be identical to wild type (Pelletier et al., 1997). Using *En-1* mutagenized Arabidopsis lines and appropriate genetic crosses, *tt6* was subsequently shown to encode F3H (Wisman et al., 1998). The *tt6(87)* allele contains a single base pair mutation (C to T) that leads to a premature stop codon and a severely truncated coding region. This is surprising because a mutation in F3H should eliminate flavonol and anthocyanidin accumulation, leading to the production of yellow seeds instead of the observed pale brown phenotype. Moreover, flavonols have been detected in *tt6(87)* plants by histochemical and biochemical analysis (Peer et al., 2001). As a result, it has been suggested that *tt6(87)* is a leaky mutation while the *En-1* and frameshift lines appear to have null phenotypes. The severity of the *TT6* truncation suggests that the enzyme is inactive, however, this has never been shown experimentally. Therefore, the possibility exists that *tt6(87)* lines have a low-level of enzyme activity.

In this study, recombinant AtF3H protein was characterized at the biochemical level. In addition, full-length cDNA clones for F3H from *Ipomoea alba*, *Ipomoea purpurea*, *Citrus sinensis*, and *Malus X domestica* were expressed in *E. coli* and their enzymatic activities confirmed for the first time. The biochemical activity of the truncated form of the protein encoded by *tt6(87)* was measured to determine whether low-level activity may account for the observed leakiness.

MATERIALS AND METHODS

Materials

All reagents were of analytical grade unless otherwise noted. L-ascorbic acid, ferrous sulfate heptahydrate, α -ketoglutaric acid (disodium salt), bovine serum albumin (BSA), (+,-) naringenin, and (+,-) dihydroquercetin (DHQ) were purchased from Sigma-Aldrich. Ampicillin trihydrate, chloramphenicol, glacial acetic acid, glycine, HEPES, potassium phosphate, and Tris free base were from Fisher Scientific. HPLC grade acetonitrile, HPLC grade methanol, and ethyl acetate were from Burdick and Jackson. DHK was from Dihydrochemical Technologies (Adelaide, Australia). Luria Broth was from Difco. Isopropyl-thio- β -D-galactoside (IPTG) was from Alexis Biochemicals. *E. coli* BL21(DE3)pLysS cells and the pET32a vector were from Novagen.

F3H Cloning

The full-length AtF3H coding region was previously cloned into pET32a (Pelletier et al., 1999). AtF3H was subcloned into pCD1 from pET32a (C.Dana and B.Winkel unpublished data). The pCD1 vector was generated by site directed mutagenesis of pET32a, creating an *NcoI* restriction site to allow for the production of clones in which the majority of expression tags in the generated protein could be removed by thrombin digestion. Full-length cDNA clones

annotated as F3H from *I. alba* (M.Rauscher, pers.comm.), *I. purpurea* (U74081), *C. sinensis* (ABO11795), and *M. domestica* (ABO74486) were identified from the NCBI database and graciously supplied by the identified sources. The coding region from each clone was amplified by Pfu-mediated PCR employing primers that incorporated restriction sites to facilitate cloning. The sequences for the upstream primers were 5'CATGCCATGGCAACTTTGTCAAC3', 5'CGGGATCCATGACGACGGTGTC3', 5'CATGCCATGGCTCCTTCAACC3', and 5'CGGGATCCATGGCTCCTCCTGC3' and for the downstream primers were 5'CGGTTCGACCTAAGCAAGAATTTGATC3', 5'CGGGATCCATGACGACGGTGTC3', 5'CGGTTCGACCTAAGCCAA-GATCTCC3', and 5'CGGTTCGACCTAAGCAAATATGTCGTC3' for *I. alba*, *I. purpurea*, *C. sinensis*, and *M. domestica* respectively. The products were digested with *SalI* and either *NcoI* (*I. alba* and *C. sinensis*) or *BamHI* (*I. purpurea* and *M. domestica*) and then subcloned into pET32a. Constructs were transformed into DH10B by electroporation and the resulting colonies screened by colony PCR (Smith et al., 1990). Purified plasmids were sequenced to confirm cloning region integrity as well as reading frame. Plasmids were subsequently used to transform BL21(DE3)-pLysS cells by heat shock and the colonies screened by small-scale test inductions as described in the Novagen pET System Manual, 8th Edition.

Site-Directed Mutagenesis of F3H

A *tt6(87)* protein expression clone was generated by introducing a single base pair (T870 to C) substitution in pCD1-AtF3H using the QuikChange system (Stratagene La Jolla, CA). Mutagenesis primers were 5'-GACATGGATTACGGTTTAGCCTGTTGAAGGAGCG-3', where the altered nucleotide is underlined, and an antisense primer with the complementary sequence. Plasmids recovered from XL1 BLUE cells were sequenced to confirm the presence of

the mutation. The clones were subsequently transformed into BL21(DE3)pLysS cells and the resulting colonies screened by small scale test inductions as for the other F3H clones.

Recombinant F3H expression

Expression was essentially as described in Pelletier, et al. (1999). Expression was induced by the addition of IPTG to 1 mM final concentration and incubation at room temperature. Expression was monitored at various time points by collecting 500 μ l aliquots of culture, lysing the cells by boiling in Laemmli buffer (Laemmli, 1970), pelleting the cellular debris, and analyzing 6 μ l of each of the resulting samples by SDS-PAGE. For standard F3H production, induction was performed for 4 h at room temperature, 250 rpm, and the cells harvested by centrifugation at 7400 x g, 4°C, for 10 min, and stored at -80°C. Frozen cells were resuspended in 3 ml of 0.2 M glycine (pH 8.5) and lysed by sonication on ice. Sonication was performed by giving a 30 sec pulse followed by a 30 sec recovery period and repeating this 6 times. The resulting cell slurry was centrifuged at 16,170 x g, 4°C, for 40 min and the supernatant used as the source of crude enzyme in activity assays.

Densitometry

A 2 μ l aliquot of each crude enzyme extract and a BSA dilution series (4,2,1,0.5, and 0.25 μ g) were fractionated on a 3% stacking/7.5% separating acrylamide gel. The samples were subsequently transferred to a 0.2 μ m pure nitrocellulose membrane (Bio-Rad) at 100V for 20 min. The membrane was stained with Ponceau S, quickly destained in distilled water until an appropriate background was achieved, and allowed to dry. The membrane was then scanned at 600 dpi resolution on an Epson Perfection 3170 Photo scanner and imported into Adobe Photoshop 7.0. The image was converted to grayscale and saved in TIFF format. Densitometry was then performed in ImageJ ver. 1.28 with calibration set to uncalibrated OD. The BSA

dilution series was used to generate a standard curve that was then used to determine the concentration of enzyme in the crude lysates. Total protein was quantified using the Bradford assay (BioRad Protein Microassay) with BSA as the standard.

F3H activity assay

Assays were based on the method developed by Britsch and Grisebach (1986). Each 100 μ l reaction contained 10 mM α -ketoglutaric acid (disodium salt), 10 mM ascorbic acid, 0.25 mM ferrous sulfate, and 50 mM glycine (pH 8.5). Naringenin was assayed in the range of 25 – 500 μ M. As an internal standard, 100 μ M DHQ was included. All flavonoid compounds were dissolved in 80% HPLC-grade methanol at a starting concentration of 10 mM. All other assay components were suspended in 50 mM glycine, pH 8.5 (or in 50 mM glycine or HEPES adjusted to pH 7.5 – 9.5 in increments of 0.25 for assays of pH optimum), degassed under vacuum for 10 min, equilibrated under N₂ for 5 min, and again degassed under vacuum for 10 min. Ascorbic acid was included in the ferrous sulfate solution to help maintain iron in the proper oxidation state. Standard assays of enzyme activity were performed at 25°C, pH 8.5, for 10 min using approximately 3.5 μ g of crude recombinant enzyme. An assay performed using a similar amount of crude thioredoxin (Trx) protein expressed from the pET32a vector without an insert was included as a negative control. Reactions were initiated by the addition of naringenin and terminated by extraction with ethyl acetate (1:1, v:v), performed by adding 200 μ l of ethyl acetate and mixing well for 1 min. Solvent layers were separated by centrifugation at 13,000 rpm in a microcentrifuge for 5 min. A 100 μ l aliquot of the organic layer was then re-extracted with 200 μ l ethyl acetate, mixed and centrifuged as before, and 200 μ l of the organic layer combined with the initial 100 μ l extract (R. Lukacin, personal communication). The solvent was evaporated in a SpeedVac under low heat. Dried samples were reconstituted in 50 μ l of 80%

methanol, mixed for 5 min, and centrifuged at 13,000 RPM in a microcentrifuge for 5 min. The samples were transferred to autosampler vials for analysis by HPLC.

HPLC analysis

Samples from the activity assays were analyzed using a Waters HPLC system consisting of two 515 series pumps, a 717plus autosampler, and a 2996 photodiode array detector, operated using Millennium 3.2 software. Samples were kept at 4°C prior to analysis, a 20 µl aliquot of each 50 µl reconstituted sample injected and fractionated at room temperature using a Nova-Pak[®] C18 3.9 X 150 mm column and a flow rate of 2.5 ml/min. Fractionation was achieved using a linear gradient of water (adjusted to pH 3 with glacial acetic acid) and HPLC-grade acetonitrile as described in Pelletier and Shirley (1996). Absorbance of the eluent was monitored from 200 to 400 nm. The resulting data were analyzed by extracting a single wavelength chromatogram at 289 nm and integrating the resulting peak areas. Concentrations of flavonoid compounds were then determined from a standard curve generated using the same HPLC conditions. Kinetic parameters were determined using Kaleidagraph v. 3.6 (Synergy Software).

RESULTS

Recombinant AtF3H was assayed as a means to investigate the biochemical properties of the enzyme. AtF3H was expressed with a N-terminal Trx fusion partner, as this was previously shown to increase protein yield (Dana and Winkel, unpublished observations). Trx-AtF3H from crude bacterial lysates was demonstrated to be active (Figure 2.3). As the Trx tag is separated from AtF3H by a flexible linker region, it is likely to have a minimal effect upon protein folding and enzyme activity. Crude lysates from Trx expressed without AtF3H in the same system were used as a negative control throughout. Removal of the Trx tag by thrombin digestion and

purification of the protein by metal affinity chromatography resulted in active protein, but activity levels were lower than with crude, uncleaved extract. Due to the labile nature of the enzyme and the observation of degradation products by SDS-PAGE, proteolysis during processing of the protein was the most likely cause of this reduced activity. All subsequent characterization was therefore performed with crude Trx-AtF3H from bacterial lysates.

The linear range for the reaction was determined by performing time course reactions with different concentrations of protein at pH 7.0 and 37°C, and was found to be approximately 10 µg of enzyme assayed for 10 min. The optimum pH for the enzyme was determined with appropriate buffers over a wide pH range (2-9), and subsequently narrowed to 0.25 pH increments in the highest area of activity (7.5-9.5). The optimum pH was shown to be in the range of 8.5 to 8.75 (Figure 2.1A), which coincides well with the value of 8.5 obtained for PhF3H, but is higher than the value of 7.5 observed for *Camellia sinensis* (Britsch and Grisebach, 1986; Punyasiri et al., 2004). AtF3H did not appear to undergo the rapid decline in activity noted for PhF3H at pH values above 8.5. Subsequent reactions were performed at pH 8.5, and the amount of protein was reduced to 3.5 µg to account for the higher activity levels under the new conditions.

To determine the effects of temperature on enzyme stability, crude lysate was incubated at various temperatures for 10 min, allowed to equilibrate to 25°C for 10 min., and assayed at 25°C for 10 min. (Figure 2.1B). The temperature optimum for the reactions was determined by performing assays at various temperatures for 10 min. AtF3H maintained stability to 30°C. Stability gradually decreased to 35°C, and greatly decreased beyond 35°C. The optimum assay temperature was 40°C. A temperature optimum has not been reported for F3H, although PhF3H has typically been assayed at 37°C. However, the stability of the enzyme is low at this

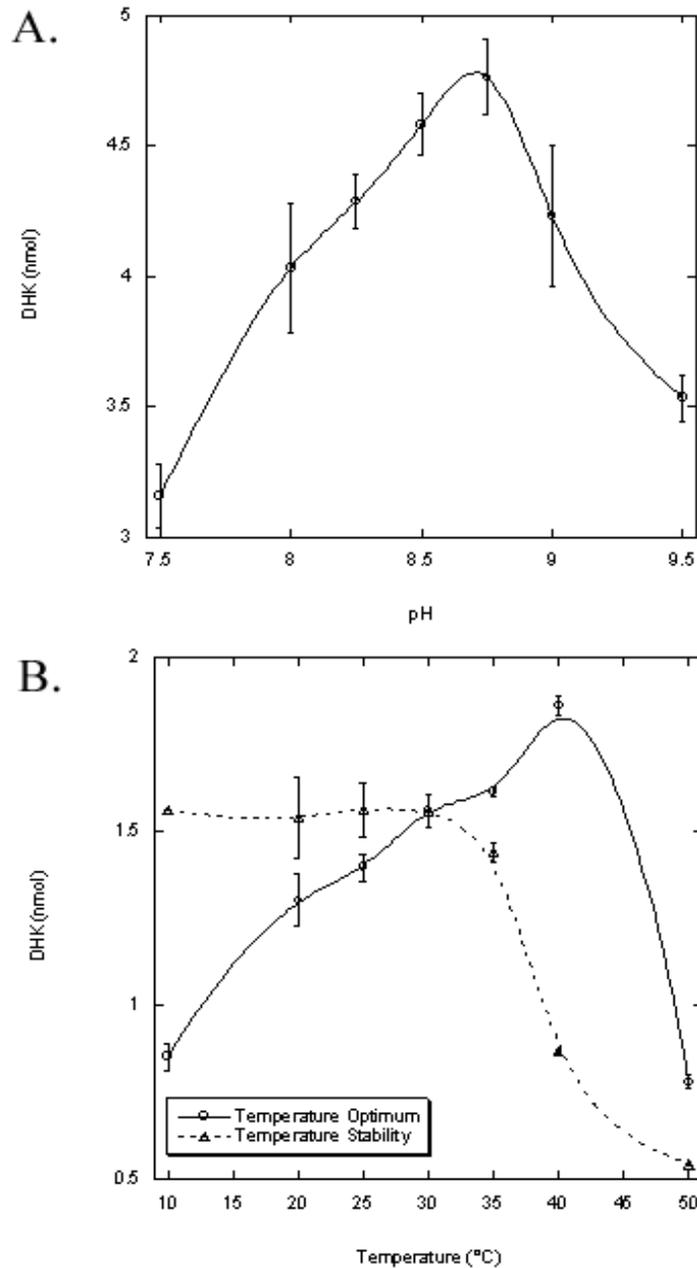


Figure 2.1. Analysis of pH and temperature effects on activity of recombinant AtF3H. Assays were performed using crude recombinant Trx-AtF3H enzyme and analyzed by HPLC as described in the Materials and Methods. Error bars represent standard error from experiments performed in triplicate. A. The pH optimum experiments were at 37°C in glycine (7.5-8.0) and HEPES (8.0-9.5) buffers. B. All temperature experiments were at pH 8.5. The temperature optimum was determined by assaying at the indicated temperature. Temperature stability was assessed by pre-incubating the enzyme at the indicated temperature and assaying at 25°C.

temperature. This suggests that at 40°C the initial activity rate is high enough to offset the loss of stability over time.

Michaelis-Menten rate constants were determined for AtF3H with respect to naringenin by saturating the system with α -KG to establish pseudo-first-order kinetics. Assays were performed in triplicate with 3.5 μ g of protein for 10 min at 25°C and pH 8.5 utilizing various concentrations of naringenin. The Michaelis-Menten equation was fit to a plot of naringenin concentration versus V_0 to determine K_m and V_{max} which were determined to be $24 \pm 3 \mu$ M and 333 ± 11 pmoles/min, respectively (Figure 2.2). The K_m agrees well with the values of 5.6 to 10 μ M determined at pH 7.0 to 8.5 and 37°C for F3H purified from *Petunia* (Britsch and Grisebach, 1986). The Arabidopsis enzyme exhibited a specific activity of approximately 1.5 mkat/kg, which compares favorably with that reported for the purified recombinant *Petunia* protein (5-10 mkat/kg), and some two orders of magnitude higher than observed for enzyme purified from *Petunia* tissue (50-100 μ kat/kg) (Britsch et al., 1992). It was observed that the supplied racemic substrate was never entirely consumed for AtF3H under conditions where the assay would be expected to reach completion (high concentration enzyme, low concentration naringenin, and 60 min. assay time). This is consistent with the enzymes being stereospecific for (2S)-naringenin as has been demonstrated for PhF3H.

The *tt6(87)* allele has been reported to be a leaky AtF3H mutation, but the source of the leakiness has yet to be determined. The *tt6(87)* allele has been sequenced and the mutation identified, but the mutant protein has never been assayed for biochemical activity leaving the possibility that low levels of activity may exist in this line. A *tt6(87)* expression clone was generated by introducing a single base pair mutation (C744 to T) by site directed mutagenesis

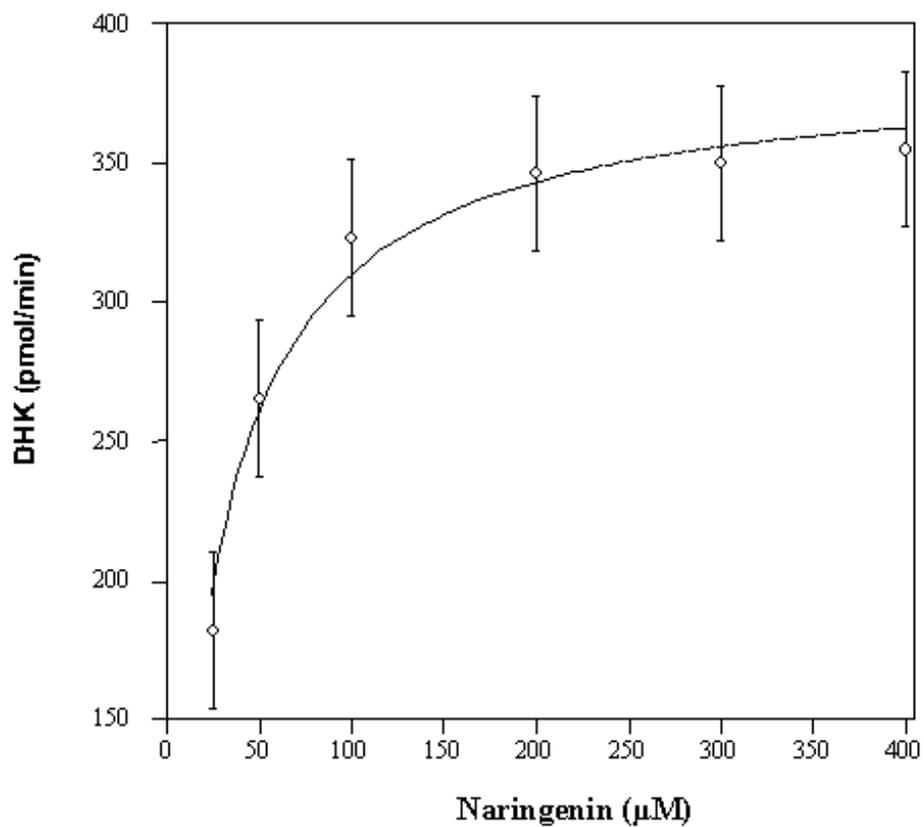


Figure 2.2. Kinetic analysis of recombinant AtF3H. Shown is a typical plot for the Trx-AtF3H reaction, plotting pmoles of DHK formed by 3.5 μg of protein in 10 min using different starting concentrations of the naringenin substrate. The K_M and V_{max} for the reaction and specific activity of the enzyme were determined directly from these plots using Kaleidagraph.

into an AtF3H expression clone. Upon induction and SDS-PAGE analysis, a band consistent with the size of the truncated protein (28 KD) was observed (data not shown). Using the conditions established for wild-type AtF3H, no enzymatic activity was observed (Figure 2.3). This is consistent with the fact that the premature stop codon eliminates a region that includes His275, a conserved Fe²⁺ coordinating residue. Mutation of the equivalent residue in PhF3H resulted in a complete loss of enzyme activity (Lukacin and Britsch, 1997). A loss of 33-34 amino acids from the C-terminus in the PhF3H proteolytic product also leads to an almost complete loss of activity (Lukacin et al., 2000), so it is not surprising the severe truncation of *tt6(87)* by 109 amino acids leads to an inactive enzyme. Residual F3H activity in the mutant line therefore cannot account for the observed leakiness.

Genes annotated as F3H for over 60 species can be identified in public databases, but only nine of these have been verified as enzymatically active. In order to assess how well sequence analysis can predict biochemical activity, the activities of enzymes from *I.alba*, *I.purpurea*, *C.sinensis*, and *M.domestica* were analyzed using the same approach employed for the Arabidopsis protein. These cDNAs were subcloned into pET32a and expressed as N-terminal Trx fusion proteins. Upon SDS-PAGE analysis of cell lysates, bands of the predicted size for the recombinant proteins were observed for each construct. However, apparent proteolytic products consistent with the size of what was previously observed in PhF3H were identified in *I. alba*, *I. purpurea*, and *C. sinensis* and *A. thaliana*. An obvious band for the proteolysis product based on estimated size was not observed in *M. domestica*. *M. domestica* groups on a different branch than the other assayed enzymes by phylogenetic analysis, but the amino acid sequence near the proposed proteolytic cleavage sites is identical in all of the proteins

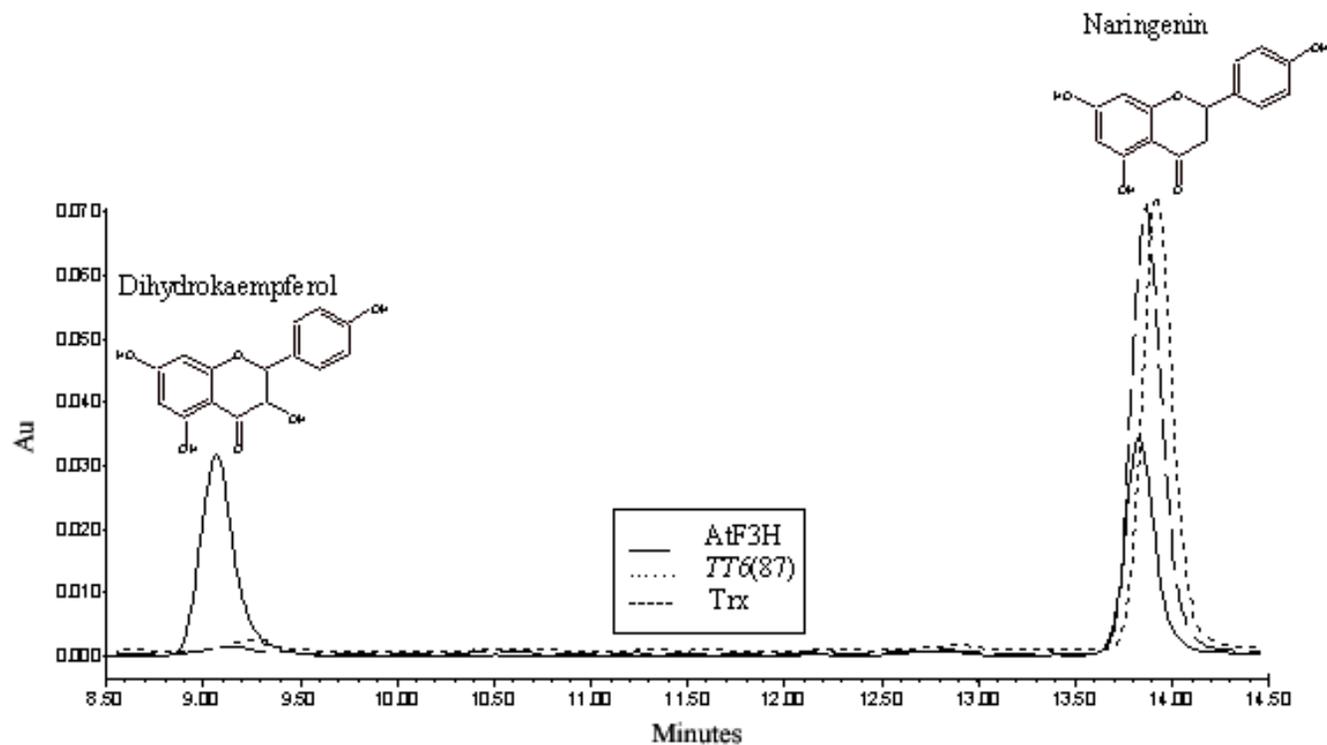


Figure 2.3. Representative HPLC chromatogram demonstrating lack of *TT6(87)* biochemical activity. Peaks are labeled with names and molecular structure of eluted compounds. Recombinant *TT6* was assayed with 3.5 μg recombinant protein at 25°C and pH 8.5 for 60 min. as determined optimal for recombinant AtF3H. *TT6* (solid line) and Trx (dashed line) negative control.

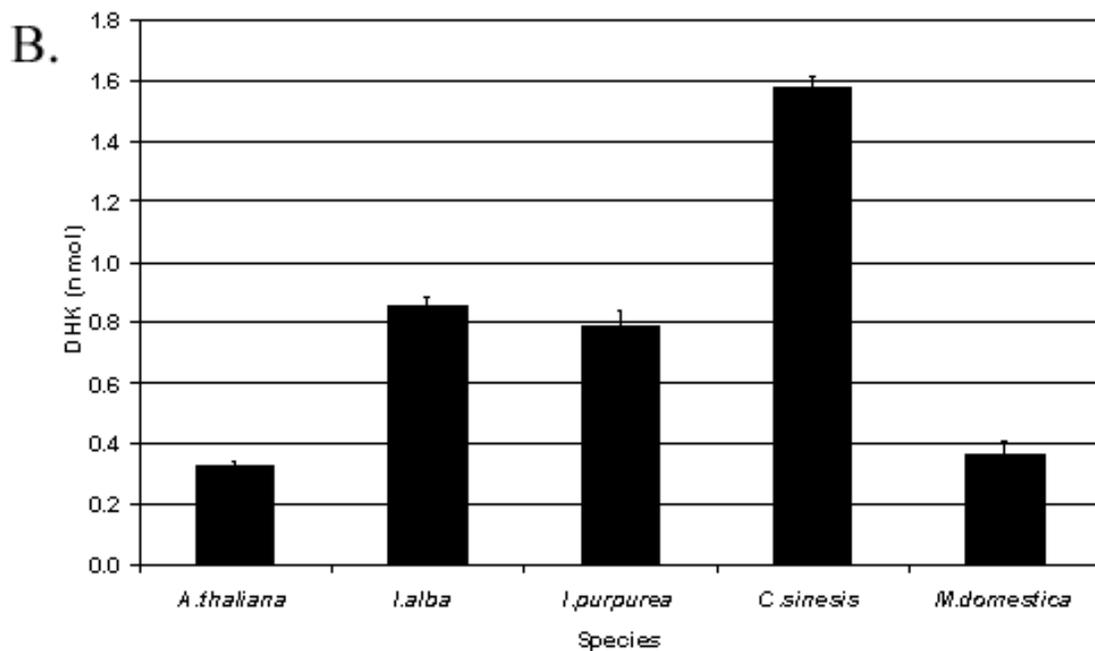
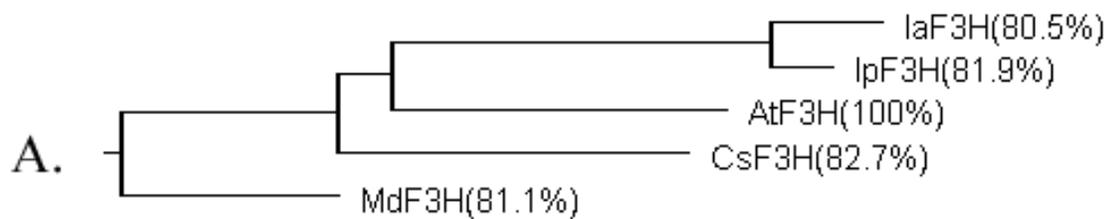


Figure 2.4. Assessment of biochemical activity for recombinant F3H from *I. alba*, *I. purpurea*, *C. sinensis*, and *M. domestica*. A. Phylogenetic tree illustrating relationships between F3H from the various species. Percent protein identities to AtF3H are given in parentheses. B. Assays were conducted in duplicate with 3.5 μ g of recombinant protein at 25°C for 10 min. DHK accumulation is plotted for each species.

(Figure 2.4A). Each enzyme was assayed in duplicate at 25°C and pH 8.5, for 10 min using 3.5 µg of protein, conditions that were optimal for AtF3H activity. All four assayed species were shown to generate DHK, albeit with different levels of activity (Figure 2.4B). The closely related IaF3H and IpF3H had levels of activity that were not substantially different. AtF3H and MdF3H also exhibited comparable levels of activity even though these enzymes are not closely related in terms of sequence. It was observed that the supplied racemic substrate was never entirely consumed for any of the enzymes as found earlier for AtF3H.

DISCUSSION

The measured biochemical characteristics of AtF3H are quite similar to those previously determined for PhF3H suggesting that the proteins are similar functionally. The temperature optimum (40°C) and pH optimum (8.5-8.75) are higher than that of the cytosol (pH 7.5) and normal growth temperature (25°C) for Arabidopsis. At physiological temperature, the enzyme maintains a balance of maximum stability and high activity levels, whereas at the optimal temperature stability is very low but activity is high enough to generate substantial product. This is not necessarily the ideal temperature *in planta*, as enzyme stability is important to maintain activity over time and for any non-enzymatic roles that the protein may serve. The optimal pH observed for *C.sinensis* was 7.5, matching the cytosolic pH exactly, while the Petunia and Arabidopsis enzymes have optima at 8.5 (Britsch and Grisebach, 1986; Punyasiri et al., 2004). The pH optimum may have structural or mechanistic implications. A pH of 8.5 would not have any obvious effects on Fe²⁺, α-KG, or naringenin stability, and would not appear to affect the protonation state of any of the proposed active site residues as compared to physiological pH. Unfortunately, a crystal structure has not yet been determined for F3H, nor for enzymes with

sufficiently high sequence identity on which to base reliable homology models. As this information becomes available, investigation of pH effects at the mechanistic level will be of great interest. Although the measured K_m value of $24 \pm 3 \mu\text{M}$ agrees well with the *Petunia* values of 5.6 to 10 μM , the value is still over two times larger. A higher K_m value would suggest that AtF3H has a lower affinity for naringenin than PhF3H. However, the differences may be due in part to the different temperatures at which the enzymes were assayed. The use of racemic naringenin due to lack of availability of (2S)-naringenin may also have contributed to the observed differences. The biochemical characteristics of AtF3H and PhF3H are quite similar, suggesting that the Trx fusion partner is not having a large impact, but it cannot be discounted that Trx could be having some effects upon activity levels.

It is clear that the leakiness observed in *tt6(87)* cannot be attributed to activity of the truncated form of F3H that is expressed in this line. It has been demonstrated that both FLS1 and ANS are promiscuous for substrates and can convert naringenin into DHK (Prescott et al., 2002; Turnbull et al., 2004), suggesting that these might be able to partially complement a deficiency in F3H activity. However, *En-1* generated and frameshift mutant lines appear to have a null phenotype, producing intensely yellow seeds (Wisman et al., 1998). Homozygous lines containing a T-DNA insertion in the AtF3H 5'UTR (Salk 023664) also have yellow seeds (D.Kwon and B.Winkel unpublished data). AtFLS1 was previously shown to prefer the unnatural (2R) naringenin, and it has been suggested that this may contribute in part to its inability to compensate in F3H mutants (Prescott et al., 2002). However, AtANS and AtFLS1 assayed by incorporation of ^{18}O into DHK were shown to have comparable activities with both naringenin enantiomers (Turnbull et al., 2004). The primary difference between *tt6(87)* and the other mutants is that it results from a premature stop codon, and should resemble the natural

protein up to the point of truncation. All of the other mutants result from insertion or frameshift and therefore contain some non-native sequence near the C-terminus of the protein. However, it is not clear how this would account for the large phenotypic difference between the mutants.

It has been suggested that suppression of the UAG stop codon in *tt6(87)* leads to the production of a full-length protein as has been identified previously in yeast and human genes (Kopczynski et al., 1992; Fearon et al., 1994; Peltola et al., 1994; Wisman et al., 1998). In eukaryotes, the requirements for termination of translation are not yet well established, however, it has been suggested that the region surrounding a stop codon can influence termination efficiency. Therefore, randomly generated stop codons introduced by mutagenesis may occur in a genetic context that allows some leakiness. However, *tt6(87)* contains no detectable full-length F3H protein based on immunoblot analysis (Pelletier et al., 1999). F3H mRNA was also not detected by northern analysis (Pelletier and Winkel, unpublished results). Therefore, stop codon suppression would not appear to be occurring at a level high enough to account for the observed leakiness. It is of note that *tt6(87)* is in a Landsberg ecotype while all of the other mutants were established in Columbia. However there are only two differences between the F3H proteins, a change of Asn to Asp at position 148 and a change of Ala to Asp at position 350, which is not present in the truncated *tt6(87)*. These small changes would not appear to account for the observed leakiness. It is however possible that ecotype differences exist in AtFLS or AtANS that could convey to them the ability to compensate in a Landsberg background. A closer examination of these enzymes in varying Arabidopsis ecotypes could provide important insights into this and other questions of interest.

F3H activity has been demonstrated previously for *Petroselinum crispum*, *Ammi majus*, *Anethum graveolens*, *Pimpinella anisum*, and *Petunia* from purified or recombinant proteins and

in crude extracts from *C. sinensis*, *C. paridisii*, *Forsythia X intermedia* and *Zea mays* (Britsch et al., 1981; Britsch and Grisebach, 1986; Rosati et al., 1998; Owens et al., 2002; Halbwirth et al., 2003; Punyasiri et al., 2004; Gebhardt et al., 2005). We have expanded upon this by demonstrating biochemical activity for the products of predicted F3H from four different species, *I. alba*, *I. purpurea*, *C. sinensis*, and *M. domestica*. All of the F3H enzymes identified by sequence homology were shown to have varying degrees of activity under the assayed conditions. This is not surprising as the homology of all of the assayed proteins with AtF3H was greater than 80%. However, there are cases where proteins identified by sequence alone have been subsequently shown to lack biochemical activity (e.g. Kim et al., 2004). It is surprising that, although all of the proteins were assayed under conditions optimized for AtF3H, this enzyme had the lowest activity. Moreover, although MdF3H did not appear to undergo proteolysis like the other F3H enzymes, activity was not higher. This would suggest that the relative specific activity of MdF3H is lower than that of the other enzymes. Thus although these enzymes share substantial sequence similarity, the existing differences appear to have an impact upon the level of activity among the enzymes. These enzymes may offer an attractive system for optimizing 2ODD enzyme activity based on structure-function relationships.

CHAPTER 3

Examination of the Flavonol Synthase Isoforms in *Arabidopsis thaliana*

Daniel K. Owens, Anne B. Alerding, and Brenda S.J. Winkel

Department of Biological Sciences and Fralin Center for Biotechnology, Virginia Tech,

Blacksburg, VA 24061-0346.

Manuscript to be submitted to Plant Physiology

INTRODUCTION

Flavonoids, and flavonols in particular, have been associated with a large variety of critical physiological processes in plants. Roles ascribed to flavonols include serving as UV sunscreens, protecting plants from oxidative stress, providing pigmentation for the attraction of pollinators and seed dispersers, contributing to male fertility in pollen tube germination, mediating photo- and gravitropism via regulation of auxin transport, and serving as defense molecules in interactions of plants with microbial pathogens, insects, and mammals (reviewed in Harborne and Williams, 2000; Winkel-Shirley, 2001, 2002; Simmonds, 2003). Flavonols are also of growing interest as phytonutrients that convey a variety of purported health benefits when consumed by animals (Rice-Evans, 2004). These compounds have identified antioxidant and anti-inflammatory properties and have been shown to be associated with a decreased incidence of coronary artery disease, stroke, lung and colorectal cancer, and rheumatoid arthritis in cohort studies (reviewed in Arts and Hollman, 2005). The flavonols are represented by three basic structures, kaempferol, quercetin, and myricetin, which are further modified by methylation, acetylation, glycosylation and other reactions to produce the myriad of bioactive flavonols present in plants.

Flavonol synthase (FLS) catalyzes the committed step in the production of flavonols, introducing a double bond into the C ring of the dihydroflavonol backbone. FLS competes for dihydroflavonol substrates with the enzyme, dihydroflavonol reductase (DFR), which leads to proanthocyanidin and anthocyanin pigments, and may thus be an important regulator of flux through this pathway. FLS activity was first described in illuminated *Petroselinum crispum* cell suspension cultures that were being tested for the activity of a related enzyme, flavanone 3-hydroxylase (F3H) (Britsch et al., 1981). Spribille and Forkman (1984) subsequently identified

and characterized FLS activity in crude extracts of flower buds from *Matthiola incana*. FLS was characterized as a member of the 2-oxoglutarate dependant dioxygenases (2-ODD) based on its requirement for alpha keto-glutarate (α -KG), Fe^{2+} , and ascorbate for optimal activity levels. The 2-ODDs represent the largest family of nonheme oxidizing enzymes, and are ubiquitous in living cells with examples having been reported in plants, animals, microorganisms, and even viruses (Schofield and Zhang, 1999). These enzymes are known to perform a variety of reactions including hydroxylation, epoxidation, desaturation, expansion, and ring formation in various metabolic pathways.

The first *FLS* gene was isolated by differential screening of a petunia cDNA library and activity was confirmed using recombinant protein produced in *Saccharomyces cerevisiae* (Holton et al., 1993). *FLS* genes have since been identified in a number of plant species, including *Solanum tuberosum* (potato) (van Eldik et al., 1997), *Petunia hybrida* (petunia) (Holton et al., 1993), *P. crispum* (parsley) (Martens et al., 2003), *Citrus unshiu* (satsuma mandarin) (Moriguchi et al., 2002), and *Arabidopsis thaliana* (Pelletier et al., 1997), based on sequence homology. These genes appear to be single copy in most plant species, including *Oryza sativa* and *Medicago truncatula*, for which full genome sequences are available. However, *Vitis vinifera* L. cultivars Shiraz and Chardonnay each contain two *FLS* genes with distinct expression patterns (Downey et al., 2003, 2004), and Southern analysis has suggested that several *FLS* genes or related sequences may be present in *C. unshiu* (Moriguchi et al., 2002). Surprisingly, the Arabidopsis genome contains six sequences with high sequence similarity to the *FLS* genes of other plant species, even though all other enzymes of the central flavonoid pathway in Arabidopsis are encoded by single genes (Shirley et al., 1995). *AtFLS1* was first described in 1997 and has been extensively characterized at both the genetic and biochemical

levels (Pelletier et al., 1997; Wisman et al., 1998; Prescott et al., 2002; Martens et al., 2003). However, there have not yet been any reports describing the biological or biochemical functions of the other five putative *AtFLS* genes.

It was previously found that *Arabidopsis* accumulates different classes of flavonols in different tissues, with quercetin representing the major flavonol in seeds, kaempferol dominating in whole flowers, and the two flavonols being present at similar levels in stamens (Shirley et al., 1995; Burbulis et al., 1996). In seedlings, high levels of quercetin are present in cotyledons, the upper hypocotyl, and at the root-shoot transition zone, while kaempferol accumulates in defined regions of the root (Murphy et al., 2000; Saslowsky and Winkel-Shirley, 2001). *AtFLS1* mutants generated with the maize transposable element *En-1* were shown to be deficient in kaempferol synthesis, but to accumulate wild type levels of quercetin (Wisman et al., 1998). This finding is consistent with the idea that *AtFLS1* has a substrate preference for dihydrokaempferol (DHK) and that one or more of the other *AtFLS* enzymes is responsible for the production of quercetin.

In this study, we examined the possibility that differential expression of *Arabidopsis* FLS isoforms with different substrate preferences determines the types and amounts of flavonols produced in response to developmental and environmental cues. Sequence analysis of *AtFLS2*, *4*, and *6* revealed that each contains a premature stop codon that eliminates at least one of the highly conserved residues required for coordinating Fe^{2+} in the 2-ODD active site (Lukacin and Britsch, 1997) and likely represent pseudogenes. *AtFLS1* enzyme activity has been previously investigated and shown to recognize both DHK and dihydroquercetin (DHQ) as substrates, although substrate preference was not determined (Wisman et al., 1998; Turnbull et al., 2004). *AtFLS1* was also shown to have promiscuous activity with naringenin, the normal substrate for

the preceding enzyme in flavonoid biosynthesis, F3H, which is also a member of the 2-ODD family.

In this study, activity assays were used to examine the relative utilization of these substrates by the three proposed functional *AtFLS* genes, *AtFLS1*, 3, and 5. A yeast two-hybrid assay was used to explore the possible interaction of these gene products with other flavonoid enzymes. Taken together, the data indicate the *AtFLS1* is the major determinant of flavonol production in Arabidopsis. The results of these studies also suggest that flavonoid enzymes such as F3'H and DFR, rather than other isoforms of FLS, control the relative amounts of quercetin and kaempferol that ultimately accumulate in different tissues.

METHODS

AtFLS cloning and FLS enzyme activity assay

A full-length *AtFLS1* clone was constructed by RT-PCR using total RNA isolated from 15-day-old Arabidopsis *Landsberg erecta* roots employing the RNeasy plant mini-kit (Qiagen). The iScript™ RT-PCR kit (Bio-Rad) was used to generate cDNA (Anne Alerding and Brenda Winkel unpublished) and *AtFLS1* sequences were amplified using the primers, 5'CGCGAATT-CATGGAGGTCGAAAGAGTCC3' and 5'CGCCTCGAGTCAATCCAGAGGAAGTTTATTG3', which incorporate *EcoRI* and *XhoI* sites, respectively (underlined). The PCR product was cloned into the corresponding sites in pET32a (Novagen).

The *AtFLS2* coding region was amplified by PCR from Arabidopsis Col EST clone SQ202h01 (Accession No. AV564339) using the primers, 5'CCGGAATTCATGGAAGTTGAGAGAGA3' and 5'CCGCTCGAGTCACTCGATAGGAAGTT3' (introduced restriction sites underlined). *AtFLS3* and *AtFLS5* were amplified by RT-PCR utilizing RNA isolated from 4-

day-old *Arabidopsis Landsberg erecta* seedlings as described in Pelletier et al. (1996) (Abey Bandera and Brenda Winkel, unpublished results). The primers 5'CCGGAATTCATGGAGATGGAGAAAAC3' and 5'CCGCTCGAGTCAGTCGAGAAGAAGCT3' were used for AtFLS3 while 5'CCGGAATTCATGGAAGAAGAGAGAGA3' and 5'CCGCTCGAGTCAGTAGACAGGAAGAT3' were utilized for AtFLS5 (introduced restriction sites underlined). All three products were cloned into the *Eco*RI and *Xho*I sites in pET32a. An *AtFLS1/AtFLS5* chimeric construct was generated by amplifying a 925 bp fragment from the *Sph*I site in pET32a through the first 120 bp in *AtFLS1* using the primers, 5'CGGCATGCAAGGAGATGG3' and 5'CGGTCGACGTTGGACCTCGGAATG3', which incorporates a *Sal*I site. The product was used to replace the corresponding *Sph*I/*Sal*I fragment in pET32a-*AtFLS5*. The sequence integrity of all clones was confirmed using the BigDye Terminator Cycle Sequencing System (ABI).

All of the constructs were transformed into BL21(DE3)pLysS cells by heat shock, and screened for protein production by small scale test induction as described in the Novagen pET System Manual (8th edition). Recombinant protein expression, activity assays, and HPLC detection were essentially as described in Owens and Winkel (submitted). Activity assays were performed in 50 mM HEPES buffer, pH 7.5, at 25°C for up to 60 min using from 3.5 to 100 µg of crude recombinant protein from lysed BL21(DE3)pLysS cells. Detection was via HPLC using a Waters 2996 photodiode array; chromatograms at OD₂₈₉ were extracted for analysis. All buffers were degassed and assay solutions were made fresh for each set of assays. The ferrous sulfate solution was always prepared in ascorbic acid to decrease oxidation of Fe²⁺.

Homology modeling

Sequence analyses were performed using Lasergene[®] (DNASTar, inc.) or Biology Workbench (<http://workbench.sdsc.edu/>). Homology models were generated for AtFLS1,

AtFLS3, and AtFLS5 based on the crystal structure of Arabidopsis anthocyanin synthase (ANS) (pdb id: 1GP4) (Wilmouth et al., 2002). The sequence of each protein was aligned with ANS and five models generated using Modeller6 essentially as described in Dana, et al. (in press). These five structures were then combined by coordinate averaging with the first structure used as the reference and overlay was on the backbone to generate a single structure. The resulting average structure was subjected to 500 steps of steepest descent minimization using the Sander module of Amber7. The structure was solvated and the net charge of the system brought to zero by the addition of Na⁺ atoms using LeaP. Equilibration was performed on the water and counter ions by molecular dynamics at constant volume for 100 ps. The solvent and counter ions as well as the entire system were each subjected to 500 steps of steepest descent minimization. All molecular dynamics calculations were performed using the AMBER 94 force-field with a time step of 2 fs and coordinates collected every 1 ps. Molecular dynamics consisted of an 80 ps heating phase to raise the temperature from 0 K to 300 K, a 100 ps constant volume equilibration, and a 1 ns constant pressure phase. All calculations were performed using up to eight processors on Virginia Tech's Laboratory for Advanced Scientific Computing and Applications Linux cluster (Anantham). Final models were generated by coordinate averaging from the last 100 ps of dynamics simulation and minimization data. The solvent and Na⁺ ion coordinates were removed from the analyzed files using the Vi text editor to improve visualization of the model. Models were visualized and analyzed using DeepView/Swiss-Pdb Viewer v3.7 sp5 and rendered with POV-ray v3.5. Structural comparisons were performed by aligning the isoform homology models using the Deep View iterative magic fit function, and calculating the corresponding RMSD values.

Yeast Two-Hybrid Analysis

AtFLS1, *AtFLS3*, *AtFLS5*, and *AtDFR* coding sequences from the *Landsberg erecta* ecotype were each amplified from pET32a constructs utilizing the upstream primers 5'CGGTCGACGATGGAGG-TCGAAAG3', 5'CGGTCGACCATGGAGATGGAG3', 5'CCGCTCGAGCATGGAAGAAGA-GAGAG3', and 5'ATGTCGACGCTGCAGGAATTCA-TG3' and downstream primers 5'AGCGGCCGCTCAATCCAGAGGAAG, 5'AGCGGCCGCT-CAGTCGAGAAGAAG3', 5'AGCGG-CCGCTCAGTAGACAGGAAG3', and 5'AGCGGCCG-CGGCACACATCTG3', respectively. Each PCR product was then digested with either *SaII* (*AtFLS1*, *AtFLS3*, and *AtDFR*) or *XhoI* (*AtFLS5*) and *NotI* then inserted into the yeast two-hybrid vectors, pBI880 and pBI881 (Chevray and Nathans, 1992). The *AtF3H* coding region was amplified from pBluescript and introduced into the *PstI* and *NotI* sites of pBI880 and pBI881 using the primers, 5'GCTGCAGATGGCTCCAGGAACTTTG3' and 5'AGCGGCCGCGAGCG-AAGATTTGGTCGAC3' (Dana and Winkel, unpublished results). Plasmids were transformed into DH10B *E.coli* cells by electroporation. The sequence integrity of all clones was confirmed using the BigDye Terminator Cycle Sequencing System (ABI). HF7c yeast cells (Feilotter et al., 1994) were transformed simultaneously with bait and prey constructs essentially as described in Kohalmi, et.al. (1998). From each transformation up to four colonies were chosen at random, grown in 5 ml of -Leu -Trp broth, and then cultured on -Leu -Trp -His solid medium at 30°C.

RESULTS

Six sequences with high homology to *FLS* genes from other plant species have been identified in the Arabidopsis genome, At5g08640 (*AtFLS1*), At5g63580 (*AtFLS2*), At5g63590 (*AtFLS3*), At5g63595 (*AtFLS4*), At5g63600 (*AtFLS5*), and At5g43935 (*AtFLS6*) (Arabidopsis Genome Initiative, 2000). One of these, At5g08640, was previously confirmed to encode an FLS enzyme based on genetic and biochemical data (Wisman et al., 1998; Prescott et al., 2002). The predicted protein sequences for all six *FLS*-like sequences cluster more closely to FLS proteins from other species than with other plant flavonoid dioxygenases based on phylogenetic analysis (Figure 3.1B). All six putative *AtFLS* genes are located on chromosome 5, with *AtFLS2-5* forming a 7.5 Kb tandem array distal from the centromere on the lower arm (Figure 3.1C). The genes within the array are no more closely related to each other than to the other two genes, suggesting that the duplication(s) leading to the amplification of this gene family is an ancient event. *AtFLS1*, 2, 3, and 5 each contain two introns, both of which correspond to previously identified conserved positions within 2-ODDs (Prescott and John, 1996). Further examination of the amino acid sequences showed that *AtFLS2*, 4, and 6 lack Fe²⁺ coordinating and α KG binding residues that are required for enzyme activity. These proteins likely represent inactive enzymes and thus further analysis focused on *AtFLS1*, 3, and 5.

To test the hypothesis that different isoforms of FLS in Arabidopsis with varying substrate specificities are responsible for the differential accumulation of flavonol endproducts, *AtFLS1*, 3, and 5 enzymes were produced in *E. coli* as Trx fusion proteins and assayed for activity with different substrates. Consistent with previous reports, the supplied DHK was converted to kaempferol completely by *AtFLS1* (Figure 3.2A) (Wisman et al., 1998; Lukacin et al., 2003), while only a portion of the supplied DHQ was converted to quercetin by this enzyme

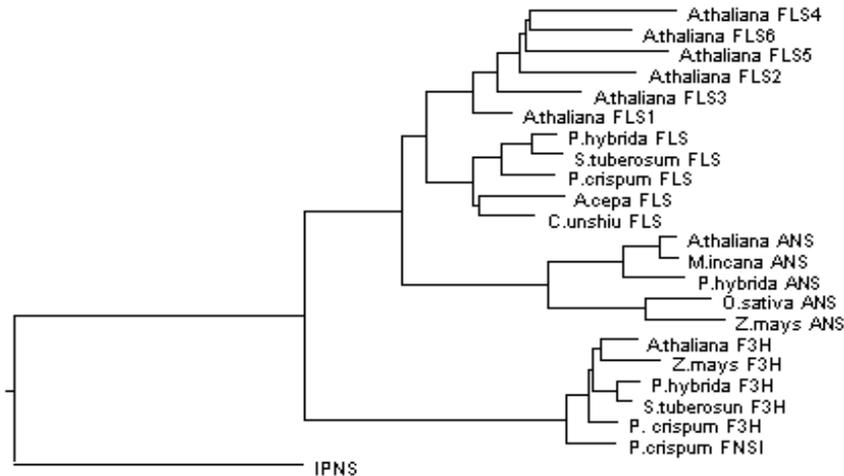
A.

```

1-----MEVE-RDQHISPFSLMAKT-----IPIDLS-----NLDEELV-FLS2
1-----MEME-KNQHISS-----LDIPVIDLS-----MPDEELV-FLS3
1-----MEVE-RDQHKPFLSLQNNKIPSSQNE-----PVVDLS-----NTMSELV-FLS4
1-----MEBE-RDHNASESSLFSLSKQLSSSTLGGSA-----VDVFWVDLS-----VSDDELV-FLS5
1-----MNVB-RDQHISPPCLLTKK-----IPVDLS-----DPSDELV-FLS6
1-----MEVE-RVQDISSSSLLTEAIFLEFIRSEKQPAITTRGPT-----PAIPVVDLS-----DPDEESV-FLS1
1-----MEVASRVQAIASLIRCMDTIPSEYIRSENEQPAITTLHGVE-----LQVFWIDLAPRVVGDGQHDHDDVEV-HieracibergiaFLS
IMKTAQSVSAT--LTMEVA-RVQAIASLIRKMDTIPSEYIRSENEQPAATLHGTV-----LQVFWIDLR-----DPDENM-PetuniaFLS
IMKTIQQSATTALTMEVA-RVQAISSITKCMDTIPSEYIRSENEQPAATLQGVV-----LEVFWIDISN-----VDDDEKL-SolanumFLS
1-----MGVE-RVQDIASATS-KDTIPVEFIRSENEQPGITTVPGTV-----LECPIDFS-----DPDEKL-RosaFLS
1-----MEVE-RVQAIASKMSRCMDTIPSEYIRSENEQPAVITMGGVV-----LQVFWIDLGS-----SMNTEENL-PetroselinumFLS
1-----MEVQ-RVQEIASLSKVIDTIPAEYIRSENEQPVISTVHGTV-----LEVFWIDLS-----DSDEKI-EustomaFLS
1-----MGVE-SV-ERERESN-EGTIPAEFIRSENEQPGITTVHGKV-----LEVPIIDFS-----DPDEKL-MalusFLS
1-----MEVE-RVQAIATLTANLGTIPPEFIRSDHERPDLTTYHGTV-----PELFWIDLA-----MSSQENV-AliumFLS
1-----MEVE-RVQAIASLSHNGTIPAEFIRSEKQPASTTYHGPA-----PEIPTIDL-----DPVQDL-CitrusFLS
1-----MAPT-RVQYVAESRPQ--TIFLEFVRVEERRPINTTDDIGLG-----RQIPVIDMCS-----LEAPELRVKT-GinkgoFLS
1-----MAEVQSVQAIASSLAL--PPEFVRESEHERPGATTRGGD-----APEIPVIDMAAPESAARVAEAREMS-CruciflS
1-----MVAVERVESLAKSGIIS--IPKEYIRPKBELESINDVLEB-----KKEDGQVPTIDLK-----NIESDDEKIR-A.thalianaANS
1-----MVAVERVESLAESEGIKS--IPKEYIRPKBELESINDVQEB-----KKEDGQVPTIDLQ-----NIESDEETR-M.incanaANS
1-----MTDAELRVEALSLSGASA--IPPEYVRPEERADLGDALALAAASODDATARI PVVDISAFD--N--DGDR-O.sativaANS
1-----MVAVVTTPSRVESLAKSGIQA--IPKEYVRQBELNGIGNIFEBEKKDEG-----PQVPTIDLK-----GIDSEKKEIR-P.hybridaANS
1-----MESSPLLQPAARVEALSLSGLSA--IPPEYVRPADERAGLGDAPDLARTDANDHTARRI PVVDISPF--LDSSSQQQQR-Z.maysANS
1-----MAPGTLTELAGEK--LNSKFVRDEDERPKVAYN-VES-----DEIPVLSLA-----GIDD-VGKR-A.thalianaF3H
1-----MAPSTLTALAQEKT--LNSKFVRDEDERPKIAYN-KES-----DEIPVLSLA-----GIDDSDVKR-P.crispumF3H
1-----MGIPVTPSTLTALAEKKT--LQTSFIRDEDERPKVAYN-QES-----NEIPIISLE-----GIDDE-TGKR-P.hybridaF3H
1-----MASTLTALANEKT--LQTSFIRDEEBERPKVAYN-KES-----DEIPVLSLQ-----GIDDI-NGRR-S.tuberosumF3H
1-----MAPVSIASVFLPTAAGEETN--VRASFVREEDERPKVPHD-RES-----DEVFWVSLG-----GIDGP--RRR-Z.maysF3H
1-----MAPTTITALAKEKT--LNLDEVRDEDERPKVAYN-QES-----NEIPIISLA-----GLDDSDGRR-P.crispumFNSI

```

B.



C.



Figure 3.1. Phylogenetic analysis of the AtFLS genes. A, Sequence alignment of the N-terminus of dioxygenases of the flavonoid pathway from diverse species where AtFLS2-6 appear to be missing a highly conserved fragment. B, Phylogenetic tree of flavonoid 2-ODD enzyme sequences from various species demonstrating that the AtFLS isoforms cluster with other FLS enzymes. C, Map of *Arabidopsis* chromosome 5 showing the arrangement of the AtFLS genes. AtFLS2-5 are clustered in a 7.5 Kb region.

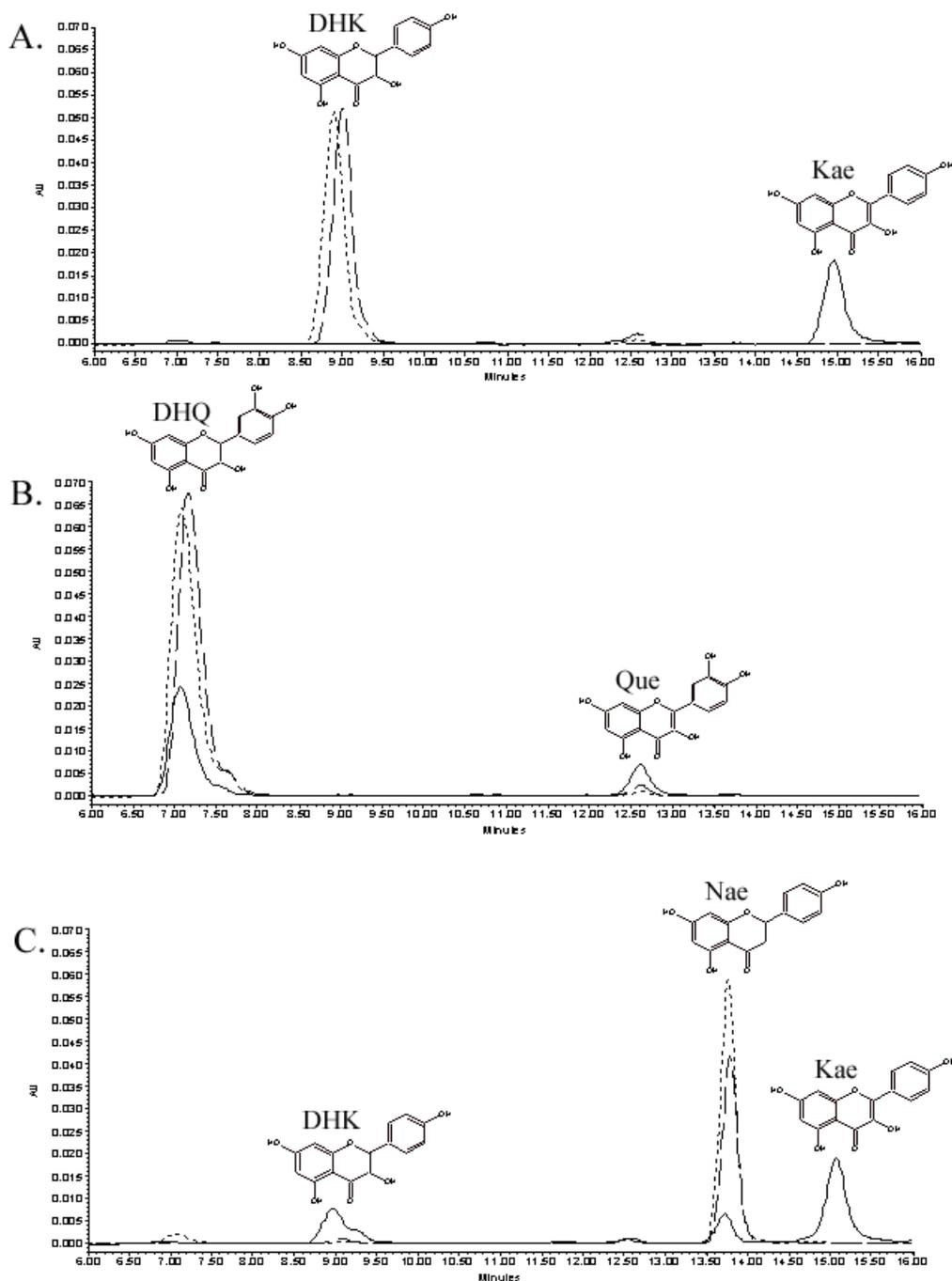


Figure 3.2. Biochemical analysis of AtFLS1, AtFLS3, and AtFLS5. AtFLS1(—), AtFLS3(----), and AtFLS5(.....) were assayed with the substrates DHK (A), DHQ (B), and naringenin (C). Resulting HPLC chromatograms are shown with peaks of interest indicated by label and structure of the corresponding compound. A background peak co-elutes with DHQ as seen in all analyzed samples, and was subtracted out of the final analysis.

(Figure 3.2B)(Turnbull et al., 2004). In addition, a portion of naringenin, normally the substrate for F3H, was converted to DHK by AtFLS1, some of which was subsequently converted to kaempferol (Figure 3.2C)(Prescott et al., 2002). DHK was utilized with the highest efficiency in these assays while, surprisingly, DHQ was used less efficiently even than naringenin. However, neither AtFLS3 nor AtFLS5 appeared to have enzyme activity with any of the substrates under a variety of assayed conditions including variations in pH, temperature, enzyme and substrate concentration, and enzyme enrichment and cleavage procedures.

A closer examination of the primary sequences of the AtFLS proteins led to the identification of an approximate 30 amino acid region that was present in AtFLS1 and all the other plant flavonoid dioxygenases examined, but missing to different degrees in AtFLS2-6 (Figure 3.1A). An arginine and glutamic acid residue in this region were strictly conserved in all of the examined dioxygenase proteins, but were completely absent in AtFLS2-6. AtFLS1 and the FLS proteins from other species appear to share conservation of a proline, also conserved in AtFLS4, and glutamic acid, also conserved in AtFLS5, with F3H that are absent in all the other AtFLS proteins. A proline is conserved in AtFLS1 and the FLS proteins from other species with F3H and FNSI, but absent in AtFLS2-6. AtFLS1 and FLS from all the other examined species contain a pair of adjacent threonines that are not present in AtFLS2-6. However, this insertion does not give any evolutionary insights as it is not clearly more similar in the clustered genes. It appears a deletion in this region may have occurred followed by rapid evolution, including the loss of some additional sequences.

In an effort to understand the observed lack of activity of AtFLS3 and AtFLS5 relative to AtFLS1, the structures of these proteins were analyzed by homology modeling based on the crystal structure (pdb id: 1GP4) of ANS from *Arabidopsis* (Wilmouth et al., 2002). AtFLS1, 3,

and 5 exhibit sufficiently high amino acid sequence identity with AtANS (37.8%, 33.9%, and 31.4%, respectively) to allow the construction of reliable homology models. The RMSD values for the homology models of AtFLS3 and AtFLS5 when compared to AtFLS1 were 1.23Å and 1.48Å respectively, indicating that the structures are quite similar in overall structure (Figure 3.3). The structure of the jellyroll core appears to be well conserved in these proteins relative to AtFLS1, but there are interesting variations in individual residues between the proteins. The Fe²⁺ coordinating residue Asp222 varies by 3.28Å from the corresponding amino acid in AtFLS3, and His220 varies by 2.55Å from the corresponding residue in AtFLS5. The rather large deviation among these critical residues AtFLS3 and AtFLS5 is intriguing in light of the observed lack of biochemical activity for AtFLS3 and AtFLS5. The largest structural differences are near the N-terminus and the largely unstructured C-terminus. An α -helix (AtFLS1 residues 5-8) is present in AtFLS5, although it varies in space from AtFLS1, but is completely absent in AtFLS3. The conserved 30 amino acid fragment in AtFLS1 encodes unstructured regions and a seven amino acid α -helix (AtFLS1 residues 26-31) at the N-terminus near the mouth of the jellyroll motif (Figure 3.3). The absolutely conserved arginine is the next to the last N-terminal residue before the helix and the absolutely conserved glutamic acid is the centermost residue in the helix. This structural element is completely absent in AtFLS3 and AtFLS5.

Initially, AtFLS1 was cloned from EST 153O10T7, which lacks the coding sequences for the 21 N-terminal amino acids. The truncation completely eliminates the first α -helix and the first seven residues in an unstructured region of the conserved 30 amino acid fragment. Interestingly, this construct had no activity with any of the tested substrates when assayed under the same conditions as the full-length AtFLS1 (data not shown). This provides additional

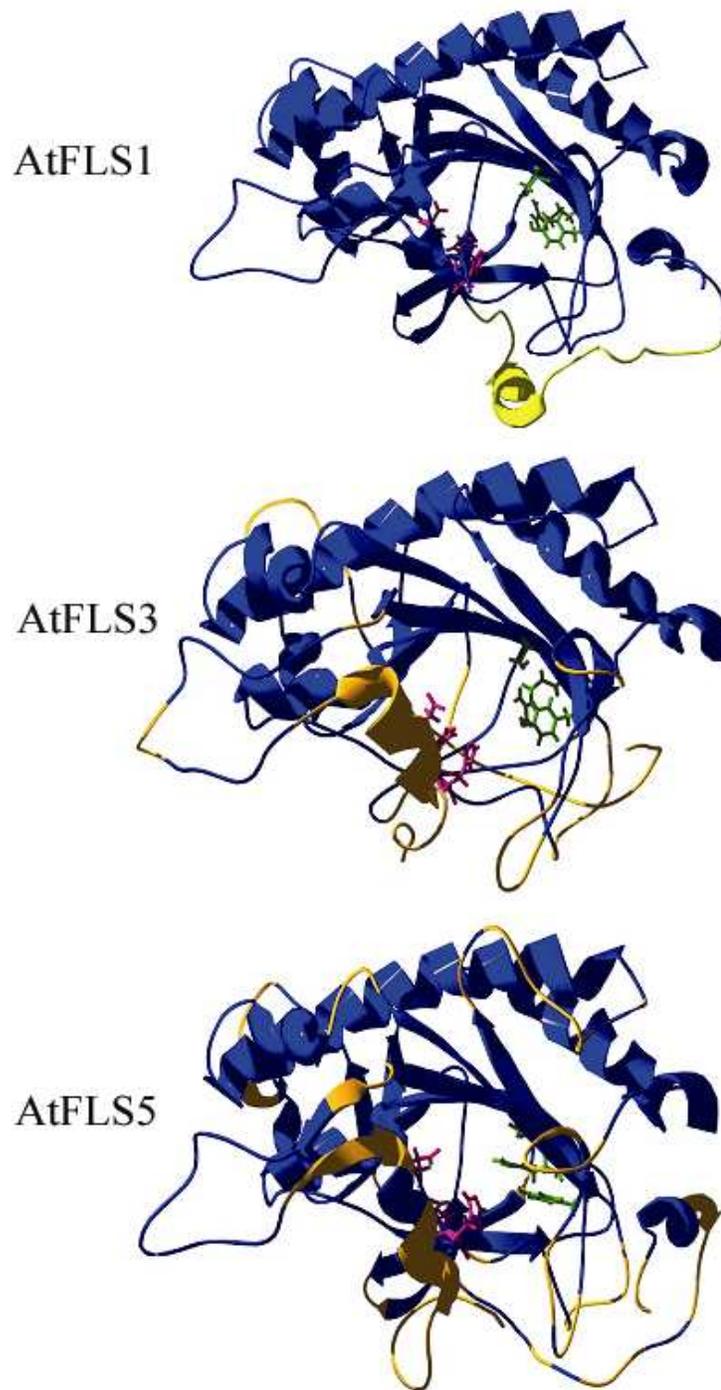


Figure 3.3. Structural comparison of AtFLS1, AtFLS3, and AtFLS5 homology models. Structures are presented in identical orientations viewed looking into the core of the jelly-roll motif. Pink represents Fe²⁺ coordinating and green α-KG binding residues as based on identity to AtANS (At4g22880) in both models. AtFLS1 contains a yellow region indicating the N-terminal fragment missing in all other AtFLS isoforms. AtFLS3 and AtFLS5 regions having RMSD values greater than 2.75Å as compared to AtFLS1 are shown in orange.

evidence that the N-terminus of AtFLS1 contains elements required for FLS enzyme activity. To test the possibility that the N-terminal region of AtFLS1 could restore the activity of the inactive AtFLS isomers, a chimeric construct was generated in which the N-terminal 30 amino acids of AtFLS5 were replaced with the first 40 amino acids of AtFLS1 (Figure 3.4A). However, the chimeric protein also had no detectable activity with any of the tested substrates (Figure 3.4B). This indicates that the 21 N-terminal amino acids of AtFLS1 are required for activity in that protein, but are not sufficient to restore activity to the inactive AtFLS5. One possible explanation for this observation is that evolutionary pressure to maintain structural integrity required for activity was lost once the N-terminal deletion occurred in AtFLS5.

The possibility that AtFLS1, 3, and 5 may serve non-enzymatic roles by acting in a proposed flavonoid biosynthetic metabolon was investigated by yeast 2-hybrid analysis. Interactions were examined against AtCHS, AtCHI, AtF3H, and AtDFR fused to either the activation or DNA binding domain of GAL4. The results of these experiments are summarized in Table 3.1. AtFLS1 interacted with AtF3H and AtDFR in both orientations. AtFLS1, 3, and 5 interacted with AtCHS fused to the GAL4 DNA binding domain, but not when fused to the activation domain. AtFLS5 interacted only with AtDFR fused to the GAL4 activation domain. These results are similar to those reported previously for AtCHS, AtCHI, and AtDFR (Burbulis and Winkel-Shirley, 1999) and suggest that AtFLS1 may function as yet another component of a large multienzyme complex. Moreover, although AtFLS3 and AtFLS5 do not have measurable enzyme activity, these proteins appear to have retained the ability to physically interact with other members of the central flavonoid pathway.

| | FLS1 | FLS3 | FLS5 |
|------------|-------------|-------------|-------------|
| CHS | +/- | +/- | +/- |
| CHI | -/- | -/- | -/- |
| DFR | +/+ | -/- | -/+ |
| F3H | +/+ | ND | ND |

Table 3.1. Yeast 2-hybrid interactions between AtFLS1, AtFLS3, AtFLS5 and CHS, CHI, DFR. AtFLS fused to activation domain/ AtFLS fused to binding domain with + indicating growth and – no growth on -His, -Leu, -Trp media. ND denotes the interaction was not determined.

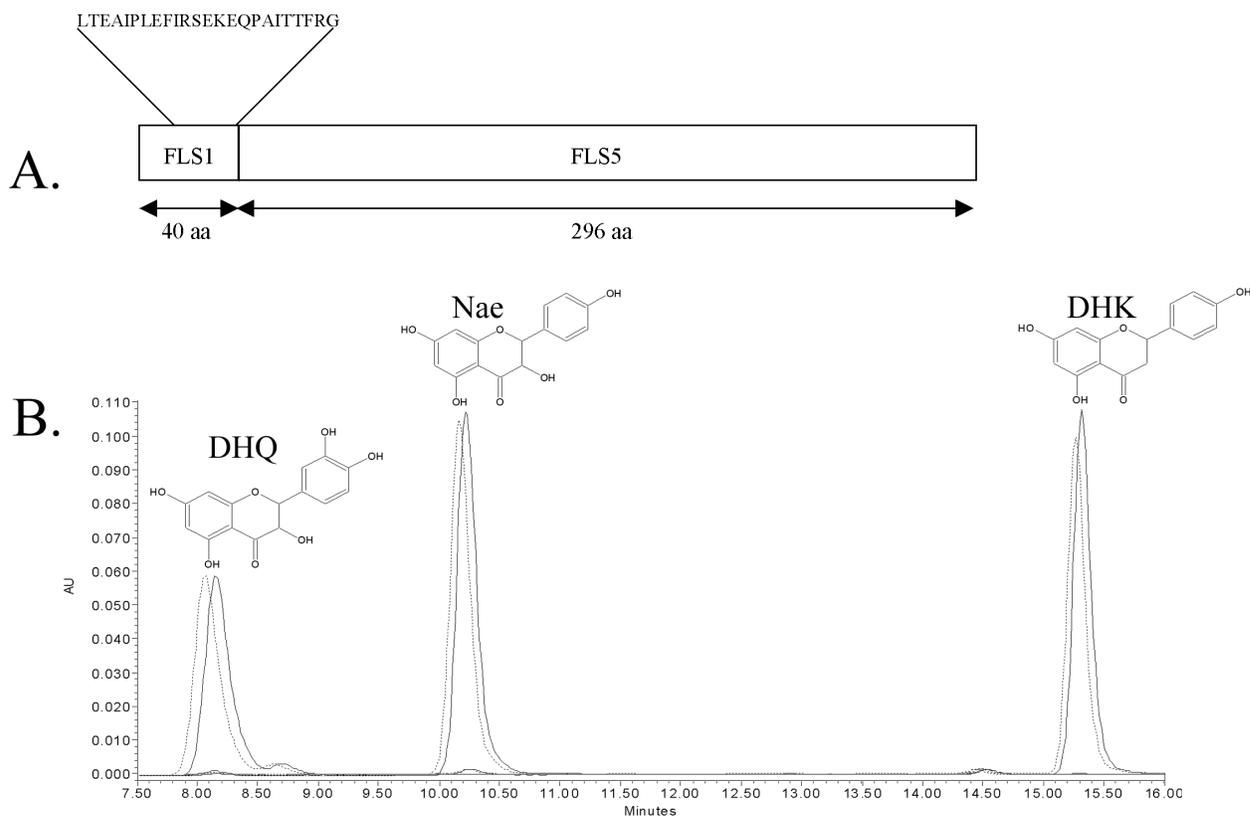


Figure 3.4. Analysis of AtFLS1/AtFLS5 chimera. A, Structure of the AtFLS1/AtFLS5 chimera formed from the 40 N-terminal amino acids of AtFLS1 and the 296 C-terminal amino acids of AtFLS5. Introduced highly conserved N-terminal flavonoid 2-ODD amino acids are indicated above the structure. B, HPLC chromatograms resulting from individual assays with DHQ, DHK, and naringenin for the AtFLS1/AtFLS5 chimera (—) and the thioredoxin negative control (----). Resulting peaks are indicated by corresponding label and structure.

DISCUSSION

The initial evidence for the presence of multiple FLS isoforms in Arabidopsis from low stringency Southern blot analysis has been confirmed with the identification of six genes having high FLS homology in the completed genome sequence. The five uncharacterized *AtFLS* genes group with the previously analyzed *AtFLS1* as well as known *FLS* genes from other organisms. This further suggested that the genes encode FLS and not a different enzyme activity. A closer examination of the genes indicated that *AtFLS2*, *4*, and *6* likely encode non-functional enzymes as each contains a premature stop codon that eliminates at least one essential iron coordinating residue.

AtFLS1, *2*, *3*, and *5* each contain three exons and two introns that are of comparable length in each gene. Exon 2 is highly conserved being 328 bp long in *AtFLS1*, *3*, and *5*. The first intron in *AtFLS2* is over 4 times longer than the first introns in the other genes. Predicted splice junctions within *AtFLS1*, *2*, *3*, and *5* are absolutely conserved corresponding to two of the four possible pairs of positions conserved among all 2-ODDs (reviewed in Prescott and John, 1996). In the Arabidopsis genome annotation, *AtFLS4* and *AtFLS6* each contain the same conserved splice sites as the other *FLS* genes, but also include odd splice sites that do not correspond to conserved positions. This prediction of splicing eliminates the premature stop codons we have identified within these genes, but also introduces an additional short intron that is not consistent with conserved 2-ODD gene structure. This suggests that the original Arabidopsis genome annotation is incorrect, likely in part due to the lack of EST sequences for these genes, which have very limited or no expression. The positioning of the genes may have evolutionary implications as *AtFLS2*, *3*, *4*, and *5* form a cluster near the bottom of chromosome 5 while *AtFLS1* and *AtFLS6* are isolated higher on the chromosome. However, *AtFLS2*, *3*, *4*, and *5*

do not group closer together than *AtFLS1* and 6 by phylogenetic analysis (Figure 3.1B) suggesting that they did not evolve as a group. It is also of note, that despite a small number of differences at the nucleotide level, *AtFLS1-6* encode virtually identical proteins between the Ler and Col ecotypes. This may imply that the catalytically inactive FLS isoforms are serving some role as a gene having no evolutionary advantage would be expected to accumulate mutations over time.

The initial hypothesis that differential expression of the AtFLS isoforms with varying substrate specificities is responsible for the previously observed differential accumulation of kaempferol and quercetin in Arabidopsis tissues appears to be incorrect based on results from the recombinant enzymes. It appears that although Arabidopsis has six genes that encode FLS enzymes, only *AtFLS1* leads to the production of an active enzyme. AtFLS3 and AtFLS5 did not have measurable enzyme activity with either the natural substrates DHK and DHQ or the promiscuous substrate naringenin under conditions in which AtFLS1 was clearly active (Figure 3.2). This is a surprising result due to the high degree of homology between these enzymes and AtFLS1, and the absolute conservation of the predicted 2-ODD Fe²⁺ coordinating and α KG binding residues in these proteins.

These results suggest that regions required for enzymatic activity within the proteins remain to be identified. AtFLS1, 3, and 5 were re-examined to determine if any clear differences could be identified at the sequence level. Overall, the majority of amino acid changes are well dispersed across the proteins. However, one clear difference was a 30 amino acid fragment that exists in AtFLS1, but was absent in all the other isoforms. Further comparisons with flavonoid dioxygenases from various plant species demonstrated that a comparable region existed in all the proteins, with two of the amino acids absolutely conserved. Further affirming the importance of

this region, a previously-cloned partial cDNA for *AtFLS1* (Pelletier et al., 1999) missing a large portion of this region was shown to be catalytically inactive. Replacement of the N-terminus of AtFLS5 with sequences from AtFLS1 did not restore enzymatic activity (Figure 3.4B), even though homology modeling suggested that the chimeric protein should undergo proper folding. Therefore the N-terminal region of FLS contains elements that should be further investigated for roles in activity, but other regions of importance remain to be identified.

AtFLS1, 3, and 5 were analyzed at the tertiary structural level by homology based modeling and molecular dynamics simulations using the crystal structure of AtANS as a template (Figure 3.3). All of the AtFLS enzymes appear to have strong conservation of the jellyroll motif with the Fe^{2+} coordinating and αKG binding residues projecting into the core of the enzyme. The greatest variability in the structures occurs at the N- and C-termini of the proteins. The 30 amino acid region in AtFLS1 that is absent in the other proteins encodes a short unstructured loop followed by an α -helix. AtFLS3 lacks the initial N-terminal α -helix observed in AtFLS1 and AtFLS5, and the RMSD values vary greatly from AtFLS1. AtFLS3 and AtFLS5 contain a much larger final C-terminal α -helix than what was observed in AtFLS1 although primary sequence variation in this region is minimal. The large α -helix at the C-terminus of AtFLS3 and AtFLS5 more closely resembles ANS than does AtFLS1. There is also variation in the unstructured regions of the proteins.

Why does Arabidopsis contain six genes for FLS if only one of the genes produces a protein with enzyme activity? One possibility is that the non-functional genes are evolutionary artifacts. For example, a gene duplication of *AtFLS1* produced a new *FLS* gene but errors in the process led to the elimination of a 30 aa fragment required for activity. Subsequent duplications of this inactive gene may have produced the remaining *FLS* isoforms, which rapidly gained a

number of mutations as the genes encode inactive proteins. Although this is an established mechanism to introduce variability in a population, the number of non-functional *FLS* genes seems quite large. Interestingly, multi-enzyme families, although not the case for most enzymes of the flavonoid pathway, are common in the relatively small, approximately 140 Mb, genome of *Arabidopsis* (Blanc et al., 2000; Arabidopsis Genome Initiative, 2000). One explanation for this phenomenon is that large-scale duplications appear to have occurred in this genome (Blanc et al., 2000).

Another consideration is that although the proteins encoded by *AtFLS2-6* are catalytically inactive several of these genes are expressed, albeit on a limited scale relative to *AtFLS1* (Alerding and Winkel, unpublished data), suggesting that they may serve non-enzymatic functions in the cell. One possibility is that they perform a structural role in the proposed *A. thaliana* flavonoid metabolon. To test this possibility, *AtFLS1*, *AtFLS3*, and *AtFLS5* were assayed in the yeast 2-hybrid system against other members of the central flavonoid biosynthetic pathway. Six interactions were identified of which four were only detectable in certain fusion combinations. Interactions have been previously observed with members of the flavonoid biosynthetic pathway in the yeast 2-hybrid system (Burbulis and Winkel-Shirley, 1999). These interactions were dependant upon whether the tested proteins were fused to the GAL4 activation or binding domain. It has been suggested that these construct specific interactions indicate a masking of domains or that interactions are dependant upon protein orientation. All of the isoforms interacted with *AtCHS* in a specific manner suggesting that this interaction domain is well conserved within the proteins. The observed interactions between *AtF3H* and *AtFLS1* may be significant in the context of metabolic channeling as they represent concurrent steps in the metabolic pathway. The observed interactions of *AtFLS1* and *AtFLS3* with *AtDFR* are

interesting as these enzymes compete for naringenin within the cell. FLS catalyzes the committed step in the synthesis of flavonols while DFR leads to the production of the proanthocyanidin or anthocyanidin pigments. This may suggest that interactions among the flavonoid enzymes directly impact accumulation of end products. It is also of note that the ability of the inactive AtFLS3 to interact with DFR could represent a method to control metabolic flux within the pathway. The possibility also exists that AtFLS3, AtFLS2, and AtFLS5 serve as non-enzymatic structural elements within the metabolon. Surface plasmon resonance is currently underway to further test isoform-specific protein interactions.

One puzzling aspect of this work that remains to be investigated is the source of differential accumulation of kaempferol and quercetin in Arabidopsis tissues. A single functional FLS enzyme has been identified that appears to have much higher activity levels with DHK, and apparently even naringenin, than DHQ. Studies of an FLS enzyme isolated from *Citrus unshui* fruits, which also has a strong preference for DHK *in vitro*, have demonstrated these tissues contain primarily quercetin-3-*O*-rutinoside (Moriguchi et al., 2002; Wellmann et al., 2002). Although in this case multiple isoforms with varying substrate preferences may still be a possibility, an alternative explanation is that the accumulation of kaempferol and quercetin is not controlled by the substrate preference of FLS, but by competition with DFR for the common substrate DHQ.

DHQ is formed from DHK by the enzyme flavanone 3' hydroxylase (F3'H) that has also been predicted to serve as a membrane anchor for the flavonoid enzyme complex (reviewed in Hrazdina and Jensen, 1992; Saslowsky and Winkel-Shirley, 2001). Due to its functionality in the flavonoid metabolon, F3'H is expected to always be present when flavonoids are being synthesized and should produce DHQ. DFR has a strong substrate preference for DHQ and little

or no activity with DHK (Forkmann and Ruhnau, 1987). If DFR and FLS are both present, DFR is likely to consume all the DHQ leading to the production of anthocyanins, and FLS could react with the remaining DHK to produce kaempferol. In tissues where DFR is not present, like roots, there would be an abundance of DHQ leading to the accumulation of quercetin.

CHAPTER 4

SUMMARY

The primary goal of this study has been to gain a deeper understanding of the 2-oxoacid dependant dioxygenase (2-ODD) enzymes that affect flavonol production in Arabidopsis. Four types of 2-ODD enzymes play essential roles in the central flavonoid biosynthetic pathway. Two of the enzymes, flavanone 3-hydroxylase (F3H) and flavonol synthase (FLS), are required for flavonol biosynthesis (Figure 1.1). As one approach to characterizing the specific biological and biochemical function of these enzymes, an activity assay employing recombinant protein expression and HPLC as a detection system was developed for F3H and adapted for use with FLS.

The kinetic properties of Arabidopsis F3H were established for the first time using the developed assay. The assay was further employed to examine the activities of F3H from *Ipomoea alba* (moonflower), *Ipomoea purpurea* (common morning glory), *Citrus sinensis* (sweet orange), and *Malus X domestica* (newton apple) that had previously been identified as encoding F3H activity based solely on sequence homology. Activity was confirmed for all investigated constructs, bringing the total number of *F3H* genes verified by direct biochemical assay to nine. Verification of biochemical activity of 2-ODD genes is an important aspect of the work that should be expanded in the future. Although the observations with the predicted F3H proteins from these species were as expected, this is not always the case. For example, the cinnamyl alcohol dehydrogenase gene family was annotated as containing seventeen members in Arabidopsis, but eight of the genes were shown to be incorrectly identified (Kim et al., 2004). There are growing numbers of putative F3H genes that remain to be investigated. It would be of great interest to expand our analysis beyond activity verification alone to more in-depth kinetic characterization. These experiments could then be combined with evolutionary comparisons

similar to that of Gebhardt, et.al. (2005) with flavonoid dioxygenases from the Apiaceae family, in which flavone synthase I was shown to have likely arisen from gene duplication of F3H.

The source of the leaky phenotype of the AtF3H mutant, *tt6(87)*, remains to be determined. The F3H activity assay was used to establish that the mutant protein does not have residual activity, as also predicted from sequence analysis, indicating that it cannot be the source of the observed phenotype. The previous suggestion that there is suppression of the premature stop codon in this allele (Wisman et al., 1998) appears implausible as no full-length AtF3H RNA (Matt Pelletier and Brenda Winkel, unpublished results) or protein (Pelletier et al., 1999) has been detected in *tt6(87)* plants. It is possible that the leakiness may be a result of the small variations in sequence among *Arabidopsis* ecotypes as the null mutants isolated by *En-1* transposon insertional mutagenesis (Wisman et al., 1998) are in the Columbia ecotype while *tt6(87)* is in Landsberg. Small differences in sequence between AtANS or AtFLS from Columbia or Landsberg may allow these enzymes to compensate for the loss of AtF3H via their identified promiscuous activity (Prescott et al., 2002; Lukacin et al., 2003; Turnbull et al., 2004) in some ecotypes but not in others. Differences among *Arabidopsis* ecotypes in herbivory (Jander et al., 2001), disease resistance (Kunkel, 1996), epicuticular wax accumulation (Rashotte et al., 1997), and trichome number and spacing (Larkin et al., 1996) have been described previously, often as the result of genetic changes in multiple genes. As one approach to addressing this question, the ANS, F3H, and FLS enzymes from various ecotypes could be compared by a combination of phylogenetic and biochemical analyses to determine their relationships and activity levels with potential substrates.

The hypothesis that differential expression of AtFLS isozymes with varying substrate specificities is responsible for the tissue-specific differences in flavonol accumulation observed

in Arabidopsis was tested. A colleague's work established that the genes have differential tissue and developmental expression patterns by promoter-GUS and semi-quantitative RT-PCR analysis. Publicly available data from MPSS and microarray databases also suggest differential patterns of expression for five of the genes, while *AtFLS6* does not appear to be expressed. However, sequence analysis of the six *AtFLS* genes revealed that three of the genes contain premature stop codons that eliminate residues essential for activity and these therefore likely represent pseudogenes. Furthermore, upon biochemical analysis it was found that only one AtFLS gene appears to encode an active enzyme, discounting the initial hypothesis that different isoforms have different enzymatic function. A closer examination of the proteins led to the identification of approximately 30 residues present in the active AtFLS1 isoform and all other examined 2-ODD enzymes, but absent in the inactive AtFLS isoforms. Analysis of an AtFLS1/AtFLS5 chimera showed that although these residues are apparently required for catalysis, they are not sufficient to restore activity to an inactive isoform. The solving of the crystal structure of the Arabidopsis 2-ODD enzyme, AtANS (Wilmouth et al., 2002), allowed for homology based modeling and molecular dynamic analysis of the AtFLS isoforms. A comparison of the models demonstrated that the greatest structural variability occurs near the N- and C-termini. One of the most striking differences was the complete lack of an α -helix corresponding to the 30 residues that are missing in AtFLS2-6 suggesting that this region might be important for enzyme activity. It was established by yeast 2-hybrid assay that AtFLS1, 3, and 5 have the ability to interact with other members of the flavonoid biosynthetic pathway; however, whether this has biological significance or represents an evolutionary artifact has yet to be determined.

The observations made with the *AtFLS* gene family were somewhat unexpected, and much work remains to be done with the system. An immediate goal is to analyze *AtFLS* gene-specific knockout plants and determine the effects on flavonol accumulation as a means to determine the *in vivo* roles of these genes. It was previously reported that an *AtFLS1 En-1* mutant line does not appear to be altered in kaempferol production as indicated by non-quantitative thin layer chromatography (Wisman et al., 1998). However, this remains to be established in a stable mutant line. Moreover, none of the other *AtFLS* genes have been characterized using a genetic approach. Each of the isoforms is represented in available T-DNA knockout collections (Rosso et al., 2003). A number of these lines have been acquired and are currently being bred to homozygosity for subsequent analysis of flavonol content. When verified homozygous lines are obtained, the initial step will be to analyze whole plant flavonol levels in each line. The protocols for this analysis are available, and the experiment should be relatively straightforward. However, the *AtFLS* genes have limited expression patterns, so it may be necessary to assay flavonol content in specific tissues. Another method to test the hypothesis that only *AtFLS1* has effects on flavonol accumulation would be to generate double mutant lines between *AtFLS1* and the other isoforms.

Another important goal that remains to be accomplished is to assess *in vivo* protein levels for these enzymes. The difficulty in achieving this lies in the lack of specific antibodies that distinguish between the isoforms. An antibody was previously generated against *AtFLS1* (Pelletier et al., 1999), but it cross-reacts with the other isoforms (Juan Stella and Brenda Winkel, unpublished results). One way to address this problem might be to generate synthetic peptides for unique regions within each isoform and use these to generate antibodies. However, enough localized variation does not exist among the proteins, except the 30 residue region

identified near the N-terminus of AtFLS1, for this to have a likelihood for success. Unless emerging technologies provide a solution, it appears there is currently no way to generate specific antibodies for the remaining isoforms that has a likelihood of success high enough to justify the accompanying cost. However, generating an AtFLS1 specific antibody would be a good preliminary step to quantifying *in vitro* FLS protein levels. AtFLS1 has demonstrated biochemical activity and appears to affect kaempferol accumulation so it would be a good initial protein for study to extend current knowledge.

An aspect of AtFLS1 enzymology that would be interesting to pursue is the further characterization of the amino terminal region of this protein. The lack of activity in an N-terminally truncated mutant of AtFLS1, together with the absence of the 30 residue segment in the inactive AtFLS2-6 isoforms, indicates that this is a critical region; however, its actual function within the enzyme remains to be identified. The 30 residue segment lies near the opening of the jellyroll core motif, and is distant from the predicted active site of the enzyme. This region is more likely important in maintaining overall protein structure than contributing directly to catalysis. Initial experiments into the significance of these residues would be best performed in AtANS as a crystal structure for the enzyme exists with substrates and co-factors bound. Generating and biochemically assaying a mutant AtANS protein with the region of interest eliminated would serve to further establish this region as significant for general dioxygenase function. An activity assay for AtANS has been developed (Turnbull et al., 2003) or the HPLC based assay for F3H and FLS could be adapted for use to eliminate the need for radioactively labeled substrates. The crystal structure for AtANS could be subjected to molecular dynamics simulations to explore the relationship between the motion of the region of interest and the predicted active site residues. The dynamics simulations performed with the

homology models of enzymatically inactive AtFLS3 and 5 would serve as a good comparison to assess the importance of the region of interest in catalysis.

The *AtFLS* system would appear to offer a unique opportunity to investigate the evolution of gene families. Our results indicate that all of the members of the *AtFLS* gene family except one have lost enzymatic function and may no longer serve an evolutionary advantage to the plant. If further evidence supports this idea, this could be an ideal model system in which to study loss of function and the subsequent divergence of genes. This work would serve well in comparisons to systems evolving new functions as an approach to develop a holistic picture of gene redundancy and evolution.

Another long-term goal that remains to be achieved is the solving of crystal structures for both F3H and FLS. The determination of these structures would be an important step in understanding the specific reaction mechanisms of these enzymes and would provide essential information for elucidating the structure and regulation of the flavonoid multienzyme complex. The residues involved in coordinating the Fe^{2+} co-factor and α -KG co-substrate have been identified and appear to be strongly conserved among 2-ODDs, but the binding of flavonoid substrates is not as clearly defined and varies among the individual enzymes. A crystal structure of FLS or F3H with flavonoid substrates bound would be an important tool to further address this question. This will be a daunting project as 2-ODD enzymes are very unstable, and AtF3H in particular is known to undergo proteolysis upon purification. However, the success solving the structure of AtANS suggests that this is an attainable goal (Wilmouth et al., 2002). AtFLS and AtF3H crystal structures would not only greatly increase our knowledge of the Arabidopsis enzymes, but could serve as templates for homology modeling and structural analysis in other systems. The development of methods to prevent AtF3H and AtFLS proteolysis and degradation

will be important to reach the overall goal, but even the generation of a crystal structure for a truncated form of either protein would provide insights into the organization and function of the proteins.

This project has led to a number of new insights into 2-ODD enzymes as well as gene families. The findings for the *AtFLS* gene family are different from what was initially anticipated and have led to a series of new questions. These questions range from structural and biochemical characteristics to evolutionary relationships and offer much opportunity for collaborations to develop a well-rounded and comprehensive view of the 2-ODD enzymes in plants and other organisms.

REFERENCES

- Arabidopsis Genome Initiative** (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**: 796-815
- Arts IC, Hollman PC** (2005) Polyphenols and disease risk in epidemiologic studies. *Am J Clin Nutr* **81**: 317S-325S
- Blanc G, Barakat A, Guyot R, Cooke R, Delseny M** (2000) Extensive duplication and reshuffling in the *Arabidopsis* genome. *Plant Cell* **12**: 1093-1101
- Britsch L, Dedio J, Saedler H, Forkmann G** (1993) Molecular characterization of flavanone 3 beta-hydroxylases. Consensus sequence, comparison with related enzymes and the role of conserved histidine residues. *Eur J Biochem* **217**: 745-754
- Britsch L, Grisebach H** (1986) Purification and characterization of (2S)-flavanone 3-hydroxylase from *Petunia hybrida*. *Eur. J. Biochem.* **156**: 569-577
- Britsch L, Heller W, Grisebach H** (1981) Conversion of flavanone to flavone, dihydroflavonol and flavonol with an enzyme-system from cell-cultures of parsley. *Zeitschrift Fur Naturforschung C-a Journal of Biosciences* **36**: 742-750
- Britsch L, Ruhnau-Brich B, Forkmann G** (1992) Molecular cloning, sequence analysis, and *in vitro* expression of flavanone 3 β -hydroxylase from *Petunia hybrida*. *J. Biol. Chem.* **267**: 5380-5387
- Brown DE, Rashotte AM, Murphy AS, Normanly J, Tague BW, Peer WA, Taiz L, Muday GK** (2001) Flavonoids act as negative regulators of auxin transport *in vivo* in *Arabidopsis thaliana*. *Plant Physiol.* **126**: 524-535
- Burbulis IE, Iacobucci M, Shirley BW** (1996) A null mutation in the first enzyme of flavonoid biosynthesis does not affect male fertility in *Arabidopsis*. *Plant Cell* **8**: 1013-1025

- Burbulis IE, Winkel-Shirley B** (1999) Interactions among enzymes of the *Arabidopsis* flavonoid biosynthetic pathway. Proc. Natl. Acad. Sci. USA **96**: 12929-12934
- Chevray PM, Nathans D** (1992) Protein interaction cloning in yeast: Identification of mammalian proteins that react with the leucine zipper of Jun. Proc. Natl. Acad. Sci. USA **89**: 5789-5793
- Dixon RA** (2004) Phytoestrogens. Annual Review of Plant Biology **55**: 225-261
- Downey MO, Harvey JS, Robinson SP** (2003) Synthesis of flavonols and expression of flavonol synthase genes in the developing grape berries of Shiraz and Chardonnay (*Vitis vinifera* L.). Australian Journal of Grape and Wine Research **9**: 110-121
- Downey MO, Harvey JS, Robinson SP** (2004) The effect of bunch shading on berry development and flavonoid accumulation in Shiraz grapes. Australian Journal of Grape and Wine Research **10**: 55-73
- Englard S, Seifter S** (1986) The biochemical functions of ascorbic acid. Annu Rev Nutr **6**: 365-406
- Faulkner IJ, Rubery PH** (1992) Flavonoids and flavonoid sulphates as probes of auxin-transport regulation in *Curcubita pepo* hypocotyl segments and vesicles. Planta **186**: 618-625
- Fearon K, McClendon V, Bonetti B, Bedwell DM** (1994) Premature translation termination mutations are efficiently suppressed in a highly conserved region of yeast Ste6p, a member of the ATP-binding cassette (ABC) transporter family. J Biol Chem **269**: 17802-17808
- Feilotter HE, Hannon GJ, Ruddell CJ, Beach D** (1994) Construction of an improved host strain for two hybrid screening. Nucl. Acids Res. **22**: 1502-1503

- Forkmann G, Ruhnau B** (1987) Distinct Substrate-Specificity of Dihydroflavonol 4-Reductase from Flowers of *Petunia-Hybrida*. *Zeitschrift Fur Naturforschung C-a Journal of Biosciences* **42**: 1146-1148
- Fritsch H, Grisebach H** (1975) Biosynthesis of cyanidin in cell cultures of *Haplopappus gracilis*. *Phytochemistry* **14**: 2437-2442
- Fritze K, Staiger D, Czaja I, Walden R, Schell J, Wing D** (1991) Developmental and UV light regulation of the snapdragon chalcone synthase promoter. *Plant Cell* **3**: 893-905
- Furuya MA, Galston AW, Stowe BB** (1962) Isolation from peas of co-factors and inhibitors of indole-3-acetic acid oxidase. *Nature* **193**: 456-457
- Galson AW** (1969) Flavonoids and photomorphogenesis in peas. *In* JB Harborne, T Swain, eds, *Perspectives in Phytochemistry*. Academic Press, New York, pp 193-204
- Gebhardt Y, Witte S, Forkmann G, Lukacin R, Matern U, Martens S** (2005) Molecular evolution of flavonoid dioxygenases in the family *Apiaceae*. *Phytochemistry*
- Halbwirth H, Martens S, Wienand U, Forkmann G, Stich K** (2003) Biochemical formation of anthocyanins in silk tissue of *Zea mays*. *Plant Science* **164**: 489-495
- Harborne JB, Williams CA** (1988) Flavone and flavonol glycosides. *In* JB Harborne, ed, *The Flavonoids, Advances in Research Since 1980*. Academic Press, New York, p 303
- Harborne JB, Williams CA** (2000) Advances in flavonoid research since 1992. *Phytochemistry* **55**: 481-504
- Hegg EL, Que L, Jr.** (1997) The 2-His-1-carboxylate facial triad--an emerging structural motif in mononuclear non-heme iron(II) enzymes. *Eur J Biochem* **250**: 625-629
- Holton TA, Brugliera F, Tanaka Y** (1993) Cloning and expression of flavonol synthase from *Petunia hybrida*. *Plant J.* **4**: 1003-1010

- Hrazdina G, Jensen RA** (1992) Spatial organization of enzymes in plant metabolic pathways. *Ann. Rev. Plant Physiol. Plant Mol. Biol.* **43**: 241-267
- Hrazdina G, Wagner GJ** (1985) Compartmentation of plant phenolic compounds; sites of synthesis and accumulation. *Annu. Proc. Phytochem. Soc. Europe* **25**: 120-133
- Hrazdina G, Wagner GJ** (1985) Metabolic pathways as enzyme complexes: evidence for the synthesis of phenylpropanoids and flavonoids on membrane associated enzyme complexes. *Arch. Biochem. Biophys.* **237**: 88-100
- Hrazdina G, Zobel AM, Hoch HC** (1987) Biochemical, immunological and immunocytochemical evidence for the association of chalcone synthase with endoplasmic reticulum membranes. *Proc. Natl. Acad. Sci. USA* **84**: 8966-8970
- Jacobs M, Rubery PH** (1988) Naturally occurring auxin transport regulators. *Science* **241**: 346-349
- Jander G, Cui J, Nhan B, Pierce NE, Ausubel FM** (2001) The TASTY locus on chromosome 1 of *Arabidopsis* affects feeding of the insect herbivore *Trichoplusia ni*. *Plant Physiol* **126**: 890-898
- Jourdan PS, McIntosh CA, Mansell RL** (1985) Naringin levels in citrus tissues .2. quantitative distribution of naringin in *Citrus paradisi* macfad. *Plant Physiology* **77**: 903-908
- Kelly CK** (1990) Plant foraging: a marginal value model and coiling response in *Cuscuta subinclusa*. *Ecology* **71**: 1916-1925
- Kempner ES, Miller JH** (1968) The molecular biology of *Euglena gracilis*. *Exp. Cell Res.* **51**: 141-149

- Kim SJ, Kim MR, Bedgar DL, Moinuddin SG, Cardenas CL, Davin LB, Kang C, Lewis NG** (2004) Functional reclassification of the putative cinnamyl alcohol dehydrogenase multigene family in *Arabidopsis*. *Proc Natl Acad Sci U S A* **101**: 1455-1460
- Kohalmi SE, Reader LJV, Samach A, Nowak J, Haughn GW, Crosby WL** (1998) Identification and characterization of protein interactions using the yeast 2-hybrid system. *Plant Mol. Biol. Manual* **M1**: 1-30
- Koornneef M** (1981) The complex syndrome of *ttg* mutants. *Arabid. Inf. Serv.* **18**: 45-51
- Kopczynski JB, Raff AC, Bonner JJ** (1992) Translational read-through at nonsense mutations in the HSF1 gene of *Saccharomyces cerevisiae*. *Mol Gen Genet* **234**: 369-378
- Kunkel BN** (1996) A useful weed put to work: genetic analysis of disease resistance in *Arabidopsis thaliana*. *Trends Genet* **12**: 63-69
- Laemmli UK** (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**: 680-685
- Landry LG, Chapple CCS, Last RL** (1995) *Arabidopsis* mutants lacking phenolic sunscreens exhibit enhanced ultraviolet-B injury and oxidative damage. *Plant Physiol.* **109**: 1159-1166
- Larkin JC, Young N, Prigge M, Marks MD** (1996) The control of trichome spacing and number in *Arabidopsis*. *Development* **122**: 997-1005
- Li J, Ou-Lee T-M, Raba R, Amundson RG, Last RL** (1993) *Arabidopsis* flavonoid mutants are hypersensitive to UV-B irradiation. *Plant Cell* **5**: 171-179
- Liu L, Gitz DCI, McClure JW** (1995) Effects of UV-B on flavonoids, ferulic acid, growth and photosynthesis in barley primary leaves. *Physiol. Plant.* **93**: 725-733

- Lukacin R, Britsch L** (1997) Identification of strictly conserved histidine and arginine residues as part of the active site in *Petunia hybrida* flavanone 3beta-hydroxylase. *Eur J Biochem* **249**: 748-757
- Lukacin R, Groning I, Pieper U, Matern U** (2000) Site-directed mutagenesis of the active site serine290 in flavanone 3beta-hydroxylase from *Petunia hybrida*. *Eur J Biochem* **267**: 853-860
- Lukacin R, Groning I, Schiltz E, Britsch L, Matern U** (2000) Purification of recombinant flavanone 3beta-hydroxylase from *Petunia hybrida* and assignment of the primary site of proteolytic degradation. *Arch Biochem Biophys* **375**: 364-370
- Lukacin R, Urbanke C, Groning I, Matern U** (2000) The monomeric polypeptide comprises the functional flavanone 3beta-hydroxylase from *Petunia hybrida*. *FEBS Lett* **467**: 353-358
- Lukacin R, Wellmann F, Britsch L, Martens S, Matern U** (2003) Flavonol synthase from *Citrus unshiu* is a bifunctional dioxygenase. *Phytochemistry* **62**: 287-292
- Martens S, Forkmann G, Britsch L, Wellmann F, Matern U, Lukacin R** (2003) Divergent evolution of flavonoid 2-oxoglutarate-dependent dioxygenases in parsley. *FEBS Lett* **544**: 93-98
- Mazza CA, Boccalandro HE, Giordano CV, Battista D, Scopel AL, Ballaré CL** (2000) Functional significance and induction by solar radiation of ultraviolet-absorbing sunscreens in field-grown soybean crops. *Plant Physiol.* **122**: 117-125
- Mo Y, Nagel C, Taylor LP** (1992) Biochemical complementation of chalcone synthase mutants defines a role for flavonols in functional pollen. *Proc. Natl. Acad. Sci. USA* **89**: 7213-7217

- Moriguchi T, Kita M, Ogawa K, Tomono Y, Endo T, Omura M** (2002) Flavonol synthase gene expression during citrus fruit development. *Physiol Plant* **114**: 251-258
- Murphy A, Peer WA, Taiz L** (2000) Regulation of auxin transport by aminopeptidases and endogenous flavonoids. *Planta* **211**: 315-324
- Ormrod DP, Landry LG, Conklin PL** (1995) Short-term UV-B radiation and ozone exposure effects on aromatic secondary metabolite accumulation and shoot growth in flavonoid-deficient *Arabidopsis* mutants. *Physiol. Plant.* **93**: 602-610
- Owens DK, Hale T, Wilson LJ, McIntosh CA** (2002) Quantification of the production of dihydrokaempferol by flavanone 3-hydroxytransferase using capillary electrophoresis. *Phytochem Anal* **13**: 69-74
- Peer WA, Brown DE, Tague BW, Muday GK, Taiz L, Murphy AS** (2001) Flavonoid accumulation patterns of *transparent testa* mutants of *Arabidopsis thaliana*. *Plant Physiol.* **126**: 536-548
- Pelletier MK, Burbulis IE, Shirley BW** (1999) Disruption of specific flavonoid genes enhances the accumulation of flavonoid enzymes and end-products in *Arabidopsis* seedlings. *Plant Mol. Biol.* **40**: 45-54
- Pelletier MK, Murrell J, Shirley BW** (1997) *Arabidopsis* flavonol synthase and leucoanthocyanidin dioxygenase: further evidence for distinct regulation of "early" and "late" flavonoid biosynthetic genes. *Plant Physiol.* **113**: 1437-1445
- Pelletier MK, Shirley BW** (1996) Analysis of flavanone 3-hydroxylase in *Arabidopsis* seedlings: Coordinate regulation with chalcone synthase and chalcone isomerase. *Plant Physiol.* **111**: 339-345

- Peltola M, Chiatayat D, Peltonen L, Jalanko A** (1994) Characterization of a point mutation in aspartylglucosaminidase gene: evidence for a read-through of a translational stop codon. *Hum Mol Genet* **3**: 2237-2242
- Prescott AG, John P** (1996) Dioxygenases: Molecular structure and role in plant metabolism. *Ann. Rev. Plant Physiol. Plant Mol. Biol.* **47**: 245-271
- Prescott AG, Lloyd MD** (2000) The iron(II) and 2-oxoacid-dependent dioxygenases and their role in metabolism. *Nat Prod Rep* **17**: 367-383
- Prescott AG, Stamford NP, Wheeler G, Firmin JL** (2002) *In vitro* properties of a recombinant flavonol synthase from *Arabidopsis thaliana*. *Phytochemistry* **60**: 589-593
- Punyasiri PA, Abeysinghe IS, Kumar V, Treutter D, Duy D, Gosch C, Martens S, Forkmann G, Fischer TC** (2004) Flavonoid biosynthesis in the tea plant *Camellia sinensis*: properties of enzymes of the prominent epicatechin and catechin pathways. *Arch Biochem Biophys* **431**: 22-30
- Rashotte AM, Jenks MA, Nguyen TD, Feldmann KA** (1997) Epicuticular wax variation in ecotypes of *Arabidopsis thaliana*. *Phytochemistry* **45**: 251-255
- Rice-Evans C** (2004) Flavonoids and isoflavones: absorption, metabolism, and bioactivity. *Free Radic Biol Med* **36**: 827-828
- Roach PL, Clifton IJ, Fulop V, Harlos K, Barton GJ, Hajdu J, Andersson I, Schofield CJ, Baldwin JE** (1995) Crystal structure of isopenicillin N synthase is the first from a new structural family of enzymes. *Nature* **375**: 700-704
- Rosati C, Cadic A, Duron M, Amiot MJ, Tacchini M, Martens S, Forkmann G** (1998) Flavonoid metabolism in *Forsythia* flowers. *Plant Science* **139**: 133-140

- Rosso MG, Li Y, Strizhov N, Reiss B, Dekker K, Weisshaar B** (2003) An *Arabidopsis thaliana* T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics. *Plant Mol Biol* **53**: 247-259
- Ryle MJ, Hausinger RP** (2002) Non-heme iron oxygenases. *Curr Opin Chem Biol* **6**: 193-201
- Saslowsky D, Winkel-Shirley B** (2001) Localization of flavonoid enzymes in *Arabidopsis* roots. *Plant J.* **27**: 37-48
- Saslowsky DE, Warek U, Winkel BS** (2005) Nuclear localization of flavonoid enzymes in *Arabidopsis*. *J Biol Chem* **280**: 23735-23740
- Schofield CJ, Zhang Z** (1999) Structural and mechanistic studies on 2-oxoglutarate-dependent oxygenases and related enzymes. *Curr Opin Struct Biol* **9**: 722-731
- Schulze-Lefert P, Dangl JL, Becker-André M, Hahlbrock K, Schulz W** (1989) Inducible *in vivo* DNA footprints define sequences necessary for UV light activation of the parsley chalcone synthase gene. *EMBO J.* **8**: 651-656
- Shirley BW** (1996) Flavonoid biosynthesis: "new" functions for an "old" pathway. *Trends Plant Sci.* **1**: 377-382
- Shirley BW, Kubasek WL, Storz G, Bruggemann E, Koornneef M, Ausubel FM, Goodman HM** (1995) Analysis of *Arabidopsis* mutants deficient in flavonoid biosynthesis. *Plant J.* **8**: 659-671
- Simmonds MS** (2003) Flavonoid-insect interactions: recent advances in our knowledge. *Phytochemistry* **64**: 21-30
- Smith M, Jessee J, Landers T, Jordan J** (1990) High efficiency bacterial electroporation: 1×10^{10} *E. coli* transformants/ μg . *Focus* **12**: 38-40

- Spribille R, Forkmann G** (1984) Conversion of dihydroflavonols to flavonols with enzyme extracts from flower buds of *Matthiola incana* R. Br. Z. Naturforsch. **39c**: 714-719
- Srere PA** (1987) Complexes of sequential metabolic enzymes. Ann. Rev. Biochem. **56**: 89-124
- Srere PA** (2000) Macromolecular interactions: tracing the roots. Trends Biochem Sci **25**: 150-153
- Stafford HA** (1974) Possible multi-enzyme complexes regulating the formation of C₆-C₃ phenolic compounds and lignins in higher plants. Rec. Adv. Phytochem. **8**: 53-79
- Stafford HA** (1990) Flavonoid Metabolism. CRC Press, Boca Raton
- Stafford HA** (1991) Flavonoid evolution: an enzymatic approach. Plant Physiol. **96**: 680-685
- Stich K, T. Eidenberger, F. Wurst, and G. Forkmann** (1992) Flavonol synthase activity and the regulation of flavonol and anthocyanin biosynthesis during flower development in *Dianthus caryophyllus* L. (Carnation). Z. Naturforsch **47c**: 553-560
- Swain SM, Singh DP** (2005) Tall tales from sly dwarves: novel functions of gibberellins in plant development. Trends Plant Sci **10**: 123-129
- Turnbull JJ, Nagle MJ, Seibel JF, Welford RW, Grant GH, Schofield CJ** (2003) The C-4 stereochemistry of leucocyanidin substrates for anthocyanidin synthase affects product selectivity. Bioorg Med Chem Lett **13**: 3853-3857
- Turnbull JJ, Nakajima J, Welford RW, Yamazaki M, Saito K, Schofield CJ** (2004) Mechanistic studies on three 2-oxoglutarate-dependent oxygenases of flavonoid biosynthesis: anthocyanidin synthase, flavonol synthase, and flavanone 3beta-hydroxylase. J Biol Chem **279**: 1206-1216

- van Eldik GJ, Reijnen WH, Ruiter RK, van Herpen MM, Schrauwen JA, Wullems GJ** (1997) Regulation of flavonol biosynthesis during anther and pistil development, and during pollen tube growth in *Solanum tuberosum*. *Plant J* **11**: 105-113
- Wellmann F, Lukacin R, Moriguchi T, Britsch L, Schiltz E, Matern U** (2002) Functional expression and mutational analysis of flavonol synthase from *Citrus unshiu*. *Eur J Biochem* **269**: 4134-4142
- Wellmann F, Matern U, Lukacin R** (2004) Significance of C-terminal sequence elements for *Petunia* flavanone 3beta-hydroxylase activity. *FEBS Lett* **561**: 149-154
- Wilmouth RC, Turnbull JJ, Welford RW, Clifton IJ, Prescott AG, Schofield CJ** (2002) Structure and mechanism of anthocyanidin synthase from *Arabidopsis thaliana*. *Structure (Camb)* **10**: 93-103
- Wingender R, Röhrig H, Hörnicke C, Schell J** (1990) *cis*-regulatory elements involved in ultraviolet light regulation and plant defense. *Plant Cell* **2**: 1019-1026
- Winkel BS** (2004) Metabolic channeling in plants. *Annu Rev Plant Biol* **55**: 85-107
- Winkel-Shirley B** (2001) Flavonoid biosynthesis: a colorful model for genetics, biochemistry, cell biology and biotechnology. *Plant Physiol.* **126**: 485-493
- Winkel-Shirley B** (2001) It takes a garden. How work on diverse plant species has contributed to an understanding of flavonoid metabolism. *Plant Physiol* **127**: 1399-1404
- Winkel-Shirley B** (2002) Biosynthesis of flavonoids and effects of stress. *Curr Opin Plant Biol* **5**: 218-223
- Wisman E, Hartmann U, Sagasser M, Baumann E, Palme K, Hahlbrock K, Saedler H, Weisshaar B** (1998) Knock-out mutants from an En-1 mutagenized *Arabidopsis thaliana*

population generate phenylpropanoid biosynthesis phenotypes. Proc. Natl. Acad. Sci. USA **95**: 12432-12437

Yao LH, Jiang YM, Shi J, Tomas-Barberan FA, Datta N, Singanusong R, Chen SS (2004)

Flavonoids in food and their health benefits. Plant Foods Hum Nutr **59**: 113-122

Zalokar M (1960) Cytochemistry of centrifuged hyphae of *Neurospora*. Exp. Cell Res. **19**: 114-132

Zerback R, Dressler K, Hess D (1989) Flavonoid compounds from pollen and stigma of

Petunia hybrida: Inducers of the *vir* region of the *Agrobacterium tumefaciens* Ti plasmid.

Plant Sci. **62**: 83-91

APPENDIX
ENZYME ASSAYS

Assays were typically performed with recombinant protein generated in the Novagen pET system from a 50 ml culture by standard induction procedures. Final cell pellets were harvested 4 hrs. post-induction by centrifugation at 7400 g, 4°C, for 10 min. and immediately stored at -80°C. Cell pellets were resuspended in 3 ml of 0.2 M cold assay buffer of the appropriate composition and pH for the enzyme being tested. The resulting cell slurry was sonicated by giving a 30 sec. pulse followed by a 30 sec. recovery period and repeating this six times. DNase was added at a concentration of 80 µg/ml and the solution incubated at 4°C for 30 min. only if the resulting solution was viscous after sonication. The lysate was centrifuged at 16,170 x g, 4°C, for 40 min and the supernatant used as the source of crude enzyme in activity assays.

Assays were performed in 100 µl reaction volume containing 10 mM α-ketoglutaric acid (disodium salt), 10 mM ascorbic acid, 0.25 mM ferrous sulfate, and 50 mM reaction buffer. In F3H reactions, 1 mM DHQ was included as an internal standard. Typical stock solutions and reaction mix composition are described in Table A.1. Reaction buffers were degassed under vacuum using a vacuum pump for 10 min, equilibrated under N₂ for 5 min, and again degassed under vacuum for 10 min. Ascorbic acid and α-ketoglutarate were prepared by dissolving in reaction buffer. Ferrous sulphate was prepared in ascorbic acid solution to reduce oxidation of the iron. All flavonoid substrates and standards were prepared in 80% methanol as they were insoluble in the employed reaction buffers. Reactions were typically started by the addition of substrate, incubated in a heat block or at room temperature, and stopped by extraction with ethyl acetate. Extractions were performed by adding 200 µl of ethyl acetate and mixing well for 1 min. Solvent layers were separated by centrifugation at 13,000 rpm in a microcentrifuge for 5

min. A 100 μ l aliquot of the organic layer was transferred to a sterile microcentrifuge tube. The reaction mix and remaining ethyl acetate was re-extracted with 200 μ l ethyl acetate, mixed and centrifuged as before, and 200 μ l of the organic layer combined with the initial 100 μ l extract. The solvent was evaporated in a SpeedVac under low heat. Dried samples were reconstituted in 50 μ l of 80% methanol, mixed for 5 min, and centrifuged at 13,000 RPM in a microcentrifuge for 5 min. Reconstituted samples were transferred to Waters[®] 8 x 40 mm total recovery autosampler vials for HPLC analysis.

Samples were kept at 4°C in the autosampler prior to injection. A 20 μ l aliquot of each 50 μ l reconstituted sample was injected and fractionated at room temperature using a Nova-Pak[®] C18 3.9 X 150 mm column at a flow rate of 2.5 ml/min. Fractionation was achieved using a linear gradient of water (adjusted to pH 3 with glacial acetic acid) and HPLC-grade acetonitrile. Acetonitrile was adjusted to 10% in the first 2 min., 50% in the following 25 min., and finally to 100% in the next 2 min. The column was washed with 100% acetonitrile for 2 min. before the next injection. Absorbance of the eluent was monitored from 200 to 400 nm. The resulting data were analyzed by extracting single wavelength chromatograms at 254 nm (flavonols) and 289 nm (flavanones and dihydroflavonols), and integrating the resulting peak areas using the Millennium 3.2 software package (Waters Corporation). Concentrations of flavonoid compounds were then determined by generating a standard curve from commercially available standards (Table A.2) using the same HPLC conditions.

| Stock Solution | Volume (μL) |
|---|--|
| 100 mM α -Ketoglutaric Acid | 10 |
| 100 mM Ascorbic Acid/ 2.5 mM Ferrous Sulphate | 20 |
| 1 mM Substrate (Nae, DHK, DHQ) | 10 |
| 1 mM Internal Standard (optional) | 10 |
| Enzyme Crude Exact | to required μg |
| 50 mM Reaction Buffer | to total |
| Total | 100 |

Table A.1. Enzyme assay reaction components.

| Flavonoid Compound | Commercial Source | Catalog Number |
|------------------------------|------------------------------|-----------------------|
| Naringenin | Sigma-Aldrich | N-5893 |
| Dihydrokaempferol | Dihydrochemical Technologies | N.A. |
| Dihydroquercetin (Taxifolin) | Sigma-Aldrich | T-4512 |
| Kaempferol | Sigma-Aldrich | K-0133 |
| Quercetin | Sigma-Aldrich | Q-0125 |

Table A.2. Commercially available flavonoid standards sources.

Curriculum Vitae

September 2005

Daniel K. Owens

Home Address

499 Newcomb St.
Apt. #3
Christiansburg, VA 24073
phone: (540) 381-5834
e-mail: dowens@vt.edu

Professional Address

Virginia Tech
107 Fralin Biotechnology Center
Blacksburg, VA 24061-0346
phone: (540) 231-9375
fax: (540) 231-7126

PERSONAL DATA

Born: April 18, 1977; Greeneville, TN
Spouse: Melissa Ann Owens
Child: Kayla Ann Owens (b. 12/17/98)

EDUCATION

B.S. East Tennessee State University -1999 (Biochemistry Concentration),
Magna Cum Laude

Ph.D. of Biology Virginia Tech – 2005 (Molecular and Cellular Biology and
Biotechnology Option)

EMPLOYMENT

August 1998 to July 2000, Laboratory Technician, East Tennessee State University
Department of Biological Sciences (with C.A. McIntosh).

July 1999 to May 2000, Laboratory Technician, East Tennessee State University
Department of Chemistry (with L.J. Wilson).

August 2000 to May 2001. Graduate Teaching Assistant, Virginia Tech Department of
Biology.

May 2001 to August 2001. Graduate Research Assistant, Virginia Tech Department of
Biology.

August 2001 to May 2002. Graduate Teaching Assistant, Virginia Tech Department of
Biology.

May 2002 to December 2002. Graduate Research Assistant, Virginia Tech Department
of Biology.

December 2002 to May 2003. Graduate Teaching Assistant, Virginia Tech Department of Biology.

May 2003 to December 2003. Graduate Research Assistant, Virginia Tech Department of Biology.

December 2003 to May 2004. Graduate Teaching Assistant, Virginia Tech Department of Biology.

May 2004 to August 2004. . Graduate Research Assistant, Virginia Tech Department of Biology.

August 2004 to December 2004. Instructor, Virginia Tech Department of Biology.

December 2004 to October 2005. Graduate Research Assistant, Virginia Tech Department of Biology.

PROFESSIONAL SOCIETIES

Sigma Xi

American Society of Plant Biologists

Phytochemical Society of North America

Beta Beta Beta Biological Honor Society (Regional Vice President, Secretary Pi Delta Chapter 1998-99)

National Eagle Scout Association

Gamma Beta Phi Honor Society

Phi Kappa Phi Honor Society

Alpha Lambda Delta Honor Society

FELLOWSHIPS AND SCHOLARSHIPS

August 1995 to May 1999, East Tennessee State University Advanced Placement Scholarship

July 2000, Michigan State University Plant Sciences Fellowship (declined acceptance)

HONORS AND AWARDS

1995 Boy Scouts of America Eagle Scout Rank

1998 ETSU Biological Sciences Dr. Herman O'Dell Award for the Outstanding Junior in Biological Sciences

2005 VT Biology Department Outstanding Teaching Award

RESEARCH GRANTS

- 1998 Faculty-Student Collaborative Research Grant “Distribution of Flavanone 3 Hydroxylase During Growth and Development of Grapefruit and Petunia Seedlings” \$800
- 2001 Sigma Xi Grant in Aid of Research “Expression and Characterization of Flavonol Synthase Isozymes” \$450
- 2001 Graduate Research Development Project "Expression and Characterization of Flavonol Synthase Isozymes in *Arabidopsis thaliana*" \$300
- 2001 Virginia Tech GSA Travel Grant Program \$100
- 2002 Graduate Research Development Project “Characterization of Flavonol Synthase Isozymes in *Arabidopsis thaliana*” \$300
- 2002 Virginia Tech GSA Travel Grant Program \$300
- 2003 Graduate Research Development Project “Examination of 2-Oxoglutarate Dependent Dioxygenases in the Central Flavonoid Biosynthetic Pathway of *Arabidopsis thaliana*” \$300
- 2004 Graduate Research Development Project “Examination of the Flavonol Synthase Isozymes in *Arabidopsis thaliana*” \$150

SEMINARS and TALKS

- 2002 Examination of 2-Oxoglutarate Dependant Dioxygenases in the Central Flavonoid Biosynthetic Pathway of *Arabidopsis thaliana*. Plant Molecular Biology Discussion Group, Blacksburg, Va.
- 2002 Examination of 2-Oxoglutarate Dependant Dioxygenases in the Central Flavonoid Biosynthetic Pathway of *Arabidopsis thaliana*. Virginia Tech Biology Department Botany Seminar Series, Blacksburg, Va.
- 2005 Characterization of 2-Oxoglutarate-Dependant Dioxygenases Leading to the Production of Flavonols in *Arabidopsis thaliana*. 2nd Annual Virginia Tech Biology Department Research Forum, Blacksburg, VA.
- 2005 Characterization of 2-Oxoglutarate-Dependant Dioxygenases Leading to the Production of Flavonols in *Arabidopsis thaliana*. Department of Biological and Health Sciences Seminar Series, Johnson City, TN.

PUBLICATIONS

Articles

Owens, D.K., T. Hale, L.J. Wilson, and C.A. McIntosh. 2002. A Sensitive Quantitative Assay of Flavanone 3 Hydroxylase by the Use of Capillary Electrophoresis. *Phytochem. Anal.* **13**: 69-74.

Owens, D.K. and B.S.J. Winkel. 2004. Kinetic Analysis of Recombinant Flavanone 3B-Hydroxylase from *Arabidopsis thaliana*. *In revision*.

Owens, D.K. and B.S.J. Winkel. 2005. Characterization of the Flavonol Synthase Isoforms of *Arabidopsis thaliana*. *In preparation*.

Abstracts Presented and Published (presenting author underlined)

Owens, D.K., T.L. Hale, L.J. Wilson, and C.A. McIntosh. 1999. Development of a Sensitive Quantitative Assay for Flavanone 3 Hydroxylase. East Tennessee State University Quillen College of Medicine 15th Annual Research Forum, Johnson City, TN.

Owens, D.K., T.L. Hale, L.J. Wilson, and C.A. McIntosh. 1999. Development of a Sensitive Quantitative Assay for Flavanone 3 Hydroxylase. Beta Beta Beta regional meeting in conjunction with The 60th annual conference of the Association of Southeastern Biologists, Wilmington, NC.

Owens, D.K., T.L. Hale, L.J. Wilson, and C.A. McIntosh. 2000. Quantification of Dihydrokaempferol Production by Flavanone 3-Hydroxytransferase Using Capillary Electrophoresis. East Tennessee State University Quillen College of Medicine 16th Annual Research Forum, Johnson City, TN.

Owens, D.K., T.L. Hale, L.J. Wilson, and C.A. McIntosh. 2000. Quantification of Dihydrokaempferol Production by Flavanone 3-Hydroxytransferase Using Capillary Electrophoresis. Annual Meeting of the Phytochemical Society of North America, Bethesda, MD.

Owens, D.K. and Brenda Winkel-Shirley. 2001. Characterization of Flavonol Synthase Isozymes in *Arabidopsis thaliana*. American Society of Plant Biologists: Plant Biology 2001, Providence, RI.

Owens, D.K. and Brenda Winkel-Shirley. 2002. Examination of 2-Oxoglutarate Dependant Dioxygenases in the Central Flavonoid Biosynthetic Pathway of *Arabidopsis thaliana*. Gordon Research Conference on Macromolecular Organization and Cell Function, Queens College, Oxford, UK.

Owens, D.K. and Brenda Winkel. 2003. Examination of 2-Oxoglutarate Dependant Dioxygenases in the Central Flavonoid Biosynthetic Pathway of *Arabidopsis thaliana*. 14th International Conference on Arabidopsis Research, Madison, WI.

Winkel, Brenda S.J., Christopher D. Dana, Daniel K. Owens, David R. Bevan, Joseph P. Noel, and Susan Krueger. 2003. Understanding Flavonoid Metabolism in Three Dimensions. 14th International Conference on Arabidopsis Research, Madison, WI.

Owens, D.K. and Brenda Winkel. 2003. Examination of 2-Oxoglutarate Dependant Dioxygenases in the Central Flavonoid Biosynthetic Pathway of *Arabidopsis thaliana*. Virginia Tech Graduate Student Assembly 19th Annual Research Symposium, Blacksburg, VA.

Owens, D.K. and Brenda Winkel. 2003. Examination of 2-Oxoglutarate Dependant Dioxygenases in the Central Flavonoid Biosynthetic Pathway of *Arabidopsis thaliana*. 1st Annual Virginia Tech Biology Department Research Forum, Blacksburg, VA.

Owens, D.K., Anne B. Walton, and Brenda Winkel. 2004. Examination of the Flavonol Synthase Isoforms in *Arabidopsis thaliana*. American Society of Plant Biologists: Plant Biology 2004, Orlando, FL.

Owens, D.K., Anne B. Walton, and Brenda Winkel. 2005. Examination of the Flavonol Synthase Isoforms in *Arabidopsis thaliana*. 2nd Annual Virginia Tech Biology Department Research Forum, Blacksburg, VA.

Alerding, Anne B., Daniel K. Owens, James H. Westwood, and Brenda S.J. Winkel. 2005. Flavonol Synthase in Arabidopsis: Gene-specific Responses to Developmental and Biotic Signals. 15th International Conference on Arabidopsis Research, Madison, WI.

TEACHING EXPERIENCE

East Tennessee State University Writing Center, Tutor

One on one tutoring on the development and construction of written assignments in various disciplines 1995

East Tennessee State University Student Support Services, Tutor

One on one tutoring in General Biology and Chemistry 1996

Virginia Polytechnic Institute Biology Department, TA,

Instructor for General Biology and Principles of Biology Laboratories. 2000-2002

Virginia Polytechnic Institute Biology Department, TA,
Preparation and teaching for Molecular Biology Laboratory 2003

Virginia Polytechnic Institute Biology Department, Instructor,
Organization and lecturing for Molecular Biology Laboratory,
supervision of 3 teaching assistants 2004

COMMUNITY SERVICE

East Tennessee State University

Adopt-a-School program; assisted a first grade teacher with her class two hours a week for the duration of a semester.

Virginia Polytechnic Institute and State University

BGSA outreach program; assisted a student teacher at Radford University by teaching a second grade class the importance of plants.

Fralin Biotechnology Center outreach programs; acted as a tour guide of the Fralin Center for elementary and middle school aged children multiple times; served on a panel reviewing presentation of data from research performed on *Arabidopsis thaliana* by high school students.

PROFESSIONAL SERVICE

Beta Beta Beta Biological Honor Society, Secretary Pi Delta Chapter 1998-99.
Beta Beta Beta Biological Honor Society, Regional Vice President 1998-99.