Molecular and Biochemical Genetics of 2-Oxoglutarate-Dependent Dioxygenases Required for Flavonoid Biosynthesis in *Arabidopsis thaliana*

Matthew K. Pelletier

**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
Biology**

**Brenda W. Shirley, Chair
Carole L. Cramer
Asim Esen
Joseph O. Falkinham
Charles L. Rutherford**

April 26, 1997

Blacksburg, Virginia

Keywords: Metabolic Regulation, Flavanone 3-hydroxylase, Flavonol Synthase, Leucoanthocyanidin Dioxygenase, Antibodies

Molecular and Biochemical Genetics of 2-Oxoglutarate-Dependent Dioxygenases Required for Flavonoid Biosynthesis in *Arabidopsis thaliana*

Matthew K. Pelletier

(ABSTRACT)

Three 2-oxoglutarate-dependent dioxygenases required for flavonoid biosynthesis were characterized in *Arabidopsis thaliana*. Genes encoding flavanone 3-hydroxylase (*F3H*), flavonol synthase (*FLS*), and leucoanthocyanidin dioxygenase (LDOX) were cloned and sequenced. The predicted proteins encoded by each of these Arabidopsis genes shared high homology with all F3H, FLS, or LDOX sequences available in Genbank. Low-stringency DNA blot analysis indicated that F3H and LDOX are encoded by a single gene in Arabidopsis, while FLS may be encoded by two or three genes.

RNA blot analysis was performed to determine the expression patterns of these three genes relative to previously-cloned genes encoding flavonoid biosynthetic enzymes. Light-induction experiments and analysis of regulatory mutants showed that the *CHS*, *CHI*, *F3H*, and *FLS1* are coordinately regulated in Arabidopsis seedlings, encode enzymes acting near the beginning of the pathway, and are therefore referred to as "early" genes. The same experiments showed that *DFR* and *LDOX* are regulated distinctly from "early" genes, share similar expression patterns in response to light, and are not expressed in the *ttg* mutant. *DFR* and *LDOX* are therefore referred to as "late" genes due to the timing of expression in response to light and the fact that they encode enzymes acting late in flavonoid biosynthesis.

To determine whether any of the previously-identified *transparent testa* mutants were defective in *F3H*, *FLS*, or *LDOX*, the chromosomal locations of these genes in the Arabidopsis genome were determined. The positions of these genes suggested that no previously-identified *tt* mutant was defective in the cloned *FLS* or *LDOX* structural genes, while *tt6* was potentially the *F3H* locus. The coding region of F3H was amplified by PCR from *tt6* genomic DNA and sequenced, and several point mutations were found in the coding region of this allele, three of which are predicted to result in amino acid substitutions.

Polyclonal antibodies were also developed using four different purified, recombinant flavonoid enzymes as antigens. These antibodies were used to determine the pattern of accumulation of flavonoid enzymes in developing seedlings. Immunoblot analysis was also performed to determine whether mutations in genes encoding specific flavonoid enzymes or an enzyme in pathways that compete for or provide substrate for flavonoid biosynthesis (mutants defective in tryptophan or ferulic acid biosynthesis) affect the levels of flavonoid enzymes. These analyses showed that mutant seedlings which lacked specific flavonoid or tryptophan biosynthetic enzymes accumulated higher steady-state levels of other enzymes in the pathway. These results suggest that the accumulation of specific flavonoid intermediates or indole can lead directly or indirectly to higher levels of flavonoid enzymes.

This work was supported by National Science Foundation grant no. DMB-9304767

Dedication

I would like to dedicate this dissertation to my wife Pam and my two sons, Ryan and Isaac. Without their love and patience, I could have never finished this work. When I was frustrated due to the many failures inevitable in science, coming home helped keep things in perspective. I'm especially thankful to Pam for listening to me complain about failed experiments, and for sacrificing to stay home with the kids so that I could finish this degree. I have truly been blessed with a wonderful wife.

Acknowledgements

I would like to thank Dr. Carole Cramer, Dr. Asim Esen, Dr. Joe Falkinham, and Dr. Charles Rutherford for serving on my advisory committee. Your support in my research endeavors and your time and efforts are greatly appreciated. I would especially like to thank Dr. Brenda Shirley for giving me the opportunity to work under her guidance these last 4 years. Thanks for the encouragement and listening ear you always provided, and for keeping me focused as I pursued my degree. I consider it a honor to have been a part of your lab. You have challenged me to push myself to the limits.

I would also like to thank the newest members of the Shirley lab, Dave Saslowsky and Mike Santos, for friendship and support. I am also thankful to Cody Cain for developing the polyclonal anti-CHI antibody used in many experiments, and for putting up with Ian and me. Special thanks go to two former undergraduates who worked with me, Jonathon Shelton and Julie Murrell for working so hard in their time here. Jonathon Shelton was responsible for the initial cloning of F3H cDNA into pBluescript, and Julie Murrell spent several, tireless months helping isolate a genomic clone encoding FLS.

Last, but not least, I would like to thank Ian Burbulis for his friendship and his ability to "fix" things that are not broken (and things that are broken), and for all the "discussions" we had ("how can you say that?!"). In addition, we have had countless, productive discussions sharing our ideas on what questions were really important for us to answer and what methods we could use to address those questions. Your comments constantly served as a challenge for me to think about what my experiments meant. Keep up the good work- it has been a great 4 years.

I would also like to acknowledge my parents, Rodney and Sandra Pelletier, for their love and support over the years. Thanks for pushing me to go to college in the first place, for financial support, and just for being great parents.

Table of Contents

Chapter 1

Literature Review

Importance of Understanding Metabolic Regulation

All living organisms from bacteria to man are able of carrying out a remarkable number of biochemical reactions that enable them to grow and survive. Many of the molecules synthesized by cells such as proteins, nucleic acids, and lipids are necessary for the cell to survive (primary metabolites) because they carry out crucial enzymatic reactions needed to provide energy, are structural components of cells, serve as signalling molecules that act in developmental programs, or carry genetic information. Other metabolites provide a selective advanatage to organisms, yet are not required for survival under ideal conditions (secondary metabolites).

Since most cells have the ability to synthesize literally thousands of different compounds, some of which would be useless or wasteful to produce in certain environmental conditions, it is crucial that the both the level and types of metabolites synthesized be tightly controlled. Likewise, the sophisticated control of the cell cycle and gene expression, both of which may be reduced to changes in metabolism, are required for the proper development and/or reproduction of organisms. Furthermore, since different compounds provide different functions in various locations within the cell, it is important that within each cell there is spatial control over metabolic pathways. Were it not for the ability of organisms to control their own metabolism in this manner, life as we know it would not exist.

In recent years, biochemical, molecular, genetic, and ultrastructural studies of cells have provided many significant insights into the mechanisms by which organisms mediate this intricate control over their own metabolism. Often, biosynthetic pathways leading to diverse products may share a common substrate(s), and *in vitro* studies using purified enzymes have shown that the activity of the first enzyme in each respective pathway may be controlled by the amount of end products, thus maintaining a balance between the diverse products. Undoubtedly, the spatial organization of eukaryotic cells plays a central role in controlling metabolism as enzymes and/or substrates available for certain compounds may be localized to a particular organelle within the cell. Interactions among enzymes within a pathway with many branch points may play a role in controlling the types and amounts of end-products produced. Yet another mechanism cells use to control metabolism is by changing the rate of transcription of genes encoding proteins that may be useful only under appropriate environmental conditions or at appropriate times in development. Some well-known examples of this type of regulation include the induction of genes encoding lactose-metabolizing enzymes in *E. coli* and the induction of pathogenesis-related genes in plants upon infection.

While research in the last few decades has greatly enhanced our knowledge of cell structure and function, our understanding is still quite limited. This is partly due to the enormous complexity of cells, and partly due to the fact that many biochemical studies have been performed *in vitro*. Biochemical studies are crucial because they provide the basis for understanding how enzymes function and their activities are modulated, and thus supply knowledge about one level of metabolic regulation. In addition, the purification of enzymes and partial peptide sequencing provide valuable information useful for cloning the genes that encode them, thus allowing subsequent molecular analysis. However, it is not always possible to apply principles learned from *in vitro* experiments to the physiological state *in vivo*, since the two environments are radically different. In contrast, genetic approaches in combination with molecular techniques allow one to analyze how the loss of one enzyme or group of compounds may affect the pools of other metabolites or the physiology of the organism *in vivo*. Thus, a combination of biochemical, genetic, and molecular tools can be used to provide important information on how organisms alter their metabolism in response to environmental or developmental cues, how changes in the flux through one pathway might affect other pathways, and about the physiological consequences of these changes. This information may be useful for understanding the function of various

metabolites *in vivo* and may also prove invaluable for successfully engineering the metabolism of plants and animals for the production of proteins or compounds useful to humans.

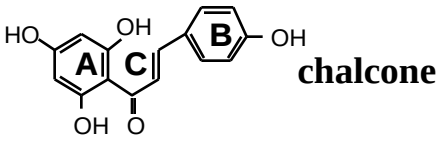
Biochemical Characterization and Functional Importance of Flavonoids

As a model for studying metabolic regulation, we have chosen the flavonoid biosynthetic pathway of *Arabidopsis thaliana*. Flavonoids are a diverse group of plant secondary metabolites that have a variety of functions in plants. One class of flavonoids, the anthocyanins, which are the primary pigments determining flower color, were used for simple, Mendelian genetic studies as early as 1900 (Stafford, 1990). The basic unit of this diverse group of compounds consists of two benzene rings (A and B) that are connected by a three carbon chain, which is closed in all flavanoids except chalcones and dihydrochalcones (Figure 1) (Stafford, 1990). The first step in flavonoid biosynthesis is the condensation of one molecule of 4-coumaroyl-CoA, which is derived from the general phenylpropanoid pathway with three molecules of malonyl-CoA, which is synthesized from the Krebs-cycle intermediate, acetyl-CoA, and CO_2. The classification of flavonoids is based primarily on the oxidation state of the C ring (Stafford, 1990). The structures and names of the various classes of flavonoids are shown in Figure 2.

While all higher plants synthesize flavonoids, the array of compounds produced differs among plant species, and the physiological roles they play may vary. For instance, the anthocyanins function to attract pollinators to flowers and seed dispersers to ripened fruit, and may also play a role in protecting plant DNA from damage by ultraviolet light (reviewed by Shirley, 1996; Stapleton and Walbot, 1994). Several studies have suggested multiple roles for at least one class of flavonoids- the flavonols. For instance, an Arabidopsis mutant that is extremely sensitive to UV-B (exhibits massive necrosis relative to wild-type) seems unaffected in the accumulation of anthocyanins and seed coat flavanoids (tannins), but lacks a glycosylated form of kaempferol (Lois and Buchanan, 1994). Similarly, kaempferol plays a role in pollen tube development in maize and petunia, but not in Arabidopsis, underscoring the fact that flavonoid functions are not necessarily conserved among plant species (Mo et al., 1992; van der Meer et al., 1992; Burbulis et al., 1996). A variety of flavonoids, including isoflavonoids, flavanones, flavonols, and 3-hydroxyflavanones are thought to act as anti-fungal agents and may also act as signalling molecules with symbiotic bacteria and fungi (reviewed by Koes et al., 1993; reviewed by Shirley, 1996).

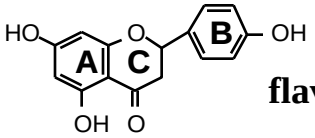
Many studies have provided insights into the site of synthesis and storage of flavonoids. Subcellular fractionation studies and immunolocalization studies have indicated that although many of the enzymes involved in flavonoid biosynthesis are “soluble”, these compounds may be synthesized by an enzyme complex loosely associated with the endoplasmic reticulum (ER) (reviwed by Wagner and Hrazdina, 1984; Hrazdina and Wagner, 1985b; Hrazdina and Wagner, 1985a; Hrazdina et al., 1986; Hrazdina et al., 1987; Hrazdina and Jensen, 1992; Burbulis *et al.*, 1996); IE Burbulis and BW Shirley, unpublished data). It is worth mentioning that two flavonoid enzymes, flavonoid 3’ hydroxylase (F3’H) and flavonoid 3’5’ hydroxylase (F3’5’H) are associated with microsomal fractions, and belong to the P450 superfamily of enzymes (Forkmann, 1980; Forkmann and Stotz, 1981; Holton et al, 1993). These enzymes may provide an “anchor” for a loosely associated enzyme complex. While flavonoids are thought to be synthesized at the ER, they are primarily stored in the plant vacuole (Hrazdina and Wagner, 1985a). A recent study of the maize mutant, *bronze2*, showed that the conjugation of glutathione to anthocyanins is required for vacuolar targeting, and a glutathione pump that transfers glutathione conjugates may be responsible for this transfer (Marrs et al., 1995). In other plant species, a mechanism involving sinapoylation of flavonoids may mediate vacuolar targeting (C. Chapple, personal communication)

Several flavonoid enzymes have been purified from one or more plant species. Among these are chalcone synthase (CHS), chalcone isomerase (CHI), flavanone 3-hydroxylase (F3H), and dihydroflavonol reductase (DFR) (reviewed by Stafford, 1990)) (Moustafa and Wong, 1967; Hahlbrock et al., 1970; Britsch and Grisebach, 1986; Fischer et al., 1988). Other enzymes such



chalcone

naringenin chalcone



flavanone

naringenin

Figure 1. Basic structure of flavonoids. The first two intermediates of flavonoid biosynthesis are shown (chalcones and flavanones). The nomenclature of flavonoids is based on the oxidation state of the C ring.

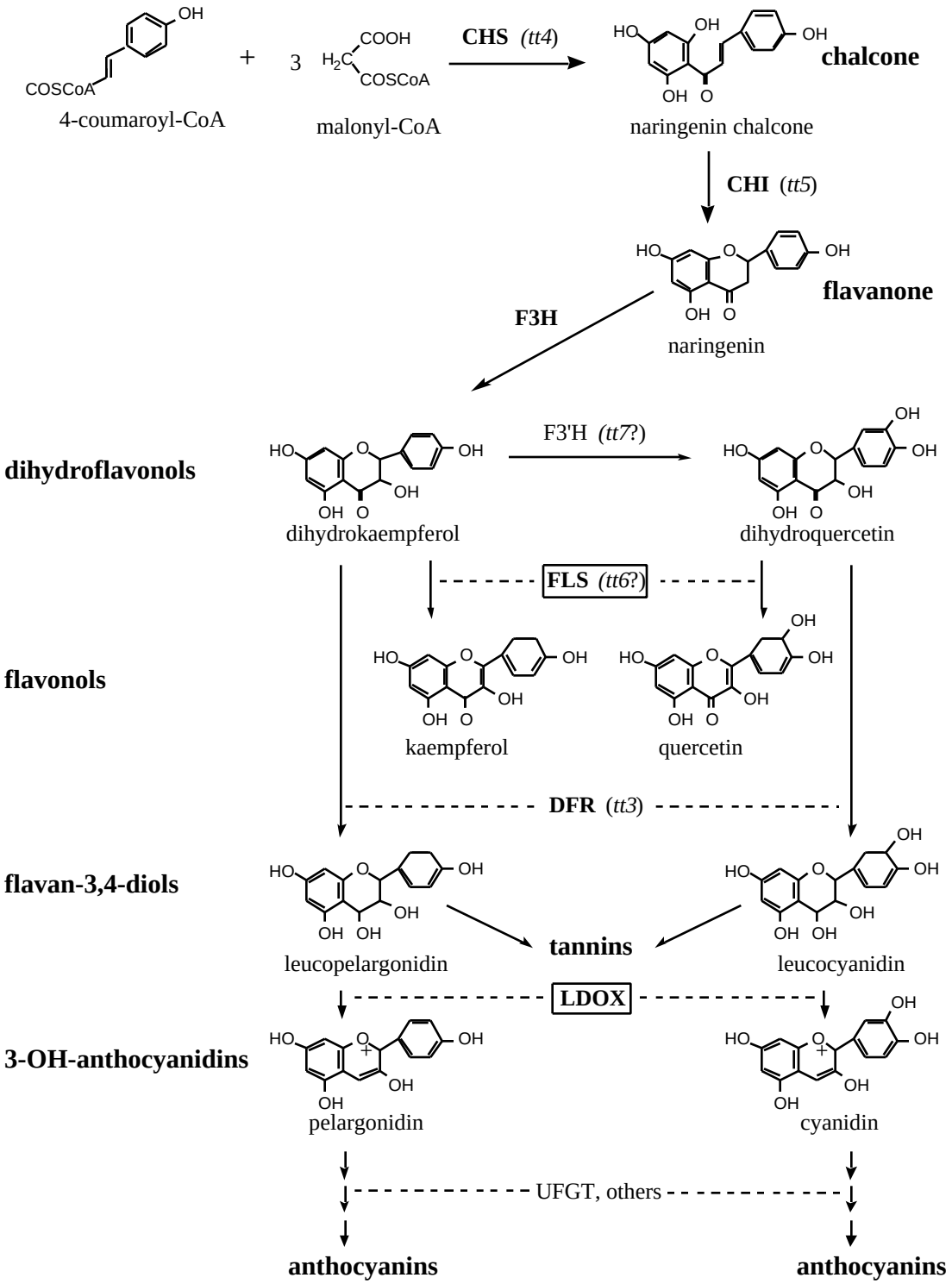


Figure 2. Schematic of the Arabidopsis flavonoid biosynthetic pathway. Enzymes for which genes have been cloned in Arabidopsis are shown in bold while those described in this study are also boxed. The names of specific compounds are given below the chemical structures while flavonoid classes are indicated in bold face.

as F3'H and flavonol synthase (FLS), although not yet purified, have been characterized according to cofactor requirements and whether the enzyme activity was found in soluble or microsomal cell fractions (Forkmann, 1980; Britsch et al., 1981; Forkmann and Stotz, 1981; Spribille and Forkmann, 1984; Forkmann et al., 1986). Of all the enzymes shown in Figure 2, all enzymes or activities except that of F3'H were classified as "soluble" enzymes. However, this is more a reflection of *in vitro* studies using enzyme preparations from homogenized tissues rather than a reflection of their *in vivo* localization. As mentioned above, several of the enzymes are believed to be loosely associated with the ER, based on cell fractionation and immunocytochemical studies.

Some flavonoid biosynthetic enzymes have varying degrees of sequence homology to enzymes from other pathways, while they do not belong to distinct enzyme families. For instance, CHS is related to the enzyme stilbene synthase (STS) which uses the same substrates as CHS to produce a stilbene rather than a chalcone (Tropf et al., 1995). Other CHS-like sequences were identified from *Gerbera hybrida*, and the substrates used by the corresponding enzymes shown to be different than chalcone synthase and stilbene synthase. These enzymes were able to use benzoyl-CoA as a substrate, but the product was not identified (Helariutta et al., 1996). In contrast, the primary amino acid sequences of CHI proteins contain no significant homology to other proteins in the Genbank database (Pelletier, unpublished results). The DFR proteins have sequence homology with 3β-hydroxysteroid dehydrogenase, an enzyme that converts pregneolone to the steroid progesterone, and it was hypothesized that these two proteins share a common ancestor (Baker et al., 1990). This is particularly intriguing since flavonoids can act as signalling molecules to affect gene expression in microorganisms and can mimic the biological effects of estrogen (reviewed by Koes *et al.*, 1993; Miksicek, 1993).

One other family of enzymes that deserves special attention due to the number of flavonoid enzymes in this group and their relevance to the current study are a subfamily of dioxygenases known as 2-oxoglutarate-dependent dioxygenases (2ODDS). This family of proteins catalyzes a variety of reactions in the primary and secondary metabolism of both plants and animals (reviewed by Prescott, 1993; reviewed by Prescott and John, 1996). These enzymes require Fe^{2+} and a reducing agent (usually ascorbate) for full activity *in vitro* and require O_2 and 2-oxoglutarate as co-substrates. Hydroxylation, desaturation, and epoxidation reactions may be catalyzed by these enzymes. 2ODDS play a fundamental role in the biosynthesis of two classes of plant hormones. A number of these enzymes are involved in the biosynthesis of gibberelins while 1-aminocyclopropane-1-carboxylate oxidase catalyzes the terminal step of ethylene biosynthesis. These enzymes also play a role in alkaloid biosynthesis. The primary amino acid identity shared by 2ODDS of different function within an individual organisms is generally on the order of 27-32%.

At least four enzymes involved in flavonoid biosynthesis are believed to be 2ODDS, including F3H, FLS, LDOX, and flavone synthase (FS). Among these four, F3H, FLS, and FS have been biochemically characterized and unequivocally identified as 2ODDS based on cofactor requirements (Forkmann, 1980; Britsch *et al.*, 1981; Forkmann and Stotz, 1981; Spribille and Forkmann, 1984; Forkmann *et al.*, 1986). LDOX is thought to belong to this family of enzymes based on sequence homology. The name LDOX is at this point in time is arbitrary. However, genetic loci associated with the conversion of flavan-3,4-diols to 3-OH-anthocyanidins were identified in maize (A2) and snapdragon (CANDI), and subsequent cloning of the genes revealed a high degree of homology to 2ODDS (Reddy and Coe, 1962; Martin et al., 1991). While the terms anthocyanidin synthase and anthocyanidin hydroxylase have been used in reference to CANDI and A2 homologues from other plant species, the name LDOX seems more appropriate and will be used throughout this manuscript.

Cloning of Genes Encoding Flavonoid Enzymes

Genes encoding several flavonoid enzymes, including CHS, CHI, F3H, FLS, LDOX, and UDPglucose 3-*O*-glucosyltransferase (UFGT) have been isolated from numerous plant species. While the methodology involved in the cloning of all of these genes will not be detailed here, the isolation of the first gene encoding each enzyme and confirmation of the enzymatic activity will be addressed. The number of genes thought to encode each enzyme in different plant species is also included.

The first CHS gene isolated was from parsley (Kreuzaler et al., 1983). A differential screening approach using a cDNA library made of poly A+ mRNA from UV-irradiated suspension cultures cells was utilized. The probes used in the hybridization were labeled RNA from light-induced or dark-grown cells, since it was known that CHS activity was greatly enhanced in response to light. The clones that specifically hybridized with RNA probes from light-grown cells were then screened with an antiserum that had been raised against a purified CHS enyzme (Kreuzaler et al., 1979). Later, in maize, transposon-tagging was used to isolate the C2 locus, which was known to encode CHS (Wienand et al., 1986). Genes encoding CHS have since been isolated from many plant species, as text-based searching of Genbank using the keyword "chalcone synthase" revealed that over 160 CHS sequences have been reported to date. CHS is encoded by a single gene in many plant species such as parsley and Arabidopsis, while in other plant species such as soybean and petunia it is encoded by a gene family of 6 to 8 members (Reif et al., 1985; Herrmann et al., 1988; Harker et al., 1990; Akada and Dube, 1995; Burbulis *et al.*, 1996).

A gene encoding CHI, the second enzyme of the flavonoid pathway, was first isolated from French bean by screening an expression library with a CHI antibody raised against the purified protein of this species (Robbins and Dixon, 1984; Mehdy and Lamb, 1987). A similar strategy using a petunia CHI antibody was used to isolate the gene from petunia, and the identity of this clone confirmed by an enzyme assay (van Tunen and Mol, 1987; van Tunen et al., 1988). Since this time, the genes encoding CHI have been reported from 17 different plant species, including maize (a monocot) (Grotewold and Peterson, 1994) and Arabidopsis (Shirley et al., 1992), which is of interest in the current study. In some plant species such as bean and Arabidopsis, CHI is encoded by a single gene while in other such as maize and alfalfa it may be encoded by two genes (Mehdy and Lamb, 1987; McKhann and Hirsch, 1994).

As with CHS and CHI, several genes encoding F3H have been isolated from plants. The gene was first cloned from snapdragon by differential screening using labeled RNA from *Del*- or *Del*+ plants (Martin *et al.*, 1991). The *Del* locus was known to control the expression of several flavonoid biosynthetic genes encoding enzymes acting after CHS in the pathway. Several clones to which *Del*+ RNA preferentially hybridized were isolated. RNA blot analysis of snapdragon lines carrying a mutation in the *INC* locus, which was known to encode F3H, with a probe derived from one of the clones revealed that this transcript was not present in the mutant, confirming the fact that the isolated gene encoded F3H (Forkmann and Stotz, 1981). Similar genes were subsequently isolated from numerous plant species. In all published reports in which the number of genes encoding F3H were investigated, it is present in one or two copies (Britsch et al., 1992; Meldgaard, 1992; Sparvoli et al., 1994; Deboo et al., 1995).

In contrast to other flavonoid enzymes, only three genes encoding FLS have been isolated to date, first from petunia and then from potato and Arabidopsis (Holton et al., 1993; Pelletier et al., 1997; van Eldik et al., 1997). FLS activity in petunia was known to be controlled by the *Fl* locus, as *flfl* petunia lines had markedly reduced FLS activity. Since the biochemical characterization of FLS had revealed that the enzyme belonged to the family of 2ODDS, Holton and colleagues amplified fragments of cDNA by PCR from *FlFl* or *flfl* lines using degenerate oligonucleotides targeted to conserved 2ODD amino acid sequences. Labeling of the respective PCR products and differential hybridization to a library from a *FlFl* line revealed several clones that hybridized more strongly to the *FlFl* -derived probe. A full-length clone was isolated and

expressed in yeast to confirm that the gene encoded FLS. Suprisingly, FLS is believed to be encoded by a small gene family in Arabidopsis, while all other flavonoid enzymes in this species are believed to be encoded by a single gene (Shirley *et al.*, 1992; Burbulis *et al.*, 1996; Pelletier and Shirley, 1996; Pelletier et al, 1997)

The first genes encoding DFR were cloned from maize and snapdragon by transposon tagging of the *A1* and *PALLIDA* loci, respectively (Martin et al., 1985; O'Reilly et al., 1985) Maize lines carrying mutations at *A1* were known to accumulate quercetin, and subsequent isolation and expression of a full-length A1 gene revealed that it encoded an enzyme capable of converting dihydroquercetin to leucocyanidin (Reddy et al., 1987; Schwarz-Sommer et al., 1987). While the enzymatic function encoded by the *PALLIDA* locus was at first unclear, sequence homology to the maize A1 protein suggested that it also encoded DFR. Genes encoding this enzyme were later isolated from petunia, barley, Arabidopsis, and *Gerbera hybrida* (Beld et al., 1989; Kristiansen and Rohde, 1991; Shirley *et al.*, 1992; Helariutta et al., 1995). The DFR protein is thought to be encoded by 3 genes in petunia, and by a single gene in barley and Arabidopsis.

The genes encoding the enzyme referred to throughout this manuscript as LDOX were originally identified as loci involved in the conversion of flavan-3,4-diols to anthocyanidins (Fig. 1) in maize (*A2*) and snapdragon (*Candi*) (Reddy and Coe, 1962; Martin *et al.*, 1991). The maize gene was cloned by transposon tagging, and particle bombardment into the maize aleurone confirmed that the cloned gene complemented the *A2* mutation (Menssen et al., 1990). Similar to the isolation of F3H from snapdragon, differential screening using a regulatory mutant (*del*) that had reduced activities of several enzymes acting late in flavonoid biosynthesis was carried out (Martin *et al.*, 1991). Linkage analysis showed that DNA homologous to one of the clones had been deleted from the *candi* mutant, and feeding experiments revealed that this mutant was blocked at an enzymatic step after DFR but before UFGT. The *Candi* clone was subsequently used to isolate homologues from apple and grape, and a petunia gene with homology to *A2* and *Candi* was isolated by differential screening (Davies, 1993; Weiss et al., 1993; Sparvoli *et al.*, 1994). A cDNA clone encoding LDOX was later isolated from Arabidopsis, and the enzyme is believed to be encoded by a single gene in this species (Pelletier *et al.*, 1997).

Mutants Defective in Flavonoid Biosynthesis

One of the reasons why the flavonoid biosynthetic pathway is a useful model pathway for studying metabolic regulation is that a variety of mutants defective in both structural and regulatory genes have been isolated from a number of plant species. The ability to isolate so many flavonoid-deficient mutants probably reflects the fact that these compounds are not required for survival in many plant species (when grown under greenhouse conditions) and that in some plants, flavonoid enzymes are encoded by multiple genes. Due to the number of mutants isolated to date, a detailed description of each mutant will not be given, except for the Arabidopsis mutants which are of direct relevance in this study.

By far, most of the flavonoid mutants that have been characterized to date are from maize, petunia, snapdragon, and Arabidopsis, as numerous regulatory and structural mutants have been isolated from each of these species. In maize, mutants defective in CHS (*c2*), DFR (*a1*), LDOX (*a2*), UFGT (*bz1*), and a glutathione S-transferase which conjugates glutathione to anthocyanins have been identified (Dooner, 1983; Dooner et al., 1985; O'Reilly *et al.*, 1985; Menssen *et al.*, 1990; Marrs *et al.*, 1995). Petunia mutants defective in CHI (*po*), F3H (*an3*), and DFR (*an6*) have been identified (Froemel et al., 1985; van Tunen and Mol, 1987; Beld *et al.*, 1989). Similarly, snapdragon mutants defective in CHS (*niv*), F3H (*inc*), DFR (*pal*), and LDOX (*can*) have been identified (Forkmann and Stotz, 1981; Martin *et al.*, 1985; Wienand *et al.*, 1986; Martin *et al.*,

1991). In addition to these, a barley mutant defective in F3H (*Ant17*) was isolated (Meldgaard, 1992).

In Arabidopsis, at least 13 loci required for flavonoid synthesis have been identified. These mutants were isolated based on the lack of brown pigments (tannins, a polymer of flavan-3,4-diols) in the seed coat (Koornneef, 1990; Shirley et al., 1995). This lack of tannins causes the seed coat to be transparent, revealing the yellow cotyledons of the developing embryo, thus causing the mutant seeds to appear yellow or pale brown rather than brown. Hence, the name *transparent testa* (*tt*) is collectively given to these mutants, although several also lack flavonoids in all other tissues as well (Shirley *et al*., 1995). While the molecular defect responsible for the *tt* phenotype has been elucidated for several of these mutants, for others it remains unknown. At least two putative regulatory mutants have been identified(*ttg* and *tt8*) that affect the accumulation of some but not all flavonoid mRNA's (Shirley *et al.,* 1995).

Three of the *tt* mutants were characterized extensively prior to the research presented in this manuscript. Two of these mutants, *tt5* and *tt3*, were shown to be defective in the CHI and DFR loci, respectively (Shirley *et al.*, 1992). The CHI gene in *tt5* was found to contain an insertion and almost the entire coding region was inverted, while in *tt3*, the DFR locus was deleted from the genome entirely. Since DFR and CHI are each encoded by a single gene in this species, *tt3* and *tt5* do not have DFR or CHI activity, respectively. In contrast to the drastic chromosomal defects in these two mutants, the 2YY6 allele of *tt4* was shown to contain a single G to A transition at the 3' acceptor site of the first intron/exon border of CHS (Shirley *et al*., 1995; Burbulis *et al.,* 1996). The abberant splicing in this mutant causes a frameshift mutation that results in a truncated protein which completely lacks the active site of CHS. Furthermore, no CHS protein was detected in this mutant and HPLC analysis revealed that no flavonoids were present. Therefore, this mutant represents a null allele of CHS. These three mutants are useful tools for determining how the loss of one flavonoid enzyme may affect the level of mRNA or protein of other pathway enzymes.

In addition to mutants defective in genes encoding flavonoid enzymes, regulatory mutants that affect the expression of various mRNA's have been identified in a number of plant species. Among these are the snapdragon mutant *del*, which is unaffected in *CHS* and *CHI* mRNA accumulation, but has drastically reduced levels of *F3H*, *DFR*, *LDOX*, and *UFGT* mRNA (Almeida et al., 1989; Martin *et al*., 1991). In maize, regulatory mutants have been isolated that affect the levels of mRNAs encoding CHS, CHI, DFR, and UFGT in a tissue-specific manner (Dooner, 1983; Cone et al., 1986; Ludwig et al., 1989; Taylor and Briggs, 1990). Similarly, the *ant13* mutant of barley is incapable of expressing *CHS*, *F3H*, or *DFR* mRNA in seeds (Meldgaard, 1992). Finally, the petunia mutants *an1, an2*, and *an11* accumulate near wild-type levels of *CHS*, *CHI*, and *F3H* mRNA but have reduced levels of *DFR* mRNA (Quattrocchio et al., 1993).

Regulation of the Flavonoid Pathway at the Level of Gene Expression

RNA blot analyses of regulatory mutants in petunia, maize, snapdragon, and Arabidopsis suggest that the flavonoid pathway may be divided into two, differentially expressed groups of genes termed "early" and "late". The early genes encode enzymes acting near the beginning of the pathway while late genes encode enzymes that catalyze terminal steps in flavonoid synthesis. However, the individual genes that make up the early or late group differs from one organism to another. For example, F3H is regulated with the early genes, CHS and CHI, in Arabidopsis seedlings and petunia flowers but with DFR (a late gene) in snapdragon flowers (Martin *et al.*, 1991; Quattrocchio *et al.,* 1993).

One of the most well-studied aspects of the transcriptional regulation of flavonoid genes is the response of the pathway to light (reviewed by Hahlbrock and Scheel, 1989). Some of the earliest experiments on the light regulation of the pathway were carried out by measuring the levels

of CHS transcription in cultured cells of parsley that were exposed to different wavelengths and dosages of light. These experiments revealed that UV light was most efficient at inducing CHS transcription, while blue light also had an effect. Experiments using Arabidopsis CHS promoter sequences fused to the β-glucuronidase (GUS) reporter gene showed that CHS expression was inducible by high-intensity white light, blue light (defined here as 400-500 nm wavelength), and UV light (defined as 320-450 nm wavelength) (Feinbaum et al., 1991). By analyzing deletion constructs of the CHS promoter fused to GUS, a DNA sequence approximately 500 bp in length was found to mediate this transcriptional regulation. Similarly, CHS, CHI, and DFR expression were shown to be inducible by blue and UV-B light in Arabidopsis (Feinbaum and Ausubel, 1988; Kubasek et al., 1992). Recently, the emerging picture of light regulation of Arabidopsis CHS has grown more complex, as the expression of this gene has been shown to be the product of complex interactions between UV-B (280-320 nm), UV-A (320-390 nm), and blue light (390-500 nm) (Fuglevand et al., 1996). Furthermore, in certain plants, the photoreceptor used to perceive the light that leads to CHS transcription is controlled by a developmental switch (Frohnmeyer et al., 1992).

The use of Arabidopsis mutants (*hy4*) that are defective in a blue light photoreceptor (CRY1) has provided further insights into the role of blue light as a signal leading to flavonoid gene transcription (Ahmad and Cashmore, 1993). Several *hy4* mutants that were grown in blue light were shown to have substantially decreased levels of flavonoid mRNA's or anthocyanins as compared to wild-type while transgenic plants overexpressing CRY1 had increased levels of flavonoid mRNA's and anthocyanins (Jackson and Jenkins, 1995; Ahmad et al, 1995; Lin et al., 1996).

In addition to light regulation, a variety of other environmental cues have been shown to induce expression of flavonoid biosynthetic genes. For instance cytokinins were shown to induce CHS, CHI, and DFR gene expression in Arabidopsis (Deikman and Hammer, 1995). Several studies have shown that sugars, in particular sucrose, and gibberlic acid induce the expression of flavonoid genes in petunia flowers and in leaves of *Camellia sinensis* and both of these compounds are required for maximal induction (Weiss and Halevy, 1989; Weiss et al., 1991; Takeuchi et al., 1994). Furthermore, flavonoid gene expression is inducible by a variety of stresses including hot or cold temperatures, mechanical wounding or treatment with elicitors (Ryder et al., 1987; Shirley and Goodman, 1993; Leyva et al., 1995) . The induction of the genes of the pathway by such a variety of stresses may be due to the roles flavonoids have in plant defense.

Project Goals

As noted above, understanding the mechanisms that regulate metabolic pathways is of fundamental importance in biology. The flavonoid pathway of *Arabidopsis thaliana* provides an excellent model for studying metabolic regulation, as several null mutants with defects in specific flavonoid genes have been identified as well as regulatory mutants affected in the expression of subsets of flavonoid genes. This allows one to examine how a mutation at different points in the pathway, and thus presumably the accumulation of different flavonoid intermediates, may affect the levels of other transcripts or proteins of the pathway. Furthermore, since a wealth of knowledge of the transcriptional regulation and biochemistry of this pathway already existed (i.e. the designation of "early" and "late" genes in some plants, the inducibility of the pathway, etc), this provides a good context into which the characterization of new genes may be placed and in which to better understand the regulation of the pathway at the protein level. *Arabidopsis thaliana* also has many advantages over other plant species, as it has one of the smallest genomes of all higher plants (~100 MB per haploid genome), produces thousands of seeds per plant in a generation of 6-10 weeks, and takes little work and space to grow (Hauge et al., 1991).

The first objective of the work presented here was to isolate and characterize genes encoding three of the four flavonoid 2ODDS (F3H, FLS, and LDOX) from Arabidopsis. Since

only genes encoding CHS, CHI, and DFR had been previously cloned, the isolation of these genes would double the total number of cloned Arabidopsis flavonoid genes and provide clones for enzymes that act sequentially in the pathway. Experiments designed to determine whether each gene was an "early" or "late gene in Arabidopsis were carried out by performing blot analysis using RNA isolated from wild-type seedlings grown in darkness and then shifted to high-intensity white light or from analysis of RNA levels in the regulatory mutants *ttg* and *tt8.* Since whether a gene in a particular tissue or plant species is "early" or "late" probably reflects differences in flavonoid function, these are important experiments. F3H may mediate a critical control point in the pathway as it is coordinately expressed with the "early" genes in some plants but with the "late" genes in others. Furthermore, characterization of FLS as an "early" or "late gene had not been previously undertaken. To determine the number of genes encoding each of these enzymes, low-stringency DNA blot analysis was performed. Finally, the chromosomal location of each 2ODD gene was determined using recombinant inbred lines that were developed by Lister and Dean (1993). This allowed comparison of the locations of these genes with those of *tt* loci, which might in turn identify *tt* mutants putatively defective in one of these three enzymes.

The second main objective of this research was to develop antibodies against all of the flavonoid enzymes for which genes had been isolated from Arabidopsis in order to characterize the pathway at the level of protein expression. This was especially important since almost all studies of the regulation of the flavonoid pathway had been performed using RNA blot analysis or promoter-reporter gene constructs, and these methods may not reveal important aspects of the post-transcriptional mechanisms by which this pathway may be regulated. Immunoblot analysis was performed using proteins from wild-type seedlings grown for 2 to 9 days in continuous white light. To address the question of whether a *tt* mutant blocked at one point in the pathway may have altered levels of other flavonoid enzymes, and to attempt to further characterize the existing collection of *tt* mutants, immunoblot analysis was also performed on 13 of these mutants. Because we are interested in metabolic regulation as a whole, we wondered if changes in the flux through pathways which provide or compete for *p*-coumaric acid, a precursor of coumaroyl-CoA one of the substrates of CHS, would affect the accumulation of flavonoid enzymes. Therefore, we examined the levels of flavonoid enzymes in mutants altered in ferulic acid or tryptophan biosynthesis to determine whether cross-pathway regulation occurs among the branches of primary and secondary aromatic amino acid metabolism.

Literature Cited

Ahmad M and Cashmore AR (1993) *HY4* gene of *A. thaliana* encodes a protein with characteristics of a blue-light photoreceptor. Nature **366**: 162-166

Ahmad M, Lin C, and Cashmore AR (1995) Mutations throughout an Arabidopsis blue-light photoreceptor impair blue-light-responsive anthocyanin accumulation and inhibition of hypocotyl elongation. Plant J. **8**: 653-658

Akada S, and Dube S (1995). Organization of soybean chalcone synthase gene clusters and characterization of a new member of the family. Plant Mol. Biol. **29**: 189-199.

Almeida J, Carpenter R, Robbins R, Martin C, and Coen E (1989). Genetic interactions underlying flower colour patterns in *Antirrhinum majus*. Genes Dev. **3**: 1758-1767.

Baker M, Luu-the V, Simard J, and Labrie F (1990). A common ancestor for mammalian 3β-hydroxysteroid dehydrogenase and plant dihydroflavonol reductase. BJ Letters **269**: 558-559.

Beld M, Martin C, Huits H, Stuitje AR, and Gerats AG (1989). Flavonoid synthesis in Petunia hybrida: partial characterization of dihydroflavonol-4-reductase genes. Plant Mol. Biol. **13**: 491-502.

Britsch L, and Grisebach H (1986). Purification and characterization of (2S)-flavanone 3-hydroxylase from Petunia hybrida. Eur. J. Biochem. **156**: 569-577.

Britsch L, Heller W, and Grisebach H (1981). Conversion of flavanone to flavone, dihydroflavonol, and flavonol with an enzyme system from cell cultures of Parsley. Z. Naturforsch **36**: 742-750.

Britsch L, Ruhnau-Brich B, and Forkmann G (1992). Molecular cloning, sequence analysis, and in vitro expression of flavanone 3 beta-hydroxylase from *Petunia hybrida*. J. Biol. Chem. **267**: 5380-5387.

Burbulis I, Iacobucci M, and Shirley B (1996). A null mutation in the first enzyme of flavonoid biosynthesis does not affect male fertility in Arabidopsis. Plant Cell **8**: 1013-1025.

Cone KC, Burr FA, and Burr B (1986). Molecular analysis of the maize anthocyanin regulatory locus C1. Proc. Natl. Acad. Sci. USA **83**: 9631-9635.

Davies K (1993). A cDNA clone for flavanone 3-hydroxylase from Malus. Plant Physiol. **103**: 291.

Deboo G, Albertsen M, and Taylor L (1995). Flavanone 3-hydroxylase transcripts and flavonol accumulation are temporally coordinate in maize anthers. Plant J. **7**: 703-713.

Deikman J, and Hammer PE (1995). Induction of anthocyanin accumulation by cytokinins in *Arabidopsis thaliana*. Plant Physiol. **108**: 47-57.

Dooner HK (1983). Coordinate genetic regulation of flavonoid biosynthetic enzymes in maize. Mol. Gen. Genet. **189**: 136-141.

Dooner HK, Weck E, Adams S, Ralston E, and Favreau M (1985). A molecular genetic analysis of insertions in the *bronze* locus in maize. Mol. Gen. Genet. **200**: 240-246.

Feinbaum RL, and Ausubel FM (1988). Transcriptional regulation of the *Arabidopsis thaliana* chalcone synthase gene. Mol. Cell. Biol. **8**: 1985-1992.

Feinbaum RL, Storz G, and Ausubel FM (1991). High intensity and blue light regulated expression of chimeric chalcone synthase genes in transgenic *Arabidopsis thaliana* plants. Mol. Gen. Genet. **226**: 449-456.

Fischer D, Stich K, Britsch L, and Grisebach H (1988). Purification and characterization of (+)-dihydroflavonol 4-reductase from flowers of *Dahlia variabilis*. Arch. Biochem. Biophys. **264**: 40-47

Forkmann G (1980). Anthocyanin biosynthesis in flowers of *Matthiola incana* flavanone 3- and flavonoid 3' hydroxylases. Z. Naturforsch. **35c**: 691-695.

Forkmann G, de Vlaming P, Spribille R, Wiering H, and Schram AW (1986). Genetic and biochemical studies on the conversion of dihydroflavonols to flavonols in flowers of *Petunia hybrida*. Z. Naturforsch **41**: 179-186.
Forkmann G, and Stotz G (1981). Genetic control of flavanone 3-hydroxylase activity and flavonoid 3' hydroxylase activity in *Antirrhinum majus* (Snapdragon). Z. Naturforsch. **36c**: 411-416.
Froemel S, de Vlaming P, Stotz G, Wiering H, Forkmann G, and Schram A (1985). Theor. Appl. Genet. **70**: 561-568.
Frohnmeyer H, Ehmann B, Kretsch T, Rocholl M, Harter K, Nagatani A, Furuya M, Batschauer A, Hahlbrock K, and Schäfer E (1992). Differential usage of photoreceptors for chalcone synthase gene expression during plant development. Plant J. **2**: 899-906.
Fuglevand G, Jackson J, and Jenkins G (1996). UV-B, UV-A, and blue light signal transduction pathways interact synergistically to regulate chalcone synthase gene expression in Arabidopsis. The Plant Cell **8**: 2347-2357.
Grotewold E, and Peterson T (1994). Isolation and characterization of a maize gene encoding chalcone flavanon isomerase. Mol. Gen. Genet. **242**: 1-8.
Hahlbrock K, and Scheel D (1989). Physiology and molecular biology of phenylpropanoid metabolism. Ann. Rev. Plant Physiol. Plant Mol. Biol. **40**: 347-369.
Hahlbrock K, Wong E, Schill L, and Grisebach H (1970). Comparison of chalcone-flavanone isomerase heteroenzymes and isoenzymes. Phytochemistry **9**: 949-958
Harker CL, Ellis THN, and Coen ES (1990). Identification and genetic regulation of the chalcone synthase multigene family in pea. Plant Cell **2**: 185-194.
Hauge B, Giraudat J, Hanley S, Hwang I, Kohchi T, and Goodman H (1991). Physical mapping of the Arabidopsis genome and its applications. Plant Mol. Biol. **2**: 239-248.
Helariutta Y, Kotilainen M, Elomaa P, Kalkkinen N, Bremer K, Teeri T, and Albert V (1996). Duplication and functional divergence in the chalcone synthase gene family of Asteraceae: Evolution with substrate change and catalytic simplification. Proc. Natl. Acad. Sci. USA **93**: 9033-9038.
Helariutta Y, Kotilainen M, Elomaa P, and Teeri TH (1995). *Gerbera hybrida* (Asteraceae) imposes regulation at several anatomical levels during inflorescence development on the gene for dihydroflavonol- 4-reductase. Plant Mol. Biol. **28**: 935-941.
Herrmann A, Schulz W, and Hahlbrock K (1988). Two alleles of the single-copy chalcone synthase gene in parsley differ by a transposon-like element. Mol. Gen. Genet. **212**: 93-98.
Holton TA, Brugliera F, and Tanaka Y (1993). Cloning and expression of flavonol synthase from *Petunia hybrida*. Plant J. **4**: 1003-1010.
Holton TA, Brugliera F, Lester DR, Tanaka Y, Hyland CD, Menting JGT, Lu CY, Farcy E, Stevenson TW, and Cornish EC (1993) Cloning and expression of cytochrome P450 genes controlling flower colour. Nature **366**: 276-279
Hrazdina G, and Jensen RA (1992). Spatial organization of enzymes in plant metabolic pathways. Ann. Rev. Plant Physiol. Plant Mol. Biol. **43**: 241-267.
Hrazdina G, Lifson E, and Weeden NF (1986). Isolation and characterization of buckwheat (*Fagopyrum esculentum* M.) chalcone synthase and its polyclonal antibodies. Arch. Biochem. Biophys. **247**: 414-419.
Hrazdina G, and Wagner GJ (1985a). Compartmentation of plant phenolic compounds; sites of synthesis and accumulation. Ann. Proc. Phytochem. Soc. Europe **25**: 120-33.
Hrazdina G, and Wagner GJ (1985b). 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, and Hoch HC (1987). Biochemical, immunological and immunocytochemical evidence for the association of chalcone synthase with endoplasmic reticulum membranes. Proc. Natl. Acad. Sci. USA **84**: 8966-70.
Jackson JA and Jenkins GI (1995). Extension-growth responses and expression of flavonoid biosynthesis genes in the Arabidopsis *hy4* mutant. Planta **197**:233-239
Koes RE, Quattrocchio F, and Mol JNM (1993). The flavonoid biosynthetic pathway in plants: function and evolution. BioEssays **16**: 123-132.
Koornneef M (1990). Mutations affecting the testa colour in Arabidopsis. Arabidopsis Information Server **27**: 1-4.
Kreuzaler F, Ragg H, Fautz E, Kuhn DN, and Hahlbrock K (1983). UV-induction of chalcone synthase mRNA in cell suspension cultures of *Petroselinum hortense*. Proc. Natl. Acad. Sci. USA **80**: 2591-2593.
Kreuzaler F, Ragg H, Heller W, Tesch R, Witt I, Hammer D, and Hahlbrock K (1979). Flavanone synthase from *Petroselinum hortense*: Molecular weight, subunit composition, size of messenger RNA, and absence of pantetheinyl residue. Eur. J. Biochem. **99**: 89-96.
Kristiansen KN, and Rohde W (1991). Structure of the Hordeum vulgare gene encoding dihydroflavonol-4- reductase and molecular analysis of ant18 mutants blocked in flavonoid synthesis. Mol. Gen. Genet. **230**: 49-59.
Kubasek WL, Shirley BW, McKillop A, Goodman HM, Briggs W, and Ausubel FM (1992). Regulation of flavonoid biosynthetic genes in germinating Arabidopsis seedlings. The Plant Cell **4**: 1229-1236.
Leyva A, Jarillo JA, Salinas J, and Martinez-Zapater JM (1995). Low temperature induces the accumulation of phenylalanine ammonia-lyase and chalcone synthase mRNAs of *Arabidopsis thaliana* in a light-dependent manner. Plant Physiol. **108**: 39-46.
Lin C, Ahmad M, and Cashmore, AR (1996). Arabidopsis cryptochrome 1 is a soluble protein mediating blue light-dependent regulation of plant growth and development. Plant J.**10**: 893-902
Lister C, and Dean C (1993). Recombinant inbred lines for mapping RFLP and phenotypic markers in Arabidopsis thaliana. The Plant J. **4**: 745-750.
Lois R, and Buchanan B (1994). Severe sensitivity to ultraviolet radiation in an Arabidopsis mutant deficient in flavonoid accumulation. Planta **194**: 504-509.
Ludwig SR, Habera LF, Dellaporta SL, and Wessler SR (1989). Lc, a member of the maize R gene family responsible for tissue-specific anthocyanin production, encodes a protein similar to transcriptional activators and contains the myc-homology region. Proc. Natl. Acad. Sci. USA **86**: 7092-7096.
Marrs KA, Alfenito MR, Lloyd AM, and Walbot V (1995). A glutathione S-transferase involved in vacuolar transfer encoded by the maize gene *Bronze-2*. Nature **3755**: 397-400.
Martin C, Carpenter R, Sommer H, Saedler H, and Coen ES (1985). Molecular analysis of instability in flower pigmentation of *Antirrhinum majus*, following isolation of the *pallida* locus by transposon tagging. EMBO J. **4**: 1625-1630.
Martin C, Prescott A, Mackay S, Bartlett J, and Vrijlandt E (1991). Control of anthocyanin biosynthesis in flowers of *Antirrhinum majus*. Plant J. **1**: 37-49.
McKhann HI, and Hirsch AM (1994). Isolation of chalcone synthase and chalcone isomerase cDNAs from alfalfa (Medicago sativa L.): highest transcript levels occur in young roots and root tips. Plant Mol. Biol. **24**: 767-77.
Mehdy MC, and Lamb CJ (1987). Chalcone isomerase cDNA cloning and mRNA induction by fungal elicitor, wounding and infection. EMBO J. **6**: 1527-1533.
Meldgaard M (1992). Expression of chalcone synthase, dihydroflavonol reductase, and flavanone-3-hydroxylase in mutants of barley deficient in anthocyanin and proanthocyanidin biosynthesis. Theor. Appl. Genet. **83**: 695-706.

Menssen A, Hohmann S, Martin W, Schnable SS, Peterson PA, Saedler H, and Gierl A (1990). The En/Spm transposable element of *Zea mays* contains splice sites at the termini generating a novel intron from a dSpm element in the *A2* gene. EMBO J. **9**: 3051-3057.
Miksicek R (1993). Commonly occurring plant flavonoids have estrogenic activity. Molecular Pharmacology **44**: 37-43.
Mo Y, Nagel C, and 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.
Moustafa E, and Wong E (1967). Purification and properties of chalcone-flavanone isomerase from soya bean seed. Phytochemistry **6**: 625-632.
O'Reilly C, Shepherd NS, Pereira A, Schwarz-Sommer Z, Bertram I, Robertson DS, Peterson PA, and Saedler H (1985). Molecular cloning of the *a1* locus of *Zea mays* using the transposable elements *En* and *Mu1*. EMBO J. **4**: 877-882.
Pelletier M, Murrell J, and Shirley B (1997). Characterization of Arabidopsis FLS and LDOX Genes: Further Evidence for Differential Regulation of "Early" and "Late" Genes. Plant Physiol. **113**: 1437-1445
Pelletier MK, and Shirley BW (1996). Analysis of flavanone 3-hydroxylase in Arabidopsis seedlings: Coordinate regulation with chalcone synthase and chalcone isomerase. Plant Physiol. **111**: 339-345.
Prescott A (1993). A dilemma of dioxygenases (or where biochemistry and molecular biology fail to meet). J. Exp. Bot **44**: 849-861.
Prescott A, and John P (1996). Dioxygenases: Molecular structure and role in plant metabolism. Annu. Rev. Plant Physiol. Plant Mol. Biol. **47**: 245-271.
Quattrocchio F, Wing JF, Leppen HTC, Mol JNM, and Koes RE (1993). Regulatory genes controlling anthocyanin pigmentation are functionally conserved among plant species and have distinct sets of target genes. Plant Cell **5**: 1497-1512.
Reddy A, Britsch L, Salamini F, Saedler H, and Rohde W (1987). The *A1* (Anthocyanin-1) locus in *Zea mays* encodes dihydroquercetin reductase. Plant Science **52**: 7-13.
Reddy G, and Coe E (1962). Intertissue complementation: a simple technique for direct analysis of gene-action sequence. Science **138**: 149.
Reif HJ, Niesbach U, Deumling B, and Saedler H (1985). Cloning and analysis of two genes for chalcone synthase from *Petunia hybrida*. Mol. Gen. Genet. **199**: 208-215.
Robbins M, and Dixon R (1984). Induction of chalcone isomerase in elicitor-treated bean cells. Comparison of rates of synthesis and appearance of immunodetectable enzyme. Eur. J. Biochem. **145**: 195-202.
Ryder T, Hedrick S, Bell J, Liang X, Clouse S, and Lamb C (1987). Organization and differential activation of a gene family encoding the plant defense enzyme chalcone synthase in *Phaseolus vulgaris*. Mol. Gen. Genet. **210**: 219-233.
Schwarz-Sommer Z, Shepherd N, Tacke E, Gierl A, Rohde W, Leclercq L, Mattes M, Berndtgen R, Peterson PA, and Saedler H (1987). Influence of transposable elements on the structure and function of the *A1* gene of *Zea mays*. EMBO J. **6**: 287-294.
Shirley B (1996). Flavonoid biosynthesis: 'new' functions for an 'old' pathway. Trends in Plant Science **1**: 377-382.
Shirley B, Kubasek W, Storz G, Bruggemann E, Koornneef M, Ausubel F, and Goodman H (1995). Analysis of Arabidopsis mutants deficient in flavonoid biosynthesis. The Plant J. **8**: 659-671.
Shirley BW, and Goodman HM (1993). An *Arabidopsis* gene homologous to mammalian and insect genes encoding the largest proteasome subunit. Mol. Gen. Genet. **241**: 586-594.
Shirley BW, Hanley S, and Goodman H (1992). Effects of ionizing radiation on a plant genome: Analysis of two Arabidopsis *transparent testa* mutations. Plant Cell **4**: 333-347.

Sparvoli F, Martin C, Scienza A, Gavazzi G, and Tonelli C (1994). Cloning and molecular analysis of structural genes involved in flavonoid and stilbene biosynthesis in grape (*Vitis vinifera* L.). Plant Mol. Biol. **24**: 743-755.
Spribille R, and Forkmann G (1984). Conversion of dihydroflavonols to flavonols with enzyme extracts from flower buds of *Matthiola incana* R. Br. Z. Naturforsch **39c**: 714-719.
Stafford H (1990). "Flavonoid Metabolism," CRC Press, Inc, Boca Raton.
Stapleton A, and Walbot V (1994). Flavonoids can protect maize DNA from the induction of ultraviolet radiation damage. Plant Physiol. **105**: 881-889.
Takeuchi A, Matsumoto S, and Hayatsu M (1994). Chalcone synthase from *Camellia sinensis*: Isolation of the cDNAs and the organ-specific and sugar-responsive expression of the genes. Plant Cell Physiol. **35**: 1011-1018.
Taylor L, and Briggs WR (1990). Genetic regulation and photocontrol of anthocyanin accumulation in maize seedlings. Plant Cell **2**: 115-127.
Tropf S, Karcher B, Schroder G, and Schroder J (1995). Reaction mechanisms of homodimeric plant polyketide synthase (stilbenes and chalcone synthase). A single active site for the condensing reaction is sufficient for synthesis of stilbenes, chalcones, and 6'- deoxychalcones. J. Biol. Chem. **270**: 7922-7928.
van der Meer IM, Stam ME, van Tunen AJ, Mol JNM, and Stuitje AR (1992). Antisense inhibition of flavonoid biosynthesis in petunia anthers results in male sterility. Plant Cell **4**: 253-262.
van Eldik G, Ruiter R, Reijnen W, Van Herpen M, Schrauwen J, and Wullems G (1997). Regulation of flavonol biosynthesis during anther and pistil development, and during pollen tube growth in *Solanum tuberosum.* The Plant J. **11**: 105-113.
van Tunen AJ, Koes RE, Spelt CE, van der Krol AR, Stuitje AR, and Mol JNM (1988). Cloning of the two chalcone flavanone isomerase genes from *Petunia hybrida*: coordinate, light-regulated and differential expression of flavonoid genes. EMBO J. **7**: 1257-1263.
van Tunen AJ, and Mol JN (1987). A novel purification procedure for chalcone flavanone isomerase from *Petunia hybrida* and the use of its antibodies to characterize the Po mutation. Arch Biochem Biophys **257**: 85-91.
Wagner G, and Hrazdina G (1984). Endoplasmic reticulum as a site of phenylpropanoid and flavonoid metabolism in *Hippeastrum*. Plant Physiol. **74**: 901-906.
Weiss D, and Halevy AH (1989). Stamens and gibberellin in the regulation of corolla pigmentation and growth in *Petunia hybrida*. Planta **179**: 89-96.
Weiss D, van Blokland R, Kooter JM, Mol JNM, and van Tunen AJ (1991). Gibberellic acid regulates chalcone synthase gene transcription in the corolla of *Petunia hybrida.* Plant Physiol. **98**: 191-197.
Weiss D, van der Luit AH, Kroon JTM, Mol JN, and Kooter JM (1993). The petunia homologue of the *Antirrhinum majus* candi and *Zea mays A2* flavonoid genes; homology to flavanone 3-hydroxylase and ethylene-forming enzyme. Plant Mol. Biol. **22**: 893-897.
Wienand U, Weydemann U, Niesbach-Kösgen U, Peterson PA, and Saedler H (1986). Molecular cloning of the *c2* locus of *Zea mays*, the gene coding for chalcone synthase. Mol. Gen. Genet. **203**: 202-207.

Chapter 2

Analysis of Flavanone 3-Hydroxylase in Arabidopsis Seedlings: Coordinate Regulation with Chalcone Synthase and Chalcone Isomerase

This work was published previously as:
Pelletier, MK, and Shirley, BW. Plant Physiol. **111**: 339-345 (1996).

Abbreviations: CHI, chalcone isomerase; CHS, chalcone synthase; DFR, dihydroflavanol reductase; DIG, digoxygenin; EST, expressed sequence tag; F3H, flavanone 3-hydroxylase; RI, recombinant inbred; RT-PCR, reverse transcription-PCR; *tt, transparent testa*; UFGT, UDPglucose 3-*O*-glucosyltransferase.

Chapter 3

Characterization of Arabidopsis FLS and LDOX Genes: Further Evidence for Differential Regulation of "Early" and "Late" Genes

This work was previously published as **M. K. Pelletier, J.R. Murrell, and B.W. Shirley.** Plant Physiology **113**: 1437-1445 (1997)

All experimental work was performed by M.K. Pelletier, except that J.R. Murrell assisted in isolating the genomic FLS clone.

Abbreviations: CHS, chalcone synthase; CHI, chalcone isomerase; DFR, dihydroflavonol reductase; DIG, digoxygenin; EST, Expressed Sequence Tag; F3H, flavanone 3-hydroxylase; F3'H, flavonoid 3'-hydroxylase; FLS, flavonol synthase; FS, flavone synthase; LDOX, leucoanthocyanidin dioxygenase; RT-PCR, reverse-transcription-PCR; *tt, transparent testa;* UFGT, UDP-flavonoid 3-O-glucosyl transferase.

The accession numbers for the sequences reported in this article are U72631 (FLS) and U70478 (LDOX).

Chapter 4

Immunoblot Analysis of Developing Arabidopsis Seedlings: Mutations in *TT3*, *TTG*, *TT5*, and *TSB1* lead to Increased Levels of Flavonoid Enzymes

M.K. Pelletier and B.W. Shirley

This chapter will be submitted to The Plant Journal

Abstract

We have developed a collection of polyclonal antibodies against the flavonoid biosynthetic enzymes chalcone synthase (CHS), chalcone isomerase (CHI), flavanone 3-hydroxylase (F3H), flavonol synthase (FLS), and leucoanthocyanidin dioxygenase (LDOX) from *Arabidopsis thaliana*. These antibodies were used to determine the steady- state levels of these enzymes in developing seedlings and in several Arabidopsis mutants. The accumulation of flavonoid enzymes over the course of seedling development correlated well with previous studies that indicated that CHS, CHI, F3H, and FLS are "early" genes while LDOX is a "late" gene. Immunoblot analysis was also performed on several mutants affected in flavonoid, tryptophan, and ferulic acid biosynthesis. Three of the *transparent testa* mutants (*tt3, tt5*, and *ttg*) that are blocked at specific enzymatic steps of flavonoid biosynthesis accumulated higher levels of the other pathway enzymes as compared to wild type, and these differences were not due to differences in the developmental state of the mutants. This suggests that the accumulation of specific flavonoid intermediates can cause an increase in the steady-state levels of proteins required for flavonoid metabolism. Furthermore, the *trp2-1* mutant, which is defective in one of the two Arabidopsis tryptophan synthase β-subunit genes, accumulated substantially higher levels of flavonoid enzymes and anthocyanins than did wild-type seedlings. The effect of *trp2-1* was not due to amino acid starvation, as this phenotype was also observed when these seedlings were grown on tryptophan-supplemented media. Finally, immunoblot and RNA blot analysis of *tt6* suggested that this mutant is defective in the Arabidopsis *F3H* gene. This was confirmed by sequence analysis of *F3H* from *tt6;* eighteen single-base-pair mutations were found in the coding region of this allele, three of which were predicted to result in radical amino acid changes.

Introduction

Flavonoids are secondary metabolites synthesized by all higher plants that are proposed to play a variety of physiological roles. These roles include protecting plants from UV light, acting as antimicrobial compounds (phytoalexins), playing a role in pollen tube maintenance, acting as signaling molecules in interactions with symbiotic microbes, and attracting pollinators (Hahlbrock and Scheel, 1989; Dooner et al., 1991; Mo et al., 1992; van der Meer et al., 1992; Koes et al., 1993; Stapleton and Walbot, 1994; Yao et al., 1995; Shirley, 1996). The flavonoid biosynthetic pathway of *Arabidopsis thaliana* serves as an excellent model to study metabolic regulation, as several of the genes encoding flavonoid enzymes have been cloned and shown to be single copy. A collection of mutants (*transparent testa* or *tt* mutants) defective in both structural and regulatory genes have been identified (Feinbaum and Ausubel, 1988; Shirley et al., 1992; Shirley et al., 1995; Burbulis et al., 1996; Pelletier and Shirley, 1996; Pelletier, 1997). Included among these mutants are a null allele of *CHS* (*tt4*-2YY6), a null allele of *CHI*(*tt5*), a null allele of *DFR* (*tt3*), and two regulatory mutants (*ttg* and *tt8*) that affect the expression of *DFR* and/or *LDOX* mRNA. These mutants, as well as several mutants from other plant species have proven to be invaluable tools for studying flavonoid metabolism, as RNA blot analyses have shown that different genetic loci control the expression of different flavonoid genes in a developmental or tissue-specific manner (Dooner, 1983; Cone et al., 1986; Ludwig et al., 1989; Taylor and Briggs, 1990; Martin et al., 1991; Meldgaard, 1992; Quattrocchio et al., 1993; Deboo et al., 1995; Shirley *et al.*, 1995; Pelletier and Shirley, 1996) (Pelletier et al, 1997). Similar studies investigating wild-type seedlings or flowers at various stages of development have helped define groups of flavonoid genes that are coordinately expressed (Jackson et al., 1992; Pelletier and Shirley, 1996).

Another useful feature of the flavonoid pathway for studying metabolic regulation is that it is related to the pathways leading to phenylalanine, tryptophan, and tyrosine biosynthesis by the

common precursor chorismate and to the ferulic acid and sinapic acid biosynthetic pathways by the common precursor *p*-coumaric acid (Figure 1). Mutants have been identified in each of these competing pathways, allowing one to easily determine how mutations in one of these three pathways might affect the levels of end products or enzymes in the other two (Chapple et al., 1992). Several studies in Arabidopsis suggest that alterations in the activity of a single enzyme in the sinapic acid, flavonoid, or tryptophan biosynthetic pathways can affect the other pathways, although the mechanism(s) involved remains unclear. For example, the Arabidopsis *trp2* mutant, which is defective in a tryptophan synthase β subunit (*TSB1*) gene, has light-brown seeds relative to wild type (Last et al., 1991). Similarly, transgenic potato expressing a tryptophan decarboxylase gene consistently exhibit reduced levels of phenylalanine and other phenolic compounds (Yao *et al.*, 1995). The Arabidopsis mutant defective in the *CHI* locus has decreased levels of sinnipate esters as compared to wild-type plants (Li et al., 1993). Finally, a recent study suggests that cinnamate 4-hydroxylase mRNA levels are elevated slightly in the *tt8* mutant of Arabidopsis (Bell-LeLong et al., 1997). Cross-regulation of metabolic pathways may be prevalent in plants, as a recent study showed that blocking histidine biosynthesis led to increased levels of mRNA for genes encoding enzymes involved in aromatic amino acids, lysine, and purine biosynthesis (Guyer et al., 1995).

To more fully investigate the regulation of the flavonoid biosynthetic pathway at the level of protein expression, we have developed antibodies against the Arabidopsis CHS, CHI, F3H, FLS, and LDOX proteins. These antibodies, along with the many mutants available in Arabidopsis provide powerful new tools to study the regulation of flavonoid biosynthesis. Immunoblot analysis using extracts from developing wild-type seedlings and from 13 different *tt* mutants, 4 tryptophan mutants, and the Arabidopsis *fah1* mutant was performed to determine how the levels of flavonoid enzymes change during seedling development and whether mutations at distinct enzymatic steps of the flavonoid, ferulic acid, or tryptophan pathways would affect the levels of these proteins. The four tryptophan mutants included in this analysis were *trp1* (defective in phosphoribosylanthranilate transferase), *trp2* (defective in one of the two tryptophan synthase β genes of Arabidopsis), and atr1d/atr2d (both of which have increased levels of *ASA1* mRNA) (Last *et al.*, 1991; Rose et al., 1992; J. Bender, John's Hopkins University, personal communication). Several mutants were found to accumulate higher levels of flavonoid enzymes and/or end products relative to wild-type seedlings. This analysis also demonstrated that *tt6* accumulated no F3H protein. Subsequent RNA blot analysis of this mutant, and sequencing of the F3H gene from *tt6* revealed that 18 point mutations were present within the coding region of this ethyl methane sulfonate (EMS)-generated mutant.

Results

Expression of Recombinant Flavonoid Enzymes in *E. coli*

In order to develop antibodies against Arabidopsis flavonoid enzymes, recombinant proteins were expressed in *E. coli*. The overexpression and purification of GST-CHI was described previously (Cain et al, submitted). Expression constructs were created by ligating cDNA sequences encoding truncated or full-length Arabidopsis CHS, F3H, FLS, DFR, and LDOX proteins in-frame with glutatione *S*-transferase or thioredoxin genes in the expression vectors, pGEX-KG or pET32a, respectively (Figure 2). Substantial amounts of each recombinant protein of the appropriate size were obtained 4 h after inducing expression with IPTG (Figure 3). However, the proportion of total protein in the soluble or insoluble fraction using this expression system varied greatly among the six recombinant enzymes, as CHI and F3H were present primarily in the soluble fraction, FLS, DFR, and LDOX were found primarily in the insoluble

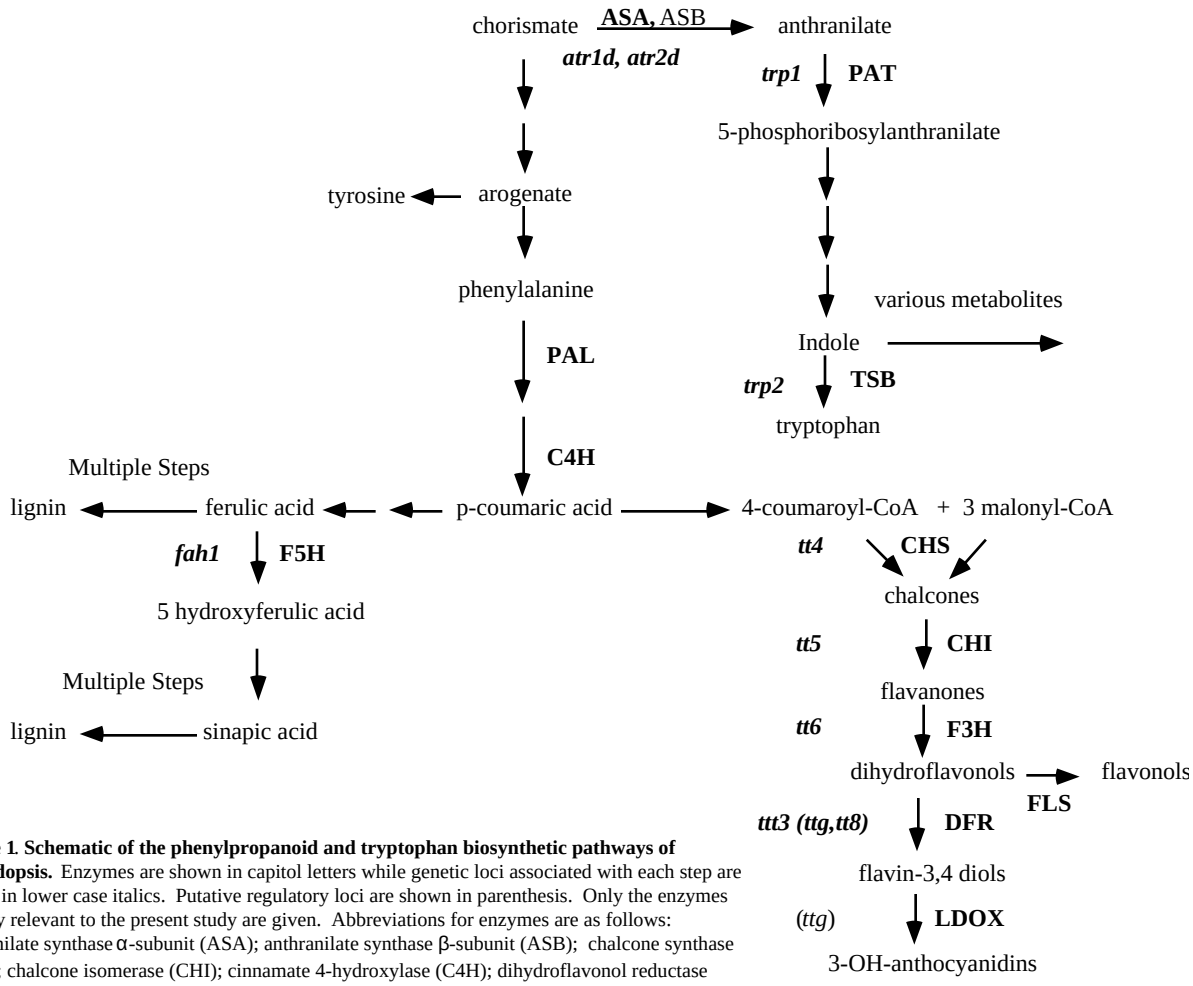


Figure 1. Schematic of the phenylpropanoid and tryptophan biosynthetic pathways of Arabidopsis. Enzymes are shown in capitol letters while genetic loci associated with each step are shown in lower case italics. Putative regulatory loci are shown in parenthesis. Only the enzymes directly relevant to the present study are given. Abbreviations for enzymes are as follows: anthranilate synthase α-subunit (ASA); anthranilate synthase β-subunit (ASB); chalcone synthase (CHS); chalcone isomerase (CHI); cinnamate 4-hydroxylase (C4H); dihydroflavonol reductase (DFR); ferulate 5-hydroxylase (F5H); flavonol synthase (FLS); flavanone 3-hydroxylase (F3H); leucoanthocyanidin dioxygenase (LDOX); phenylalanine ammonia lyase (PAL); tryptophan synthase β-subunit (TSB).

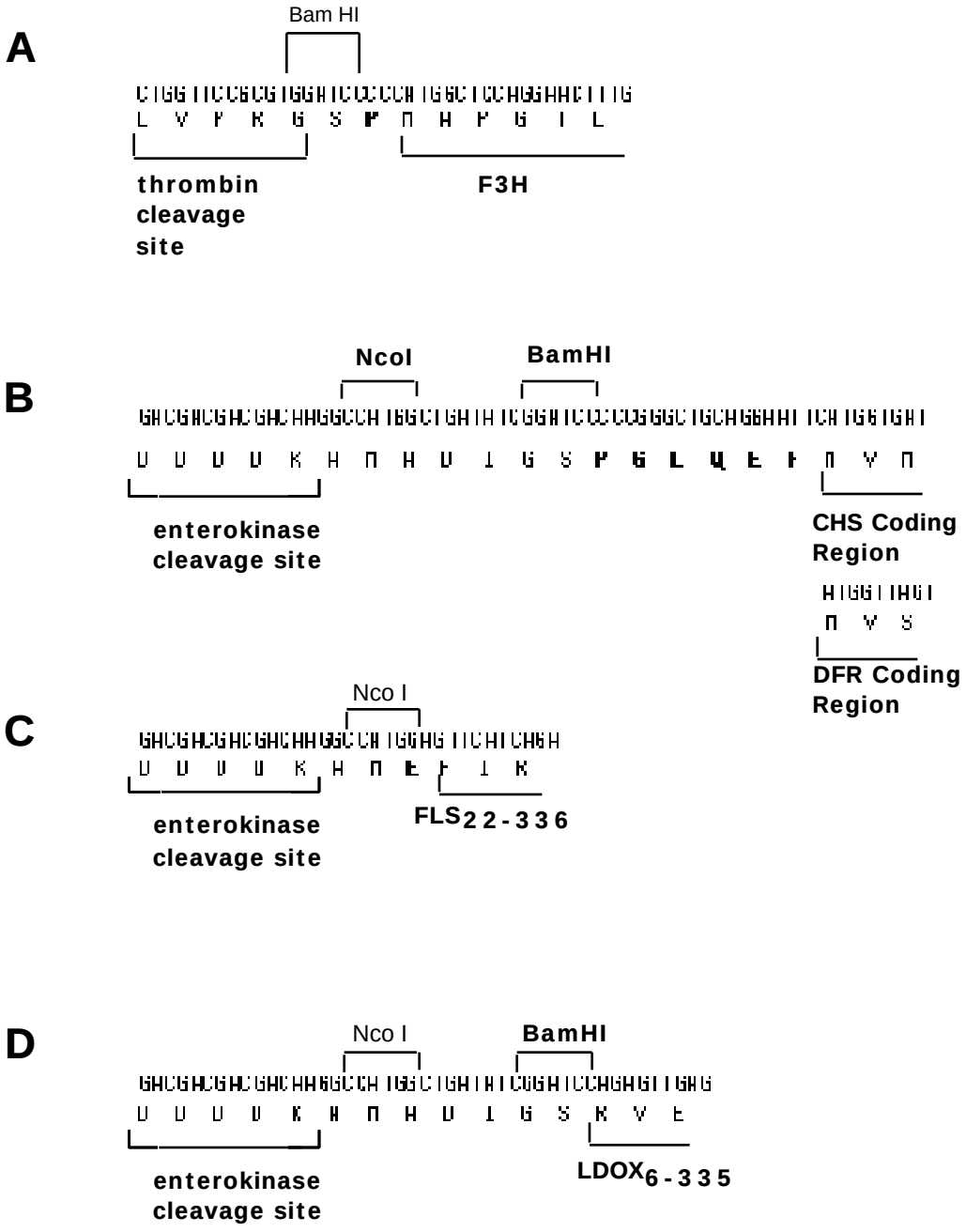


Figure 2. Diagram showing insertion sites of Arabidopsis cDNAs into expression vectors.
A. The F3H coding region was isolated a *BamHI-EcoRI* fragment from pF3H.CR and ligated into the corresponding sites of pGEX-KG.
B. The CHS and DFR coding regions were isolated as *BamHI/XhoI* or *SalI* fragments from pCHS.CR and pBluescript-DFR.CR and ligated into the corresponding sites of pET32a.
C. A cDNA encoding FLS (amino acids 22-336) was synthesized by PCR using pFLS.cDNA as template. The product was digested with *NcoI* and *SalI* fragment ligated into the corresponding sites of pET32a.
D. A cDNA encoding LDOX (amino acids 6-335) was synthesized by RT-PCR, digested with *BamHI,* and ligated into the corresponding site of pET32a.

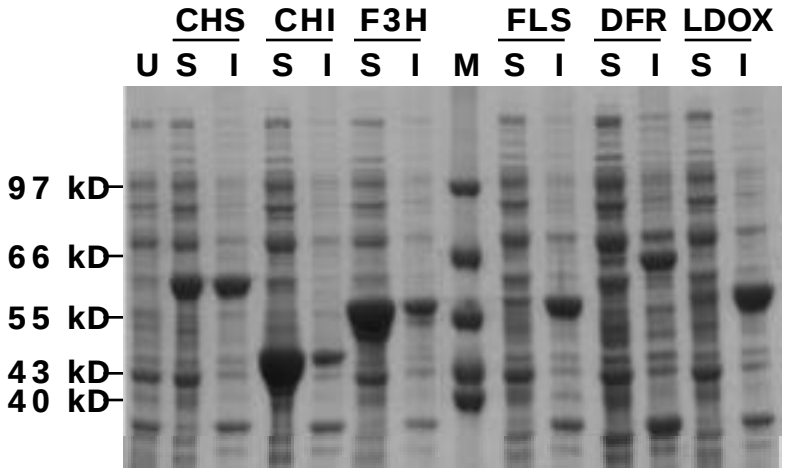


Figure 3. Expression levels and solubility of flavonoid enzymes as fusions to thioredoxin in *E. coli.* Total protein was extracted from a BL21 (DE3) culture before induction of expression by IPTG (U). Soluble (S) and Insoluble (I) protein was extracted from BL21(DE3 cells) expressing each construct 4 hours after addition of IPTG to the culture medium. Proteins were electrophoresed on a 10% SDS polyacryllamide gel,, and the gel was subsequently stained with Coomassid Blue to visuallize proteins.

fraction, and CHS was found in nearly equal amounts in the two fractions (Figure 3). It should be noted that the thioredoxin-CHS fusion was significantly more soluble than the a GST-CHS fusion previously expressed (Cain et al, submitted). While CHI and F3H shown in this figure were fusions to thioredoxin, similar results were obtained when these proteins were fused to GST. Each recombinant protein was purified from the appropriate soluble fraction by affinity chromatography and used to immunize chickens.

Immunoblot Analysis of Developing Arabidopsis Seedlings

While the expression of *CHS, CHI, and DFR* in developing Arabidopsis seedlings has been studied previously by RNA blot analysis, very little was known about the expression patterns at the protein level (Kubasek et al., 1992; Cain et al., submitted). Therefore, we performed immunoblot analysis using the antibodies developed against the recombinant proteins described above to determine the steady-state level of CHS, CHI, F3H, FLS, and LDOX protein present at days two through nine of seedling development. In these immunoblots, the purified anti-CHS IgY antibodies react with a single protein of the appropriate molecular weight, and these proteins were not detected when extracts were reacted with preimmune serum or the secondary antibody alone (data not shown). In contrast, the purified anti-CHI and anti-LDOX IgY antibodies reacted with one or more other proteins (data not shown; Cain et al, submitted). However, it is simple to determine which of the cross-reacting proteins correspond to CHI and LDOX, since Arabidopsis mutants affected in these two proteins (*tt5* and *ttg* respectively*)* lack the corresponding proteins of the appropriate molecular weight (Shirley *et al*., 1992; Pelletier et al., 1997) (described below). Similarly, the anti-F3H IgY and anti-FLS IgY reacted with two proteins in these extracts. The lower molecular weight bans probably represent relatively stable degradation products of these two proteins, since neither of these proteins was detected when extracts were reacted with preimmune serum or a secondary antibody alone and the migration position of the higher molecular weight protein was consistent with the predicted molecular weights of F3H and FLS1 . Furthermore, upon preparation of crude extracts, F3H from petunia has been reported to degrade into a relatively stable, 38 kD product (Britsch et al, 1992). High levels of CHS and CHI protein were present during the first days of seedling development (Figure 4). The steady-state levels of these proteins decreased significantly at day 5, and these proteins were barely detectable by day 6. In contrast, F3H and FLS proteins were readily detectable through day 6, and small amounts of protein were present through day 9 (Figure 4). Finally, LDOX protein seemed to accumulate slightly later than the other enzymes (by day 3), and persisted at significant levels through day 8 or 9 of development. Immunoblots were performed using serum or purified IgY from chickens immunized with recombinant DFR, but no proteins were detected in the plant extracts (data not shown).

Immunoblot Analysis of Mutants Altered in Enzymes of the Flavonoid, Tryptophan, and Ferulic Acid Pathways

To determine whether any mutants are affected in the steady-state levels of CHS, CHI, F3H, FLS1, or LDOX protein, immunoblot analysis was performed on 13 *transparent testa* mutants, *trp1-100, trp2-1, atr1d, atr2d,* and *fah1* mutant seedlings. This analysis was performed using protein extracts from 4-day-old seedlings, as all flavonoid enzymes tested were expressed at high levels at this stage of development (Figure 4). The Landsberg ecotype was used as the wild-type control for all of the mutants except for 2YY6 and *trp2-1,* which were isolated from a Columbia parental background (Last *et al*., 1991; Shirley *et al*., 1995). As described previously, no CHS protein was detected in the *tt4* (2YY6) mutant, which is defective in the Arabidopsis *CHS* gene and no CHI protein was detected in the *tt5* mutant, which is defective in the *CHI* gene

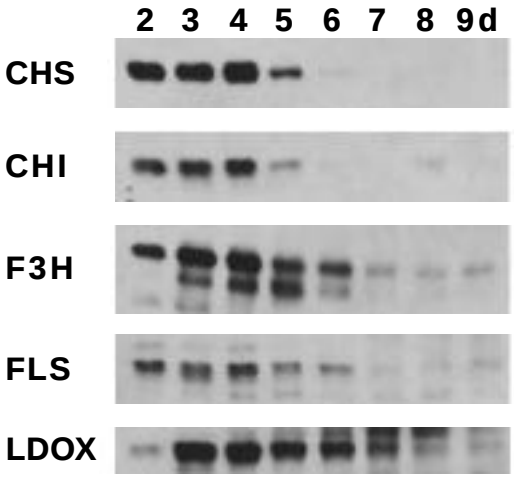


Figure 4. Analysis of flavonoid enzyme levels in developing Arabidopsis seedlings. Protein was extracted from wild-type (La) seedlings grown in continuous white light that were harvested at days two through nine of development. Immunoblot analysis was carried out using antibodies against CHS, CHI, F3H, FLS, or LDOX.

(Burbulis *et al.*, 1996; Cain et al, submitted) (Figure 5). Likewise, no LDOX protein was detected in the *ttg* mutant, which was previously shown to lack *LDOX* mRNA (Pelletier, 1997). However, we were suprised to find that no F3H protein could be detected in the *tt6* mutant. This result suggests that the *TT6* locus encodes the F3H enzyme or a factor regulating its expression, as this is the only protein that was significantly reduced in this mutant. The lower molecular weight band detected when extracts were reacted with anti-F3H IgY probably represents a degradation product, as neither of these protein species are detected in *tt6* and F3H is encoded by a single gene in Arabidopsis (Pelletier and Shirley, 1996). Several of the *tt* mutants, including *tt3, tt4, tt5,* and *ttg* appeared to have moderately elevated levels of several flavonoid enzymes, as did the *trp2-1* mutant. The three other tryptophan mutants and *fah1* did not significantly alter the steady state levels of flavonoid enzymes. In addition, in *tt11*, a newly-identified locus required for flavonoid biosynthesis in Arabidopsis that has not yet been characterized in detail (M. Koornneef, personal communication), F3H, FLS, and LDOX levels were significantly reduced relative to wild type. The differences uncovered here are unlikely to be the result of unequal loading of samples, since each experiment was repeated at least twice and a cross-reacting protein detected with the LDOX IgY antibodies was constant in all samples (data not shown).

While the immunoblot analysis of the mutants described above suggested that defects in a particular flavonoid enzyme or regulatory protein could result in increases in the steady-state levels of the other, wild-type proteins in the pathway, we reasoned that these differences might have been due to slight differences in the developmental states of the mutants, and not the direct result of the genetic lesions. Therefore, the immunoblot analysis was repeated using protein extracts isolated from *tt3, tt4, tt5, tt6, tt8,* and *ttg* seedlings harvested at days four through six of development. These mutants were chosen for further analysis primarily because the molecular nature of the defect is precisely known (Figure 1), making it possible to examine the effects that the loss of a particular enzyme has on the other proteins in the pathway. This experiment was repeated twice, and several differences in the steady-state levels of flavonoid enzymes between the mutants and wild-type seedlings grown under identical conditions were consistently observed. Three mutants (*tt3, tt5,* and *ttg*) were found to accumulate substantially higher steady-state levels of all flavonoid enzymes examined, except for those proteins previously shown to be defective in each mutant (Figure 6). While the results obtained with *tt4, tt6,* and *tt8* were not as striking, moderate differences were observed. For instance, *tt4* consistently accumulated moderately higher levels of F3H and LDOX proteins, but did not seem to affect CHI or FLS protein levels. Similarly, CHS and CHI appeared to be at moderately higher levels in *tt6* at days 4 and 5, but had decreased to wild-type levels by day 6. The results obtained with *tt8* were less clear, although these seedlings consistently had lower levels of all enzymes examined by day 6 of development.

Further Analysis of F3H in *tt6*

The results of the immunoblot analysis described above suggested that *TT6* was the *F3H* locus or encoded a regulator specifically required for F3H protein accumulation. A series of experiments was therefore carried out to distinguish between these possibilities. First, blot analysis was performed on RNA extracted from *tt6* seedlings. The amount of *F3H* mRNA was found to be significantly reduced in the *tt6* mutant (although of wild-type size) while *FLS* mRNA levels were similar to that of wild-type seedlings (Figure 7). This phenotype is similar to that of *tt4* (2YY6), which has reduced levels of *CHS* mRNA and no detectable protein (Burbulis *et al.*, 1996). The 2YY6 mutant contains a G to A transition at the 3' end of the intron that results in an improperly spliced *CHS* mRNA and a severely truncated protein. To determine whether any mutations were present in the *F3H* gene of *tt6*, the coding region of *F3H* was amplified from this mutant by PCR using *pfu* polymerase, subcloned into pBluescript, and sequenced. Eighteen point mutations were found in *F3H* from *tt6* (Figure 8). These mutations were not randomly distributed, as most occurred in three relatively short segments of the gene (6 substitutions in a 27 bp segment,

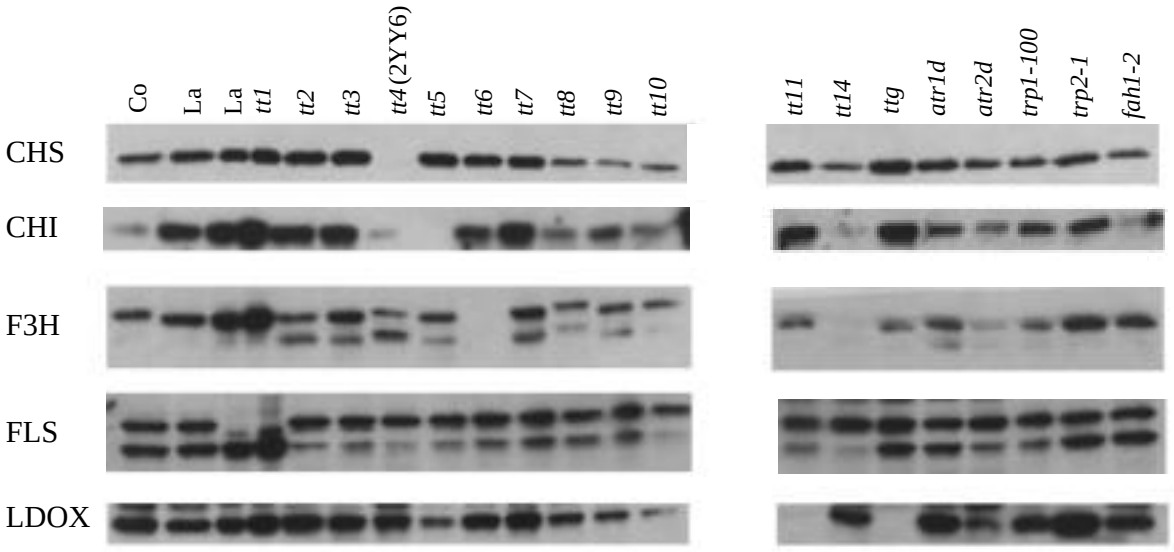


Figure 5. Analysis of flavonoid enzyme levels in Arabidopsis mutants. Immunoblot analysis was performed using 48 µg of protein from 4-d-old seedlings grown in continuous white light. The genotype of the seedling is indicated at the top of each lane. The Landsberg (La) sample shown on the left contained protein isolated from the same experiment as all mutants shown except *tt1*, while the La sample in the third lane served as a control for this sample. A Columbia (Co) sample was included as a control for *tt4* and *trp2-1*.

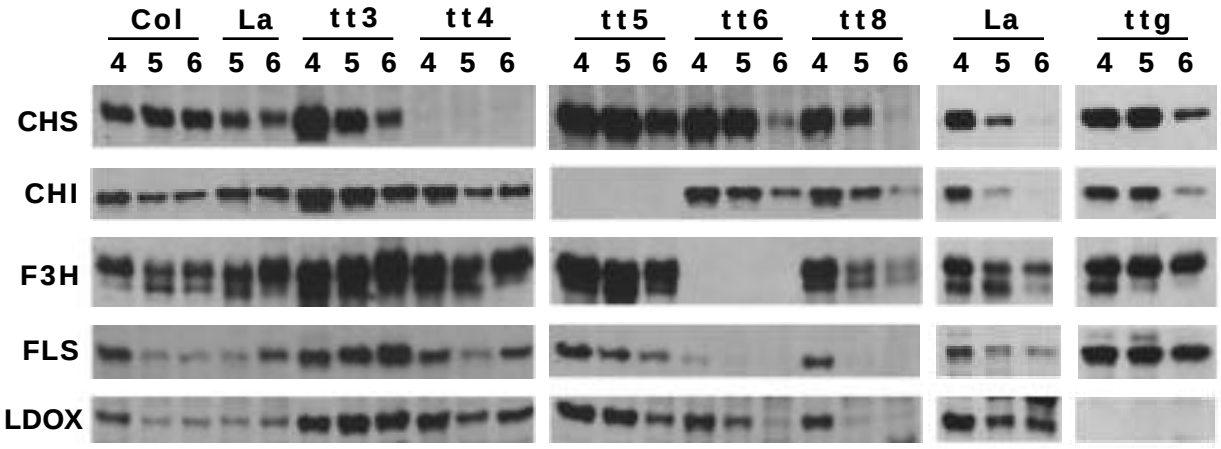


Figure 6. Analysis of flavonoid enzyme levels in selected *tt* mutants at different developmental stages. Immunoblot analysis was performed using 48 µg protein taken from 4-, 5-, and 6-d-old seedlings grown in continuous white light. The genotype of each sample is indicated at the top of each lane. The Landsberg (La) sample shown on the right served as the control for the *ttg* samples.

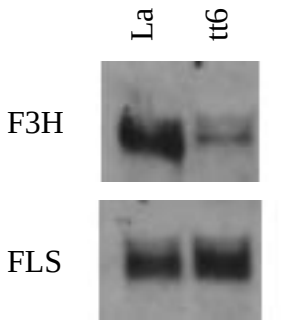


Figure 7. Accumulation of F3H and FLS mRNA in *tt6*. RNA blot analysis was performed using 10 µg of total RNA isolated from 3-d-old wild-type or *tt6* seedlings grown in continuous white light.

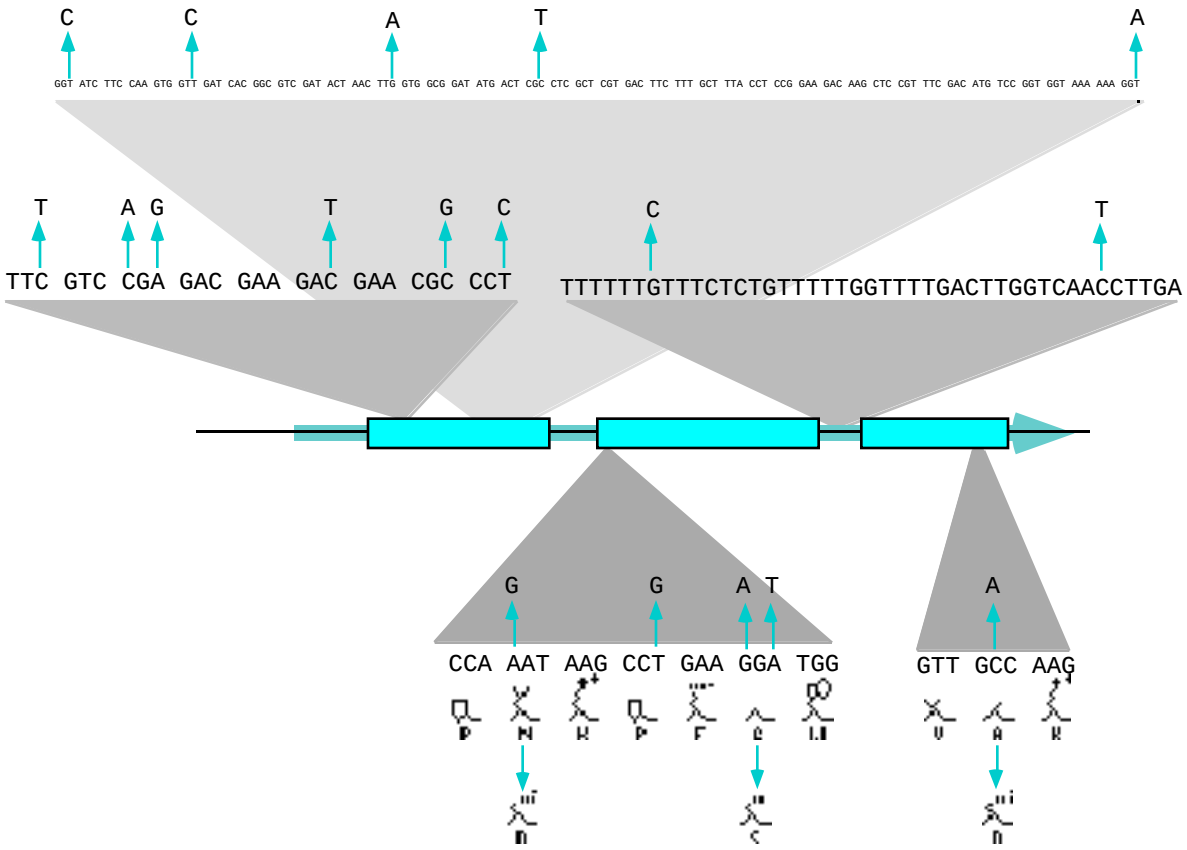


Figure 8. Schematic of the structure of F3H in *tt6*. The three exons of F3H are shown as shaded boxes. The mutational "hot spots" within the F3H coding region are shown as triangles. The DNA sequence substitutions found in F3H from *tt6* are shown above the wild-type sequence. The predicted changes in amino acid sequence are shown below the DNA sequence.

5 substitutions in a 132 bp segment, and 4 subsitutions in a 18 bp segment). However, only three of the 18 mutations were predicted to result in amino acid substitutions in the protein (Asn_{148} to Asp, Gly_{152} to Ser, Ala_{350} to Asp). Importantly, the three substitutions predicted to result in amino acid changes were confirmed by sequencing a second, independently-generated PCR product. These findings strongly suggest that *tt6* is the *F3H* locus. While the number of mutations found in this gene were surprising, EMS-induced mutations have been reported to occur in clusters in other studies (Okagaki et al., 1991; Caetano-Anolles et al., 1993).

trp2-1 Seedlings Accumulate Higher Levels of Flavonoid Enzymes

Of all the tryptophan mutants included in the preliminary immunoblot analysis (Figure 5), only *trp2-1* seemed to be affected in the levels of flavonoid enzymes as compared to wild-type seedlings. We reasoned that these differences may have been due to developmental differences between *trp2-1* and Columbia, especially since under the light conditions used in these experiments *trp2-1* seedlings are auxotrophic for tryptophan. It seemed likely that the observed increase in the levels of flavonoid enzymes was due to amino acid starvation, since expression of flavonoid genes is inducible by a variety of stresses (Feinbaum and Ausubel, 1988; Last *et al.*, 1991; Kubasek *et al.*, 1992; Shirley and Goodman, 1993; Leyva et al., 1995). Alternatively, it seemed possible that the *trp2-1* mutation resulted in the accumulation of a compound that was directly or indirectly responsible for the increased levels of flavonoid enzymes. Since the *trp2-1* mutant is reported to accumulate high levels of indole relative to wild-type plants, we attempted to mimic this effect by growing wild-type seedlings on ms-sucrose plates supplemented with indole (Normanly et al., 1993). Immunoblot analysis was performed on protein extracts taken from 4- or 5- day-old Columbia and *trp2-1* seedlings grown on ms-sucrose plates in the presence or absence of tryptophan or indole. The steady state levels of CHS, CHI, F3H, and LDOX were substantially higher in all cases in the *trp2-1* mutant, regardless of the day seedlings were harvested or whether they were grown in the presence or absence of tryptophan (Figure 9). The results obtained using FLS IgY were less clear, as the lower molecular weight protein detected with this antibody appeared to be slightly more abundant in *trp2-1*, but the cross-reacting protein of higher molecular weight was absent in *trp2-1* by day 5 of development. In contrast, the increased levels of flavonoid enzymes observed in the *trp2-1* mutant was not consistently observed in Columbia seedlings grown on indole-containing media (Figure 9).

While performing immunoblot analysis on *trp2-1*, we noticed that the mutant seedlings appeared to accumulate higher levels of pigments relative to wild-type controls. Therefore, in an effort to correlate the increased levels of several flavonoid enzymes in this mutant with increased levels of end products, the anthocyanin content of *trp2-1*seedlings grown on ms-sucrose media with or without added indole or tryptophan was measured and compared to wild type. *trp2-1* seedlings consistently accumulated approximately twice the level of anthocyanins as compared to wild type seedlings, regardless of the composition of the growth media (Figure 10). Interestingly, a small decrease in anthocyanin levels were observed in wild-type seedlings grown on tryptophan-supplemented media and a substantial decrease was observed on those grown on indole-supplemented (Figure 10).

Discussion

Antibody development and Immunoblot Analysis of Developing Seedlings

We have expressed six different flavonoid enzymes from Arabidopsis (CHS, CHI, F3H, FLS, DFR, and LDOX) as fusions to either GST or thioredoxin and used the purified recombinant proteins to generate antibodies in chicken (Figures 2 and 3; Cain et al, submitted). The IgY fractions were then used to characterize the expression of these proteins in Arabidopsis seedlings. While the cross-reactivity of the antibodies varied, it was possible to determine what signal

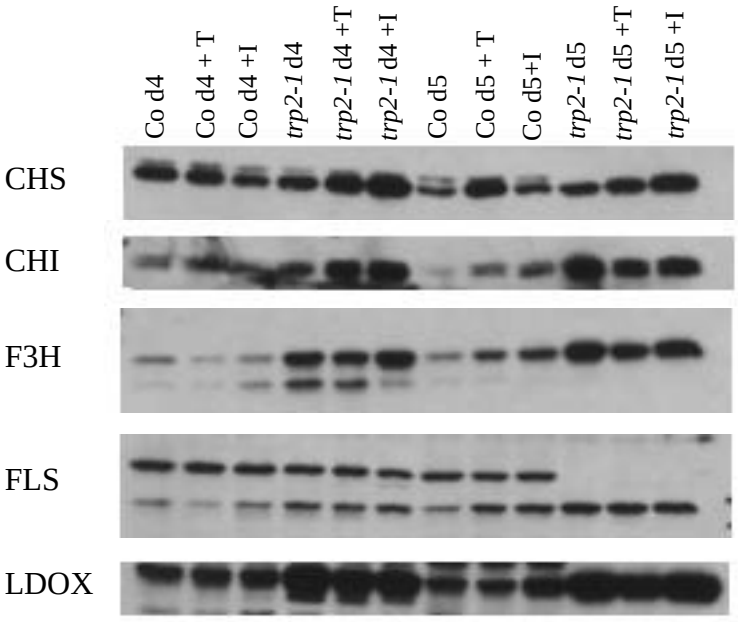


Figure 9. Effect of tryptophan and indole on flavonoid enzyme levels in wild-type and *trp2-1* seedlings. Immunoblot analysis was performed using 48 µg protein from 4- and 5-d-old Columbia (Co) and *trp2-1* seedlings grown on ms-sucrose plates with 50 µM tryptophan (T) or 50 µM indole (I).

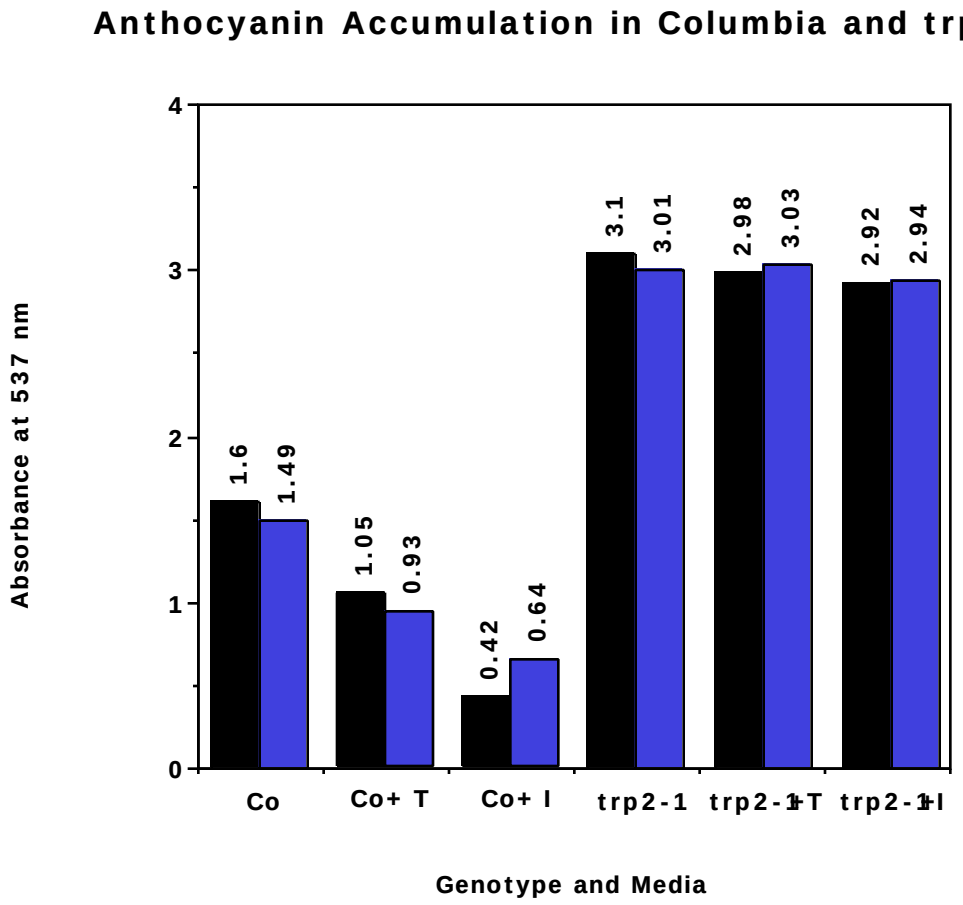


Figure 10. Effect of tryptophan and indole on anthocyanin levels in wild-type and *trp2-1* seedlings. Anthocyanins were extracted from equivalent numbers of 4 or 5-d-old Columbia (Co) or *trp2-1* seedlings grown in continuous white light using acidic methanol (1% v/v). Two independent sets of seedlings of each genotype were grown on unsupplemented media and media supplemented with 50 µM tryptophan (T) or 50 µM indole.

corresponded to the appropriate protein by using mutants deficient in CHS, CHI, F3H, and LDOX. In the case of FLS IgY, two cross-reacting proteins were detected that were not observed when either preimmune serum or a secondary antibody alone were used. Therefore, these two proteins probably represent different FLS proteins since this enzyme appears to be encoded by a small gene family in Arabidopsis (Pelletier et al., 1997)

Immunoblot analysis of developing seedlings revealed that CHS and CHI proteins accumulated to the highest levels early in seedling development and were the first two proteins to diminish in abundance as seedlings aged (Figure 4). Similarly, F3H and FLS protein levels were high early in young seedlings, yet persisted longer as than CHS and CHI. In contrast, the LDOX protein began to accumulate to high levels approximately 1 d after the first four enzymes, and was still present through day 9 of development. We have previously shown that CHS, CHI, F3H, and FLS represent "early" genes in Arabidopsis seedlings, while LDOX represents a "late" gene, and the accumulation of these proteins over the course of seedling development correlates well with these studies (Pelletier and Shirley, 1996; Pelletier et al., 1997). This regulation may reflect a need for different classes of flavonoids at various stages of seedling development.

F3H is Defective in *tt6*

When the location of the Arabidopsis *F3H* locus was determined, it was found to be present on chromosome 3 near the position of *tt6*, suggesting that this mutant may be defective in the *F3H* gene (Pelletier and Shirley, 1996). It was subsequently suggested that the correct position of *tt6* was on chromosome 5 (J. Campanella and C. Town, personal communication). This, combined with the fact that thin-layer chromatography analysis of extracts from *tt6* had shown that flavonols were not present in this mutant (although pigments are present) led us to believe that *TT6* was not the *F3H* locus (Shirley *et al.*, 1995). However, the data in this report show that *tt6* is defective in the *F3H* gene, and we are currently analyzing the F2 progeny of a cross between *tt6* and *cer7*/*gl1* lines or *lu* lines to independently establish whether this locus is on chromosome 3 or 5.

Several lines of evidence strongly suggest that *TT6* is the *F3H* locus. First, immunoblot analysis of the *tt6* mutant revealed that no F3H protein was present, while the levels of all other flavonoid enzymes were not significantly affected (Figures 5 and 7). Second, blot analysis using RNA isolated from this mutant showed that F3H mRNA was significantly reduced compared to wild-type seedlings, while FLS mRNA expression was unaffected (Figure 7). Finally, amplification of the F3H locus from *tt6* and subsequent sequencing of the gene revealed that at least 18 single base pair substitutions are found in this gene as compared to the wild-type Landsberg sequence (Pelletier and Shirley, 1995). These substitutions are unlikely to be due to errors which occurred during PCR amplification, as the sequence obtained from an independently-derived PCR product (to confirm the mutations predicted to result in amino acid changes) was identical to the initial sequence. Furthermore, these mutations are not due to ecotype differences since *tt6* was isolated in a Landsberg background and the sequence of *FLS* from *tt6* was identical to wild type (Koornneef, personal communication; Pelletier and Shirley, 1996). We were surprised by the large number of substitutions found in this gene, as well as by the fact that only three of these mutations resulted in amino acid changes, since most other studies of EMS -induced mutants report one or a few changes in any given locus. However, at least two other studies suggest that mutational "hot-spots" may exist for EMS-induced damage in plant genomes (Okagaki *et al.*, 1991; Caetano-Anolles *et al.*, 1993). The three mutations of consequence in *F3H* from *tt6* are predicted to result in the substitution of Asp for Asn_{148}, Ser for Gly_{152}, and Asp for Ala_{350}. Two of these three substitutions introduce residues with altered charge and therefore could result in improper folding, leading to the degradation of this protein, or to loss of reactivity with the polyclonal anti-F3H IgY. It is also possible that the F3H mRNA produced in this mutant is degraded or not translated, since the RNA level is reduced in *tt6*, and the sequence of the 5'

untranslated region of *F3H* from *tt6* was not determined. Interestingly, although no F3H protein is detected in this mutant and the enzyme appears to be encoded by a single gene in Arabidopsis, colored compounds are still present in *tt6* seedlings (data not shown). However, at least one plant species (sorghum) is reported to accumulate 3-deoxyanthocyanidins, a class of colored flavonoids with antimicrobial activity (Hipskind et al., 1996). Since a functional F3H protein is not required to synthesize these compounds, it is possible that 3-deoxyanthocyanidins are the colored pigments observed in *tt6* seedlings. HPLC analysis will be done in collaboration with Ralph Nicholson (Purdue University) to determine whether this is the case.

Effects of Mutations in Primary and Secondary Aromatic Amino Acid Metabolism on Flavonoid Enzyme Accumulation

The antibodies described above were also used to determine the steady-state level of flavonoid enzymes present in 13 different *tt* mutants, 4 tryptophan mutants, and the ferrulate pathway mutant *fah1*(Figure 5). Immunoblot analysis of these mutants yielded several interesting findings. For instance, several of the *tt* mutants and the *trp2-1* mutant seedlings appeared to accumulate substantially higher levels of flavonoid enzymes as compared to wild type. In addition, *tt11* seemed to have significantly reduced levels of F3H, FLS, and LDOX.

To determine whether the differences in several of the mutants were due to differences in the developmental state of the seedlings, the levels of flavonoid enzymes were determined in *tt3*, *tt4* (2YY6), *tt5*, *tt6*, *tt8*, and *ttg* seedlings harvested at days 4-6 of germination (Figure 7). This experiment revealed that *tt3*, *tt5*, and *ttg* seedlings consistently accumulated higher levels of all flavonoid enzymes examined (other than the one specifically affected in each mutant) compared to wild type, while small differences in the timing or amount of flavonoid enzyme accumulation were observed in *tt4*, *tt6*, and *tt8* seedlings.

These data raise some intriguing new questions regarding the regulation of the flavonoid biosynthetic pathway. For instance, what is the mechanism responsible for increased steady-state protein levels in *tt3*, *tt5*, and *ttg*? Does it involve increases in the transcription of these genes, more efficient translation of the mRNA present, or increases in protein stability? Previous studies showed that protoplasts expressing the reporter gene chloramphenicol acetyl transferase (CAT) under the control of a CHS promoter had approximately 5-fold CAT activity when the culture medium contained the phenylpropanoid intermediate p-coumaric acid (Figure 1) as compared to control protoplasts (Loake, 1991). However, based on earlier studies, it seems unlikely that the mechanism responsible for the differences in protein levels noted in this study involves transcriptional regulation, since no such increases in mRNA levels were observed in *tt3*, *tt5*, or *ttg* (Shirley *et al.*, 1995). However, this possibility can not be ruled out completely as this study only examined seedlings at day 3 of development. Why do mutations in *TT3* (CHI), *TT5* (DFR), and *TTG* (a probable regulatory locus affecting *DFR* and *LDOX* mRNA expression) lead to increases in all protein levels while mutations in *TT4* (CHS), *TT6* (F3H), and *TT8*(a probable regulatory mutant with reduced DFR mRNA levels) have less of an effect? The *tt5* mutant is expected to accumulate chalcones, while *tt3* and *ttg* are expected to accumulate dihydroflavonols or flavonols, since these mutants are unlikely to contain any functional CHI or DFR proteins, respectively (Figure 1). Thus, it is possible that accumulation of specific flavonoid intermediates may lead to an increase in the steady state levels of flavonoid enzymes. Interestingly, there is evidence to suggest that the stability of enzymes can be controlled by pathway intermediates. A study in yeast on the stability of HMG-Co reductase (HMGR), an enzyme involved in isoprenoid and sterol biosynthesis, strongly suggested that the half life of this enzyme is controlled at least in part by the amount of a as-yet-undetermined pathway intermediate present (Hampton and Rine, 1994). In this study, blocking the flux through early steps in the pathway (but not later enzymatic steps) was

shown to lead to an increase in the half-life of HMGR. It is therefore possible that changes in the flux through the flavonoid pathway may result in differences in the stability of flavonoid enzymes, resulting in the altered steady-state levels observed in these experiments. Further experiments must be carried out to determine whether this is a mechanism operating in the flavonoid pathway. While it is unknown why mutations in *TT4*, *TT6*, and *TT8* do not have as significant an effect as do the other three mutants examined, the fact that mutations in one of two sequential enzymes in this pathway can have a very different physiological effect has been demonstrated previously, since *tt5* had significantly lower levels of sinipate esters, while *tt4* mutants had slightly elevated levels of these compounds as compared to wild type (Li *et al.*, 1993).

Because in our initial experiments *trp2-1* seemed to have higher steady-state levels of all flavonoid enzymes examined, this mutant was also studied in more detail. An increase in the steady-state levels of CHS, CHI, F3H, and LDOX proteins was observed in the *trp2-1* mutant, even when it was grown on tryptophan-containing media, indicating that this increase was not merely the result of amino acid starvation. Interestingly, the amount of FLS1 present in *trp2-1* was only slightly elevated, while a cross-reacting band of higher molecular weight was actually undetectable in five-day-old seedlings. Consistent with the immunoblot analysis, *trp2-1* seedlings accumulated approximately twice the level of anthocyanins than the wild-type control grown on unsupplemented media, regardless of the media they were grown on. In contrast, wild-type seedlings grown on indole or tryptophan-supplemented media accumulated significantly less anthocyanin than those grown on unsupplemented media.

The results of the experiments using *trp2* seedlings are intriguing for a number of reasons. As mentioned in the introduction, expression of a tryptophan decarboxylase enzyme in potato leads to a decrease in the levels of phenylalanine and other phenolics (Yao *et al.*, 1995). Presumably, this is because tryptophan, which is thought to inhibit its own synthesis, is unable to accumulate in the plants, leading to more substrate being used by anthranilate synthase, and less being converted to phenylalanine (Figure 1) (Bentley, 1990). In light of this study and the fact that *trp2-1* seeds are pale brown relative to wild type, we were suprised to find an increase in flavonoid enzymes and end products in developing seedlings of this mutant (Last *et al.*, 1991). Since this mutant is reported to accumulate indole at levels approximately 31 times that of wild-type plants (Normanly *et al.*, 1993), it is possible that indole or an indole-derived compound is responsible for the increase in flavonoid enzymes observed in these experiments (Normanly *et al.*, 1993). Interestingly, 13 different alleles of *trp2* (including *trp2-1*) were previously shown to contain 2 to 6-fold the steady-state levels of ASA protein in wild type, although the mechanism underlying this increase remains unknown. Wild-type seedlings grown in the presence of 50 μM indole did not, however, accumulate increased levels of flavonoid enzymes (data not shown), and in fact contained significantly less anthocyanins than those grown on unsupplemented media (Figure 10). However, it is impossible to determine whether feeding indole to wild-type seedlings faithfully mimics the physiological state of *trp2-1* with regard to *in vivo* indole concentration. Therefore, although we have shown that *trp2-1* accumulates significantly higher levels of several flavonoid enzymes as well as anthocyanins, the mechanism responsible for this phenomenon is yet to be elucidated.

Materials and Methods

Cloning of Arabidopsis Flavonoid Genes into Bacterial Expression Vectors

The F3H coding sequence was amplified by RT-PCR as described previously (Shirley and Hwang, 1995), using 10 µg total RNA extracted from 3-d-old Landsberg seedlings grown in continuous white light. This reaction was performed using F3HCRS1 (5'ATGGCTCCAGGAACTTT3') and F3HCRA1 (5'CTAAGCGAAGATTTGGT3') (synthesized by Gibco-BRL, Grand Island, NY) and an annealing temperature of 49 °C. The product of this reaction, a cDNA containing the entire coding region of F3H, was ligated into the *SmaI* site of pBluescript to create pF3H.CR. The resulting plasmid was digested with *BamHI* and *EcoRI*, and the F3H coding region ligated into expression vector pGEX-KG (Pharmacia Piscataway NJ) to create pGEX-F3H.CR (Figure 1).

The F3H coding region was excised from pF3H.CR with *NcoI* and *EcoRI* to create pET32aF3H.CR. Similarly, a portion of the Arabidopsis LDOX gene (encoding amino acids 6-335) was synthesized by RT-PCR using total RNA as described above. In this reaction, the LDOXs1 (5'CGCGGATCCAGAGTTGAGAGTCTAGC 3') and LDOXa1 (5' CGCGGATCCTTAAGCAAAAGTCCGTGGAG 3') primers were used at an annealing temperature of 51° C. A DNA fragment of the expected size was purified from a 0.8% agarose/TBE gel using a Qia-Quick Gel Extraction Kit (Qiagen, Chatsworth, CA), digested with *BamHI*, and ligated into pET32a to create pET32a-$LDOX_{6-335}$. A portion of an Arabidopsis FLS gene corresponding to amino acids 22-336 was amplified by PCR from pFLS.cDNA (Pelletier et al., 1997) using FSSEN1 (5'CATGCCATGGAGTTCATCAGATCAGAG3') and T7 primers and an annealing temperature of 48° C. This product was digested with *NcoI* and *SalI* and ligated into pET32a to create pET32a-FLS_{22-336}. DNA fragments encoding full-length Arabidopsis CHS and DFR proteins were excised from pCHS.CR and pBluescript-DFR.CR using *BamH*I and *Xho*I or *Sal*I, respectively (Pelletier and Shirley, 1996). After gel-purification as described above, these fragments were ligated into pET32a to create pET32a-CHS.CR and pET32a-DFR.CR.

Expression and Purification of Recombinant Flavonoid Enzymes from *E. coli*

BL21DE3 cells containing pLyseS (Novagen, Madison WI) and one of the recombinant pET32a constructs described above were grown in LB broth (500 mL) containing 100 µg/mL ampicillin and 34 µg/mL chloramphenicol at 37° C with vigorous shaking (250 rpm) to an $O.D._{550}$ of 0.5-0.7. Expression of the recombinant protein was then induced by adding isopropyl thio-β-D-galactoside (IPTG) to 250 µM, and the cultures were transferred to room temperature and grown for an additional 4 h. The cells were then centrifuged at 3300g at 4° C for 10 minutes, resuspended in ice-cold lysis buffer (20 mM Tris, 100 mM NaCl) using 2/5 of the original culture volume and spun as before. The cell pellets were then placed at -70° C for at least one hour.

The frozen cell pellets containing recombinant proteins were crushed using a glass rod and resuspended in buffer (3.5 mL per g cell pellet). Upon lysis, PMSF (phenylmethylsulfonyl fluoride, Sigma, St. Louis MO) was added to 1 mM, DNaseI (Boehringer Manheim, Indianapolis IN) was added to 40 µg/mL, $MgCl_2$ was added to 10 mM, and the cells were incubated on ice until the solution was no longer viscous (30-60 minutes). The lysate was then spun in a Beckman TLA 100 rotor at 20,000 rpm for 10 minutes at 4° C to pellet insoluble proteins, and the supernatant was spun as before but at 40,000 rpm. After this high-speed spin, the supernatant contained the

soluble protein. The insoluble proteins were then solubilized from the pellet by boiling for 10 minutes in Laemlli buffer as described below for SDS-PAGE analysis or were solubilized with 8M urea (DFR) for purification. Total soluble protein (13 mL) containing recombinant CHS, FLS or LDOX was incubated with 1mL of Talon metal affinity resin (Clontech, Palo Alto, CA) for 1 h at room temperature. The resin was then washed four times with 10 mL lysis buffer for 10 minutes at room temperature, and the recombinant protein was eluted from the resin in 1 mL fractions using lysis buffer which also contained 100 mM EDTA and/or 200 mM imidazole. All other aspects of the purification protocol were carried out according to the manufacturer's recommendation. Expression and purification of the GST-F3H fusions were carried out as described previously (Cain et al., submitted).

Antibody Production

Approximately 1 mg of thioredoxin-CHS and 200 µg each of purified GST-F3H, thioredoxin-FLS and thioredoxin-LDOX were sent to Cocalico Biologicals (Reamstown PA) for the development of polyclonal antibodies in chicken. IgY was purified from egg yolks from immunized chickens using a commercially available kit (Eggstract kit) (Promega, Madison WI) according to the manufacturer's recommendations. Purified IgY was resuspended in PBST (0.5% Tween 20) containing 0.02% sodium azide and aliquots were stored at -80°.

Plant Growth, Protein and Anthocyanin Extraction, and Immunoblot Analysis

Arabidopsis seedlings were grown on MS-sucrose plates in continuous white light (120-150 µEinsteins) as described previously (Pelletier and Shirley, 1996). Proteins were extracted by grinding seedlings to a fine powder with a mortar and pestle in liquid nitrogen, and adding 500 µL of Laemli buffer (Laemlli, 1970) per gram of plant tissue. Phenylmethylsulfonylfluoride (PMSF) was then added to 1 mM, samples were placed on ice until all seedlings were harvested and the slurry was then boiled for 10 min. The extract was centrifuged at room temperature at 17000g two times to pellet insoluble material. Proteins were quantified using the Bio-Rad microassay procedure, using bovine serum albumin (Sigma) to generate a standard curve. All samples were then diluted to either 3 or 4 µg/µL in Laemlli buffer and 12 µL of extract was loaded on a 10% SDS-PAGE for immunoblot analysis.

Gel electrophoresis and transfer of proteins was carried out using the Miniprotean II system (BioRad, Hercules CA). After electrophoresis for 1 to 2 h at 100 V, proteins were transferred to 0.45 micron nitrocellulose (BioRad) at 100 V for 25 to 40 minutes. The resulting nitrocellulose filters were incubated in blocking solution (5% carnation instant dry milk in PBST) for 30 to 60 min at room temperature with shaking (100 rpm). The filters were then incubated with primary antibodies in fresh blocking solution for 1 h with agitation (100 rpms). Working dilutions for each antibody were as follows: anti-CHS IgY (1:200); anti-CHI IgY (1:100); anti-F3H IgY (1:200); anti-FLS IgY (1:75); anti-LDOX IgY (1:75). The filters were washed in large volumes of PBST 3X for 5 to 15 minutes and then incubated with goat anti-chicken IgY-horseradish peroxidase conjugate (Jackson Immunologicals) diluted 1:5000 in blocking solution for 1 h. Filters were washed in with PBST as described above and detection of proteins was carried out using a chemiluminescent substrate (ECL, Amersham) according to the manufacturer's protocol.

RNA extraction and blot analysis were carried out as described previously (Pelletier et al, 1997). Anthocyanins were extracted and quantified using a modification of the method of Rabino and Mancinelli by crushing seedlings in 3 µl of acidic methanol (1% v/v) per mg tissue and incubating for 1 h at room temperature (1986). The debris were removed by centrifugation for 5 min at 2,000 rpms, and the supernatant extracted with 2 volumes of chloroform to remove

chlorophyll and its degredation products. The absorption of each solution at 530 nm was then determined spectrophotometrically.

Sequence Analysis of the F3H Gene of *tt6*

Genomic DNA was isolated from *tt6* seedlings using a modification of the method of Dellaporta (1983). Approximately 200 mg of tissue was ground in 500 µL of CTAB buffer (2% w/v CTAB, 100 mM Tris (pH 7.5), 20 mM EDTA, 1.4 M NaCl, and 1% w/v pvp) in a microfuge tube and then incubated for an hour at 65 °C. This solution was then extracted twice with chloroform-isoamyl alcohol (24:1), 0.6 volumes of room temperature isopropanol was added to the aqueous phase, and the DNA was pelleted at 17,000g for 5 minutes. The pellet was then washed with 500 µL of 80% ethanol and 15 µL of 3M NaOAc and incubated on ice for 30 minutes. The DNA was then pelleted, washed twice with 70% ethanol, dried, and resuspended in 50 µL of TE buffer.

The coding region of F3H was amplified by PCR using *pfu* polymerase (Stratagene) and 0.1 µl of *tt6* genomic DNA using F3Htt6s1 (5'CCG GAA TTC TCA GCT ACC ACT CTC TC) and F3Htt6a1 (5' CCG GAA TTC CAA AAC AGA ACC AAC GC) primers. For the first cycle, denaturation was carried out at 94° C for 1 min and an annealing temperature of 50° C was used. For all other cycles denaturation was for 30 sec and the annealing temperature was 54° C, and a total of 45 cycles were run. This PCR product was digested with *EcoRI* and *HindIII* and the resulting fragments ligated into the corresponding sites of pBluescript KS (Stratagene) for sequencing as described previously (Pelletier and Shirley, 1996). An independent reaction using the same primers was sequenced directly as described previously (Pelletier and Shirley, 1996).

Literature Cited

Bell-LeLong D, Cusumano J, Meyer K, and Chapple C (1997). Cinnamate-4-hydroxylase Expression in Arabidopsis. Plant Physiol. **113**: 729-738.

Bentley R (1990). The shikimate pathway - A metabolic tree with many branches. Crit Rev Biochem Mol Biol **25**: 307-384.

Burbulis I, Iacobucci M, and Shirley B (1996). A null mutation in the first enzyme of flavonoid biosynthesis does not affect male fertility in Arabidopsis. Plant Cell **8**: 1013-1025.

Caetano-Anolles G, Bassam BJ, and Gresshoff PM (1993). Enhanced detection of polymorphic DNA by multiple arbitrary amplicon profiling of endonuclease-digested DNA: identification of markers tightly linked to the supernodulation locus in soybean. Moll Gen. Genet. **241**: 57-64.

Chapple CC, Vogt T, Ellis BE, and Somerville CR (1992). An Arabidopsis mutant defective in the general phenylpropanoid pathway. Plant Cell **4**: 1413-24.

Cone KC, Burr FA, and Burr B (1986). Molecular analysis of the maize anthocyanin regulatory locus C1. Proc. Natl. Acad. Sci. USA **83**: 9631-9635.

Deboo G, Albertsen M, and Taylor L (1995). Flavanone 3-hydroxylase transcripts and flavonol accumulation are temporally coordinate in maize anthers. Plant J. **7**: 703-713.

Dellaporta SL, Wood J, and Hicks JB (1983). A plant DNA minipreparation: Version II. Plant Mol. Biol. Reporter **1**: 19-21.

Dooner HK (1983). Coordinate genetic regulation of flavonoid biosynthetic enzymes in maize. Mol. Gen. Genet. **189**: 136-141.

Dooner HK, Robbins TP, and Jorgensen RA (1991). Genetic and developmental control of anthocyanin biosynthesis. Annu. Rev. Genet. **25**: 173-199.

Feinbaum RL, and Ausubel FM (1988). Transcriptional regulation of the *Arabidopsis thaliana* chalcone synthase gene. Mol. Cell. Biol. **8**: 1985-1992.

Guyer D, Patton D, and Ward E (1995). Evidence for cross-pathway regulation of metabolic gene expression in plants. Proc. Natl. Acad. Sci. USA **9295**: 4997-5000.

Hahlbrock K, and Scheel D (1989). Physiology and molecular biology of phenylpropanoid metabolism. Ann. Rev. Plant Physiol. Plant Mol. Biol. **40**: 347-369.

Hampton R, and Rine J (1994). Regulated degradation of HMG-CoA reductase, an integral membrane protein of the endoplasmic reticulum, in yeast. The Journal of Cell Biology **125**: 299-312.

Hipskind J, Goldsbrough P, Urmeev F, and Nicholson R (1996). Synthesis of 3-deoxyanthocyanidin phytoalexins in sorghum does not occur via the same pathway as 3-hydroxylated anthocyanidins and phlobaphenes. Maydica **41**: 155-166.

Jackson D, Roberts K, and Martin C (1992). Temporal and spatial control of expression of anthocyanin biosynthetic genes in developing flowers of Antirrhinum majus. Plant Journal **2**: 425-434.

Koes RE, Quattrocchio F, and Mol JNM (1993). The flavonoid biosynthetic pathway in plants: function and evolution. BioEssays **16**: 123-132.

Kubasek WL, Shirley BW, McKillop A, Goodman HM, Briggs W, and Ausubel FM (1992). Regulation of flavonoid biosynthetic genes in germinating Arabidopsis seedlings. The Plant Cell **4**: 1229-1236.

Laemlli U (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227**: 680-685.

Last RL, Bissinger PH, Mahoney DJ, Radwanski E, and Fink G (1991). Tryptophan mutants in Arabidopsis: the consequences of duplicated tryptophan synthase beta genes. Plant Cell **3**: 345-58.

Leyva A, Jarillo JA, Salinas J, and Martinez-Zapater JM (1995). Low temperature induces the accumulation of phenylalanine ammonia-lyase and chalcone synthase mRNAs of *Arabidopsis thaliana* in a light-dependent manner. Plant Physiol. **108**: 39-46.

Li J, Ou-Lee T, Raba R, Amundson R, and Last R (1993). Arabidopsis flavonoid mutants are hypersensitive to UV-B irradiation. Plant Cell **5**: 171-179.
Loake G, Choudhary A, Harrison M, Mavandad M, Lamb C, and Dixon R (1991). Phenylpropanoid pathway intermediates regulate transient expression of a chalcone synthase gene promoter. Plant Cell **3**: 829-840.
Ludwig SR, Habera LF, Dellaporta SL, and Wessler SR (1989). Lc, a member of the maize R gene family responsible for tissue-specific anthocyanin production, encodes a protein similar to transcriptional activators and contains the myc-homology region. Proc. Natl. Acad. Sci. USA **86**: 7092-7096.
Martin C, Prescott A, Mackay S, Bartlett J, and Vrijlandt E (1991). Control of anthocyanin biosynthesis in flowers of *Antirrhinum majus*. Plant J. **1**: 37-49.
Meldgaard M (1992). Expression of chalcone synthase, dihydroflavonol reductase, and flavanone-3-hydroxylase in mutants of barley deficient in anthocyanin and proanthocyanidin biosynthesis. Theor. Appl. Genet. **83**: 695-706.
Mo Y, Nagel C, and 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.
Normanly J, Cohen JD, and Fink GR (1993). Arabidopsis thaliana auxotrophs reveal a tryptophan-independent biosynthetic pathway for indole-3-acetic acid. Proc. Natl. Acad. Sci. USA **90**: 10355-9.
Okagaki R, Neuffer M, and Wessler S (1991). A deletion common to two independently derived *waxy* mutations of maize. Genetics **128**: 425-431.
Pelletier MK, and Shirley BW (1995). A genomic clone encoding flavanone 3-hydroxylase (Accession No. U33932) from Arabidopsis thaliana. Plant Physiol. **109**: 1125.
Pelletier MK, and Shirley BW (1996). Analysis of flavanone 3-hydroxylase in Arabidopsis seedlings: Coordinate regulation with chalcone synthase and chalcone isomerase. Plant Physiol. **111**: 339-345.
Pelletier M, Murrell J, and Shirley B (1997). Characterization of Arabidopsis FLS and LDOX Genes: Further Evidence for Differential Regulation of “Early” and “Late” Genes. Plant Physiol. **113:** 1437-1445
Quattrocchio F, Wing JF, Leppen HTC, Mol JNM, and Koes RE (1993). Regulatory genes controlling anthocyanin pigmentation are functionally conserved among plant species and have distinct sets of target genes. Plant Cell **5**: 1497-1512.
Rabino I, and Mancinelli A (1986). Light, Temperature, and Anthocyanin Production. Plant Physiol. **81**: 922-924.
Rose A, Casselman A, and Last R (1992). A phosphoribosylanthranilate transferase gene is defective in blue fluorescent *Arabidopsis thaliana* tryptophan mutants. Plant Physiol. **100**: 582-592.
Shirley B (1996). Flavonoid biosynthesis: 'new' functions for an 'old' pathway. Trends in Plant Science **1**: 377-382.
Shirley B, Kubasek W, Storz G, Bruggemann E, Koornneef M, Ausubel F, and Goodman H (1995). Analysis of Arabidopsis mutants deficient in flavonoid biosynthesis. The Plant J. **8**: 659-671
Shirley BW, and Goodman HM (1993). An *Arabidopsis* gene homologous to mammalian and insect genes encoding the largest proteasome subunit. Mol. Gen. Genet. **241**: 586-594.
Shirley BW, Hanley S, and Goodman H (1992). Effects of ionizing radiation on a plant genome: Analysis of two Arabidopsis *transparent testa* mutations. Plant Cell **4**: 333-347.
Shirley BW, and Hwang I (1995). The interaction trap: In vivo analysis of protein-protein associations. Methods in Cell Biology **49**: 401-416.
Stapleton A, and Walbot V (1994). Flavonoids can protect maize DNA from the induction of ultraviolet radiation damage. Plant Physiol. **105**: 881-889.

Taylor L, and Briggs WR (1990). Genetic regulation and photocontrol of anthocyanin accumulation in maize seedlings. Plant Cell **2**: 115-127.
van der Meer IM, Stam ME, van Tunen AJ, Mol JNM, and Stuitje AR (1992). Antisense inhibition of flavonoid biosynthesis in petunia anthers results in male sterility. Plant Cell **4**: 253-262.
Yao K, De Luca V, and Brisson N (1995). Creation of a metabolic sink for tryptophan alters the phenylpropanoid pathway and the susceptibility of potato to *Phytophthora infestans*. Plant Cell **7**: 1787-1799.

Chapter 5

Conclusions

Three genes encoding 2-oxoglutarate-dependent dioxygenases were isolated from *Arabidopsis thaliana* and characterized at the genetic, molecular, and biochemical level. Low-stringency DNA blot analysis indicated that F3H and LDOX are encoded by single genes while FLS is probably encoded by two or three genes. F3H was shown to be encoded by the *TT6* locus, based on comparison of map positions obtained for the locus and the cloned gene, RNA blot analysis of *F3H* in *tt6*, immunoblot analysis, and sequencing of the allele from *tt6*. Furthermore, RNA blot analysis was used to study the expression of flavonoid genes in seedlings, and it was shown that *CHS*, *CHI*, *F3H*, and *FLS1* are "early" genes, while *DFR* and *LDOX* are "late" genes in Arabidosis. Recombinant CHS, F3H, FLS, DFR, and LDOX were overexpressed in *E. coli*, purified, and used to develop polyclonal antibodies in chicken. Immunoblot analysis of Arabidopsis seedlings showed that the distinction of "early" and "late" genes could be seen at the level of protein accumulation. Furthermore, mutants defective in the *TT3*, *TT5*, *TTG*, or *TSB1* loci accumulated higher levels of flavonoid enzymes than did wild-type seedlings.

There are several important ways in which the molecular characterization of genes encoding F3H, FLS, and LDOX from Arabidopsis builds on the work of others. For instance, *F3H*, although it had been cloned from a number of plant species was thought to be pivitol in the regulation of the pathway, since it was an "early" gene in petunia flowers but a "late" gene in snapdragon. We showed that in Arabidopsis seedlings, as in petunia flowers, *F3H* is an "early" gene, although the physiological significance of this finding remains elusive. Isolating and characterizing a gene encoding FLS was important for several reasons. First, only two other genes had been isolated previously, and thus we have provided another tool that may be useful to isolate homologues from other plant species. Second, we were the first to provide evidence from any plant species that FLS is an "early" gene. It is also interesting that of a total of six genes encoding flavonoid enzymes isolated to date from Arabidopsis, only FLS appears to be present in more than one copy. Further studies need to be done to isolate these genes and characterize the expression patterns in various tissues using gene-specific probes. This would provide important information on where each isoform is expressed and may give clues as to the substrate preferences of the individual enzymes, since different flavonols accumulate in different Arabidopsis tissues. The finding that *TTG* is required for accumulation of *DFR* and *LDOX* mRNA, but that *TT8* is required only for DFR accumulation was surprising. Since *TT8* is not the structural gene encoding DFR and this locus is known to contain highly reduced levels of DFR mRNA, it was thought that *tt8* was a regulatory mutant that affected accumulation of all "late" genes. Our results suggest that *TT8* may encode a regulatory factor that specifically modulates DFR mRNA accumulation. Isolation, sequencing and functional analysis of *TT8* would clarify the mechanism by which this occurs.

The antibodies we have developed against Arabidopsis CHS, CHI (developed by Cody Cain), F3H, FLS, and LDOX represent the most comprehensive collection of flavonoid-specific antibodies in any plant species to date and provide a powerful new tool to examine the the regulation and localization of the pathway using immunoblot analysis and immunocytochemical techniques. The characterization of the specificity of these antibodies presented in this dissertation and optimization of their use lays the groundwork for future studies. These antibodies will be used to determine the localization of flavonoid enzymes *in vivo*. In addition, a variety of methods using the recombinant proteins in conjunction with these antibodies are currently being used to establish whether flavonoid enzymes interact specifically with one another *in vitro*. There are also plans to develop enzyme assays to determine whether these proteins are catalytically active.

Examining the regulation of the flavonoid pathway at the level of protein expression is important, since most studies have been performed using RNA blot analysis, and our results indicate that this pathway may also be regulated at the post-transcriptional level. For example, we have shown that mutations at *TT3*, *TT5*, and *TTG* lead to an increased accumulation of flavonoid enzymes, while previous RNA blot analysis had revealed no such differences. Further studies will be required to determine the mechanism of this upregulation. One important step towards this goal is to determine which and to what level flavonoid intermediates are accumulating in the *tt* mutants.

This will become feasible in the near future, as we have just acquired a new high-performance-liquid-chromotography (HPLC) system. Once the intermediates that accumulate in these mutants are identified, wild-type seedlings could be fed these compounds to attempt to establish phenocopies of each mutatation. Upon the development of the enzyme assays mentioned above, it would also be of value to determine whether the increases in flavonoid enzyme levels in these mutants lead to an increase in the amount of activity present per unit protein.

As outlined in the introduction, understanding the diverse mechanisms by which cells mediate control over their own metabolism is of fundamental importance in biology. Flavonoid biosynthesis represents an attractive model in which to study this regulation, due to the dispensability of the pathway and the large collection of mutants that are available. Our studies have built on decades of previous research as to the function and control of this pathway by showing that in addition to the transcriptional regulation previously observed, post-transcriptional regulation also occurs. Furthermore, our results suggest that cross-pathway regulation may exist between primary (tryptophan) and secondary (flavonoid) metabolic pathways. This regulation has important implications for attempts aimed at altering metabolism in transgenic organisms**.** Finally, future studies on the localization and structural organization of the putative flavonoid enzyme complex will likely yield insights into other mechanisms by which cells mediate metabolic control.

April 1997

Curriculum vitae

Matthew Kent Pelletier

Born: Kankakee, IL, May 15, 1970

Office Address: Department of Biology
Virginia Polytechnic Institute and State University
Derring Hall 5028
Blacksburg, VA 24061-0406;
phone: (540)231-9375; fax : (703) 231-9307

Home Address: 1224 University City Blvd D37
Blacksburg, VA 24060
phone (540) 552-5317

Education:

1989-1993 B.S. Biology, Liberty University

1993-present PhD., Biology, **in progress**, Virginia Polytechnic Institute and State University. Expected Completion May 1997

Professional Positions and Activities:

Research Assistant, Department of Biology, Virginia Polytechnic and State University

1993 Summer and Fall
1994 Spring and Summer
1995 Spring and Summer
1996 Spring, Summer, and Fall

Teaching Assistant, Department of Biology, Virginia Polytechnic and State University

Fall 1994: Principles of Biology
Fall 1995: Molecular Biology Lab

Guest Lecturer for General Biology, Fall 1995

Virginia Tech Botany Seminar Series Speaker Selection Committee, 1996-1997

Reviewed a manuscript for Plant Physiology in October 1996

Guest Lecturer for General Biology, Fall 1996

Supervision of Undergraduate Research Projects

Jonathon Shelton Fall 1994/Spring, Summer 1995 "Demonstration of Interactions between Chalcone Synthase and Chalcone Isomerase using the Yeast Two-Hybrid System"

Julie Murrell Spring 1996 "Cloning and Characterization of Flavonol Synthase from *Arabidopsis thaliana*"

Seminars:

Plant Molecular Biology Discussion Group, May 11, 1995

Virginia Tech Botany Seminar Series, "Analysis of flavanone 3-hydroxylase in Arabidopsis Seedlings" February 7, 1996

Liberty University, "Analysis of flavanone 3-hydroxylase in Arabidopsis Seedlings" February 16, 1996

Plant Molecular Biology Discussion Group, March 5, 1996

University of Chicago, "Characterization of F3H, FLS, and LDOX from Arabidopsis: Coordinate Regulation of "Early" and "Late" Genes (post-doctoral interview) ; August, 1996

Boyce Thompson Institute, Cornell University, "Characterization of F3H, FLS, and LDOX from Arabidopsis: Coordinate Regulation of "Early" and "Late" Genes (post-doctoral interview) ; September, 1996

Abstracts:

Johnson, C.S., Burbulis, I.E., **Pelletier, M.K**., Cain, C.C., and Shirley, B.W. (1993) Protein-Protein Interactions between the First Two Enzymes of Flavonoid Biosynthesis. Fifth International Conference on Arabidopsis Research, August 19-22, Columbus OH.

Pelletier, M.K., Burbulis, I.E., Cain, C.C., Iacobucci, M., and Shirley, B.W. (1994) Protein-Protein Interactions between Chalcone Synthase and Chalcone Isomerase from Arabidopsis. Fourth International Congress of Plant Molecular Biology, June 19-24, 1994, Amsterdam, The Netherlands.

Pelletier, M.K. and Shirley, B.W. (1995) Cloning and Characterization of an Arabidopsis Flavanone 3-Hydroxylase Gene. Sixth International Conference on Arabidopsis Research, June 7-11, Madison Wisconsin

Pelletier, M.K. and Shirley, B.W. (1996) Flavonol synthase and leucoanthocyanidin dioxgenase genes provide further evidence for distinct regulation of "early" and "late" flavonoid biosynthetic steps in Arabidopsis. Plant Molecular Biology Gordon Conference July 21-26, 1996.

Publications:

Pelletier, M.K. and Shirley, B.W. (1995) A Genomic Clone Encoding Flavanone 3-hydroxylase (Genbank U33932) from *Arabidopsis thaliana.* **Plant Physiology**.109: 1125

Pelletier, M.K. and Shirley, B.W. (1996) Analysis of Flavanone 3-Hydroxylase in Arabidopsis Seedlings: Coordinate Regulation with Chalcone Synthase and Chalcone Isomerase. **Plant Physiology**. 111: 339-345

Burbulis, I.E., **Pelletier, M.K.**, Cain, C.C., and Shirley, B.W. (1996) Are Flavonoids Synthesized by a Multi-Enzyme Complex? Bulletin of the Southern Association of Agricultural Scientists. 9: 29-36

Pelletier, M.K., Murrell, J.R., and Shirley, B.W. (1997) Characterization of Arabidopsis FLS and LDOX Genes: Further Evidence for Differential Regulation of "Early" and "Late" Genes **Plant Physiology**. 113: 1437-1445

Pelletier, MK and Shirley, BW. (1997) Immunoblot Analysis of Arabidopsis Seedlings: Mutations in *TT3*, *TTG*, *TT5*, and *TSB1* lead to Increased Levels of Flavonoid Enzymes. To be submitted to The Plant Journal

Awards and Honors:

Who's Who Among Students In American Universities and Colleges, 1992-1993, Liberty University

Outstanding Biology Education Student, Liberty University, 1993.

Grant Proposals:

Sigma Xi, Grant in Aid of Research. Determining Interaction Domains between Chalcone Synthase and Chalcone Isomerase from *Arabidopsis thaliana*. M.K. Pelletier, principal investigator. Date submitted: October 25, 1993 Funded: Spring 1994 Amount: $450 +$450 match from Biology Department

Pending:

American Cancer Society Post-Doctoral Fellowship. "Analysis of Pathogen Defense Responses in Mutants Defective in Free Radical Production and Detoxification". Submitted March 1, 1997

National Institutes of Health. "Analysis of Pathogen Defense Responses in Mutants Defective in Free Radical Production and Detoxification". Submitted April 5, 1997