Plant Physiol. Biochem., 1998, **36** (1-2), 135-144

Flavonoid and flavonol glycoside metabolism in *Arabidopsis*

Terrence L. Graham

Department of Plant Pathology, Ohio State University, 201 Kottman Hall, 2021 Coffey Road, Columbus, OH 43210, USA. (Fax 1-(614) 292 4455; E-mail graham.1@osu.edu)

Abstract The molecular genetic analysis of *Arabidopsis* promises to add greatly to our understanding of secondary product pathways and their biological roles. This review focuses on flavonoid metabolism in *Arabidopsis*, a particularly relevant pathway for such analysis because it is so highly conserved in plants. Several of the biosynthetic enzymes of the phenylpropanoid and flavonoid pathways have been cloned and/or mapped to specific loci. Recent advances concerning the regulation of anthocyanin metabolism by light and by possible transcriptional activators (Myc- and Myb-like elements) are discussed as well as the possible roles of flavonols in cold acclimation, ultraviolet light protection and male fertility. The use of HPLC profiling for the discovery of additional metabolic and regulatory mutants is described. The identification of the major flavonol glycosides in wild-type *Arabidopsis* is given and the partial characterization of two new flavonol glycoside mutants discussed. The biochemical genetics for flavonol glycoside formation in soybean is presented as an example, and the high degree of specificity of the glycosyl transferases and the interesting diversity of the resulting flavonol glycosides are discussed. Hypotheses regarding the potential biological activities of the flavonoid conjugates and the possible regulatory roles of highly aglycone-specific glucosidases are presented. © Elsevier, Paris.

Key words Flavonoid, flavonol, anthocyanin, glycosides, mutant, HPLC, *Arabidopsis*.

Abbreviations ACC, acetyl CoA carboxylase; 4CL, 4-coumaroyl-CoA ligase; CHI, chalcone isomerase; CHS, chalcone synthase; DFR, dihydroflavonol reductase; FST, flavonol sulphotransferase; F5H, ferulate 5-hydroxylase; F3H, flavanone 3-hydroxylase; F3'H, flavonoid 3'-hydroxylase; Glc, glucose; PAL, phenylalanine ammonia-lyase; Rha, rhamnose.

Introduction

Plant secondary metabolites are both extremely numerous and diverse. Although the general pathways of secondary product metabolism are conserved over a wide range of plants, the great diversity is the result of specialized biosynthetic pathways. The predominant secondary metabolites so far identified in *Arabidopsis* are the flavonoids, the hydroxycinnamic acid esters, the glucosinolates, the brassinosteroids and the indole phytoalexins. Given the diversity of secondary metabolism, what can study of these various secondary metabolites in *Arabidopsis* provide us with? First of all, it can provide invaluable information on the metabolic pathways, developmental regulation and biological functions of highly conserved secondary products, such as the flavonoids. Secondly, investigation of somewhat less common metabolites, such as the glucosinolates and indoles will still have relatively wide relevance and application, especially within the brassica. Finally, since it is currently impossible to exhaustively study the molecular genetics (let alone the biochemistry) of all plants, *Arabidopsis* can provide potentially valuable paradigms which will facilitate research on the regulation of secondary metabolism in even the most specialized pathways in other plants.

An excellent review of secondary metabolism in *Arabidopsis* was recently published (Chapple *et al.*, 1994) as well as a review on the characterization of the transparent testa (*tt*) mutants (Shirley *et al.*, 1995). A review on the brassinosteroids appears in this volume (Szekeres and Koncz, 1998). Because of their general interest in a wide range of plants, I have cho-

sen to focus on recent developments in the flavonoid area. In addition to providing a short general update on flavonoid work in *Arabidopsis*, I will highlight recent collaborative work in our lab and others on the characterization of some of the flavonoid mutants derived from HPLC screens. In addition, the biochemical genetics of flavonol glycoside formation in soybean is described as an illustration of the complexity of these metabolites and the possible biological and regulatory implications of conjugation are discussed. The biological roles of flavonoids in plants has been the subject of a very recent review by Shirley (1996) and will be only briefly alluded to as relevant to the text below.

Characterization of flavonoid biosynthetic enzymes and mutants

The general ring structure and numbering scheme for the flavonoids is shown in figure 1. The B ring of

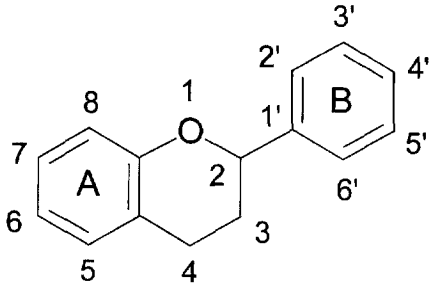

Figure 1. *A general structure and numbering system for the flavonoids.*

the flavonoids is derived from phenylpropanoid metabolism, while the A ring is the result of the head-to-tail condensation of 3 acetate molecules derived from malonyl-CoA. In isoflavonoids, the B ring is covalently linked at the 3 rather than the 2 position. Of the mutants in *Arabidopsis* that potentially affect flavonoid metabolism, those of particular importance are the transparent testa (tt) mutants. These mutants are linked to flavonoid metabolism due to the fact that the flavan-3,4-diols are precursors of the condensed tannins, which impart testa coloration. Eleven such mutants exist (Koornneef, 1990; Shirley *et al.*, 1995) and the functions of the corresponding genes have been identified or hypothesized for several.

The flavonoid biosynthetic mutants identified to date include mutants deficient in the early enzymes of flavonoid metabolism, chalcone synthase (CHS) and chalcone isomerase (CHI). CHS and CHI are both encoded by single-copy genes corresponding to the *TT4* and *TT5* loci respectively (Feinbaum and Ausubel, 1988; Shirley *et al.*, 1995). Dihydroflavonol reductase (DFR), a key enzyme in shunting flavonols into the anthocyanidins, is also a single copy gene in *Arabidopsis* and corresponds to the *TT3* locus (Shirley *et al.*, 1995). Yeast two hybrid analysis has shown that CHS, CHI and DFR are involved in protein-protein interactions (Burbulis I. E. and Shirley B. W., unpublished). It is hypothesized that these enzymes form a complex associated with the endoplasmic reticulum which facilitates channeling of intermediates in the pathway. A defect in flavonoid 3'-hydroxylase (F3'H) activity, which converts the flavonol kaempferol to quercetin, is thought to be associated with the *tt7* mutation (Koornneef *et al.*, 1982). Flavanone-3-hydroxylase (F3H) has also been examined and is also a single copy gene (Pelletier and Shirley, 1996). Immunoblot analysis of *tt6* and sequencing of the F3H gene from the mutant suggest that it is defective in F3H due to the presence of several critical single base pair mutations (Pelletier M. K. and Shirley B. W., unpublished). An additional gene, with homology to plant flavonol sulphotransferases (FSTs) has been shown to be induced by salicylic acid, jasmonic acid and incompatible infections (Lacomme and Roby, 1996). These interesting enzymes, first identified in *Flaveria* sp., catalyze the transfer of the sulphate group from 3'-phosphoadenosine 5'-phosphosulfate to the 3-hydroxyl group of flavonol aglycones to form flavonol sulphate esters. Finally, two copies of acetyl-CoA caboxylase (ACC), the enzyme forming malonyl-CoA from acetyl-CoA, and the source of the acetyl groups of the A ring of flavonoids, are present in *Arabidopsis* (Yanai *et al.*, 1995). In parsley, this enzyme is coordinately regulated with other flavonoid pathway enzymes (Hahlbrock, 1981).

Genes encoding several enzymes of phenylpropanoid metabolism have also been identified. The mutant *fah1* is defective in accumulation of sinapic acid metabolites; *FAH1* has been shown to encode ferulate-5-hydroxylase (F5H), an enzyme which defines a new cytochrome p450 subfamily, CYP84 (Meyer *et al.*, 1996). Phenylalanine ammonia-lyase (PAL, Ohl *et al.*, 1990; Wanner *et al.*, 1995) and 4-coumaroyl-CoA ligase (4CL, Lee *et al.*, 1995) have also been described. While PAL represents a multigene family with at least 3 members, 4CL is again a single copy

gene. These various enzymes and the corresponding loci, if known, are summarized in table 1.

Table 1. *Phenylpropanoid and flavonoid genes identified in* Arabidopsis.

Gene	Copies	Mutant
Acetyl CoA carboxylase (ACC)	2	-
Phenylalanine ammonia-lyase (PAL)	3	-
Ferulate-5-hydroxylase (F5H)	?	*fah1*
4-Coumarate Co-A ligase (4CL)	1	-
Chalcone synthase (CHS)	1	*tt4*
Chalcone isomerase (CHI)	1	*tt5*
Flavonoid-3'-hydroxylase (F3'H)	?	*tt7*
Flavonol-3-hydroxylase (F3H)	1	*tt6*
Flavonol sulfotransferase (FST)	?	-
Dihydroflavonol reductase (DFR)	1	*tt3*

The Bronze genes in maize offer additional examples of genes which may be identified in *Arabidopsis* sp. The Bronze-1 gene encodes a UDPG-flavonol-3-O-glucosyltransferase (Dooner *et al.*, 1991). The Bronze-2 gene in corn is genetically linked to the last step in anthocyanin synthesis in maize, and has recently been shown to encode a glutathione S-transferase that glutathionates the anthocyanins allowing their uptake into the vacuole by the glutathione pump (Marrs *et al.*, 1995).

Light regulation of flavonoid metabolism in *Arabidopsis*

Light induced developmental expression of phenylpropanoid and flavonoid enzymes leads to anthocyanin production in *Arabidopsis* seedlings. The induction of these genes occurs in three coordinate groups which correlate to their relative position in the metabolic pathways. PAL is expressed before CHS and CHI, which in turn are expressed before DFR (Kubasek *et al.*, 1992). This implies distinct regulatory mechanisms for these genes. F3H, which is coordinately expressed with CHS and CHI in *Arabidopsis*, is coordinately expressed with different subsets of genes in different species (Pelletier and Shirley, 1996), suggesting a point of differential regulation in various plants.

The *HY4* gene encodes the cryptochrome (CRY1) protein, a flavin-type blue light receptor which mediates inhibition of hypocotyl elongation. The *hy4* mutation leads to disruption of both inhibition of hypocotyl elongation and anthocyanin accumulation (Jackson and Jenkins, 1995). Engineered point mutations in CRY1 show similar effects (Ahmad *et al.*, 1995). Conversely, overexpression of CRY1 results in increased CHS expression, anthocyanin accumulation and hypersensitivity to blue, UV-A and green light (Lin *et al.*, 1996). Further experiments with the long-hypocotyl mutant *hy4*, suggest that UV-B and UV-A/blue light act through separate but synergistic pathways in inducing CHS expression (Fuglevand *et al.*, 1996). Only the UV-A/blue light response involves the CRY1 photoreceptor. Studies with the *fhy1* mutant, which is deficient in a downstream component of the phytochrome A pathway, and with a phytochrome A deficient mutant, suggest that red-light induction of CHS is mediated by phytochrome A, not phytochrome B, and that phytochrome A is not a component of the blue-light signalling pathway (Barnes *et al.*, 1996). The independence of the UV-B and UV-A/blue light pathways from the phytochrome pathway is supported in additional trangenic and physiological studies (Batschauer *et al.*, 1996; Christie and Jenkins, 1996). Several negative regulatory genes (*COP*, *DET* and *FUS*) act at a nuclear level downstream from the photoreceptors (Chory *et al.*, 1996). Mutations in *DET1* or *DET2* cause anthocyanin accumulation in dark grown plants, although they appear to act on different pathways (Chory and Peto, 1990; Chory *et al.*, 1991). Similarly, mutations in at least nine of the *FUSCA* genes cause anthocyanin accumulation in the absence of light (Misera *et al.*, 1994). The *FUSCA* genes appear to be upstream of the *TT* and *TTG* genes and downstream of phytochrome. An additional negative regulator of light expression is the *ICX1* gene. Mutations at this locus also show increased chalcone synthase expression in response to light (Jackson *et al.*, 1995).

Regulation of flavonoid genes during development and stress

Several mutations affecting the regulation of flavonoid pathways have also been identified. Mutants *ttg* and *tt8* are both hypothesized to be regulatory in nature. In both cases, CHS and CHI expression is similar to that in wild-type plants, whereas expression of the DFR gene is reduced or absent. Moreover, in the pleiotropic mutant *ttg*, both flavonoid synthesis and trichome development is disrupted. Since seven independent alleles at this locus all show the same dual phenotype, it seems likely that *ttg* encodes a regulatory factor. It has been hypothesized that *ttg* may encode a Myc-like regulator that would control both flavonoid

and trichome phenotypes by interaction with different Myb-like factors, including the *gl1* gene which is involved in trichome development (Chapple *et al.*, 1994; Shirley *et al.*, 1995). This notion is supported by the fact that both phenotypes are restored in the ttg mutation by transgenic expression of the Myc-like maize transcriptional regulator, *R* (Lloyd *et al.*, 1992, 1994). Interestingly, the *C1* maize transcriptional activator (a Myb factor) had no effect alone, but in the presence of *R* caused anthocyanin production in tissues normally not expressing the pigment (Lloyd *et al.*, 1992).

Although flavonoids are necessary for pollen tube growth in a number of plants, two recent studies suggest that flavonoids may not be required for Mendelian male fertility in *Arabidopsis*. A mutation at the 3' splice acceptor site of CHS completely suppressed formation of flavonoids in *Arabidopsis* flowers with no effect on male fertility (Burbulis *et al.*, 1996). Moreover, the CHS negative *tt4* mutant, which is devoid of flavonols in the reproductive organs, also set seed and demonstrated normal pollen tube formation (Ylstra *et al.*, 1996). Flavonoids may not represent a universal prerequisite for fertility. It is interesting, however, that the flavonoids involved in fertility are the 3-O-glycosides, whereas, as noted below, the 7-O-glycosides predominate in *Arabidopsis*. Fourteen independent *coi1* mutants, insensitive to the phytotoxin coronotine, are male sterile, insensitive to methyl jasmonate, resistant to coronotine-producing bacterial pathogens and fail to accumulate anthocyanins in response to methyl jasmonate or coronotine (Feys *et al.*,1994). It is suggested that methyl jasmonate may act as a signal involved in anthocyanin formation and male fertility.

Although PAL and CHS mRNAs accumulate in *Arabidopsis* in response to low temperatures in a light dependent manner (Leyva *et al.*, 1995), examination of two tt mutants, unable to accumulate anthocyanins, suggest that the pigments are not needed for freezing tolerance. On the other hand, four freezing intolerant mutations (*sfr3, sfr4, sfr6* and *sfr7*) reduce or eliminate the accumulation of anthocyanins during cold acclimation (Mckown *et al.*, 1996). These may represent yet another set of factors modulating anthocyanin accumulation and/or linked phenotypes.

Although flavonoid mutants are hypersensitive to UV-B irradiation (Li *et al.*, 1993; Lois and Buchanan, 1994), mutants *fah1* and *tt5*, deficient in sinapate ester and flavonoid formation respectively, were used to show that although both classes of compounds may function in protection against UV-B injury, the former are more effective (Landry *et al.*, 1995).

Although a number of papers have appeared on genes regulated during infection of *Arabidopsis*, these have focused almost entirely on non-flavonoid responses and will not be reviewed here. The only noteworthy exception is the FST induction alluded to above.

HPLC metabolic profiling as a strategy for flavonoid mutant selection

Thus, although a number of *Arabidopsis* flavonoid mutants are available and have begun to yield important information towards our understanding of these pathways, nearly all of these mutants were discovered indirectly, through seed color or light response phenotypes. The development of a high resolution HPLC metabolic profiling protocol allowed direct screening for metabolic mutants (Graham, 1991). Although such an approach is somewhat time/labour intensive for a routine screen, by profiling many metabolites in each chromatogram one is examining a very large number of potential phenotypes simultaneously. The flavonoids represent an excellent example of such a use of HPLC, in that they are precursors for many metabolites, and are themselves subject to both polymerization and the formation of a complex series of glycosides. Thus, not only would biosynthetic and regulatory mutations be uncovered, but mutations in the glycosyl transferases and glycosidases involved in the steady state levels of the various conjugates would be found as well. Another obvious advantage of this type of approach is that the screen itself provides clues to the biochemical phenotype that can facilitate further characterization.

The HPLC profiles of various *Arabidopsis* organs and tissues are quite distinct. Rosette leaves were chosen for our screening efforts since they displayed a rich array of metabolites, many of which overlapped those in other tissues (Graham, 1991; Mittal, 1996). Rosette leaves are also conveniently harvested early in growth and their removal has little effect on the vigor or subsequent flowering of the plant. Profiling of leaves from two week old plants was carried out as described by Graham (1991), adapted in that a linear gradient of 0-55% acetonitrile over 22 min was used.

An important limitation of this more rapid HPLC profiling procedure is that some metabolites are not completely resolved and shifts in metabolites can cause their overlap with other existing metabolites.

Thus, though highly effective as an initial screen, the protocol must be followed up by other biochemical or chemical analyses. Once a putative mutant is identified, longer and higher-resolution HPLC runs can be made. Since it is relatively easy to check the nature and purity of peaks by complete ultraviolet scans as they elute, this is a powerful secondary approach that can also be employed. The use of additional methods of separation such as capillary electrophoresis, separation by HPLC using different supports and mass spectroscopy are also powerful complementary methods. Finally, the importance of isolation and chemical analyses of individual peaks is discussed below.

The major peaks in HPLC profiles of *Arabidopsis* correspond to a complex array of flavonol glycosides

The partial characterization of the six major peaks in the HPLC chromatogram of the Col-0 ecotype is shown in table 2 (Mittal, 1996; Mittal S., Graham T. L., Davis K. R., Thomas-Oates J., Brull L., and Carlson R. W., unpublished). The identities are the result of purification of the peaks and their hydrolyses products and subsequent ultraviolet, nuclear magnetic resonance and fast atom bombardment mass spectral analyses. In all cases, the sugars appear to be linked in the α configuration and are linked to the aglycone via an 7-O-glycosidic bond. Thus, the major ultraviolet-absorbing metabolites in *Arabidopsis* are flavonol glycosides. These results support and extend previous work by Koornneef *et al.* (1982) and Li *et al.* (1993). Analysis of samples at mutiple wavelengths or by wavelength substraction could be a powerful way of uncovering additional metabolites of interest. Interestingly, the peak at 14.1 was only seen in certain samples (Mittal, 1996). The reason for this, and its metabolic origin, are not completely clear. Mass spectral data suggested that it was an isomer of Rha-Glc-Rha-Kaempferol. The identities of the metabolites suggest that rhamnose may be the preferred sugar attached to the 7 position of the aglycone and that further glucose or rhamnose residues are then added to form the di- and tri-glycosides.

Table 2. *Identification of the major HPLC peaks in Col-O.*

Retention time (*min*)	Compound
11.6	Rha-Glc-Rha-Quercetin
12.1	Rha-Glc-Rha-Kaempferol
12.8	Glc-Rha-Quercetin
13.6	Glc-Rha-Kaempferol
	Rha-Rha-Quercetin
14.1	Isomer of Rha-Glc-Rha-Kaempferol
14.6	Rha-Rha-Kaempferol

A variety of metabolic and regulatory mutants in flavonol metabolism are selected by HPLC

Using the HPLC profiling protocol to screen ethane methyl sulphonate (EMS) mutagenized Col-O seed (Graham, 1991), we initially identified about 25 putative mutants in flavonoid metabolism. Of these, one mutant, C1123, was of particular interest and is discussed in detail below. In a second effort, over 5,000 T-DNA tagged plants (Ws ecotype), representing 1,000 independently transformed lines, were screened by the same HPLC protocol (Mittal, 1996). The profiles of wild type Ws tissues are nearly identical to those of Col-0. Twenty putative mutants were found in this second screen. In addition, preliminary identifications of the peaks in wild-type and several mutant lines were undertaken. A brief summary of these various findings is presented here.

Of the mutants from the transposon tagged Ws plants, two showed a complete loss of several specific compounds with no additional changes, two accumulated novel compounds with reductions in others, two showed enhanced accumulation of existing compounds and fourteen showed very low accumulations of nearly all compounds with possible alterations in specific compounds (Mittal, 1996). These observations confirmed the usefulness of the approach and the expectation that the screen would pick up different classes of mutants that reflect the nature of various potential changes in the overall metabolic grid. The specifics for a few mutants are given below. Identification of the major HPLC metabolites allowed the further characterization of the biochemical phenotypes of several of the mutants.

Among the mutants, one showed greatly enhanced levels of 6 relatively non-polar peaks (late in the HPLC chromatogram) and had an early flowering (2 weeks as opposed to 4 - 5 weeks) phenotype (Mittal, 1996). Consistent with early flowering, the mutant also had fewer rosette leaves than wildtype. The new peaks in this mutant may correspond to flavonoid aglycones, but this has not been confirmed chemically. Thus, this mutant may be deficient in overall conjugation of the flavonols. Two mutants with specific losses in quercetin conjugates were seen (Mittal, 1996). In

one of these, at least two of the quercetin conjugates (Rha-Glc-Rha- and Glc-Rha-) were completely missing from the HPLC profile, and it is possible that it is completely devoid of quercetin. If so, it could be deficient in F3'H and related to *tt7*. In the other mutant, designated *aip*, the loss of quercetin conjugates was further confirmed by ultraviolet, nuclear magnetic resonance and mass spectral analyses. In *aip*, discussed further below, several new kaempferol conjugates also appeared, suggesting that the mutation may be more complex and/or regulatory in nature.

Fourteen mutants were identified which showed very low accumulations of all peaks, with changes in the relative amounts of specific compounds (Mittal, 1996). It is perhaps not surprising that these mutants were the most abundant, since a block in the expression or regulation of any gene upstream of flavonoid metabolism could lead to such changes. The relationship of these mutants to the *tt* mutants has not yet been pursued. Future work with these mutants should perhaps start with genetic crosses to allow their placement into complemetation groups.

Genetic and biochemical analysis of the flavonol mutants, C1I23 and *aip*

The mutants C1I23 and *aip* were of particular interest since they also displayed interesting developmental phenotypes. These mutants thus may be either regulatory in nature or they may point to specific roles of flavonol glycosides in these particular developmental processes.

The identification of the major metabolites in the mutant C1I23 were similarly pursued through ultraviolet, nuclear magnetic resonance and mass spectral analyses and they were found to be generally the same as those in Col-0 (Mittal, 1996; Mittal S., Graham T. L., Davis K. R., Thomas-Oates J., Brull L. and Carlson R. W., unpublished). However, C1I23 was confirmed as showing an altered glycosylation pattern. The predominant phenotype involves alterations in the triglycosides. While Rha-Glc-Rha-Kaempferol predominates over Rha-Glc-Rha-Quercetin in the wild-type, the opposite is true in the mutant. Parallel shifts are not apparent in the other kaempferol or quercetin glycosides, suggesting that the mutation (*a*) differentially affects terminal glycosylation of the aglycones or (*b*) that there is simply a greater net relative accumulation of quercetin, and the triglycoside is the preferred conjugate. In addition, two novel metabolites are present in C1I23 at high levels compared to wild type, where they are below the limits of detection. These are not flavonols and remain unidentified. Genetic analysis of the mutant demonstrated that the phenotype was conferred by a single nuclear recessive gene (Mittal, 1996). Due to the multiple changes in the profile, it seems likely that the mutation is not solely affecting a glucosyl transferase, but some earlier or more global function. As expected, the changes seen in rosette leaves of C1I23 are reflected in differences in the profiles of some other organs as well (Mittal, 1996). Some organs show near wild-type profiles, some show changes chemically parallel to those in the rosette, while others appear to involve different metabolites from those seen in the rosette leaves and will need to be further analyzed. Again, this evidence suggests a more global target for the mutation. C1I23 thus demonstrates the usefulness of HPLC to detect multiple alterations in the metabolic grid due to a single gene mutation. To facilitate future positional cloning of C1I23, it was mapped to chromosome 2 (Mittal, 1996).

A consistent developmental phenotype of C1I23 was the formation of exactly double the number of rosette leaves prior to flowering, which was subsequently delayed. Which of the shifts in metabolites alluded to above, if any, relate to this phenotype is unknown. However, a preliminary HPLC analysis of other late flowering mutants led to the discovery that the *fb* mutation (Koornneef *et al.*, 1983) has an altered metabolic profile compared to its parent ecotype Ler that is similar to the differences between C1I23 and Col-0 (Mittal, 1996). Whether these mutations are genetically related or allelic remains to be examined, but it further suggests that C1I23 may be a regulatory mutant.

The quercetin conjugate deficient mutant, designated as *aip*, was also further examined. This mutant demonstrated an alteration in the phyllotaxy pattern in the inflorescence, hence *A*ltered *I*nflorescence *P*attern. The cauline leaves and siliques formed in whorls in this mutant, rather than the normal alternate arrangement in the wild type (Mittal, 1996). Biochemically, the *aip* mutant completely lacked the quercetin conjugates (Rha-Glc-Rha-, Glc-Rha- and Rha-Rha-) and showed instead the presence of the normal and several novel kaempferol conjugates (Glc-Glc-Rha-, Glc-Rha-Glc-, Rha-Glc-Glc-, Rha-Glc-, Glc-Glc-Glc- and Glc-Glc-), which are absent or below the limits of detection in the wild type. There is considerable overlap between these new peaks and those found in the wild

type, however, and the overall profiles between mutant and wild type were surprisingly similar given these dramatic differences. The differences were fully apparent only after full characterization of the isolated peaks by ultraviolet, nuclear magnetic resonance and mass spectral analyses. This confounded simple HPLC comparisons of other organs, where the compounds were not actually isolated and characterized.

Genetically, the *aip* phenotype was also conferred by a single, nuclear recessive gene. The absence of quercetin in *aip* suggests a possible block in the conversion of kaempferol to quercetin, at first suggesting a parallel to the *T* locus in soybean or the *TT7* locus in *Arabidopsis*. However, the dramatic changes in glycosylation patterns of kaempferol suggest a more complex or regulatory basis for this mutant. Rhamnose is no longer the only sugar in direct linkage to the aglycone, and new di- and tri-glucosides have appeared. A genomic library of the *aip* mutant has been constructed and is being screened for the T-DNA tag to isolate possible clones.

Flavonol glycoside biochemical genetics

The genes involved in the formation of the flavonol 3-O-glycosides in soybean have been identified through classical genetics. Although the flavonols in *Arabidopsis* appear to be the 7-O-glycosides, soybean provides a very useful illustration to demonstrate the complexity of conjugates that can occur from relatively simple genetics.

Nine glycosides each of kaempferol or quercetin were identified by Buttery and Buzzell (1975) in soybean. They are all 3-O-glycosides. An additional flavonol aglycone, isorhamnetin, has been identified in certain soybean lines (Le-Van, N. and Graham, T. L., unpublished). Although closely related metabolically to quercetin, the genetics of its formation or glycosylation have not been elucidated. The hydroxylation of the B ring of kaempferol to form quercetin requires the *T* gene (Buttery and Buzzell, 1973). Thus, *T* is possibly a homolog to the *TT7* gene in *Arabidopsis*. The genes for glycosylation of kaempferol or quercetin are the same for both aglycones. The monoglucoside K5 is found in all soybean cultivars and both K5 and Q5 are found in all cultivars carrying the *T* gene. Four additional non-allelic dominant genes give rise to the various diglycosides and triglycosides by simple additions to this monoglucoside. These include glucosyl transferases that transfer glucose in β linkage to the 6 position (*FG1*) or the 2 position (*FG3*) of the monoglucoside and rhamnosyl transferases that transfer rhamnose in an α linkage to either the 6 position (*FG2*) or the 2 position (*FG4*) of the monoglucoside. Taken together, the *T* and various *FG* genes can give rise to over 32 flavonoid phenotypes.

Another gene *wm*, which is associated with magenta flower colour, dramatically reduces flavonol glycoside levels in leaves and flowers, but does not affect the formation of the aglycones in pod pubescence (Buttery and Buzzell, 1987). The magenta flower color is in fact thought to be due to the lack of flavonol glycosides in the otherwise purple flower background caused by anthocyanins (Buzzell *et al.*, 1977). Thus the *WM* locus seems to regulate the organ specific formation or glycosylation of the flavonols.

In summary, the genetics of flavonol glycoside formation in soybean generally support the notion that addition of O-glycosidic bonds to flavonoids occurs after synthesis of the aglycone, and that specific sugar linkages are added in single steps, each conferred by a different gene. It is very intriguing that such specificity and complexity in glycosylation patterns exist. Are specific biological functions associated with these genes? In this respect, it is interesting that soybean lines carrying K9 have several novel characteristics, including reduced numbers of stomata, especially on the upper leaf surface (Buttery *et al.*, 1992). As noted above, initial results with the HPLC mutant screens in *Arabidopsis* are also very promising in that several show developmental phenotypes.

Possible regulatory roles of flavonoid specific glycosidases

Flavonoid and isoflavonoid aglycones have been implicated in a wide range of highly specific biological phenomena such as, virulence and nodulation gene induction, parasite-specific chemotaxis, the regulation of auxin transport, flowering, and biotic and abiotic stress resistance (Shirley, 1996). In our own lab, the accumulation of isoflavone conjugates at points distant from the infection site has been proposed as a means to increase the defense potential of soybean tissues (Graham, 1995), a phenomenon sharing some parallels to systemic acquired resistance (SAR) in other plants. Some of these conjugates (*e.g.* those of daidzein) play a direct role in defense. However, we have recently shown that the release of free genistein or kaempferol specifically activates a peroxidase-like (Type II) NADH oxidase involved in the programming of phytoalexin elicitation competency and thus defense

potential in soybean (Graham, 1995; Graham T. L. and Graham M. Y., unpublished). While the isoflavone genistein predominates in seedling tissues, the flavonol kaempferol predominates in mature tissues. Other soybean metabolites do not activate this enzyme. Thus, the conjugation and subsequent release of flavonoids from their aglycones may play both direct and regulatory roles in defense. Interestingly, one of the early genes triggered in response to the SAR-related signal salicylic acid in tobacco is a UDPG-glucose: flavonoid glucosyltransferase (Horvath and Chua, 1996). Although it has been postulated that this enzyme may function in the protection of cells against the damaging effects of light, it seems possible that kaempferol conjugation could play a role in defense activation parallel to the present hypothesis for soybean.

The conjugation and release of various aglycones from their conjugates could thus represent a key point for regulation of a number of highly specific processes. A common criterium for such a regulatory phenomenon is that it show a high degree of specificity. Thus, if they play a regulatory role, the enzymes responsible for release of flavonoid aglycones should show specificity towards their conjugates and would have to be in the proper compartment for such a role. Such enzymes have now been characterized in a number of systems. Barz and coworkers isolated both an isoflavone specific esterase (Hinderer *et al.*, 1986) and glucosidase (Hösel and Barz, 1975) from chickpea. Likewise, we have recently purified an isoflavone specific glucosidase from soybean (Hsieh M. C. and Graham T. L., unpublished). The soybean enzyme has somewhat different properties than that from chickpea and is able to directly cleave the glucose residue from the 6"-malonylated isoflavones, thus obviating the need for a separate esterase.

The high degree of specificity possible in glucosidases is indicated by the fact that both the soybean and chickpea isoflavone-specific glucosidases seem to have specificity towards isoflavone-7-O-glucosides. Neither has substantial activity on isoflavones with other substitution patterns, nor any of the flavonol glycosides tested to date. Thus, as with the genes for the glycosyl transferases, a series of aglycone specific hydrolases may exist and play a central role in the regulation of the activities of secondary products. Another glycosidase with a particularly high degree of specificity is involved in hydrolysis of the lignin precursor, coniferin (Dharmawardhana *et al.*, 1995). Other developmentally related or defense related glycosidic conjugates exist, such as conjugates of cytokinins, auxins, and those of the SAR-related signal, salicylic acid. We are currently cloning the various members of the soybean glucosidase gene family and have also identified a series of expressed sequence tags in *Arabidopsis* with high homology to glucosidases. Determination of the function of these various enzymes will be quite interesting as many of them may play regulatory roles. This work, and parallel research on the glucosyl transferases, should also very effectively complement further investigations on the flavonol and other glycoside mutants.

Conclusion

Rapid headway has been made in deciphering the flavonoid pathways in *Arabidopsis*. The application of additional mutant screening protocols, such as HPLC, is leading to the definition of new mutants. Other approaches, such as applying the yeast two hybrid system to identify protein-protein interactions with existing enzymes and regulatory factors may also yield useful information. The complexity of the flavonoid glycosyl conjugates and their potential biological function deserves particular attention. Especially intriguing are the potential regulatory roles of the specific glucosyltransferases and glycosidases involved in their formation, release and turnover.

(Received June 6, 1997; accepted July 8, 1997)

References

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

Barnes S. A., Quaggio R. B., Whitelam G. C. and Chua N. H., 1996. Fhy1 defines a branch point in phytochrome A signal transduction pathways for gene expression. *Plant J.*, **10**, 1155-1161.

Batschauer A., Rocholl M., Kaiser T., Nagatani A., Furuya M. and Schaefer E., 1996. Blue and UV-A light-regulated CHS expression in *Arabidopsis* independent of phytochrome A and phytochrome B. *Plant J.*, **9**, 63-69.

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

Buttery B. R. and Buzzell R. I., 1973. Varietal differences in leaf flavonoids of soybeans. *Crop Sci.*, **13**, 103-106.

Buttery B. R. and Buzzell R. I., 1975. Soybean flavonol glycosides: identification and biochemical genetics. *Can. J. Bot.*, **53**, 219-224.

Buttery B. R. and Buzzell R. I., 1987. Leaf traits associated with flavonol glycoside genes in soybean. *Plant Physiol.*, **85**, 20-21.

Buttery B. R., Gaynor J. D., Buzzell R. I., Mactavish D. C. and Armstrong R. J., 1992. The effects of shading on kaempferol content and leaf characteristics of five soybean lines. *Physiol. Plant.*, 86 , 278-284.

Buzzell R. I., Buttery B. R. and Bernard R. L., 1977. Inheritance and linkage of a magenta flower gene in soybean. *Can. J. Genet. Cytol.*, **19**, 749-751.

Chapple C.C.S., Shirley B.W., Zook M., Hammerschmidt R. and Somerville S.C., 1994. Secondary metabolism in *Arabidopsis*. In *Arabidopsis*, Somerville C. and Meyerowitz E., ed., Cold Spring Harbor Laboratory Press, 989-1030.

Chory J. and Peto C. A., 1990. Mutations in the *DET1* gene affect cell-type-specific expression of light-regulated genes and chloropast development in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA*, **87**, 8776-8780.

Chory J., Nagpal P. and Peto C. A., 1991. Phenotypic and genetic analysis of *det2*, a new mutant that affects light-regulated seeding eevelopment in *Arabidopsis*. *Plant Cell*, **3**, 445-460.

Chory J., Chatterjee M., Cook R. K., Elich T., Franauser C., Li J., Nagpal P., Neff M., Pepper A., Poole D., Reed J. and Vitart V., 1996. From seed germination to flowering, light controls plant development via the pigment phytochrome. *Proc. Natl. Acad. Sci. USA*, **93**, 12066-12071.

Christie J. M. and Jenkins G. I., 1996. Distinct UV-B and UV-A.blue ligth signal transduction pathways induce chalcone synthase gene expression in *Arabidopsis* cells. *Plant Cell*, **8**, 1555-1567.

Deikman J. and Hammer P. E., 1995. Induction of anthocyanin accumulation by cytokinins in *Arabidopsis thaliana*. *Plant Physiol.*, **108**, 47-57.

Dharmawardhana D. P., Ellis B. E. and Carlson J. E., 1995. A beta-glucosidase from lodgepole pine xylem specific for the lignin precursor coniferin. *Plant Physiol.*, **107**, 331-339.

Dooner H. K., Robbins T. P. and Jorgensen R. A., 1991. Genetic and developmental control of anthocyanin biosynthesis. *Annu. Rev. Genet.*, **25**, 173-199.

Feinbaum R. L. and Ausubel F. M., 1988. Transcriptional regulation of the *Arabidopsis thaliana* chalcone synthase gene. *Mol. Cell. Biol.*, **8**, 1985-1992.

Feys B. J. F., Benedetti C. E., Penfold C. N. and Turner J. G., 1994. *Arabidopsis* mutants selected for resistance to the phytotoxin coronatine are male sterile, insensitive to methyl jasmonate and resistant to a bacterial pathogen. *Plant Cell*, **6**, 751-759.

Fuglevand G., Jackson J. A. and Jenkins G. I., 1996. UV-B, UV-A, and blue light signal transduction pathways interact synergistically to regulate chalcone synthase gene expression in *Arabidopsis*. *Plant Cell*, **8**, 2347-2357.

Graham T. L., 1991. A rapid, high resolution high performance liquid chromatography profiling procedure from plant and microbial secondary metabolites. *Plant Physiol.*, **95**, 584-593.

Graham T. L., 1995. Cellular biochemistry of phenylpropanoid responses of soybean to infection by *Phytophthora sojae*. In *Handbook of phytoalexin metabolism and action*, Danile M. and Purkayastha R. P., ed., Marcel Dekker, New York, 85-116.

Hahlbrock K., 1981. Flavonoids. In *The biochemistry of plants: A comprehensive treatise*, 7, Stumpf P. K. and Conn E. E., ed., Academic Press, New York, 425-455.

Hinderer W., Kîster J. and Barz W., 1986. Purification and properties of a specific isoflavone 7-O-glucoside-6"-malonate malonylesterase from roots of chickpea (*Cicer arietinum* L.). *Archiv. Biochem. Biophys.*, **248**, 570-578.

Horvath D. M. and Chua N.-H., 1996. Identification of an immediate-early salicylic acid-inducible tobacco gene and characterization of induction by other compounds. *Plant Mol. Biol.*, **31**, 1061-1072.

Hîsel W. and Barz W., 1975. β-Glucosidases from *Cicer arietinum L.* : Purification and properties of isoflavone-7-O-glucoside-specific β-glucosidases. *Eur. J. Biochem.*, **57**, 607-616.

Jackson J. A. and Jenkins G. I., 1995. Extension-growth responses and expression of flavonoid biosynthesis genes in the *Arabidopsis hy4* mutant. *Planta*, **197**, 233-239.

Jackson J. A., Fuglevand G., Brown B. A., Shaw M. J. and Jenkins G. I., 1995. Isolation of *Arabidopsis* mutants altered in the light-regulation of chalcone synthase gene expression using a transgeneic screening approach. *Plant J.*, **8**, 369-380.

Koornneef M., 1990. Mutations affecting the testa color in *Arabidopsis*. *Arabidopsis Information Service*, **28**, 1-4.

Koornneef M., Luiten W., de Vlaming P., and Schram A. W., 1982. A gene controllling flavonoid-3'-hydroxylation in *Arabidopsis*. *Arabidopsis Information Service*, **19**, 113-115.

Koornneef M., van Eden J., Hanhart C. J., Stam P., Braaksma F. J. and Feensta W. J., 1983. Linkage map of *Arabidopsis thaliana*. *J. Hered.*, **74**, 265-272.

Kubasek W. L., Shirley B. W., McKillop A., Goodman H. M., Briggs W. and Ausubel F. M., 1992. Regulation of flavonoid bisynthetic genes in germinating *Arabidopsis* seedlings. *Plant Cell*, **4**, 1229-1236.

Lacomme C. and Roby D., 1996. Molecular cloning of a sulfotransferase in *Arabidopsis thaliana* and regulation during development and in response to infection with pathogenic bacteria. *Plant Mol. Biol.*, **30**, 995-1008.

Landry L.G., Chapple C.C.S. and Last R. L., 1995. *Arabidopsis* mutants lacking phenolic sunscreens exhibit enhanced ultraviolet-B injury and oxidative damage. *Plant Physiol.*, **109**, 1159-1166.

Lee D., Ellard M., Wanner L.A., Davis K. R. and Douglas C.J., 1995. The *Arabidopsis thaliana* 4-coumarate:CoA ligase (4CL) gene: Stress and developmentally regulated expression and nucleotide sequence of its cDNA. *Plant Mol. Biol.*, **28**, 871-884.

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

Li J., Ou Lee T. M., Raba R., Amundson R. G. and Last R. L., 1993. *Arabidopsis* flavonoid mutants are hypersensitive to UV-B irradiation. *Plant Cell*, **5**, 171-179.

Lin C., Ahmad M. and Cashmore A. R., 1996. *Arabidopsis* cryptochrome 1 is a soluble protein mediating blue-light-dependent regulation of plant growth and development. *Plant J.*, **10**, 893-902.

Lloyd A. M., Walbot V. and Davis R. W., 1992. *Arabidopsis* and *Nicotiana* anthocyanin production activated by maize regulators *R* and *C1*. *Science*, **258**, 1773-1775.

Lloyd A. M., Schena M., Walbot V. and Davis R. W., 1994. Epidermal cell fate determination in *Arabidopsis*: Patterns defined by a steroid-inducible regulator. *Science*, **266**, 436-439.

Lois R. and Buchanan B. B., 1994. Severe sensitivity to ultraviolet radiation in an *Arabidopsis* mutant deficient in flavonoid accumulation: II. Mechanism of UV-resistance in *Arabidopsis*. *Planta*, **194**, 504-509.

Marrs K. A., Alfenito M. R., Lloyd A. M. and Walbot V., 1995. A glutathione S-transferase involved in vacuolar transfer encoded by the maize gene Bronze-2. *Nature*, **375**, 397-400.

Mckown R., Kuroki G. and Warren G., 1996. Cold responses of *Arabidopsis* mutants impaired in freezing tolerance. *J. Exp. Bot.*, **47**, 1919-1925.

Meyer K., Cusumano J. C., Somerville C. and Chapple C.C.S., 1996. Ferulate-5-hydroxylase from *Arabidopsis thaliana* defines a new family of cytochrome P450-dependent monooxygenases. *Proc. Natl. Acad. Sci. USA*, **93**, 6869-6874.

Misera S., Mueller A. J., Weiland-Heidecker U. and Jurgens G., 1994. The *FUSCA* genes of *Arabidopsis*: Negative regulators of light response. *Mol. Gen. Genetics*, **244**, 242-252.

Mittal S., 1996. Biochemical, genetic and molecular analysis of the phenylpropanoid pathway in *Arabidopsis thaliana*. *Ph. D. Thesis*, The Ohio State University.

Ohl S., Hederick S.A., Chory J. and Lamb, C. J., 1990. Functional properties of a phenylalanine ammonia-lyase promoter from *Arabidopsis*. *Plant Cell*, **2**, 837-848.

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 Physiol.*, **111**, 339-345.

Sharma Y. K. and Davis K. R., 1994. Ozone-induced expression of stress-related genes in *Arabidopsis thaliana*. *Plant Physiol.*, **105**, 1089-1096.

Sheahan J.J., 1996. Sinapate esters provide greater UV-B attenuation than flavonoids in *Arabidopsis thaliana* (Brassicaceae). *Am. J. Bot.*, 83, 679-686.

Shirley B. W., 1996. Flavonoid biosynthesis: new' functions for an old' pathway. *Trends Plant Sci.*, **1**, 377-382.

Shirley B.W., Kubasek W. L., Storz G., Bruggemann E., Koornneef M., Ausubel F. M. and Goodman H.M., 1995. Analysis of *Arabidopsis* mutants deficient in flavonoid biosynthesis. *Plant J.*, **8**, 659-671.

Szekeres M. and Koncz C., 1998. Biochemical and genetic analysis of brassinosteroid metabolism and function in *Arabidopsis*. *Plant Physiol. Biochem.*, **36**, 145-155.

Wanner L.A., Li G., Ware D., Somssich I. E. and Davis K. R., 1995. The phenylalanine ammonia-lyase gene family in *Arabidopsis thaliana*. *Plant Mol. Biol.*, **27**, 327-338.

Yanai Y., Kawasaki T., Shimada H., Wurtele E. S., Nikolau B. J. and Ichikawa N., 1995. Genomic organization of 251 kDa acetyl-CoA carboxylase genes in *Arabidopsis*: Tandem gene duplication has made two differently expressed isozymes. *Plant Cell Physiol.*, **36**, 779-787.

Ylstra B., Muskens M. and Van Tunen A. J., 1996. Flavonols are not essential for fertilization in *Arabidopsis thaliana*. *Plant Mol.Biol.*, **32**, 1155-1158.