Biochemical and Molecular Characterization of a Novel Dioxygenase, Flavonol 6-Hydroxylase, Involved in Flavonol Biosynthesis from *Chrysosplenium americanum*

Dominique Anzellotti

A Thesis

in

The Department

of

Biology

Presented in Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy at

Concordia University

Montréal, Québec, Canada

May 2004

Your file Votre référence
ISBN: 0-612-96938-X
Our file Notre référence
ISBN: 0-612-96938-X

ABSTRACT

Biochemical and Molecular Characterization of a Novel Dioxygenase, Flavonol 6-Hydroxylase, Involved in Flavonol Biosynthesis from *Chrysosplenium americanum*

Dominique Anzellotti, Ph.D.

Concordia University, 2004

Plants synthesize a vast array of natural products, collectively known as secondary metabolites, of which flavonoids constitute an important group. Flavonoid compounds are ubiquitous in Nature and play important roles in plant physiology, development and the plant's interaction with its environment. They contribute to flower color, ultraviolet protection, cell signalling and defence against plant pathogen attack, in addition to their potential benefits to mammalian organisms. In most cases, structure-activity relationships have been observed and, therefore, modification of the substitution pattern of certain compounds may enhance the quality and/or performance of plants through metabolic engineering.

Chrysosplenium americanum, a semi-aquatic weed, elaborates a variety of highly methylated flavonoids that are synthesized by regio-specific enzymes, and are believed to protect the plant against microbial attack. Several of the substitution reactions involved in the proposed pathway of their biosynthesis previously have been elucidated. One of the enzymes involved at the branchpoint of this pathway, the flavonol 6-hydroxylase (F6H) has been the primary focus of this research. The novelty of this enzyme lies in the fact that it is the only reported α-ketoglutarate-dependent dioxygenase that introduces a

phenolic hydroxyl group on the flavonoid ring; more specifically at position 6 of trimethylquercetin, which is subsequently methylated. The dioxygenase nature of the enzyme has been established and the hydroxylation reaction has been characterized at the biochemical level, including its physicochemical properties and kinetic mechanism. The F6H protein was purified to near homogeneity in its functional state, through an original strategy involving conventional chromatography and affinity columns, and partial amino acid sequences of the purified protein were obtained.

The microsequence information obtained was used as a guide in primer design for the cloning of the *F6H* cDNA from a *C. americanum* library that was screened using a PCR-based strategy. A cDNA fragment was subsequently isolated and cloned and its gene product was characterized at the biochemical and molecular levels. The protein predicted from the *F6H* cDNA clone contains the conserved motifs for this class of enzymes, as well as the microsequences obtained from the purified plant protein. In addition, the clone exhibits F6H activity when expressed in both prokaryotic and eukaryotic expression systems. At the molecular level, *F6H* is present as a single copy in the *C. americanum* genome, and contains two introns of approximate length and position comparable to a particular class of dioxygenases, suggesting an evolutionary relationship. This is supported by the results of a phylogenetic analysis of *F6H* and related genes, where *F6H* clusters with flavanone 3-hydroxylase, a dioxygenase considered to be an "early gene" with respect to its induction given its location in the flavonoid biosynthetic pathway and expression pattern. The functional significance of the gene encoding *F6H* is discussed in relation to the metabolic engineering of flavonoid biosynthesis in plants.

DEDICATION

To my parents for

their love and support.

To my grandparents for

their encouragement.

ACKNOWLEDGMENTS

There is no way for me to adequately express my thanks and gratitude to those who have contributed to my graduate studies. I would like to thank my fellow graduate students and colleagues in the Biology and Biochemistry departments who have provided not only invaluable insights and discussions, but also encouragement as well. I would particularly like to thank Dr. Ingrid Muzac, Kevin Lam and Damiano Ferraro for all their help and attention, technical and otherwise. In addition, I would like to express my gratitude to Dr. Christian NDong and Jean-Benoit Frenette-Charron for their trust, support and encouragement.

My committee members, Dr. Luc Varin and Dr. Partick Gulick, have always been available for discussions and provided useful advice whenever necessary. I am grateful to them for their attention and the generous use of their laboratory resources. Dr. Selvadurai Dayanandan, who I consider an honorary committee member, has been extremely helpful, patient and understanding. I appreciate his insightful advice and contributions, as well as his enthusiasm.

Most importantly, I would like to express my sincere gratitude to my supervisor, Dr. Ragai K. Ibrahim, who has been an excellent role model for a young scientist. He has provided me with opportunities and challenges that have helped me to grow as a researcher throughout my graduate studies. I am deeply appreciative of his generosity, guidance and support.

I am very grateful to Jarred Chicoine for his support, both intellectually and emotionally, encouragement and assistance. I would also like to thank my family and friends who have been a constant source of encouragement. Lastly, I would like to acknowledge the support of the granting agencies, NSERC, FCAR and NRCan.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ACCO	aminocyclopropane carboxylate oxidase
AdoMet	S-adenosyl-*L*-methionine
ANS	anthocyanidin synthase (LDOX, leucoanthocyanidin dioxygenase)
cF6H	F6H cDNA
CHI	chalcone isomerase
CHS	chalcone synthase
D4H	desacetoxyvindoline 4-hydroxylase
DFR	dihydroflavonol reductase
F3H	flavanone 3-hydroxylase
F6H	flavonol 6-hydroxylase
FLS	flavonol synthase
FPLC	fast protein liquid chromatography
FS	flavone synthase
GA(s)	gibberellin(s)
GA2H	gibberellin 2-hydroxylase
GA3H	gibberellin 3-hydroxylase
GA7Ox	gibberellin 7-oxidase
GA20Ox	gibberellin 20-oxidase
gF6H	genomic F6H
H6H	hyoscyamine 6-hydroxylase
HPLC	high performance (pressure) liquid chromatography
IFS	isofavone synthase
IPNS	isopenicillin N-synthase
αKDD	α-ketoglutarate-dependent dioxygenase
α-KG	α-ketoglutarate
LOX	lipoxygenase
ORF	open reading frame
P4H	prolyl 4-hydroxylase
HPPD	hydroxyphenylpyruvate dioxygenase
PMSF	phenylmethylsulfonylfluoride
rF6H	recombinant F6H
rF6Hb	recombinant F6H expressed in bacteria (*E. coli*)
rF6Hy	recombinant F6H expressed in yeast (*P. pastoris*) - intracellular
rF6Hys	recombinant F6H expressed in yeast (*P. pastoris*) - secreted
TriMeQ	3,7,4'-trimethylquercetin
TriMeQg	3,7,4'-trimethylquercetagetin

A. INTRODUCTION

Plant secondary metabolites function primarily as defence compounds and signaling molecules. Flavonoids, one of the largest groups of secondary metabolites, play important roles in the physiology, ecology and defence of plants, such as flower pigments, ultraviolet protectants and antimicrobial compounds, to mention a few (Bohm, 1998 and references therein). This wide range of activities is the result of their structural diversity that encompasses a variety of functional group substitutions (Ibrahim and Anzellotti, 2003).

Chrysosplenium americanum (Saxifragaceae), a semi-aquatic weed, accumulates a variety of tetra- and penta methylated flavonol glucosides that are substituted at positions 3,6,7,2′,4′ and/or 5′ of the flavonol ring (Collins *et al.*, 1981; (De Luca and Ibrahim, 1985a). Their biosynthesis from the parent aglycone, quercetin is catalyzed by a number of stepwise, substrate-specific, position-oriented *O*-methyltransferases (De Luca and Ibrahim, 1985a; Khouri *et al.*, 1986) and two distinct *O*-glucosyltransferases (Latchinian *et al.*, 1987; Latchinian and Ibrahim, 1989) in a highly ordained sequence (Figure 1).

During the course of their biosynthesis, the partially methylated intermediate, 3,7,4′-trimethylquercetin is converted to 3,7,4′-trimethylquercetagetin after hydroxylation at position 6 by an α-ketoglutarate-dependent dioxygenase (αKDD) flavonol 6-hydroxylase, and subsequently *O*-methylated at that position. 3,7,4′-Trimethylquercetagetin serves as a substrate for further substitution reactions leading to the formation of pentamethylated flavonol glucosides. In addition, it has been proposed

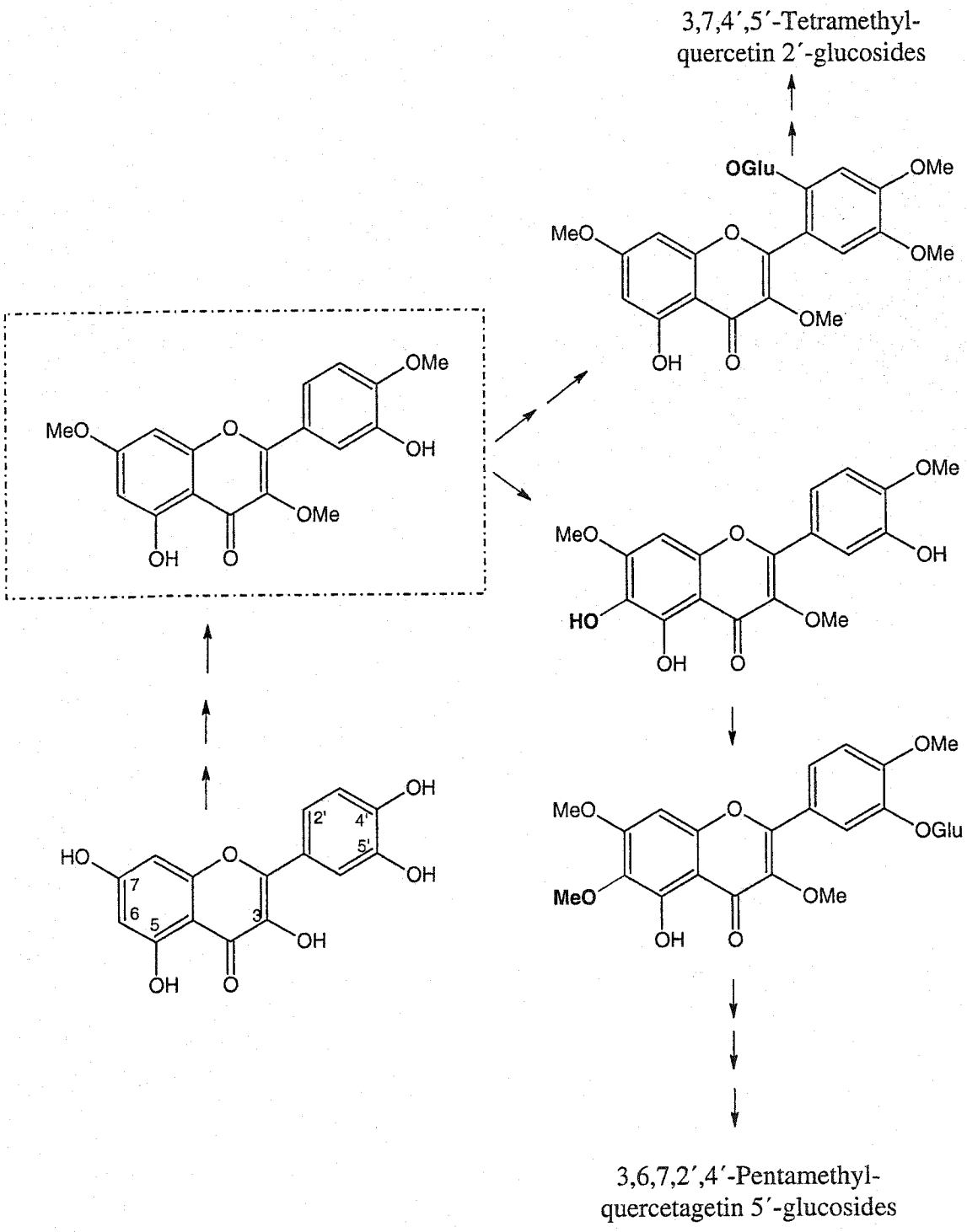


Figure 1: Polymethylated flavonol biosynthesis in *C. americanum*

that flavonols are biosynthesized through a channelled enzymatic pathway in *C. americanum* (Ibrahim, 1992), and that flavonol 6-hydroxylase may be associated with such enzyme aggregates in order to efficiently catalyze the production of highly substituted flavonol derivatives.

Plant αKDDs (EC 1.14.11.-) constitute a class of hydroxylases composed of non-heme, iron-containing cytosolic enzymes that utilize an α-keto-acid as a cosubstrate and require ascorbate as a reducing agent for the reactive iron moiety. This widespread class of enzymes has been implicated in a variety of plant metabolic pathways, including the biosynthesis of some amino acids, hormones, signaling molecules and a variety of secondary metabolites (Prescott and John, 1996). The hydroxylation at position 6 of flavonols is of particular interest, given that such aromatic hydroxylations are usually carried out by cytochrome P450-dependent monooxygenases. Notably however, there are several examples where a particular hydroxylation reaction may be carried out by either class of oxygenase enzymes. Apart from hydroxylation reactions, αKDDs catalyze desaturations, expansions, epoxidations and ring closures, depending on the substrate and the source organism. Given the variety of reactions catalyzed by αKDDs, it has been difficult to elucidate a common reaction mechanism or intermediate that would allow a more accurate categorization of such enzymes. As a result of their multifunctional nature, as well as the fact that they are involved in various metabolic pathways among all biological systems, recent research has focused on understanding the effects of αKDD activity and the evolution of their genes.

As each plant species can produce only a particular set of flavonoid molecules, it is not uncommon to find a plant possessing a high nutritional value, but low resistance to

pathogen attack. Likewise, the medicinally active flavonoids found in numerous species are only produced under certain environmental conditions or in minute quantities, thus rendering their recovery from such plants quite expensive and time-consuming. In many cases, it is not only the class of compound, but also the type and degree of modification of the parent molecule that determines its effectiveness. Our understanding of how and why a particular plant elaborates these compounds at the biochemical and molecular levels can be used to enhance the performance of nutritively valuable species, and contribute to the development of productive alternatives to these rare plants as the sole sources of important drugs.

B. AIM OF WORK

One of the missing steps in the pathway of polymethylated flavonol biosynthesis in *C. americanum* is the hydroxylation step at position 6 of the partially methylated intermediate, 3,7,4′-trimethylquercetin. Therefore, the studies included in this dissertation will mainly focus on the biochemical and molecular aspects of this novel enzyme, the flavonol 6-hydroxylase (F6H), as described below.

1. **To purify the enzyme protein to near homogeneity with the ultimate goal of obtaining internal peptide sequence information for the cloning of its gene.** A purification scheme involving both conventional and original affinity chromatographic techniques will be employed. These experiments also will allow us to obtain relatively purified protein for kinetic measurements and to assess the cross-reactivity of antibodies generated against a heterologous dioxygenase, thus indicating the presence of putatively shared epitopes and motifs.

2. **To investigate the F6H enzyme protein at the biochemical level with regards to its physicochemical characteristics and kinetic properties.** Data obtained from these experiments will establish the dioxygenase nature of the F6H, as well as provide information as to the cofactor requirements, potential inhibitors, substrate preference and reaction order for the enzyme

reaction. This information then can be compared to that reported for other proteins of this class in order to characterize the reaction mechanism.

3. **To isolate a cDNA clone encoding the F6H for further biochemical and molecular studies of the recombinant protein.** This clone will be expressed in a heterologous expression system in order to produce sufficient amounts of purified, enzymatically active protein for further analyses. The labile nature of the native plant protein and its low abundance render its purification and kinetic studies tedious and labour-intensive. These obstacles may be avoided with the use of a recombinant protein exhibiting F6H activity. Different expression systems will be tested in order to determine the conditions amenable to recombinant F6H activity.

4. **Genomic analyses will provide insight into the organization of the gene coding for this dioxygenase.** Certain αKDDs are encoded by multigene families, while others are present as a single copy; this will be determined for the *F6H* in *C. americanum*. The number, position and sequence of potential introns will be determined and used for comparison with other genes of this class.

5. **Phylogenetic analyses of the *F6H* gene and related dioxygenases involved in plant secondary metabolism.** This will provide an insight into the evolution of such dioxygenases, as well as the relatedness of *F6H* to other

genes. Inferences can be made regarding any shared motifs and how they may affect the reaction mechanism, especially if a particular sequence is closely related to *F6H*, and if such mechanistic information is available.

6. **A preliminary study of closely related *Chrysosplenium* species to assess the presence of potential F6H homologs.** A biochemical characterization of the enzymes responsible for the hydroxylation of flavonols produced by such species will provide information regarding the existence of potential homologs. The existence of F6H homologs is significant since the type of reaction catalyzed, *i.e.*, the introduction of a hydroxyl group on an aromatic ring system, is uncommon for a dioxygenase. Therefore, analysis of any such homologs will provide a valuable insight as to how this reaction is catalyzed in contrast to other dioxygenase-catalyzed reactions.

The elucidation of F6H activity is significant for several reasons. First, it identifies and characterizes a novel enzyme activity, not only for its preferred substrate, but also for the type of reaction catalyzed by this category of enzymes. Second, it will provide further evidence in support of the proposed biosynthetic pathway of highly substituted flavonoids in *C. americanum*. In addition, it may advance our understanding of this particularly novel hydroxylation step, with respect to its biochemical requirements and its order in the biosynthetic sequence.

C. LITERATURE REVIEW

C.1 Functions of Flavonoids in Nature

Plants produce a diverse array of metabolites that do not directly participate in primary metabolism. These secondary metabolites appear to function primarily in enhancing plant defense and reproductive success. Flavonoids constitute one of the major classes of secondary metabolites that play important roles in plant growth, development and defense. This group of metabolites, consisting of over 5000 differently substituted compounds, owe their numerous physiological activities to the extensive structural diversity they exhibit (Croteau *et al.*, 2000).

Flavonoids include the chalcones, flavanones, flavones, isoflavones, flavonols and anthocyanidins as the major subclasses. As a result of their structural diversity, they exhibit a wide spectrum of activities. The conjugated nature of the flavonoid moiety allows for the absorption of ultraviolet radiation in the range of 280-320 nm. In fact, it has been postulated that their ultraviolet B-filtering nature may have served as the driving force for their selection and played a role in the evolution of land plants (Stafford, 1991). Also related to their ability to absorb ultraviolet radiation, flavonoids used as nectar guides can attract insect pollinators. Additionally, anthocyanins serve as visible flower pigments that also attract insect and bird pollinators (Saito and Harborne, 1992). A widespread effect of flavonol aglycones on pollen germination and pollen tube growth has been established in several, but not all, plant species. The absence of such compounds, resulting from a deficiency in the enzymes responsible for their biosynthesis or substitution, results in a reduced seed set or male sterility (Taylor and Jorgensen,

1992). Interactions between soil microorganisms and leguminous plant species can be mediated by root flavonoids. These interactions can be quite complex and include stimulation of bacterial growth, spore germination, hyphal growth and branching (Pueppke, 1996).

Insect attack often leads to physical injury, thereby providing a means of entry for pathogenic microorganisms. The amounts of certain flavonoid compounds are increased in response to such an attack, thus serving as physical barriers to prevent or retard microbial invasion. Flavonoids also can serve to protect the plant against viral infection, herbivores and other invading plants. This may be accomplished through different means; by synthesizing phytoalexins, compounds produced *de novo* in response to infection, whereas allelopathy is a phenomenon whereby a compound serves as a signal to discourage invasion by another organism. Lastly, numerous therapeutic uses for flavonoids have been demonstrated as anti-microbial, anti-tumorigenic and anti-viral compounds, as well as modulators of the immune and inflammatory responses (Middleton and Kandaswami, 1994). One example of their antiviral activity is produced through the weakening of interactions between virus coat proteins, as well as increasing viral susceptibility to host RNases (French and Towers, 1992). Other important roles of flavonoids can be found in several specialized reviews (Rice-Evans and Miller, 1996; Bohm, 1998; Hollman and Katan, 1998; Dixon and Steele, 1999).

C.2 Flavonoid Biosynthesis

The basic structure of a flavonoid consists of two aromatic rings, A and B, which are joined by a three-carbon unit forming the heterocyclic ring C. The aromatic ring A is

derived from three acetate units through polyketide synthesis, whereas rings B and C originate from a phenylpropanyl residue derived from L-phenylalanine through the shikimate pathway (Figure 2). Briefly, flavonoid biosynthesis begins with the stepwise condensation of one molecule of 4-coumaroyl CoA with three units of malonyl CoA, catalyzed by chalcone synthase (CHS), the first committed step in flavonoid biosynthesis. This results in the formation of a naringenin chalcone, the first stable intermediate in flavonoid biosynthesis. The latter is converted to a flavanone by chalcone isomerase (CHI), resulting in *2S*-naringenin with a heterocyclic ring C. Flavanones represent important branching-point intermediates in flavonoid metabolism, although they may accumulate in different substituted forms. They are sequentially oxidized to give rise to different flavonoid subclasses distinguished according to the oxidation level of the heterocyclic ring C.

Conversion of flavanones to dihydroflavonols, by hydroxylation of position 3, is catalyzed by flavanone 3β-hydroxylase (F3H), and the subsequent desaturation of the C2-C3 bond by flavonol synthase (FLS) results in the formation of flavonols. Flavanones also serve as substrates for flavone synthase (FS), which desaturates the C2-C3 bond to produce flavones by a mechanism different from that of FLS (Turnbull *et al.*, 2004). FS is found predominantly in plants as a membrane-associated monooxygenase (FSII); however, certain members of the Apiaceae, such as parsley, catalyze this reaction by a soluble αKDD-dependent enzyme (FSI). The colored pigments anthocyanidins, are produced from dihydroflavonol precursors through the sequential activities of two enzymes, dihydroflavonol reductase (DFR) and anthocyanidin synthase (ANS). F3H, FLS, ANS have been shown to be

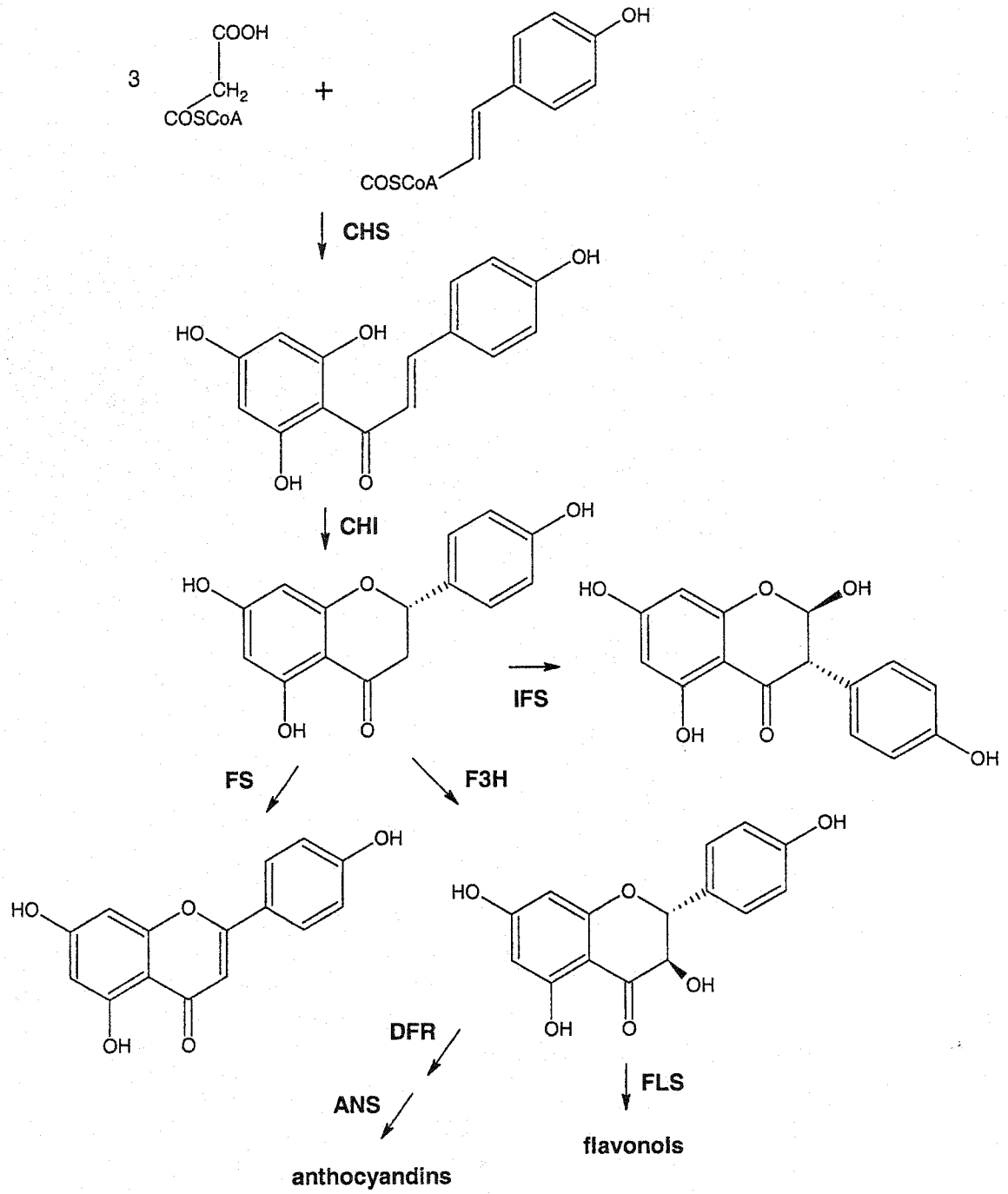


Figure 2: Flavonoid biosynthetic pathway

αKDD-dependent enzymes in numerous plant species and participate in the modification of the flavonoid C ring. The enzyme reactions involved in flavonoid biosynthesis have been reviewed (Forkmann and Heller, 1999) and are depicted in Figure 2. Subsequent tissue- and species-specific substitution reactions, such as hydroxylations, methylations, sulfations, glycosylations, acylations and prenylations result in the vast array of structurally diverse groups of flavonoids (Ibrahim and Anzellotti, 2003).

C.3 Polymethylated Flavonol Biosynthesis in *C. americanum*

C. americanum accumulates a variety of polymethylated flavonol glucosides derived from the parent aglycone, quercetin. The final products found in this tissue are the 2′- and 5′-glucosides of 3,7,4′,5′-tetramethylquercetin and 3,6,7,2′,4′,-pentamethylquercetagetin, respectively, whereas the low (mono-, di- and tri-) methylated intermediates do not accumulate in this plant.

The initial steps of the biosynthetic pathway (Figure 1) involve the sequential *O*-methylation of quercetin to produce 3,7,4′-trimethylquercetin by three successive position-specific *O*-methyltransferases (De Luca and Ibrahim, 1985a; Khouri *et al.*, 1986; Ibrahim *et al.*, 1987). At this point two separate hydroxylation reactions take place, one of which results in the formation of trimethylquercetagetin (trimethyl-6-hydroxyquercetin), the substrate for subsequent methylation and glucosylation reactions. The second hydroxylation takes place at position 2′ of the trimethylated quercetin intermediate. *O*-Glucosylation of positions 2′ and 5′ is catalyzed by two separate glucosyltransferases (Latchinian *et al.*, 1987; Latchinian and Ibrahim, 1989). Given the absence of lignified tissues, that would otherwise serve as a protective barrier, this semi-

aquatic plant may compensate for this susceptibility by accumulating highly methylated metabolites believed to protect the tissue against pathogen attack and UV damage.

The fact that *C. americanum* produces 6-hydroxylated flavonol derivatives, makes it a good model for the investigation of their biosynthesis. Hydroxylation at position 6 is an uncommon substitution reaction that occurs in this plant only after partial methylation of the flavonoid ring system. Whereas, hydroxylation of an aromatic carbon is often catalyzed by cytochrome P450-dependent monooxygenases, given the polymethylated nature of the substrate, a dioxygenase-dependent activity is not unexpected, since plant monooxygenases seem to intervene at an early, rather than later, stage in biosynthesis. In addition, the biosynthesis of polymethylated flavonols in this tissue has been thoroughly investigated, and a plausible pathway has been proposed (Figure 1). There is significant information regarding the biochemistry and localization of most enzymes involved in this pathway (Ibrahim, 1992), thus providing additional insight as to how and why 6-hydroxyflavonols are produced in this organism.

C.4 Oxygenases

Oxygenases play a significant role in the metabolism of biological compounds. Monooxygenases, also referred to as cytochrome P450-dependent, mixed-function oxidases incorporate one atom of molecular oxygen into the substrate, while the second atom is reduced to water. These microsomal enzymes require the presence of a reductant, commonly NADPH, for the reaction to proceed (Nozaki, 1979). On the other hand, dioxygenases are cytosolic enzymes that catalyze the incorporation of both oxygen atoms into their substrates (De Carolis and De Luca, 1994). Dioxygenases are involved in

several types of reactions, namely ring cleavages, double hydroxylations, α-ketoglutarate-requiring reactions and fatty acid metabolism (De Carolis and De Luca, 1994). Plant dioxygenases can be divided into two main categories: lipoxygenases (LOX) and α-ketoglutarate dependent dioxygenases (αKDDs). The former enzymes are involved in unsaturated fatty acid oxygenation, and participate in the production of jasmonic and abscisic acids from fatty acid hydroperoxides and carotenoids, respectively (Gardner, 1991). In addition to participating in flavor and aroma compound production, LOX activity has been associated with the loss of essential nutrients through oxidation (Donnelly and Robinson, 1995).

Plant αKDDs constitute a class of hydroxylases composed of non-heme, iron-containing soluble enzymes. This widespread class of enzymes has been implicated in a variety of plant metabolic pathways, including the biosynthesis of some amino acids, hormones, signaling molecules and a variety of secondary metabolites (Prescott and John, 1996). Apart from hydroxylation reactions, αKDDs catalyze desaturations, expansions, epoxidations, and ring closures, depending on the substrate and the source organism (Prescott, 1993).

C.5 Reaction Mechanism of αKDDs

αKDDs are characterized by their requirement for an α-keto-acid as a cosubstrate, typically α-ketoglutarate (α-KG), which is oxidatively decarboxylated to succinate in the process of the hydroxylation reaction (Figure 3; De Carolis and De Luca, 1994). However in certain cases, the keto-group can be provided by the primary substrate,

thereby eliminating the need for a cosubstrate (Fernandez-Maculet *et al.*, 1993; Garcia *et al.*, 1997).

$$\alpha\text{-KG} + O_2 + S \xrightarrow[\text{ascorbate}]{Fe^{+2}} S\text{-}OH + CO_2 + \text{succinate}$$

Figure 3: Reaction catalyzed by αKDDs

Nevertheless, all enzymes in this class exhibit an absolute requirement for ferrous ions which, through the generation of a ferryl oxidant intermediate, activate the otherwise stable dioxygen for incorporation into various substrates (Figure 4; Blanchard *et al.*, 1983). In order to maintain the iron in the catalytically active state, ascorbate serves as a reducing agent in most αKDD reactions. At the same time, ascorbate can act as the cosubstrate, causing an uncoupling of the hydroxylation reaction from the decarboxylation of α-KG, more often than not resulting in the formation of a dead-end complex and inactivation of the enzyme. This may be one of the main reasons for the labile nature of these enzymes.

αKDDs participate in the production or modification of various plant metabolites such as gibberellins, ethylene, alkaloids and flavonoids, among others. Given the dissimilar nature of such substrates, it is difficult to accurately predict the participation of αKDDs in various pathways. In general, compounds rich in oxygen do not serve as good substrates for such oxygenases as they do not require further activation, with several notable exceptions. In addition, αKDDs are typically involved in hydroxylation reactions at the later stages of a given pathway, whereas cytochrome P450 mixed-function monooxygenases are usually implicated in earlier biosynthetic steps. In addition, there are cases where the same reaction is catalyzed by both αKDDs and

cytochrome P450 monooxygenases (Prescott and John, 1996). It is clear, however, that the role of the dioxygenase is to harness the oxidative power of ferrous ions in order to efficiently oxygenate a given substrate, while minimizing the production of harmful intermediates and by-products such as peroxides or free radicals. The use of α-KG allows for the control of electron flow and the trapping of the oxygen adduct, thereby allowing the catalysis of a variety of reaction types. Nevertheless, potential α-KG-dependent oxygenation in a given biosynthetic pathway is often difficult to accurately predict, since the nature of αKDD substrates is very dissimilar, ranging from amino acids to phenolic compounds.

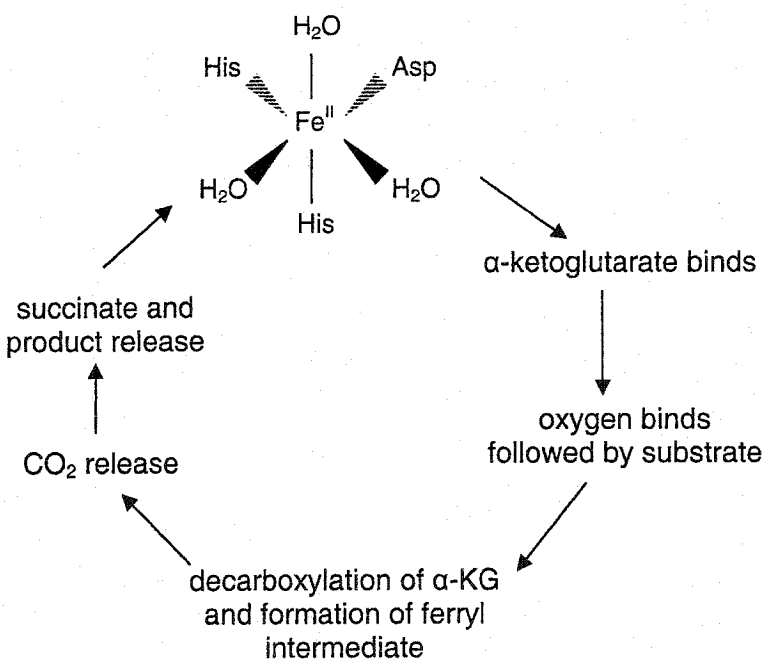


Figure 4: Simplified reaction mechanism of αKDDs (adapted from Prescott, 2000)

C.6 Plant αKDDs

C.6.1 Gibberellin Biosynthesis

Gibberellic acids (GAs), commonly known as gibberellins (Figure 5), regulate several aspects of plant growth and development, such as seed germination, shoot growth, flower induction, fruit set and fruit development (Phillips *et al.*, 1995). αKDDs catalyze several of the later steps of GA biosynthesis.

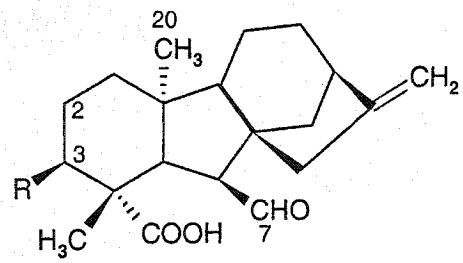


Figure 5: Basic structure of gibberellic acids

The most physiologically active GAs, hydroxylated at the C-3 position, are the result of GA 3β-hydroxylase (GA3H) activity. This enzyme utilizes various differentially-modified GA substrates and is subject to feedback inhibition (Lange, 1997). Although GA3H from *Phaseolus vulgaris* exhibits the typical αKDD cofactor requirements, it does not utilize α-KG as a cosubstrate. Rather, α-KG acts as an activating cofactor while ascorbate assumes the role of a secondary oxygen acceptor (Lange *et al.*, 1994). A gene encoding GA3H has been isolated and characterized from *Arabidopsis thaliana* (*GA4*), and a semi-dwarf mutant (*ga4*) has been identified as being deficient in GA3H activity and, thus exhibits reduced levels of C-3-hydroxylated GA derivatives (Chiang *et al.*, 1995). This mutant is the result of a single amino acid change (Cys→Tyr) resulting in a protein with little or no activity. In addition, overexpression of

ga4 transcripts was observed, thus indicative of feedback regulation. This was confirmed when exogenous application of active GAs resulted in a decrease in *ga4* expression. These experiments demonstrate that GA3H is under the control of molecular regulatory mechanisms, as well as direct regulation through product inhibition.

The inactivation of GAs results from the activity of GA 2β-hydroxylase (GA2H; GA 2β-oxidase) (Martin *et al.*, 1999). Therefore, the ratio of active to inactive GAs is determined, in large part, by the combined activity of two αKDDs, GA3H and GA2H. Multiple forms of GA2H have been detected in various species with different substrate specificities, suggesting the existence of a multigene family encoding GA2Hs that are differentially responsive to various regulatory mechanisms (Thomas *et al.*, 1999). The activities of three *GA2H* gene products from *Pisum sativum* have been confirmed, and pea seeds and young pea plants carrying the *slender* (*sln*) mutation exhibit no GA2H activity (Martin *et al.*, 1999). As a result of a single base deletion, *sln* produces a truncated, inactive protein. Given that the remaining GA2H isozymes do not complement this phenotype at this stage, it suggests that other members of this gene family are subject to differential regulation.

The above GA hydroxylases incorporate a hydroxyl group at a specific position; however, other GA αKDDs catalyze the oxidation of their substrate, through different mechanisms, and have thus been termed oxidases. Two other GA αKDDs have been identified that serve as major branch points in the biosynthetic pathway, and are subject to complex regulatory mechanisms (Lange, 1997). GA 7-oxidase (GA7Ox) catalyzes the conversion of GA12 aldehyde to a carboxylated derivative, that serves as a precursor to subsequent αKDDs (Lange, 1997). This reaction also can be catalyzed by a cytochrome

P450-dependent monooxygenase in *Cucurbita maxima* endosperm (Prescott and John, 1996). The existence of such parallel pathways is a characteristic feature not uncommon to αKDDs, and often leads to the production of differentially modified metabolites.

GA 20-oxidase (GA20Ox) catalyzes successive oxidations of a variety of GA substrates, ultimately resulting in the production of C-19 GAs through the loss of the C-20 carbon unit (Phillips *et al.*, 1995). The flexibility of the active site allows for differential substrate binding and sequential catalysis leading to the production of slightly different GA structures. GA20Ox regulation has been studied in detail, and this enzyme has been determined to be the primary target of feedback inhibition for this pathway (Phillips *et al.*, 1995). Mapping studies indicate that the *GA5* locus of *A. thaliana* encodes a GA20Ox (Coles *et al.*, 1999). Genomic analyses have demonstrated that the *Arabidopsis GA5,* as well as its pumpkin counterpart, belong to multigene families. Overexpression of *GA5* resulted in taller plants that flowered earlier, whereas antisense suppression produced *Arabidopsis* plants with shorter hypocotyls (Coles *et al.*, 1999). The above-mentioned genes are being considered as potential targets for the manipulation of plant development through transgenics.

Other αKDDs involved in GA biosynthesis are the 12β-oxidase and 13β-oxidase from *C. maxima* (Hedden *et al.*, 1984; Lange *et al.*, 1994), however less work has been carried out on these enzymes in comparison with the other GA hydroxylases and no cDNA clone has been characterized to date. However, these reactions serve as further examples of parallel pathways since the same reaction also is catalyzed by a P450 monooxygenase.

C.6.2 Ethylene Biosynthesis

The hormone ethylene is produced upon wounding, senescence and fruit ripening (Liu *et al.*, 1997). Aminocyclopropane 1-carboxylate oxidase (ACCO) catalyzes the final step in ethylene biosynthesis; however, instead of using α-KG as the cosubstrate, ascorbate is oxidized to dehydroascorbate during the oxidation reaction (Dong *et al.*, 1992) similar to the reaction mechanism of GA3H (Figure 6), and is therefore not a true αKDD.

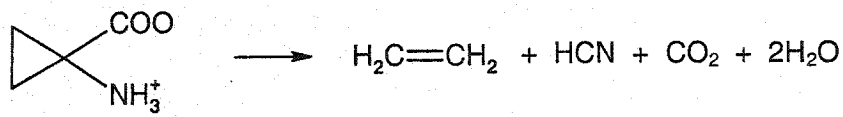


Figure 6: Reaction catalyzed by ACCO

A unique characteristic of ACCO is its requirement for activation by carbon dioxide believed to occur through the carbamylation of an amino acid residue (Prescott and John, 1996). In addition, ACCO is oxidatively self-inactivated by the production of hydrogen peroxide, generated from the auto-oxidation of ascorbate by oxygen, and through a non-oxidative mechanism resulting from its binding with ferric ions and ascorbate in a dead-end complex (Prescott, 2000).

Despite these biochemical differences, *ACCO* from *Helianthus* and *Nicotania* have been found to exhibit relatively high amino acid sequence identity (~28-37 %) to other plant αKDDs (Lincoln *et al.*, 1987). As evidenced from reporter gene expression experiments, *ACCO* is differentially-regulated, depending on the cell type, developmental stage and the nature of stress stimuli, as would be expected for an enzyme involved in the production of an critical signaling molecule (Blume and Grierson, 1997). Experimental

evidence suggests that a multigene family in sunflower seeds and tobacco encodes ACCO, explaining the intricate regulation of this gene (Liu *et al.*, 1997).

Genetic studies identified a number of genes involved in the ethylene signal transduction pathway, which is similar to that of prokaryotic and other eukaryotic cells (Ecker, 1995). Therefore, manipulation of these genes can alter the plant's response to this signaling molecule, thus providing new tools for the agriculture industry to modify fruit development and leaf senescence.

C.6.3 Alkaloid Biosynthesis

Apart from the many physiological roles of alkaloids in plants, several compounds are of particular importance, because of their use in cancer therapy and a variety of pharmaceutical activities (St-Pierre *et al.*, 1999). Whereas the initial stages of biosynthesis of one such metabolite, the tropane alkaloid scopolamine, are largely unknown, the final step in its production has been elucidated in *Hyoscyamus niger* roots and demonstrated to be catalyzed by hyoscyamine 6β-hydroxylase (H6H) as shown in Figure 7 (Hashimoto and Yamada, 1987) .

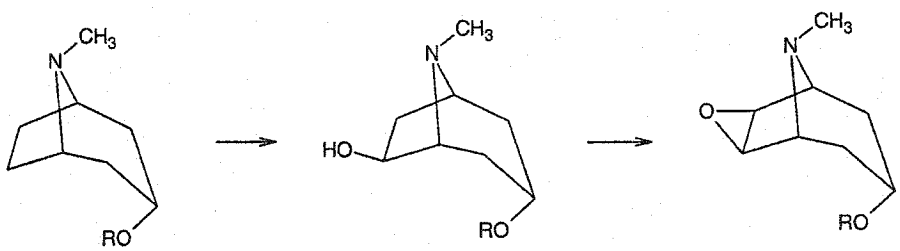


Figure 7: Reaction catalyzed by H6H

H6H catalyzes the conversion of hyoscyamine to scopolamine by a hydroxylation at position 6β to form 6-hydroxyhyoscyamine followed by 7β-dehydrogenation to 6,7-dehydrohyoscyamine *in vitro* and epoxidation to the final metabolite (Matsuda *et al.*, 1991). The *H6H* gene has been isolated and characterized from *Hyoscyamus* and found to exhibit an overall amino acid identity (~25-48 %) to other plant αKDDs depending on the species, mainly within the conserved motifs and the iron-binding residues (Matsuda *et al.*, 1991; Suzuki *et al.*, 1999). The gene is abundantly transcribed only in root tissue and in auxin-free culture media, however it is not yet clear whether *H6H* is solely affected, or the entire biosynthetic pathway is downregulated (Hashimoto and Yamada, 1987) in other tissues. *H6H* was expressed in *Atropa belladonna* and resulted in leaves with elevated scopolamine content, thereby allowing the production of this otherwise scarce compound in readily accessible tissues (Suzuki *et al.*, 1999). Southern blots of *Hyoscyamus* genomic DNA have illustrated that the gene coding for H6H is present as a single copy. Genomic analysis demonstrated that *H6H* gene organization is similar to the organization of genes coding for other αKDDs with respect to intron numbers and positions (Suzuki *et al.*, 1999).

Desacetoxyvindoline 4-hydroxylase (D4H) is involved in the biosynthesis of the dimeric anti-cancer drugs, vinblastine and vincristine in *Catharanthus roseus* (De Carolis and De Luca, 1993). One of the final step in this pathway (Figure 8) is catalyzed by D4H, present in low abundance as a soluble monomer of ~ 45 kDa, that exhibits typical αKDD cofactor and cosubstrate requirements (De Carolis and De Luca, 1993).

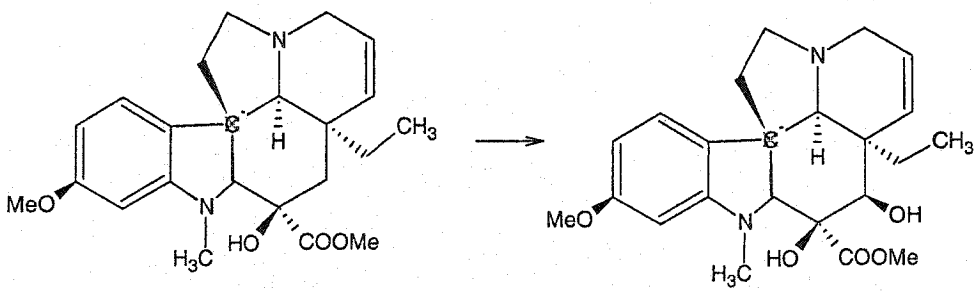


Figure 8: Reaction catalyzed by D4H

The reaction mechanism has been extensively studied and determined to follow a sequential order of substrate binding and product release, typical of this category of enzymes. This enzyme also is susceptible to induction by light (De Carolis and De Luca, 1993) and methyl jasmonate (Vazquez-Flota and De Luca, 1998) and mostly is expressed in leaf tissue. Interestingly, the enzyme expressed in etiolated seedlings is inactive until exposure to light, suggesting the existence of a post-transcriptional control mechanism (Vazquez-Flota *et al.*, 1997).

The *D4H* gene is present as a single copy and the deduced amino acid sequence shows overall sequence similarity (~30 %) to other plant αKDDs, particularly *H6H* (~33 %), including the conserved motifs (Vazquez-Flota *et al.*, 1997). Also, intron/exon junctions are present in regions of high conservation, suggesting that related plant dioxygenases may have evolved from a common ancestral gene. Attempts to date, to produce vinblastine in tissue culture have failed, since the enzymes involved in the earlier steps of its biosynthesis are absent in this *in vitro* system (Prescott and John, 1996).

C.6.4 Flavonoid Biosynthesis

One of the best characterized reactions implicated in this pathway involves the conversion of (*2S*)-naringenin to dihydrokaempferol by the flavanone 3-hydroxylase (F3H; Figure 9; Britsch, 1990).

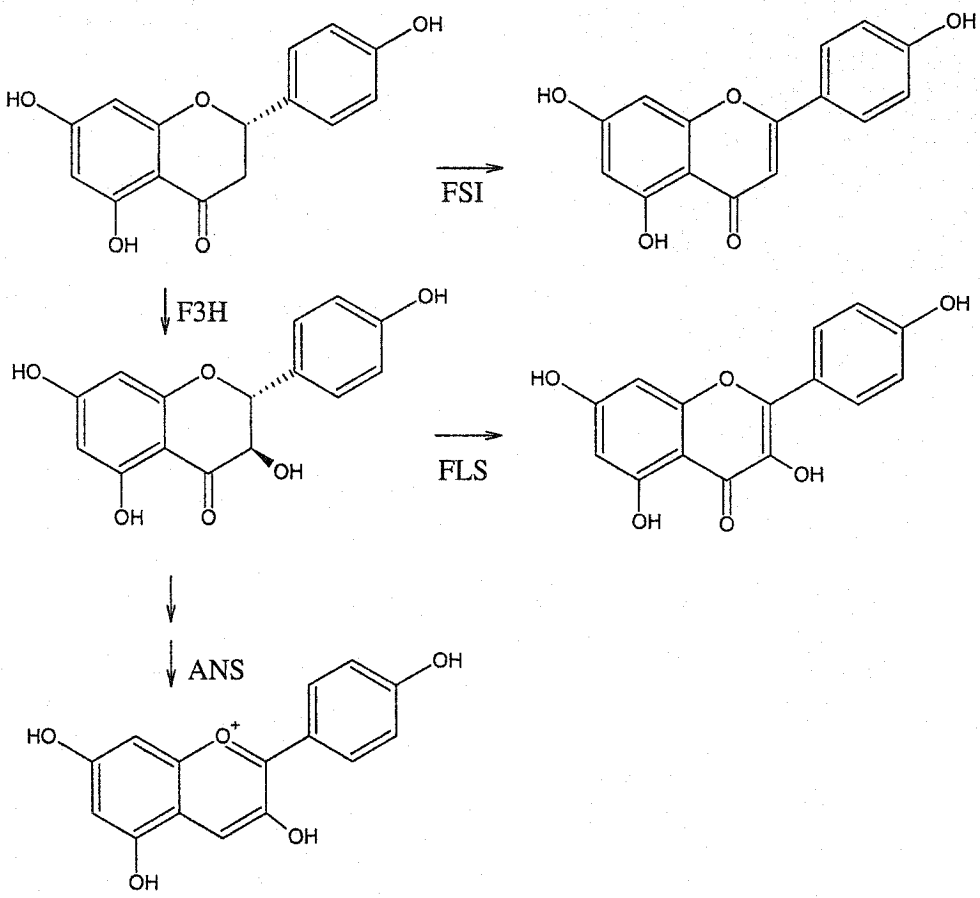


Figure 9: αKDDs involved in flavonoid biosynthesis

The *Petunia hybrida* enzyme, a soluble monomer of ~42 kDa, is highly stereospecific with regard to substrate specificity (Britsch *et al.*, 1992), since modification of the B-ring results in reduced enzyme activity. F3H is not very

susceptible to product inhibition; however, pyridine dicarboxylate compounds act as strong inhibitors as they chelate iron at the active site. This class of compounds has been found to act as general inhibitors of most αKDDs, suggesting the existence of a common reaction mechanism. The F3H reaction has been studied in detail in several species, and the gene encoding F3H in *Petunia* has been cloned and characterized (Britsch *et al.*, 1992). As with the other genes, *F3H* exhibits a certain degree of sequence similarity to other αKDDs, particularly in the conserved regions. A conserved hydrophobic, leucine-rich motif also has been identified that may play a role in intracellular protein-protein interactions, although this has yet to be examined.

Sequence alignments of different plant αKDDs have revealed eight amino acids that are strictly conserved within the group; some of which comprise the cosubstrate binding site (Arg-Xaa-Ser), others are necessary for iron chelation (two His, one Asp), whereas the remaining residues are required for general catalytic activity (Britsch *et al.*, 1992). Site-directed mutagenesis at any of these sites reduced or abolished F3H activity (Lukacin and Britsch, 1997). Through genetic complementation and sequencing of transposon-mutagenized *A. thaliana,* it was demonstrated that F3H is encoded by the *TT6* locus (Wisman *et al.*, 1998). Mutant *tt6* plants do not accumulate flavonoids and generate a higher proportion of phenylpropanoids, due to metabolite redirection. Southern blots indicate that the *Petunia F3H* and *Arabidopsis TT6* genes are present as a single copy in their respective organisms (Britsch *et al.*, 1992; Wisman *et al.*, 1998).

Flavone synthase I (FSI) catalyzes the C2-C3 desaturation of naringenin to apigenin in parsely suspension cultures (Britsch, 1990). This enzyme seems to be restricted to members of the Apiaceae, has not been found in other organisms to date and

its gene was recently cloned (Martens *et al.*, 2001). It is believed that a hydroxylated intermediate is involved in the reaction mechanism, followed by dehydration. This enzyme has a monooxygenase counterpart (FSII) that has been detected in all flavone-containing flowers studied thus far (Prescott and John, 1996).

Flavonol synthase (FLS) also is a desaturase that catalyzes the double bond formation at C2-C3 of dihydroflavonols to generate flavonols and has been studied at the biochemical and molecular levels in several species (Holton *et al.*, 1993; van Eldik *et al.*, 1997). Its physicochemical and kinetic properties have not been examined in detail, although results show that its biochemical characteristics are typical of αKDDs with regards to cosubstrate and cofactor requirements (Holton *et al.*, 1993; Prescott, 2000). The multi-gene family encoding FLS isoforms result in flexibile substrate specificities. *A. thaliana* expressing a mutant *fls* gene produced significantly lower levels of kaempferol and resulted in alteration of flower color (Holton *et al.*, 1993). However, the seed color was not changed as would be expected with *transparent testa* (*tt*) mutants. It had previously been thought that *tt6* coded for FLS, however, knock-out mutants have demonstrated that FLS is not encoded by the *tt6* locus (chromosome #3), but is in fact located on a separate chromosome (#5) in *A. thaliana* (Wisman *et al.*, 1998). In addition, the *FLS* gene contains two introns located at highly conserved positions, and Northern analysis suggested that it is an "early" gene with respect to light induction (Pelletier *et al.*, 1997).

Characterization of a fourth αKDD involved in flavonoid metabolism arose from studies on maize *a2* mutants, where a deficiency in the anthocyanidin biosynthetic pathway was observed from the modification of pigmentation (Menssen *et al.*, 1990).

Anthocyanidin synthase / leucoanthocyanidin dioxygenase (ANS/LDOX) was originally classified as an αKDD based on sequence conservation; however, it was only characterized as an αKDD after its cDNA was expressed and the recombinant protein was characterized biochemically (Saito *et al.*, 1999). ANS catalyzes the conversion of leucoanthocyanidins to 2-flaven-3,4-diols, which under acidic conditions are transformed to anthocyanidins. ANS genes have been cloned from several species, and found to be present as single gene copies that share amino acid identity (~31-38 %) with other flavonoid dioxygenases (Martin *et al.*, 1991; Davies, 1993; Sparvoli *et al.*, 1994; Weiss *et al.*, 1993). Expression studies have confirmed that ANS is a "late"-expressing gene in the pathway, which is to be expected given its role in anthocyanidin biosynthesis (Pelletier *et al.*, 1997).

A number of 6- and 8-hydroxyflavonols, such as quercetagetin (6-hydroxyquercetin), gossypetin (8-hydroxyquercetin) and herbacetin (8-hydroxykaempferol) as well as their derivatives occur naturally in plants. A microsomal preparation of *Tagetes patula* flowers has been reported to catalyze the 8-hydroxylation of quercetin, but not of its methyl derivatives, and required NADPH and molecular oxygen for enzymatic activity (Halbwirth *et al.*, 1999). Further biochemical characterization has confirmed that this reaction is catalyzed by a cytochrome P450-dependent, mixed-function monooxygenase. A flavanone 6-hydroxylase isolated from elicitor-induced soybean has been characterized at the biochemical and molecular levels and shown to be a P450 monooxygenase (Latunde-Dada *et al.*, 2001). In contrast, the flavonol 6-hydroxylase of *Chrysosplenium americanum,* an αKDD that utilizes partially methylated flavonols, but not the parent aglycone quercetin, is the subject of this work.

C.6.5 Other αKDDs

A number of other αKDDs involved in plant metabolism have been studied to varying degrees. Prolyl 4-hydroxylase (P4H), which has been extensively studied in mammalian systems, catalyzes the introduction of a hydroxyl group into peptide side chains that are subsequently incorporated into collagens in mammals, and into cell walls in plants (Kivirikko and Pihlajaniemi, 1998). Unfortunately, much less information is available on the plant-derived enzyme. In both plants and mammals, the enzyme is an α_2/β_2 heterotetramer; the α subunit possesses the active site required for iron binding, whereas the enzyme protein disulfide isomerase (PDI) is the β monomer and is found associated with the ER and Golgi in plants (John and Bulleid, 1994; Wojtaszek *et al.*, 1999). This enzyme was one of the first αKDDs characterized, that exhibited the typical cofactor and cosubstrate requirements, and was thus considered to be a type member for this category of enzymes, regardless of its tetrameric nature and low sequence identity to other αKDDs (Myllyharju and Kivirikko, 1997). There have been several ESTs identified in plants that exhibit high sequence similarity to the α subunit of mammalian P4H, however no plant gene has been cloned to date. A putative β subunit was cloned in alfalafa, although it is not yet known whether this PDI forms part of P4H (Shorrosh and Dixon, 1991).

A number of other genes exhibiting similarity to other αKDD genes have recently been cloned, although the functions of the proteins they encode have yet to be elucidated. In one case involving the plant response to iron deficiency, *Ids2* and *Ids3* have been isolated and their gene products determined to be involved in the production of mugineic acids from methionine (Nakanishi *et al.*, 1993; Okumura *et al.*, 1994). When a plant is

subjected to iron deficiency, the roots secrete mugeneic acids to chelate metal ions; however, there is no direct evidence of the involvement of these enzymes in the biosynthesis of such compounds.

Hydroxyphenylpyruvate dioxygenase (HPPD) has been extensively studied in mammals and *Pseudomonas*, however little is known about the plant protein even though it is believed to be the site of action of a bleaching herbicide. This enzyme activity, which is involved in the production of homogentisate, has been studied in carrot and corn seedlings (Garcia *et al.*, 1997). Given that HPPD decarboxylates the primary substrate and does not exhibit a requirement for α-KG, it can not be considered a true αKDD, although it exhibits sequence conservation with other αKDDs at the C-terminal end, suggesting that this region may be involved in cofactor binding.

C.7 Classification of αKDDs

As of yet, there is no clear definition of what is a dioxygenase, other than the fundamental incorporation of two oxygen atoms into substrate molecules. However, from the examples cited above, there is a discrepancy in definition even at this most basic level. Before the characterization of αKDD clones, classification was a simple matter; if the enzyme required the cofactors typical of the reaction (α-KG, Fe^{+2}, ascorbate), it was considered a member of this class. αKDDs catalyze a variety of reactions, such as typical hydroxylations (P4H, F3H, D4H), desaturations (FLS, FSI), oxidation (IPNS, ACCO, GA-oxidases) and epoxidations (H6H). In addition, while most of these enzymes utilize α-KG as a co-substrate, others use alternate keto-groups. The explanation initially offered for the grouping of enzymes with dissimilar activities into a single category arose

from the possibility that all reactions possessed hydroxylated intermediates (Prescott, 1993), thus justifying their classification, however this now seems unlikely. Gene sequence analyses have demonstrated that biochemical and molecular classifications do not correlate well. Although the degree of identity between isozymes may be significant (~85 %), the different types of αKDDs show little sequence similarity (25-30 %) at both the amino acid and nucleotide levels. In addition, in order to accurately align sequences, enzymes should be of similar sizes, and this poses a problem because molecular weights range from 26 kDa (GA2H) to 86 kDa (amino acid hydroxylases) and not all enzymes are monomeric in their functional state. In certain cases, proteins with high biochemical similarity exhibit low molecular identity, and *vice versa* (Myllyla *et al.*, 1992). In the former case, this may be evidence of convergence, while the latter can perhaps be explained by a common reaction mechanism. Furthermore, it has been difficult to elucidate the precise reaction mechanism of αKDDs because of their labile, multi-functional nature.

Nevertheless, apart from classification based solely on biochemical analyses of this class of enzymes, recent developments in the molecular characterization of novel genes show that they possess sequence similarity to biochemically characterized αKDDs. However, an interesting fact has become evident; certain members exhibiting some sequence similarity catalyze unrelated chemical reactions, while other members that exhibit little or no sequence similarity to one another demonstrate high biochemical similarity (Prescott, 2000). The mechanistic and evolutionary implications of such discrepancies have yet to be elucidated. The reactions catalyzed by αKDDs are complex and precise intermediate stages remain to be elucidated; therefore, a reaction only reflects

the overall result, potentially masking any commonalities in the reaction mechanism. There may be a common mode of action for all αKDDs, however, the mechanistic information required to postulate a plausible reaction model is not yet available. In addition, although there are common structural motifs and evidence of conserved genomic organization, whether this is the result of convergence and/or evolution from a common ancestor remains unclear.

C.8 Protein Structure

Bacterial αKDDs have been studied extensively and are well characterized with respect to protein structure. The structure of prokaryotic dioxygenases can be extrapolated to plant αKDDs, since their gene sequences exhibit a low, but significant degree of overall similarity (~24 %), and certain conserved motifs and residues are present in all their subsets. Crystallographic studies of *Aspergillus nidulans* isopenicillin *N*-synthase (IPNS), catalyzing the ring closure of *L*-δ-(α-aminoadipoyl)-*L*-cysteinyl-*D*-valine to isopenecillin N (Figure 10), have shown a β-sheet jellyroll motif common to viral capsid proteins (Roach *et al.*, 1995). This motif comprises the core of the enzyme, which is thought to be stabilized by a leucine zipper structure that forms part of an exposed α-helix. The ferrous ion is located in a buried active site, for the isolation of reactive oxoferryl species (Ming *et al.*, 1991), a feature common to all αKDDs. The iron binding site exhibits a distorted octahedral geometry with two water molecules, an aspartate, a glutamine and two histidine residues at the C-terminus, serving as ligands for the ferrous ion.

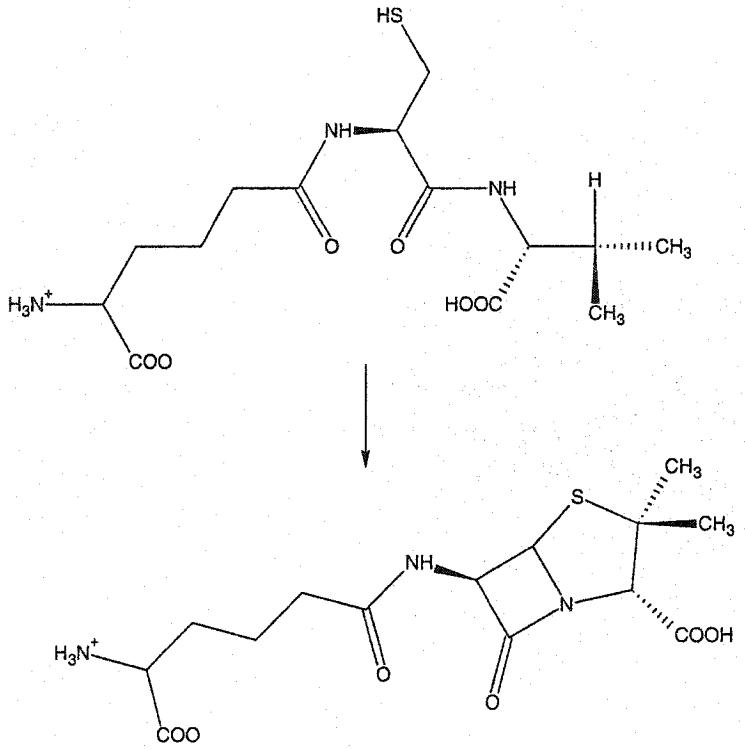


Figure 10: Reaction catalyzed by IPNS

One water molecule and the Gln residue are displaced upon substrate binding (Ming *et al.*, 1991). However, oxygen is incorporated into water and not into the final product; therefore, IPNS is a dehydrogenase/desaturase with oxygen acting as the hydrogen acceptor.

Confirmation of histidine involvement in iron binding and catalysis resulted from diethylpyrocarbonate (DEPC) treatment of the enzyme, virtually abolishing its activity. This work correlated well with the results obtained with other αKDDs, which demonstrated that DEPC treatment resulted in modifications to the conserved histidines that prevented iron association (Lukacin and Britsch, 1997). Although all residues and motifs present in IPNS are not strictly conserved in all αKDDs, there is a certain degree of conservation of the residues involved in ligand binding and catalysis. In most cases,

one Asp and two or three His residues are conserved, along with one Arg residue believed to anchor α-KG (Roach *et al.*, 1995; Lukacin and Britsch, 1997).

C.9 Gene Expression and Evolution

The expression of the genes coding for these enzymes is dissimilar. Multigene families encode certain classes of enzymes, while other genes are found only as single copies. Thus far, it seems as though multigene families encode the αKDD proteins involved in hormone biosynthesis, allowing for the temporal, spatial and developmental regulation required (Balague *et al.*, 1993; Drory *et al.*, 1993; Chiang *et al.*, 1995; Phillips *et al.*, 1995). Genes involved in the early stages of phenylpropanoid biosynthesis typically comprise multigene families as well, namely chalcone synthase (Koes *et al.*, 1989), chalcone isomerase (van Tunen *et al.*, 1988), dihydroflavonol reductase (Beld *et al.*, 1989) and coumarate-CoA ligase (Cukovic *et al.*, 2001). However, the αKDDs investigated to date are present in a single copy, with the exception of *FLS* in *A. thaliana* (Pelletier *et al.*, 1997), a feature that may facilitate the investigation of regulatory mechanisms affecting these genes.

This leads to the subject of αKDD evolution. Aside from biochemical characterizations, the affiliations that have been made thus far involve the degree of sequence similarity and the number and positions of introns. We should keep in mind that the degree of identity within groups, or of isozymes, is variable. In addition, among plant αKDDs with different functions there is ~24-34% amino acid identity. Therefore, any combination yielding higher values may be an indication of relatedness. We also must consider that enzyme activities present in most species, including those of a lower

classification order, likely evolved much earlier than those present in a few higher plants. Researchers believe that the likely order of evolution begins with *F3H* (bryophytes and ferns), followed by *FLS* (bryophytes and ferns) to GA hydroxylases (ferns) and finally to *ACCO* (gymnosperms) and *ANS* (gymnosperms and angiosperms) (Osborne, 1989; Stafford, 1991).

Intron position is interesting, because researchers have found that there are primarily three introns at four varying positions (Prescott and John, 1996), suggesting that the ancestral genes may have contained three introns at positions 1, 3, and 4, given that these are the most common intron locations (Figure 11). However, a survey of *A. thaliana* dioxygenases indicated the presence of 8 to 10 introns and just as many positions (Prescott, 2000). This has yet to be investigated in other αKDDs.

Phylogenetically, αKDDs constitute a superfamily with numerous subgroups, one of which is involved in flavonoid biosynthesis. It is evident that those belonging to a single pathway are located in a related subgroup, perhaps the result of gene duplication and divergence. Given that the entire series of genes has yet to be cloned and biochemically characterized from a single plant species, any assumptions on the ancestry of these genes based on the compilation of αKDDs originating from different organisms would not be valid, unless it was clear that orthologs and not paralogs were being compared. However, the core structural genes involved in the flavonoid biosynthetic pathway from maize and snapdragon have been cloned and characterized (Rausher *et al.*, 1999).

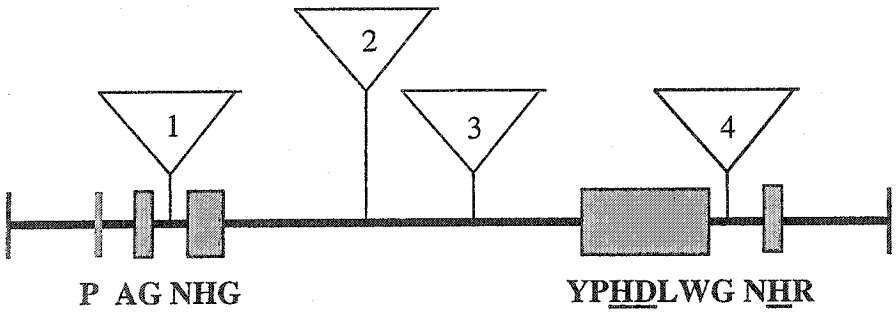


Figure 11: αKDD genomic organization and conserved motifs

Rausher and colleagues (1999) also have isolated homologous genes from the morning glory and have reported on the similarity and possible evolution of these genes. They believe that genes acting upstream in the pathway have evolved at a significantly slower pace than downstream genes. This is likely due to the fact that they participate in numerous biochemical pathways. In addition, the regulatory genes involved in this pathway have evolved at a faster pace than the structural genes they affect. These results suggest that the evolution of flavonoid biosynthesis, and ultimately flower coloration, is mediated more by regulatory than structural genes. In general, the degree of similarity within subgroups is higher than between different subgroups, although the values for the latter are variable. In addition, the genes encoding proteins in the same pathway, although not in the same subgroup, show expectedly higher identity than unrelated genes.

Surprisingly, in numerous cases, both monooxygenases and dioxygenases catalyze similar reactions in different pathways. There must be an evolutionary explanation for this redundancy. It is believed that due to the architecture of their active sites, αKDDs are more versatile enzymes with respect to the range of possible substrates they accept (Hegg and Que, 1997). More specifically, the range of co-ordination

positions of the amino acid residues at the catalytic site allows for more flexibility in dioxygenases than the single, possible co-ordination site in cytochrome P450-dependent monooxygenases (Hegg and Que, 1997; Schofield *et al.*, 1997). In other words, the same dioxygenase can catalyze the conversion of several substrates, and it may be for this reason that researchers believe this class of enzymes has participated in the elaboration of numerous biosynthetic pathways throughout evolution.

D. MATERIALS AND METHODS

D.1 Plant Material

Chrysosplenium americanum (Saxifragaceae), Schwein ex Hooker, was collected from St. Anicet, Province of Québec, and was maintained in the greenhouse under conditions simulating its natural habitat regarding light intensity, temperature and humidity.

Chrysosplenium alternifolium (Saxifragaceae) was purchased from Lost Horizons Nursery, Province of Ontario, and was maintained under greenhouse conditions.

D.2 Chemicals and Buffers

Flavonol substrates were from our laboratory collection. All other chemicals were of analytical reagent grade. Silica TLC plates (Kieselgel 60 / Kieselgur F_{254}) were from Merck.

The following buffers were used for purification of the native protein: **A**, 100 mM Tris-HCl (pH 7.3), 10 mM DTT, 10 mM ascorbate, 0.5 mM PMSF, 10% glycerol (v/v); **B** (assay buffer), 50 mM Tris-HCl (pH 7.3), 10 mM DTT, 150 mM NaCl, 10% glycerol (v/v); **C**, as in B but with 500 mM NaCl; **D**, 50 mM Tris-HCl (pH 7.3), 15 mM NaCl; **E**, 25 mM bis-Tris-iminodiacetic acid (pH 7.1), 10 mM DTT, 10 mM NaCl, 10% glycerol (v/v), 5% betaine (w/v); **F**, Polybuffer 74 (1:15, v/v pH 4.0), 10 mM NaCl, 10% glycerol (v/v), 5% betaine (w/v); **G**, 100 mM potassium phosphate (pH 7.6), 5 mM EDTA, 10 mM diethyldithiocarbamate, 14 mM β-mercaptoethanol, 0.5 mM PMSF, 10 % glycerol (v/v); **H**, 25 mM Bis-Tris (pH 7.1), 14 mM β-mercaptoethanol, 10% glycerol

(v/v), 5% betaine (w/v); **I**, Polybuffer 74 (1:15, v/v pH 4.0), 10 mM KCl, 10 % glycerol (v/v), 5% betaine (w/v). All FPLC buffers were filtered, degassed and stored at 4°C.

The following buffers were used for the molecular characterization of the cDNA clone and the recombinant protein: **J** (wash solution I), 2X SSC, 0.1% SDS (w/v); **K** (wash solution II), 0.1X SSC, 0.1% SDS (w/v); **L** (DNA extraction buffer), 100 mM Tris/HCl (pH 7.4), 50 mM EDTA, 500 mM NaCl, 14 mM β-mercaptoethanol; **M** (TE), 10 mM Tris-HCl (pH 8.0), 1 mM EDTA; **N** (RNA extraction buffer), 50 mM Tris-HCl (pH 9.0), 100 mM NaCl, 20 mM $Mg(OAc)_2$, 0.5 M Sucrose, 5 mM DTT, 20 mM EDTA, 1% SDS; **O** (blocking solution), 5% SDS (w/v), 125 mM NaCl, 25 mM sodium phosphate (pH 7.2); **P** (blot washing solution), 0.5% SDS (w/v), 12.5 mM NaCl, 2.5 mM sodium phosphate (pH 7.2)

D.3 Purification and Biochemical Characterization of the F6H

D.3.1 Preparation of α-Ketoglutarate Affinity Gel

Epoxy-activated Sepharose (Pharmacia), used for the coupling of ligands with hydroxyl groups, was swollen and washed with water. Coupling was achieved using 250 μmol α-KG/ml of wet beads according to the manufacturer's instructions. It should be noted that the affinity support had the α-ketoglutarate binding randomly via the hydroxyls of the carboxy groups.

D.3.2 Preparation of Immunoaffinity Gel

Anti-D4H (polyclonal antibody raised against the desacetoxyvindoline 4-hydroxylase purified from *Catharanthus roseus;* Vazquez-Flota *et al.*, 1997) crude anti-

serum was affinity purified. Briefly, the crude serum was brought to 40% ammonium sulfate saturation and the precipitate was collected by centrifugation at 12 000 *g* for 10 min. The pellet was desalted against Tris-buffered saline (TBS; 20 mM Tris/HCl pH 7.3, 500 mM NaCl). The protein fraction was loaded onto a Hi-Trap Protein-A Sepharose column (Pharmacia), pre-equilibrated with TBS, and the IgGs were eluted according to the manufacturer's instructions. The antibody-coupled gel was prepared by a direct coupling method. The IgG-purified anti-D4H serum was incubated with Protein-A Sepharose beads (Pharmacia) at 2 mg antibody per mL of wet beads for 1 hour at room temperature with agitation. The beads were washed twice with 200 mM sodium borate (pH 9.0) and resuspended in the same buffer containing 20 mM dimethylpimelimidate. After mixing for 30 min, the coupling reaction was stopped by first washing the beads in 200 mM ethanolamine (pH 8.0), followed by a 2-hr incubation at room temperature in 200 mM ethanolamine with gentle shaking. The beads were resuspended in TBS and the efficiency of coupling was verified by boiling samples before and after the coupling reaction in Laemmli's buffer (Laemmli, 1970), and electrophoresing the samples in a 12% SDS-PAGE gel. After staining with Coomassie, the absence of heavy-chain bands at 55 kDa indicated efficient coupling.

D.3.3 Immunotitration

Aliquots (70-150 µL) of protein samples from various purification steps were incubated with increasing amounts of purified IgG's overnight at 4°C. The immune complexes were precipitated by centrifugation at 15 000 *g* for 25 min at 4°C, and F6H

activity was assayed in the supernatant. Control experiments were carried out using equivalent amounts of non-immune serum for comparison.

D.3.4 Extraction and Purification of F6H

All steps were carried at 4°C unless otherwise specified. Shoot tips were frozen in liquid nitrogen and pulverized together with fine sand (10% w/v) and polyvinylpolypyrrolidone (PVPP; 10%, w/w), in order to remove phenolic substances, before being homogenized with buffer A (1:4, w/v). The homogenate was filtered through nylon mesh, and the filtrate centrifuged at 15 000 *g* for 10 min. The supernatant was treated with solid ammonium sulfate, and the protein that precipitated between 35% and 80% salt saturation was collected by centrifugation at 18 000 *g* for 15 min.

(a) Gel Filtration

The ammonium sulfate pellet was solubilized in a minimal volume of buffer B, applied onto a preparative Superose 12 column (16 mm x 300 mm) previously equilibrated with buffer B, and the protein was eluted with the same buffer at a flow rate of 1 mL/min. Two-mL fractions were collected and assayed for F6H activity using the direct enzyme assay (see below).

(b) Anion-Exchange Chromatography

The Superose 12-active fractions were loaded onto a Mono Q HR5/5 column pre-equilibrated with buffer B, and the protein was eluted with a linear (0 % to 100 %) gradient of buffer C in buffer B at a flow rate of 0.5 mL/min. One-mL fractions were collected and assayed for F6H enzyme activity.

(c) α-Ketoglutarate Affinity Chromatography

The Mono Q-active fractions were diluted to 15 mM NaCl and loaded onto a α-ketoglutarate-Sepharose column pre-equilibrated with buffer D. The protein was eluted step-wise with 10, 25 and 50 mM α-KG in buffer D, and one-mL fractions were assayed for enzyme activity.

(d) Immunoaffinity Chromatography

The Superose 12-active fractions were loaded onto an IgG-crosslinked Protein A-Sepharose column pre-equilibrated in buffer B, and allowed to stand overnight at 4°C. After washing three times with buffer B, the proteins were eluted twice with 5 vol. of 100 mM Tris-HCl buffer (pH 4.0), and the eluate adjusted to pH 7.0 before assaying for F6H activity.

(e) Chromatofocusing

To determine the apparent *pI* of F6H, the active Mono Q fractions were desalted on a PD-10 column, and loaded onto a Mono P HR5/20 column (5 mm i.d. x 200 mm bed height) pre-equilibrated with buffer E. The column was developed with buffer F at a flow rate of 0.5 mL/min, and 1-mL fractions were assayed for enzyme activity.

D.3.5 Extraction and Purification of Flavonol 6-*O*-methyltransferase (F6OMT)

The method outlined by Khouri and Ibrahim (1987) was used with minor modifications. Briefly, *C. americanum* shoots were homogenized with buffer G (1:4, w/v) and PVPP (10 %, w/w). The homogenate was filtered through a nylon mesh and the filtrate centrifuged for 15 min at 20 000 *g*. The supernatant was stirred with AG1-X2 resin (Bio-Rad; 1:10, w/v) pre-equilibrated with 0.1 M potassium phosphate buffer (pH

7.6) then filtered. This extract was subjected to ammonium sulfate precipitation, and the fraction that precipitated between 35-70 % salt saturation was collected by centrifugation and resuspended in a minimal volume of buffer H. After gel filtration on a Superose 12 column using the same buffer, the F6OMT-active fractions were pooled and subjected to chromatofocusing on a Mono P column pre-equilibrated with buffer H. The column was developed with 50 ml of buffer I at a flow rate of 0.5 mL/min, and one-mL fractions were collected into tubes containing 100 µL of 500 mM Tris-HCl (pH 8.0). Fractions were assayed for the Mg^{2+}-dependent F6OMT activity (Khouri and Ibrahim, 1987), and the active fractions were pooled and used for the F6H indirect enzyme assay.

D.3.6 Direct Enzyme Assays

F6H activity was assayed using 2.5 to 5 µM flavonol substrate, 10 mM α-ketoglutarate, 10 mM ascorbate, 0.25 mM ferrous sulfate, 100 mM Tris-HCl (pH 8.3) and up to 100 µg protein in a final volume of 100 µL. After incubation for 1 hr at 30°C, the reaction was stopped with 20 µL 6 M HCl and the products were extracted with 500 µl ethyl acetate. After centrifugation, the organic phase was collected and lyophilized. The residue was redissolved in 120 µL methanol and analyzed by HPLC on an analytical RP C18-Silica column (Merck), using an isocratic elution protocol consisting of methanol:water:acetic acid (75:24:1; v/v/v). Alternatively, an analytical RP C18-Silica Symmetry column (Waters) was used to chromatograph the reaction products and reference samples using a gradient elution protocol using 20 % methanol in 1 % acetic acid for 5 min, followed by a linear gradient to 75 % methanol and 1 % acetic acid in 35 min; maintained for 10 min followed by re-equilibration to start conditions. Identity of

the reaction products was confirmed by co-elution with reference compounds and by using a F6OMT-coupled assay, as well as LC-MS analysis using APCI (atmospheric pressure chemical ionization) under the following conditions, Source Temp., 550°C; Capillary Temp., 150°C; Sheath/Aux Gas Flow, 80 units/10 units; Width/Energy, 5 amu 22 %, sid 50 %.

D.3.7 Indirect Enzyme Assays

The F6OMT-coupled assay was used to verify the position of hydroxylation catalyzed by F6H activity. The hydroxylase reaction was allowed to proceed as described in the direct enzyme assay before being stopped with 1 mM EDTA. This was followed by the addition of 10 mM Mg^{2+} (to overcome chelation by EDTA), together with 0.025 μCi [^{14}C]-*S*-adenosyl-*L*-methionine (AdoMet; 59 μCi/μmol) and 0.4 to 1.0 μg of the Mono P-purified F6OMT, in a final volume of 200 μL. The reaction was allowed to proceed for 30 min at 30°C, stopped with 20 μL 6 M HCl, and the methylated products were extracted with 500 μL of a mixture of ethyl acetate:benzene (1:1, v/v). An aliquot of the organic phase was counted for radioactivity, and the remainder was co-chromatographed with reference compounds on TLC silica plates, using toluene:ethylformate:*n*-butyl acetate:formic acid (25:50:23:2, v/v/v/v) as solvent, then autoradiographed.

D.3.8 Kinetic Analysis

Kinetic data were analyzed using the derivations for ter-reactant systems stated in (Segel, 1975). Briefly, substrate interaction (1) and product inhibition (2) data were fitted to the following equations:

(1)

$$\frac{v}{V_{max}} = \frac{\frac{[A]\,[B]\,[C]}{K_a K_b K_c}}{1 + \frac{[A]}{K_a} + \frac{[A]\,[B]}{K_a K_b} + \frac{[A]\,[B]\,[C]}{K_a K_b K_c}}$$

(2)

$$\frac{v}{V_{max}} = \frac{[B]}{K_{mB}\left(1 + \frac{K_{ia}}{[A]}\right)\left(1 + \frac{K_{ib}K_{mc}}{K_{mb}[C]}\right) + [B]\left(1 + \frac{K_{ma}}{[A]} + \frac{K_{mc}}{[C]} + \frac{K_{ip}K_{mr}K_{mc}[Q]}{K_{ic}K_{mp}K_{iq}K_{ir}}\right)}$$

where A, B, C, are the varied substrates, P, Q, R are the reaction products. K_a, K_b, K_c are respective limiting Michaelis constants, V_{max} is the maximum velocity and K_{ia}, K_{ib}, K_{ic}, and K_{ip} K_{iq} K_{ir} are the interaction terms.

The data from kinetic studies are presented as double reciprocal plots which were fitted by linear regression analysis. The type of interaction or inhibition was suggested by the pattern obtained from these plots and parameter values were calculated accordingly using the above equations, as well as from slope and intercept replots. Several replicates (n = 5 to 8) for each data point were performed to ensure an acceptable standard deviation (*ca.* 10 %). The errors in substrate and inhibitor concentrations were assumed to be constant at all concentrations.

In order to vary the oxygen concentrations, reacti-vials containing Superose 12-purified F6H were sealed, placed on ice and flushed with N_2 for 30 min prior to the addition of substrates, which were further flushed with N_2 for 10 min. The appropriate oxygen concentration introduced into each vial was based on the % by volume in the incubation atmosphere (headspace). The K_m for oxygen is expressed in μM unit on the

basis of the known solubility of oxygen at 37°C, with a saturating concentration of 960 μM.

D.3.9 SDS-Polyacrylamide Gel Electrophoresis and Western Blotting

SDS-PAGE was performed according to the method of Laemmli (1970) using molecular weight markers (Bio-Rad) for calibration under denaturing conditions. After electrophoresis, proteins were stained with Coomassie Brilliant Blue (R-250) or silver reagent (Bio-Rad). The gels were submitted to a semi-dry eletrophoretic transfer onto nitrocellulose membranes, and probed with the IgG-purified anti-D4H (1:500) and alkaline phosphatase-conjugated goat-anti-rabbit IgG's (1:5000; Bio-Rad). Alkaline phosphatase activity was revealed in the dark using nitroblue tetrazolium and 5-bromo-4-chloro-3-indolyl phosphate (Bio-Rad) as substrates. Alternatively, blots were probed with anti-His (1:5000; Pharmacia) and HRP-conjugated goat-anti-mouse IgGs (1:3000; Pharmacia) or IRDye800-conjugated goat-anti-mouse IgGs (1:5000; Rockland Immunochemicals). All detection procedures were carried out according to the manufacturers' instructions.

D.3.10 Protein Determination

The protein concentration was estimated using the Bio-Rad protein assay based on the method of Bradford (1976) according to the manufacturer's specifications, with bovine serum albumin as the standard protein.

D.4 Cloning and Characterization of Recombinant F6H

D.4.1 Protein Microsequencing

The ligand-affinity purified F6H protein preparations were subjected to trypsin digestion and sequencing in order to obtain internal peptide information to be used as part of the cloning strategy. This analysis was performed at the Harvard Microchemistry Facility using microcapillary, reverse-phase HPLC nano-electrospray tandem mass spectrometry (MS/MS) on a Finnigan LCQ quadrupole ion trap mass spectrometer. The sequences were the result of a single microcapillary reverse-phase HPLC run, directly coupled to the nano-electrospray ionization source of an ion trap mass spectrometer. This instrument configuration is capable of acquiring individual sequence (MS/MS) spectra on-line at high sensitivity (<<<10 fmol) for multiple peptides in the chromatographic run. These MS/MS spectra, also referred to as CID sequence of fragmentation spectra, are then correlated with known sequences using the algorithm Sequest developed at the University of Washington (Eng *et al.*, 1994); and programs developed by (Chittum *et al.*, 1998). MS/MS peptide sequences then are reviewed by a scientist for consensus with known proteins, and the results are manually confirmed for fidelity.

D.4.2 Primers

The primers derived from the microsequence information obtained from the native plant protein were designated Micro1 and Micro2 according to the microsequence used. Degenerate primers were used in order to account for codon degeneracy in the peptide sequences. Amplification reactions involving inosine-containing degenerate primers were not successful in that most reactions failed to produce discrete bands upon

visualization after gel electrophoresis. When smears were observed, the fragments were randomly cloned and sequenced but resulted from non-specific cross-reactivity of the primers since none of the fragments obtained exhibited homology to any known αKDD-like sequence.

Micro1For - 5′-GCTGGATCCTCCTTCATATA-3′;

Micro1Rev- 5′-(GC)GATAC(AT)GG(TC)AA(TC)(CG)TTAG(ATGG(CT)TATAC-3′;

Micro2For-5′-AATTGTTGAAGCATGTGAAG-3′;

Micro2Rev-5′-(TC)GGGGTTAG(AG)AG(TC)GT(AC)CG(AG)AA-3′;

T3 + 6 - 5′- GCT CGA AAT TAA CCC TCA CTA AAG GG - 3′

T7 + 6 - 5′- GAA TGG TAA TAC GAC TCA CTA TAG GGC G- 3′

ConsHydFor - 5′- CA(AG) (GA)T(TA) GTT (GA)AT CAT GGG (GA)TT- 3′

ConsHydRev - 5′- TGG (TC)T(GC) TGG ACA TGG TGG GTA- 3′

F6H·1For - 5′- AAG AGG ATG ATG ACC ATG GA- 3′

F6H·1Rev - 5′- TTC TTC CTA GGT AGG CAA TA - 3′

Frag2For - 5′- TGA ATT GTA ATA CGA CTC ACT TAT AGG - 3′

Frag2Rev - 5′- GAT GGG TGG TAC AGG TCT CG - 3′

F6HfullFor - 5′- CAA GAG AAG ACT CTT AAC TCT AGA - 3′

F6HfullRev (GSP1) - 5′- GAG GTA AGC CTC CTC AGC GAT GAT CTC - 3′

GSP2 - 5′- GTA GGA GAT CCC CTA GTA GGT CTT CTG TAT C - 3′

FNestFor- 5′- GGA ACT GAT GCT GCT GCT AGA - 3′

FNestRev- 5′- AGC ATC AAG AGC GTT AAG AGC - 3′

F6HCloneFor- 5′- GAT AAG GAT CCG AAT CCA AGA GTG ACA - 3′

F6HCloneRev- 5′- CAG CCA AGC TTC TTA AGC AAG AAT TTC - 3′

PicZF6HFor- 5′-ATC GGT ACC TCA ATA ATG TCT ATT CCA AGA GTG ACA-3′

PicZF6HRev- 5´- GCC GCC GCG GCT AGC AAG AAT TTC CTC - 3´

PicZαF6HFor- 5´- GTC GGT ACC TCG ATT CCA AGA GTG ACA - 3´

PicZαF6HRev - 5´- GCC GCC GCG GCG CTA GCA AGA ATT TCC TC - 3´

5'AOX1 site - 5´- GAC TGG TCC CAA TTG ACA AGC - 3´

3'AOX1 site - 5´- GCA AAT GGC ATT CTG ACA TCC - 3´

InvF6HFor - 5´- GGA TGG ATC CTT CTT CAT ATC GGA - 3´

InvF6HRev - 5´- TCC ATG CTC ATC ATC GAT CTC TCC GAT - 3´

ADA1 - 5´- GTA ATA CGA CTC ACT ATA GGG CAC GCG TGG TCG ACG GCC CGG GCT GGT - 3´

ADA2 - 5´- P-ACC AGC CCG GAM7 - 3´

AP1 - 5´- GTA ATA CGA CTC ACT ATA GGG C - 3´

AP2 - 5´- ACT ATA GGG CAC GCG TGG T - 3´

D.4.3 Screening of the *C. americanum* cDNA Library

A previously constructed *C. americanum* cDNA Lambda UniZap XR library (Gauthier *et al.*, 1996) was screened using three separate methods in attempts to isolate a putative full-length F6H clone.

(a) Immunological screening

The cDNA library was screened by immunodetection as described in (Sambrook *et al.*, 1989) with anti-desacetoxyvindoline 4-hydroxylase immuno sera (1:5000) as the primary antibody. Colorimetric detection was carried out using the alkaline phosphatase-conjugated anti-rabbit IgG antibody as the secondary antibody according to manufacturer's instructions.

(b) PCR-based screening.

The cDNA library was screened by Polymerase Chain Reaction (PCR) as described in (Friedman *et al.*, 1990). After optimization of aliquot size, 5 μL of the cDNA library (representing 1 X 10^6 pfu) were used per reaction. Primer combinations: i) Micro1Rev and Micro2For; ii) Micro1For and T3; iii) Micro1Rev and T7; iv) Micro2For and T3; v) Micro2Rev and T7. Cycling programs were as follows: Denaturation, 94°C, 40 sec; Annealing, 42 - 48°C, 1 min (dependent on the T*m* of the reaction primers); Extension, 72°C, 2 min; repeated for a total of 30 cycles with a final extension time of 10 min. PCR products were verified by PCR with different primer sets (ConsHydFor and ConsHydRev), or with nested primers (FNestFor and FNestRev), when possible. Following gel electrophoresis, selected PCR fragments were isolated and purified using the DNA isolation from agarose gel kit (Millipore). The amount of DNA recovered was quantified by visualization, after ethidium bromide staining on agarose gels and comparison with molecular weight markers (1 kbp Ladder, MBI Fermentas). An optimal amount of PCR fragment was ligated to the pGEM-T vector using the pGEM-T cloning kit (Promega) and transformed into *E. coli* strain XL1-Blue by heat shock (Sambrook *et al.*, 1989). Inserts were subjected to DNA sequencing in order to characterize the product isolated from the PCR-based screen.

(c) Hybridization screening

The bacteriophage cDNA library was screened using an oligonucleotide probe obtained from an initial screen of the same library by PCR as described in Seidman (1994). The probe consisted of a 602-bp fragment (F6H·1) containing the microsequence obtained from the native protein, as well as areas of similarity with known αKDD sequences.

Detection was carried out using IRDye technology and the Odyssey workstation provided by Li-Cor. Briefly, approximately one million plaques were plated on NZY medium and subsequently transferred to charged nylon membranes for hybridization to the terminally, IRdye800-labeled probe. Post-transfer, DNA was fixed to membranes by UV crosslinking using a Stratalinker. Wet membranes were prehybridized with UltraHyb-OS solution (Ambion) at 42°C for 2 h. The probe was labeled according to manufacturer's instructions (Li-Cor) and diluted in a fresh aliquot of UltraHyb-OS solution, which contains 50 % formamide, for hybridization at 42°C overnight (~16 h). Membranes were washed twice with buffer J (wash solution I) at 45°C for 5 min, followed by two more washes with buffer K (wash solution II) at 45°C for 15 min prior to detection.

D.4.4 DNA Sequencing

Amplified DNA was purified using a Qiagen PCR purification kit, and electrophoresed on 1% agarose gel with ethidium bromide (0.33 μg/mL) at 7 v/cm for 40 min. Amplified fragments were visualized and documented using a GeneSnap 4.00-Gene Genius Bio Imaging System (Syngene). The digital image files were analyzed using Gene Tools software from Syngene. The amount of DNA was estimated using a 1kbp DNA Ladder (MBI Fermentas). The purified, amplified DNA was directly sequenced using ABI Big Dye™ Terminator v. 3.0 and 3.1 Cycle Sequencing Ready Reaction kit (diluted 1: 4) and an ABI310 automated genetic analyzer (Applied Biosystems, Foster City CA). Each region was sequenced between 2 to 4 times. The thermal cycling profile of sequencing reactions were: Denaturation, 94°C, 10 sec; Annealing, 50°C, 5 sec; Extension, 60°C, 4 min; repeated for a total of 25 cycles.

Alternatively, DNA sequencing was carried out by the Centre for Structural and Functional Genomics (Concordia University) and/or Sheldon Biotechnology Centre (McGill University) using the dideoxy-mediated chain termination method (Sanger *et al.* 1977).

D.4.5 Plant Genomic DNA Isolation

Genomic DNA was extracted by grinding approximately 3 g of tissue to a fine powder in liquid nitrogen with subsequent transfer to a sterile 50 mL conical tube. Fifteen mL of buffer L (DNA extraction buffer) and 2 mL of 10 % SDS (w/v) were added to the powder and vigorously agitated for 5 min prior to incubation at 65 °C for 10 min. Starches and proteins were precipitated by the addition of 5 mL of 5 M potassium acetate with incubation on ice for 20 min and centrifugation for 20 min at 12 000 *g*. The supernatant was filtered through Miracloth and the nucleic acids were precipitated with 10 mL isopropanol and incubation at -20 °C for 1 h. The pellet was recovered by centrifugation at 12 000 *g* for 20 min. The pellet was dissolved in 4 mL buffer M prior to the addition of 10 μL of RNase A (10 mg/mL) and subsequent incubation at 37°C for 30 min. The DNA preparation was purified by repeated extractions with phenol:chloroform:isoamyl alcohol (phenol:$CHCl_3$:*iso*AmOH; 25:24:1, v/v/v) followed by precipitation with 0.1 vol of 3 M sodium acetate (pH 5.2) and 2.5 vol of 95 % cold ethanol. The pellet was recovered by centrifugation at 12 500 *g* for 30 min, washed with 70 % cold ethanol and resuspended in TE (10 mM Tris-HCl, pH 8.0, 1 mM EDTA) or water.

D.4.6 RNA Extraction

RNA was isolated from shoot tissue according to a modified method of Ko *et al.* (1990). Plant material was ground to a fine powder in liquid nitrogen and transferred to a sterile tube. Freshly prepared buffer N (RNA extraction buffer) was added to the powder together with one vol of phenol:$CHCl_3$:*iso*AmOH (25:24:1), agitated vigorously for 20 min, and centrifuged for 15 min at 4 000 *g*. The aqueous phase was further extracted with phenol:$CHCl_3$:*iso*AmOH (25:24:1), and the nucleic acids precipitated with 0.6 vol isopropanol for 30 min on ice, followed by centrifugation at 12 000 *g* for 10 min. The pellet was resuspended in RNase-free water and RNA was precipitated overnight at 4°C with LiCl at a final concentration of 2 M. The pellet was collected by centrifugation at 4000 *g* for 20 min, and resuspended in 500 µL RNase-free water. Nucleic acids were precipitated overnight at -20 °C with 0.1 vol sodium acetate (3 M, pH 5.2) and 2.5 vol of 95 % ethanol. The pellet was collected by centrifugation at 12 500 *g* for 20 min, washed with 70 % ethanol and resuspended in 20 µL RNase-free water. RNA was treated with the DNA-free kit (Ambion) to eliminate DNA contamination, according to manufacturer's instructions and quantified spectrophotometrically or by agarose gel electrophoresis.

D.4.7 Reverse Transcription – Polymerase Chain Reaction (RT-PCR)

Reverse transcription was carried out by mixing 2.5 µg RNA with 3 µL oligo-dT (300 ng/µL), heating to 70°C for 5 min and cooling to room temperature. A cocktail was

prepared composed of 2.5 μL 10X RT buffer, 0.8 μL dNTPs (25 mM), 0.6 μL RNasin (RNase inhibitor), 0.25 μL BSA (10 mg/mL) and 2 μL MMLV RT (250 U/μL) in a total volume of 10 μL (MBI Fermentas). This cocktail was added to the RNA mixture and incubated at 37°C for one h. The enzyme was inactivated by incubation at 65°C for 20 min. Two μL of the RT reaction products were used as template for subsequent PCR using F6H-specific primers in combination with an oligo-dT primer.

D.4.8 Inverse Polymerase Chain Reaction (Inverse PCR)

Inverse PCR following the protocol outlined by Sambrook and Russell (2001) was used in an attempt to isolate the remaining 5′-end of the *F6H* gene. Primers (InvF6HFor and InvF6HRev) were designed to the outermost regions of the F6H·1 fragment. This region was selected for primer design since any fragments amplified using this technique could be verified for conformity in overlapping regions with the cDNA fragment isolated from the library screens. The restriction enzymes used for the digestion of genomic DNA, *Bam*HI, *Eco*RI or *Hin*dIII, did not cleave within the F6H.1 sequence. A total of 5 μg of genomic DNA was digested overnight at 37°C in a total volume of 45 μL. The digested DNA was successively extracted with phenol:$CHCl_3$:*iso*AmOH (25:24:1) and then with $CHCl_3$:*iso*AmOH (24:1). DNA was precipitated with 0.1 vol sodium acetate and 2.5 vol 95 % ethanol. The DNA was recovered by centrifugation, and dissolved in 10 mM Tris-HCl (*pH* 8.0). Digested DNA was subsequently ligated at 15°C overnight at concentration ranges of 0.1 to 1 μg/ml in a total volume of 100 μL, in order to maximize monomeric recircularization. The ligated DNA was extracted and purified as stated above, dissolved in 50 μL water and subjected

to PCR, along with positive and negative controls. Thermal cycling was carried out using 10 μL of the ligated template DNA and 20 pmole of each primer under the following conditions: Denaturation, 94°C, 45 sec; Annealing, 52-57°C, 45 sec; Extension, 72°C, 3 min; repeated for a total of 30 cycles, with a final extension time of 10 min. PCR products were analyzed by electrophoresis and often resulted in the presence of discrete bands, however smears were occasionally observed depending on the restriction fragment used and the annealing temperature. Fragments were gel-purified and cloned into the pGEM-T cloning vector for DNA sequencing as described previously.

D.4.9 GenomeWalker

Another attempt to isolate the remaining 5′-end of the F6H·1 fragment employed a modified GenomeWalker technique from Clonetech. *C. americanum* genomic DNA (2.5 μg) was digested separately with five blunt-end restriction enzymes (*Dra*I, *EcoR*V, *Sca*I, *Sma*I, *Ssp*I) overnight at 37°C in a volume of 100 μL. After complete digestion was verified by agarose gel electrophoresis of an aliquot, the remainder of the digested DNA was purified with phenol:$CHCl_3$:*iso*AmOH (25:24:1) and precipitated with sodium acetate and ethanol. The pellet was solubilized in 20 μL TE and DNA was quantified. A total of 5 ligations were set up with adaptors. Each ligation contained 100 ng of digested DNA, 25 μM ADA1, 25 μM ADA2, 1.6 μL 10X ligation buffer and 3U T4 DNA ligase. The reactions were incubated at 16°C overnight and stopped by incubation at 70°C for 5 min. To each tube, 72 μL TE was added and mixed. The products of these ligation reactions were used as primary PCR templates. Primary PCR reactions consisted of 1 μL of the ligation reaction products, 10 μM AP1 and 10 μM GSP1 in addition to typical PCR

reagents in a total volume of 50 µL. Thermal cycling was carried out as follows for the first round: Denaturation, 94°C, 45 sec; Annealing, 60°C, 30 sec; Extension, 72°C, 3 min; repeated for a total of 30 cycles, with a final extension time of 10 min.

This was followed by a second round of amplifications where primary PCR products were used as template for secondary PCR. In this case, 1 µL of the primary PCR reaction was added to 10 µM AP2, 10 µM GSP2 and typical PCR reagents, in a total volume of 50 µL. Thermal cycling for this round was carried out as follows: Denaturation, 94°C, 30 sec; Annealing, 60°C, 45 sec; Extension, 72°C, 3 min; repeated for a total of 30 cycles, with a final extension time of 10 min. PCR products were analyzed by agarose gel electrophoresis and compared with appropriate controls. Putatively positive fragments subsequently were cloned into the pGEM-T vector for DNA sequencing.

D.4.10 Isolation of Genomic *F6H* Clone

To isolate the genomic region coding for the F6H, genomic DNA (0.1 to 1.0 µg) was used as a template for PCR using the primers, F6HfullFor and F6HfullRev (20 µM each). Thermal cycling for this round was carried out as follows: Denaturation, 94°C, 45 sec; Annealing, 58°C, 1 min; Extension, 72°C, 3 min; repeated for a total of 30 cycles, with a final extension time of 10 min. PCR products were analyzed by agarose gel electrophoresis and compared with appropriate controls. Fragments were subsequently cloned into the pGEM-T vector for DNA sequencing.

D.4.11 Southern Blot Analysis

C. americanum genomic DNA was digested and analyzed according to Sambrook and Russell (2001). For each restriction enzyme, 15 μg of DNA was digested at 37°C overnight in a volume of 45 μL. After ensuring complete digestion with an aliquot of each sample, the remainder was subjected to electrophoresis at 1 V/cm on a 0.7 % agarose gel cast in 1X TAE. After denaturation, the DNA was transferred under alkaline conditions by capillary action overnight to a charged-nylon membrane using buffer L. Once transfer was complete, the membrane was neutralized and the DNA was subjected to UV-crosslinking.

The membrane was hybridized using the protocol stated previously. The probe was labeled using the BioPrime Kit (Amersham) following the manufacturer's protocol. This system is designed for the preparation of biotinylated nucleotide probes for blotting with a Streptavidin-conjugated detection system, in this case the IRDye800-conjugated Streptavidin (Rockland Immunochemicals). The probe detection was carried out according to the protocol suggested by Li-Cor for biotinylated probes. Briefly, the membrane was pre-hybridized using ULTRA-hyb OS at 42°C for 2-3 h. Hybridization was carried out with denatured, biotinylated probe solubilized in a fresh aliquot of ULTRA-hyb OS overnight at 42°C. The membrane was washed twice with Wash Solution I at 45°C for 5 min, followed by two more washes with wash solution II at 45°C for 15 min. The membrane was blocked with buffer O (blocking solution) for 1 h at room temperature. The blot was probed with IRDye800-labelled Streptavidin (1:10 000) in blocking solution for 1 h at room temperature. The membrane was washed with buffer P (blot washing solution) three times, 15 min each, prior to detection.

D.4.12 Cloning Strategies

(a) Construction of pTrc-His-cF6H

The cDNA fragment encoding the near-full length F6H was cloned into an *E. coli* expression system (pTrc-His; Invitrogen) The fragment was amplified by PCR using gene-specific primers, F6HCloneFor and F6HCloneRev, containing *Bam*HI and *Hin*dIII recognition sites, respectively. The DNA insert is positioned downstream and in-frame with a sequence that encodes the $(His)_6$ tag and an enterokinase cleavage recognition site. The pTrc-His-F6H construct was transformed into Top 10 cells by heat shock, and selected on ampicillin (50 μg/mL) containing medium. Recombinant protein expression was carried out according to manufacturer's instructions.

(b) Construction of pPicZ/Zα-cF6H-His

Heterologous expression in the *Pichia pastoris* expression system (EasySelect pPicZ-His; Invitrogen) can be either intracellular or secreted. Secretion of the recombinant protein requires the presence of an in-frame signal peptide. In this case, the $(His)_6$ tag is fused to the C-terminal end of the recombinant protein. The fragment was amplified by PCR using gene-specific primers, PicZF6HFor and PicZF6HRev or PicZαF6HFor and PicZαF6HRev for intracellular and secreted systems, respectively, and containing *Kpn*I and *Sac*II recognition sites. For the pPicZ construct the initiation ATG was part of the yeast consensus sequence (AATAATGTCT). For the pPicZα construct, the insert was cloned in-frame with the N-terminal signal sequence and C-terminal tag. The pPicZ-F6H and pPicZα-F6H constructs were transformed into *Pichia* cells by electroporation according to manufacturer's instructions. Putative multi-copy recombinants were selected by plating cells on media containing increasing concentrations of Zeocin (500,

1000, 2000 μg/mL). *Pichia* transformants were analyzed according to their Mut (methanol utilization) phenotype on minimal media (according to protocol) Transformation of the yeast strain (X-33) with linearized constructs favors single cross-over recombination at the *AOX1* locus, therefore most of the transformants should be Mut^+, however with the presence of the *AOX1* sequences in the plasmid, there is a chance that recombination will also occur in the 3´ *AOX1* region, disrupting the wild-type *AOX1* gene and creating Mut^S transformants. That is, loss of the *AOX1* gene alters the cells ability to metabolize methanol. Testing cell growth on minimal media containing dextrose or methanol will allow the confirmation of the Mut^+ phenotype, where there is little growth with dextrose but normal growth on methanol-containing media. PCR was also used to confirm the presence of the insert.

D.4.13 Preparation of Protein Extracts from Heterologous Expression Systems

(a) Recombinant F6H expression in E.coli

Recombinant protein production in *E. coli* strain Top 10 was induced by the addition of 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) for 4 h at 37°C in LB broth containing 100 μg/mL ampicillin. Bacterial extracts were prepared according to the manufacturer's protocol for purification of recombinant $(His)_6$-tagged proteins on Ni-NTA resin (Qiagen). Briefly, after resuspension of the cell pellet and treatment with lysozyme, cells were subjected to sonication, and lysates were cleared by centrifugation at 10 000 *g* for 20 min. The supernatant then was assayed for F6H enzyme activity or purified on Ni-NTA resin. Given that Tris-based buffers nor the reducing agent DTT could be used, , the affinity purified protein fraction, which was eluted with buffer containing 250 mM

imidazole, was immediately subjected to buffer exchange on a PD-10 column against buffer B. F6H activity was assayed as previously described using the direct enzyme assay, and the reaction products were analyzed by HPLC.

(b) Recombinant F6H expression in P. pastoris

Recombinant protein expression in *P. pastoris* is dependent on the *AOX1* gene promoter, with methanol as the inducer. Recombinant F6H production was induced by the addition of 0.5 % methanol (in the absence of glucose that acts as a repressor) every 24 h over a period of 3 to 4 days, for both intracellular and secretory systems. A buffered culture medium (containing 100 mM potassium phosphate, pH 8.0) was used for cell growth and protein induction in order to enhance protein stability and limit enzyme inactivation. In the case of intracellular recombinant protein production, cells were lysed using glass beads and extracts were prepared according to manufacturer's instructions. Secreted proteins were collected and concentrated by ammonium sulfate precipitation, and the protein that precipitated between 35 to 70 % saturation was pelleted and used for activity determinations and subsequent purification. Cleared extracts were subjected to purification on Ni-NTA resin as previously described.

D.4.14 Purification of Recombinant Proteins

The recombinant protein obtained from both prokaryotic and eukaryotic expression systems was further purified by gel filtration on Superose 12 (16 mm x 300 mm) or Superdex 75 (10 mm x 290 mm) resin. In both cases, the sample was applied

onto either of the columns, previously equilibrated with buffer B, and the protein was eluted with the same buffer at a flow rate of 1 mL/min. Two-mL fractions were collected and assayed for F6H activity using the direct enzyme assay.

D.4.15 Cleavage of Polyhistidine Tag from Fusion Proteins

In order to determine the effect of the $(His)_6$ tag on protein activity, the tag was cleaved from the -pTrc-His-cF6H protein using enterokinase (Gibco) according to manufacturer's instructions. Enterokinase cleaves the fusion product at the junction of the two moieties.

D.4.16 Phylogenetic Analyses

The 15 genes that were selected for analysis from the PubMed and GenBank databases (http://www.ncbi.nlm.nih.gov) included αKDDs where biochemical information is available. The amino acid sequences were aligned using multiple sequence alignment programs, such as CLUSTALW (Thomson and Gibson, 1994) at http://www.ebi.ac.uk/clustalw, using PHYLIP output format. Upon completion of sequence alignment, the data were transferred into a data exploration program, MacClade 4.03 (Maddison and Maddison, 2000), and into a phylogenetic tree building and analysis program, PAUP 4.0, beta 4 (Swofford, 2000). The alignment was subjected to visual inspection and manual editing in MacClade 4.03 before tree reconstruction.

The optimality criterion employed for the distance method was Neighbor-Joining (Saitou and Nei, 1987). The aligned amino acid sequences were analyzed using a heuristic search algorithm with 1000 random addition sequences, and the tree bisection-

reconnection branch (TBR) swapping option. All characters are weighted equally, gaps treated as missing characters, and multi-state taxa interpreted as uncertain. The robustness of the tree branches was tested using bootstrapping analysis with heuristic search based on 1000 replicates and utilizing simple sequence addition and TBR swapping option and the distance measure was set to be equal to mean character difference.

E. RESULTS

E.1 Purification and Biochemical Characterization of the F6H

E.1.1 Purification of the F6H

C. americanum shoot tips were used as the protein source since they exhibited flavonol 6-hydroxylase activity (Anzellotti and Ibrahim, 2000) when incubated with the α-KG-dependent dioxygenase cofactors. Mature leaf and stem tissues exhibited a fraction of the F6H activity observed in shoot tips, ca. 23% and 10%, respectively. This hydroxylation reaction did not occur in the root tissue, or when either tissue was assayed in the presence of monooxygenase cofactors, such as NADPH and FAD.

Using ammonium sulfate precipitation followed by size exclusion chromatography on Superose 12 and ion-exchange on Mono Q columns, the F6H was purified 538-fold with a specific activity of 97.1 pkat/mg and 0.63% recovery (Table 1). Affinity chromatography on an α-KG-Sepharose column resulted in a protein preparation of a high purity as evidenced by SDS-PAGE (Figure 12), although it exhibited a significantly lower specific activity than the Mono Q-purified fraction. In this case, competitive elution with the co-substrate may have resulted in substrate inhibition of the enzyme activity, although dialysis of the eluted protein preparation did not further enhance enzyme activity. Immunoaffinity chromatography could only be achieved using the Superose 12 preparation, and the protein eluted from this column exhibited a lower specific activity than the Mono Q fraction. This is likely the result of the acidic elution conditions, although immediate neutralization of the eluate did not enhance enzyme activity. On the other hand, no protein could be recovered when the Mono Q-purified

TABLE 1

Purification of the flavonol 6-hydroxylase

The crude protein extract was successively purified using the indicated steps and F6H activity was determined using the direct enzyme assay.

Step	Total protein *mg*	Specific activity *pkat/mg*	Total activity *pkat*	Purification *–fold*	Recovery *%*
Crude	113.0	0.18	20.34	-	100
$(NH_4)_2SO_4$	29.9	0.25	7.48	1.4	26.5
Superose 12	4.7	8.53	40.1	47.4	4.16
Mono Q	0.71	97.1	68.9	539	0.63
α-KG-Sepharose	0.04	23.4	0.94	130	0.04
α-D4H-Sepharose [a]	0.41	18.7	7.67	104	0.36

[a] The Superose 12 fraction was used for the immunoaffinity chromatography step on a desacetoxyvindoline 4-hydroxylase (D4H) antibody-conjugated Sepharose column.

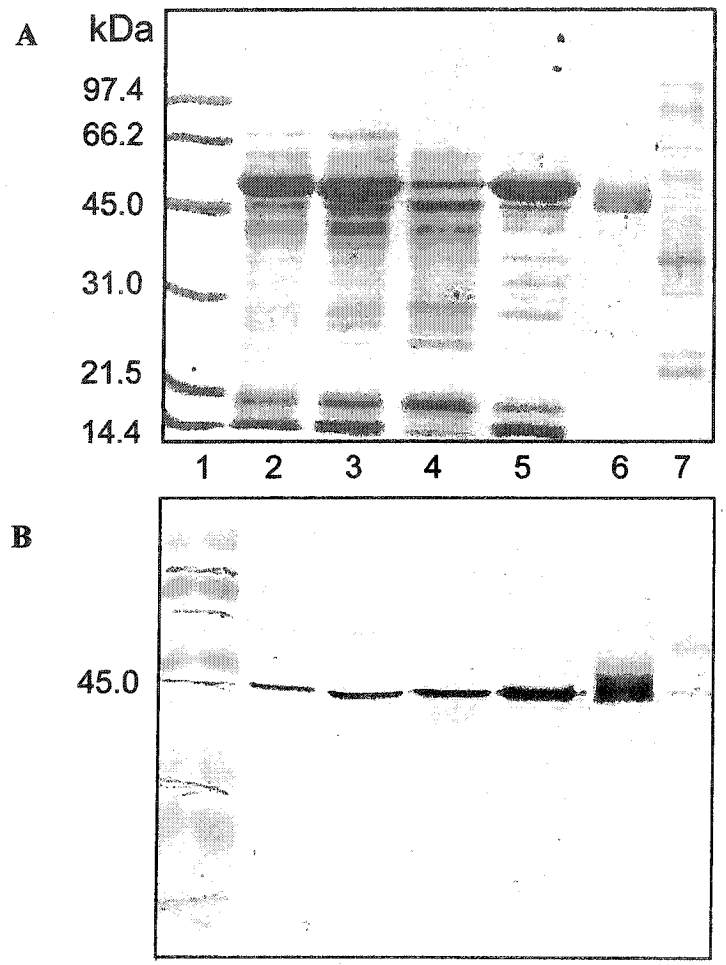


Figure 12: SDS-PAGE (A) and Western blot (B) of F6H at different stages of purification

1, Mr markers

2, Crude extract (10 μg)

3, Ammonium sulfate pellet (10 μg)

4, Superose 12 fraction (10 μg)

5, Mono Q fraction (5 μg)

6, α-KG-Sepharose fraction (2 μg)

7, Anti-D4H-Sepharose fraction (5 μg)

sample was subjected to immunoaffinity purification. The use of either affinity column resulted in the elimination of most contaminating proteins as demonstrated from SDS-PAGE profiles (Figure 12), however since the ligand-affinity technique is more specific in its interaction than the heterologous antibody, the α-KG-sepharose resin allowed for the isolation of a relatively homogenous F6H preparation and may be the preferred method for purification of this class of proteins. Formation of the enzyme reaction product was linear with Superose 12 protein concentrations of 0.5 μg to 80 μg, and with time up to 90 min.

E.1.2 Identification of the Enzyme Reaction Product

The enzyme reaction product of the Superose-12 purified enzyme preparation was identified using several techniques.

(1) HPLC analysis was used to determine retention times (*Rt*) for a set of flavonoid standards (individually and in combination) using both the isocratic and gradient elution methods outlined previously. Their *Rt* values (Figure 13) and UV absorption spectra (Figure 14) were determined and compared with those of the F6H enzyme reaction product. The peak corresponding to the putative 6-hydroxylated product was collected from several assays and re-chromatographed on HPLC, separately and in the presence of a reference sample of 6-hydroxy-3,7,4′-trimethylquercetin (3,7,4′-trimethylquercetagetin; TriMeQg). Upon isocratic elution, the enzyme reaction product exhibits an *Rt* value of 3.2 min (Figure 13B) and a λ_{max} of 254, 279, 345 nm (Figure 14B), as compared with the respective values, 5.2 min and 254, 355 nm for the substrate,

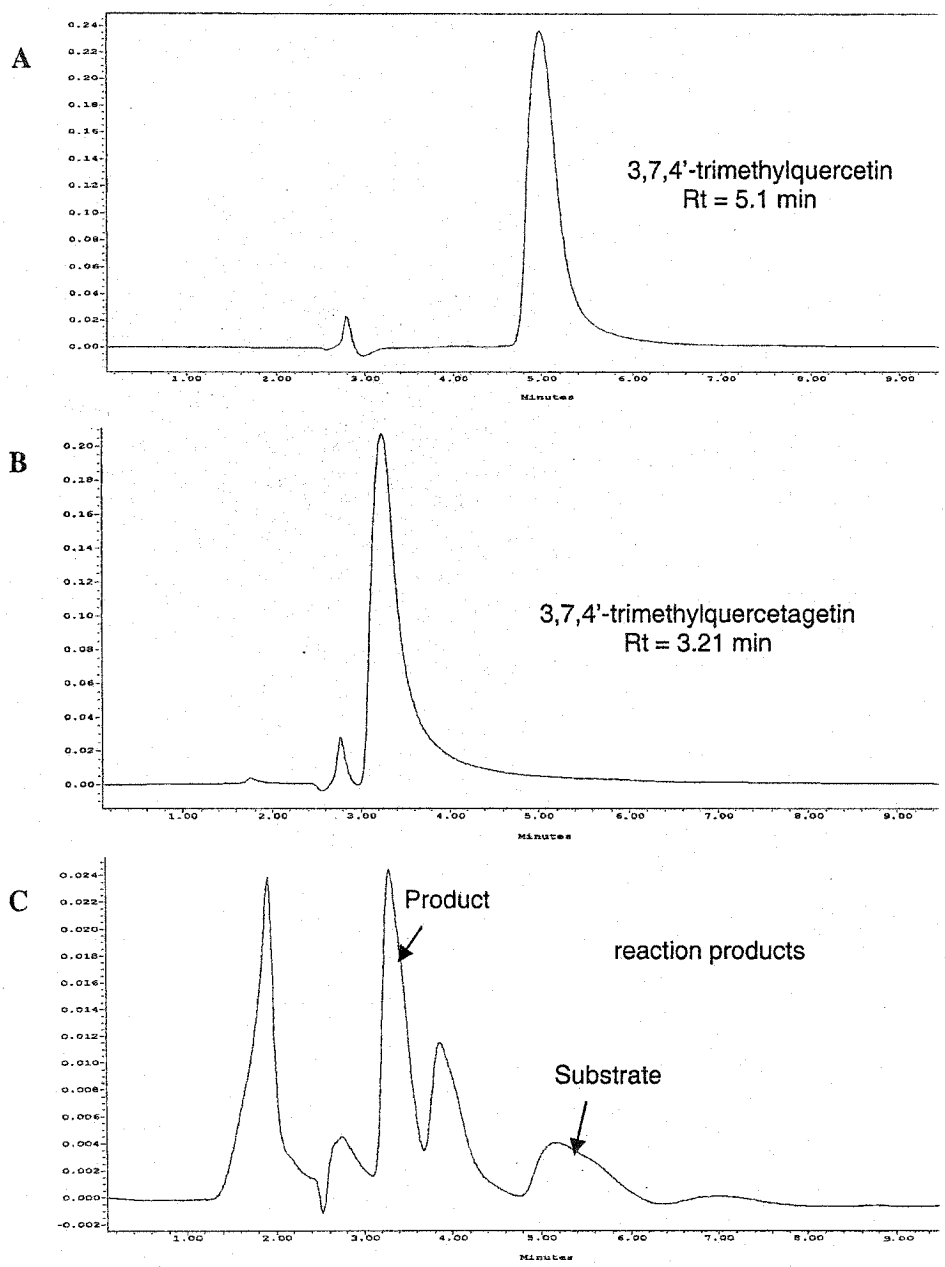


Figure 13: Identification of the F6H reaction product

(A) HPLC profile of 3,7,4´-trimethylquercetin (TriMeQ) standard using the isocratic elution method; solvent system: methanol-water-acetic acid (75:24:1; v/v/v).
(B) HPLC profile of the 3,7,4´-trimethylquercetagetin (TriMeQg) standard.
(C) An enzyme assay profile showing a major reaction product and the residual substrate. Additional peaks represent endogenous metabolites present in the Superose 12-purified fraction which were also observed in control reactions.

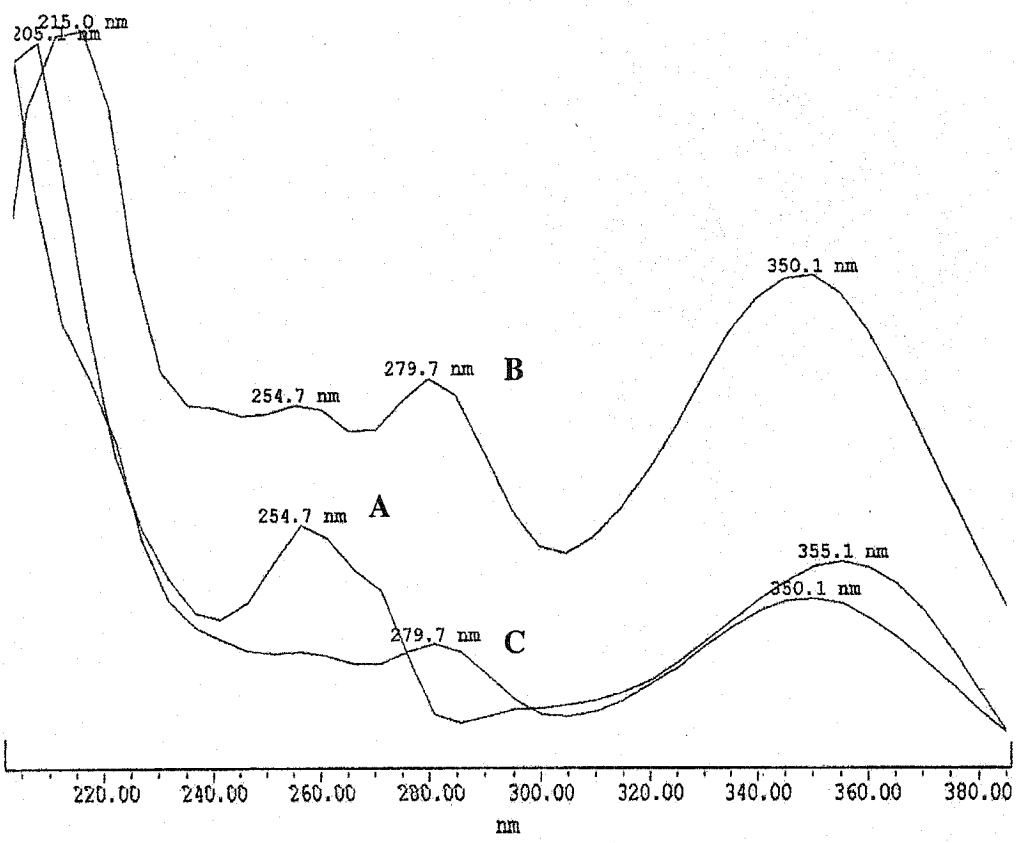


Figure 14: UV-Absorption maxima of the F6H substrate (A), product (B) and TriMeQg reference sample (C)

(A) TriMeQ standard (*Rt* = 5.2 min).
(B) TriMeQg standard (*Rt* = 3.2 min).
(C) F6H reaction product (*Rt* = 3.2 min).

3,7,4′-trimethylquercetin (TriMeQ; Figures. 13, 14). These two parameters determined for the reaction product are identical to those of the reference compound, TriMeQg.

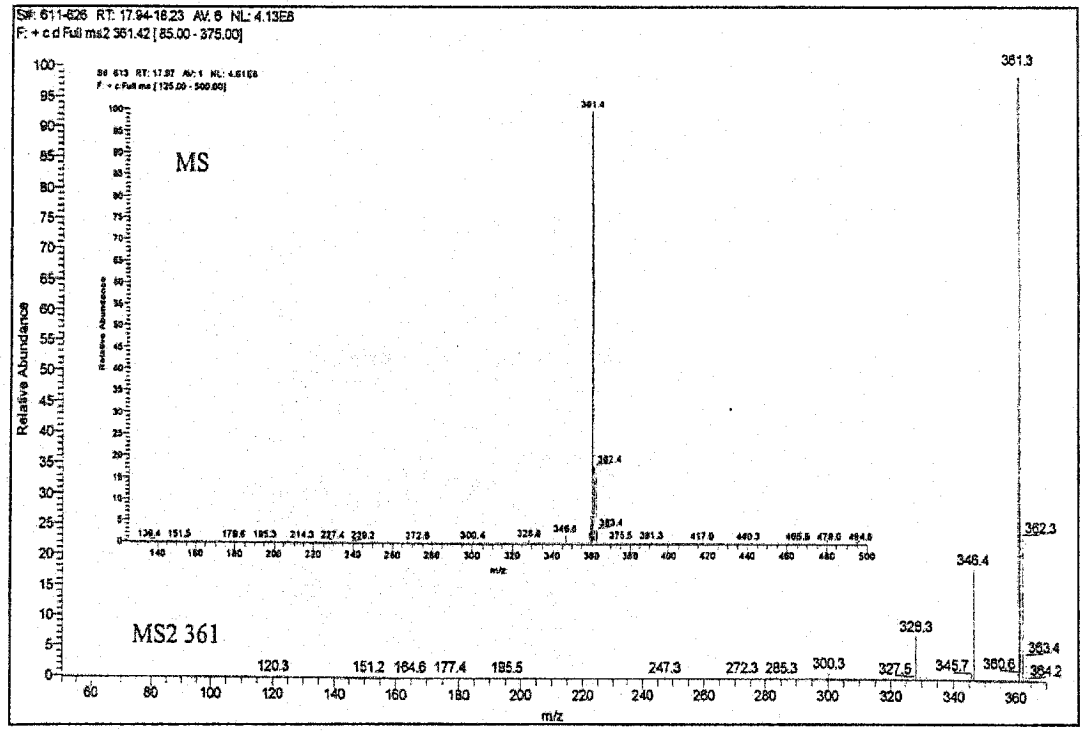
(2) The HPLC peak corresponding to the enzyme reaction product was subjected to LC-MS analysis that produced an *m/z* signal of 361.4 as expected for TriMeQg, before fragmentation of the ring system, and an *m/z* fragmentation pattern typical of TriMeQg (Figure 15).

(3) In order to verify the position of hydroxylation on the enzyme reaction product, an indirect enzyme assay was carried out, in which the F6H reaction was coupled with a flavonol 6-*O*-methyltransferase (F6OMT) activity obtained from the same tissue. Previously, F6OMT has been characterized biochemically, and was shown to be the sole Mg^{2+}-dependent enzyme of the *Chrysosplenium* OMTs (De Luca and Ibrahim, 1985a). The F6H reaction product, TriMeQg obtained from several direct F6H assays was coupled to the Mg^{2+}-dependent F6OMT reaction in the presence of [^{14}C]AdoMet, and chromatographed on TLC. Control enzyme assays carried out in the absence of Mg^{2+} and with the F6H flavonol substrate separately, resulted in endogenous product formation only (Figure 16, lane 5), thereby demonstrating the specificity of the enzyme reaction to position 6 of the flavonol substrate. These results clearly indicate the specific methylation of the 6-hydroxyl group of TriMeQg by a F6OMT, the only Mg^{2+}-dependent enzyme in the stepwise methylation of *Chrysosplenium* flavonoids.

E.1.3 Substrate Specificity

The substrate specificity of the F6H was tested using the Superose 12-purified enzyme preparation with various flavonol substrates (Table 2) known to occur in *C.*

A



B

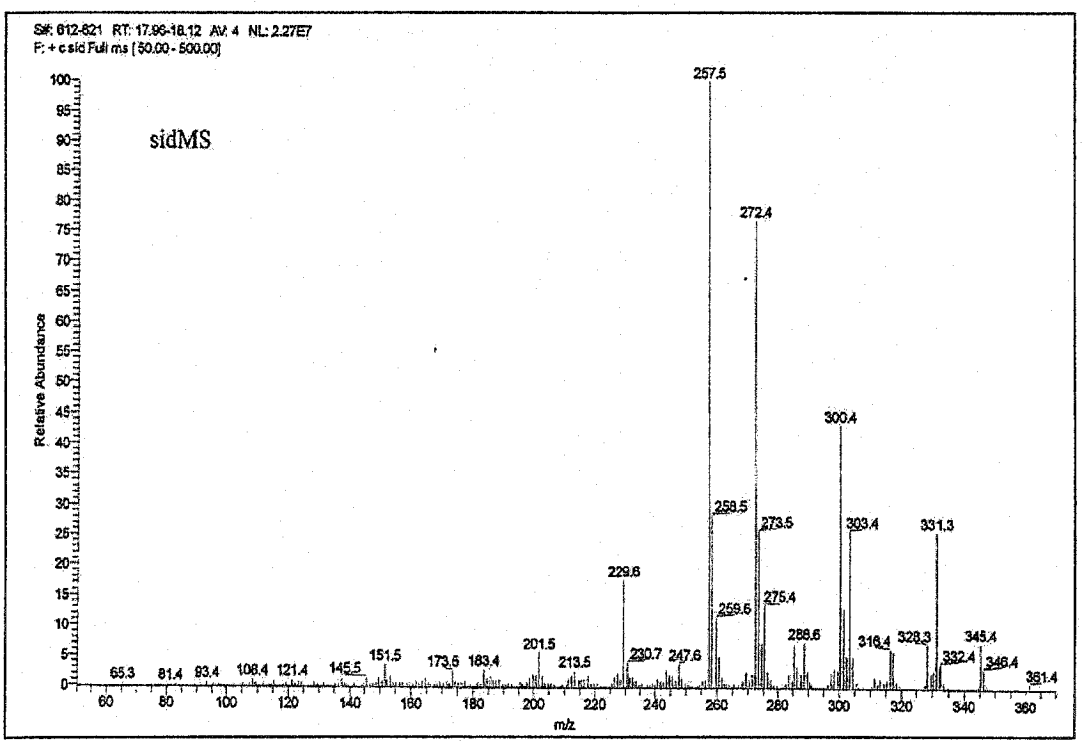


Figure 15: Mass spectrum of F6H reaction product and 3,7,4´-TriMeQg

(A) Molecular ion (*m/z*) for the F6H reaction product
(B) Molecular ions (*m/z*) for fragmentation products of 3,7,4´-TriMeQg

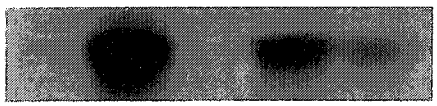


Figure 16: Autoradiogram of the F6OMT-coupled reaction products

1, F6OMT + TriMeQg + [^{14}C]AdoMet

2, F6OMT + TriMeQg + [^{14}C]AdoMet + Mg^{2+}

3, F6OMT + [^{14}C]AdoMet

4, F6H (with required cofactors and substrates) + F6OMT + [^{14}C]AdoMet + Mg^{2+}

5, F6H (with required cofactors and substrates) + F6OMT + [^{14}C]AdoMet

TABLE 2

Substrate specificity of the flavonol 6-hydroxylase

The Superose 12 fraction was assayed with 5 μM of the indicated substrates using the direct enzyme assay. [a]

Substrate [b]	Relative activity [c] (%)
3,7,4´-Trimethylquercetin	100
3,7-Dimethylquercetin	77
3,7,3´,4´-Tetramethylquercetin	70
3-Methylquercetin	18
Quercetin	10
Kaempferol	7
Myricetin	4

[a] Other substrate concentrations tested were; 1, 2.5, 5, 10 and 20 μM. The results obtained were similar for 1, 2.5, 5 and 10 μM. Assays carried in the presence of 20 μM flavonol substrate exhibited reduced activities and slightly different specificities, resulting from substrate inhibition.

[b] No F6H activity was detected with naringenin, eriodyctyol, apigenin, luteolin, gossypetin or quercetagetin when used as substrate, regardless of concentration. Rhamnetin (7-*O*-methylquercetin, 14 % at 20 μM), tamarexitin (4´-*O*-methylquercetin, 10 % at 50 μM) and isorhamnetin (3´-*O*-methylquercetin, 8 % at 50 μM)

[c] Estimated as 8.13 pkat/mg for 100% activity

americanum, mainly quercetin and its methylated derivatives. The results indicate that the hydroxylase reaction exhibits a strong preference for 3,7,4′-TriMeQ, followed, in descending order, by 3,7-dimethyl- ≈ 3,7,3′,4′-tetramethyl- >> 3-methylQ, as substrates. The aglycones, quercetin, kaempferol, myricetin, quercetagetin and gossypetin were not accepted as substrate to a significant extent thus substantiating the specificity of 6 hydroxylation for methylated flavonols.

E.1.4 Immunotitration

Whole rabbit anti-D4H anti-serum (Vazquez-Flota *et al.*, 1997) did not significantly inhibit F6H activity when added to either the crude or partially purified enzyme preparations, in comparison with the non-immune serum. On the other hand, the IgG-purified anti-serum resulted in a dose-dependent inhibition of enzyme activity with preparations from different stages of purification, relative to the non-immune serum, with a maximum inhibition of 60 % for the Superose 12-purified fraction (Figure 17). The observed difference in the level of reduction of F6H activity in the crude preparation, as compared with the Superose 12-purified sample may be attributed to the stabilizing effect of elevated protein concentrations in the former sample. In addition, extraneous proteins may have cross-reacted with the antiserum, thereby titrating out anti-D4H IgGs.

E.1.5 pH Optimum and Apparent *pI*

The optimum pH for the hydroxylation reaction was determined using the Superose 12-purified fraction in the presence of appropriate buffers (Figure 18). The enzyme exhibited a sharp optimum at pH 8.0 in 50 mM Tris-HCl buffer, which decreased

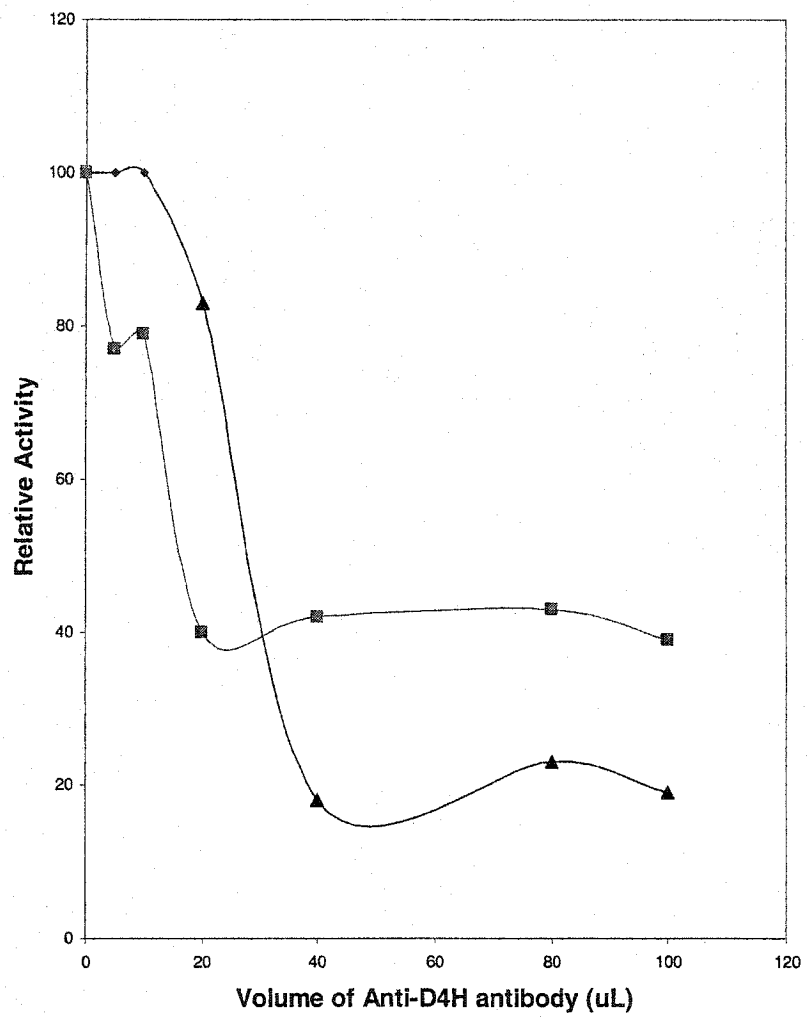


Figure 17: Immunotitration of F6H activity using IgG-purified anti-D4H serum

The relative activity was determined in comparison to parallel experiments using pre-immune serum using the direct enzyme assay.
Protein content of both sera was kept constant at 0.18 mg/ml.
■, ammonium sulfate pellet; ▲, Superose 12 fraction.

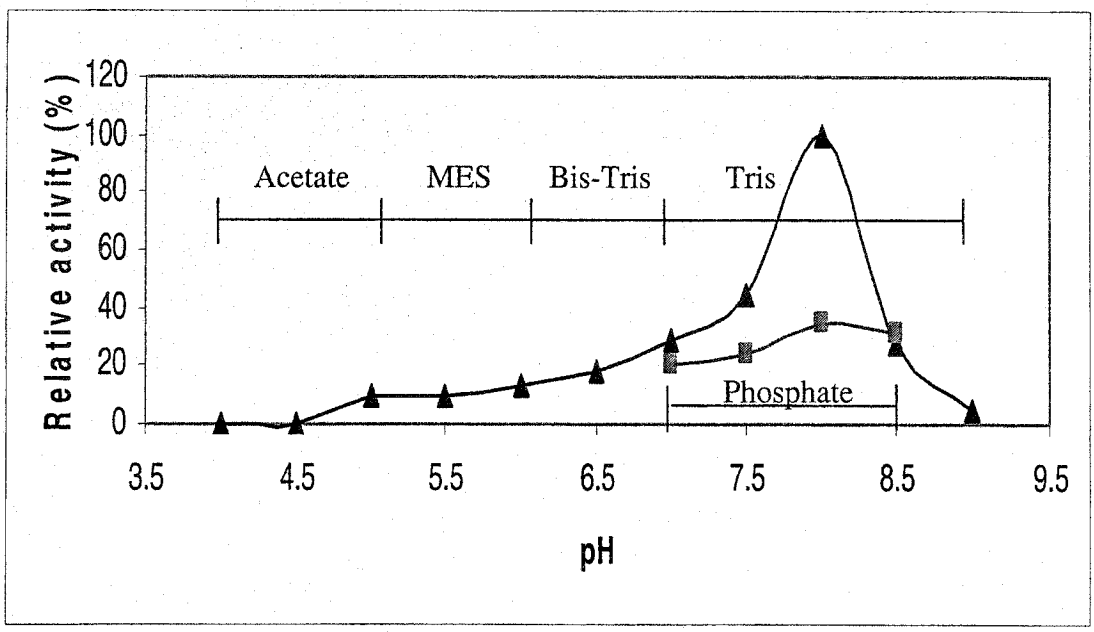


Figure 18: pH Dependence of the F6H reaction

Control activity was determined using the direct enzyme assay in the presence of 50 mM of the indicated buffers, αKDD cofactors and 5 μM 3,7,4′-TriMeQ.

by 55% and 72 % at pH 7.5 and 8.5, respectively. However, it was later determined that a purification scheme utilizing buffers at a slightly reduced pH (~7.5) resulted in a preparation with a more stable enzyme activity. Chromatofocusing of the Superose 12-purified fraction on a Mono P column under the stated conditions gave an apparent *pI* value of 5.1, indicating that F6H is an anionic protein under physiological conditions (Figure 19). This chromatographic step was not included in the purification procedure since it significantly affected enzyme stability, in that F6H activity was lost within 24 hours of elution from the Mono P column, likely due to changes in pH during chromatography.

E.1.6 Molecular Mass

The molecular mass of the 6-hydroxylase, as determined by SDS-PAGE and probing with the anti-D4H antibody, is 45 kDa (Figure 12). It compares well with the ~ 43 kDa estimated by gel filtration (Figure 20B), and indicates that the F6H is a monomer. There is a small peak of F6H activity at a predicted molecular mass of ~91 kDa (Figure 20B), which suggests dimerization of F6H monomers. There has been some evidence of dimerization of αKDDs, likely as a result of the purification process, although these enzymes are believed to function as monomers *in vivo* (Britsch *et al.*, 1992; Ruetschi *et al.*, 1993; Fukumori and Hausinger, 1993; Eichhorn *et al.*, 1997). Nevertheless, the fraction with highest F6H activity (Figure 20B) represents the monomeric 43 kDa form of the enzyme.

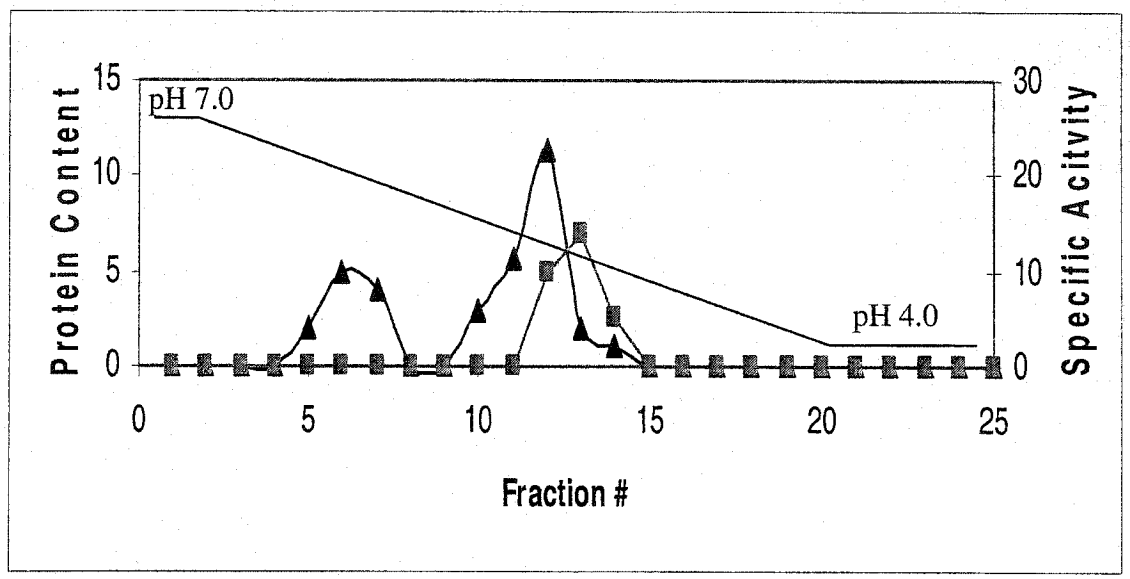


Figure 19: Chromatofocusing of F6H on a Mono P column

F6H activity was determined using the direct enzyme assay.
(▲) Protein content (µg/ml)
(■) Specific Activity (pkat/mg)

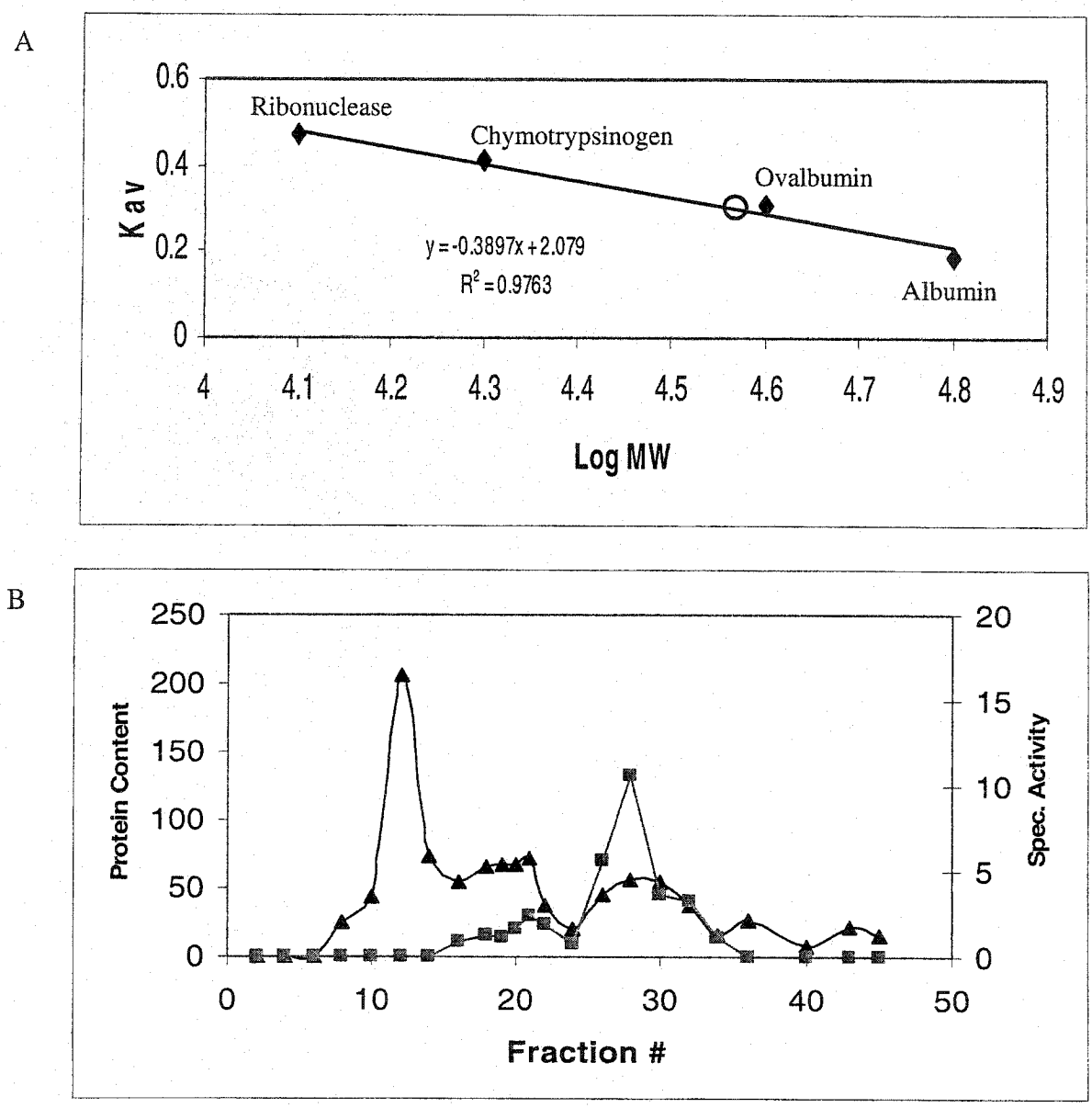


Figure 20: Gel filtration chromatography on Superose 12 resin

F6H activity was determined using the direct enzyme assay.
(A) Calibration curve for the Superose 12 column using M_r standards.
(B) Superose 12 fractionation of F6H activity
(▲) Protein content (µg/ml)
(■) Specific activity (pkat/mg)

E.1.7 Cofactor Dependence

Maximum enzyme activity was obtained upon incubation of the assay mixture with 0.25 mM ferrous ions, 10 mM α-KG and 10 mM ascorbate as cofactors, in the presence of 2.5 and 5 μM TriMeQ as the flavonol substrate (Table 3). However, Fe^{2+} had the greatest effect on enzyme activity and enzyme reactivation in both the crude and partially purified samples. This property was confirmed by the complete inhibition of enzyme activity upon incubation with 5 mM EDTA.

The oxygen dependence of the hydroxylation reaction was assessed by pre-incubation of the enzyme preparation with a mixture of glucose (1mM) and glucose oxidase (10 U) for 20 min prior to the addition of the required F6H cofactors; where little product was formed under these conditions. Similarly, there was no detectable enzyme activity under anaerobic conditions (saturation with nitrogen gas; data not shown). On the other hand, an oxygen-generating system involving the addition of catalase (1 mg/mL) resulted in a 1.5-fold increase in enzyme activity of partially purified preparations, which may have resulted from the removal of peroxides by catalase. However, to determine whether this stimulation was the direct result of catalase activity, and not due to its effect as a stabilizing protein, a control assay containing the enzyme preparation in the presence of bovine serum albumin (1 mg/mL) was carried out in parallel, but did not significantly affect F6H activity (data not shown).

E.1.8 Effect of Metal Ions

Various cations were tested for their ability to substitute for ferrous ions in the F6H reaction (Table 4). Of the cations tested (Co^{2+}, Cu^{2+}, Ni^{2+}, Mn^{2+}, Zn^{2+} and Mg^{2+}),

TABLE 3

Cofactor dependence of the flavonol 6-hydroxylase [a]

Additions	Relative Activity (%)
Cofactors	
Control [b]	0
α-Ketoglutarate (10 mM)	62
Fe^{2+} (0.25 mM)	75
Ascorbate (10 mM)	46
α-Ketoglutarate, Fe^{2+}, Ascorbate	100
Other Additions	
Control [c]	100
Catalase (1mg/mL)	154
Glucose oxidase (10 units) + 1 mM glucose	15
EDTA (1 mM / 5 mM)	47 / 0
Succinate (10 mM / 20 mM)	65 / 23
Fe^{3+} (0.25 mM / 0.5 mM)	71 / 44

[a] The direct enzyme assay was used to determine F6H activity.

[b] Assays were carried out using the Superose 12-purified fraction in the presence of 50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 10 mM DTT, 10% glycerol (v/v) and 2.5 μM 3,7,4´-trimethylquercetin.

[c] As above, but in the presence of 10 mM α-ketoglutarate, 0.25 mM Fe^{2+} and 10 mM ascorbate (100% relative activity = 8.53 pkat/mg).

TABLE 4

Effect of metal cations on flavonol 6-hydroxylase activity

Metal [a]	Relative Activity [b] (%)
Mn^{2+}	17
Cu^{2+}	5
Ni^{2+}	0
Co^{2+}	0
Zn^{2+}	0
Mg^{2+}	0
Fe^{3+}	70

[a] All cations were in the form of sulfate salts, and assayed at a concentration of 0.25 mM. No activity was observed at 0.1 mM concentrations.

[b] The Superose 12-purified fraction was used in the presence of all required cofactors and substrates. (Control activity, 100% relative activity = 9.1 pkat/mg).

at 0.1 mM and 0.25 mM concentrations, Mn^{2+} was the only divalent ion that resulted in any significant (17% of control) F6H activity. On the other hand, incubation of the assay mixture with 0.25 mM ferric sulfate resulted in 70% of control activity, which was abolished in the absence of ascorbate (data not shown). This suggests the involvement of ascorbate in the reduction of ferric ions.

E.1.9 Protein Stability

Given the fact that enzymes of this class are susceptible to both oxidative and non-oxidative inactivation, maintaining the enzyme preparations and chromatography buffers under nitrogen gas saturation helped to preserve protein activity. At the same time, ferrous ions, which significantly contribute to F6H activity, also can serve as a source for inactivation. This is due to the oxidation of Fe^{2+} to Fe^{3+} ions in the presence of hydrogen peroxide, commonly known as Fenton chemistry. In addition, αKDDs have been shown to catalyze the decarboxylation of the cosubstrate, α-KG, without concomitant hydroxylation of the primary substrate, with ascorbate acting as the oxygen acceptor (Myllyla *et al.*, 1984). This uncoupling results in an alternate inactivation mechanism that leads to the formation of a dead-end complex. Given that α-KG is a metabolite common to several biosynthetic pathways, its presence cannot be avoided in crude and partially purified protein extracts, but can be limited by further purification of the sample.

The instability of this enzyme posed a significant problem for its purification and biochemical characterization, although the addition of the protease inhibitor PMSF, as well as the reducing agents, DTT and ascorbate during extraction improved enzyme stability. Nevertheless, more than 50 % of the enzyme activity of the Superose 12-

purified fraction was lost upon storage for 5 days at -20°C even in the presence of 20% glycerol as a stabilizing agent, whereas storage at 4°C under similar conditions resulted in a gradual loss of F6H activity to only 20 % within one week. In addition, sequential chromatography of the enzyme preparation resulted in a progressively more unstable F6H activity. The addition of cofactors to the storage buffer did not significantly enhance protein stability of either the Superose 12- or Mono Q-purified preparations but in fact, interfered with sample detection on FPLC when included in the chromatography buffers.

E.1.10 Kinetic Analysis

The kinetic constants of F6H were determined by the direct enzyme assay using the Superose 12-purified preparation as an appropriate source of abundant and stable enzyme protein, and values that fell within an acceptable range of standard deviation were used in the analysis. The flavonoid product formation was linear with reaction time, and the amount of protein added. When the flavonol substrate was varied, higher concentrations (>10 μM) resulted in substrate inhibition (data not shown), which occurs when the substrate molecule binds with an improper enzyme form, most likely after the release of the hydroxylated flavonol product. This phenomenon also may be due to the low solubility of the substrate at elevated concentrations. This was not the case with α-KG since saturation was obtained without inhibition. The replots of the kinetic data gave K_m values of 0.27 μM, 60 μM and 5.26 μM for TriMeQ, α-KG and oxygen, respectively, and a V_{max} of 9.7 pkat/mg for the conversion of TriMeQ to its 6-hydroxy derivative, TriMeQg.

E.1.11 Initial Velocity Patterns

Since the F6H reaction exhibited Michaelis-Menten kinetics and inhibition was observed with the flavonol substrate, initial velocity patterns were obtained using a flavonol concentration below 10 μM. Detailed kinetic studies were carried out by varying the concentration of one substrate in the presence of different fixed concentrations of the second substrate, with that of the third substrate remaining constant (Figures 21, 22 and Table 5). Lineweaver-Burke plots of various concentrations of TriMeQ and different fixed concentrations of α-KG with constant O_2 gave an intersecting pattern (Figure 21A), indicating a reversible connection between the two substrates. On the other hand, when the α-KG was varied with different fixed concentrations of TriMeQ under saturating oxygen, parallel lines were obtained (Figure 21B), suggesting that saturation with oxygen severed the connection between the remaining substrates, tentatively placing oxygen between α-KG and TriMeQ in the binding order. Reciprocal plots of various oxygen concentrations at different fixed concentrations of α-KG or TriMeQ yielded converging patterns (Figure 22A to C).

E.1.12 Product Inhibition Patterns

The order of substrate binding and product release was determined from product inhibition studies (Figure 23 and Table 6). The rate equations derived for a ter ter sequential mechanism predict that the last product released acts as a competitive inhibitor of the reaction with respect to the first substrate, since they bind to the same enzyme form. Non-competitive inhibition patterns are expected for all other substrate-product combinations. Succinate competitively inhibited the reaction with respect to α-KG, but

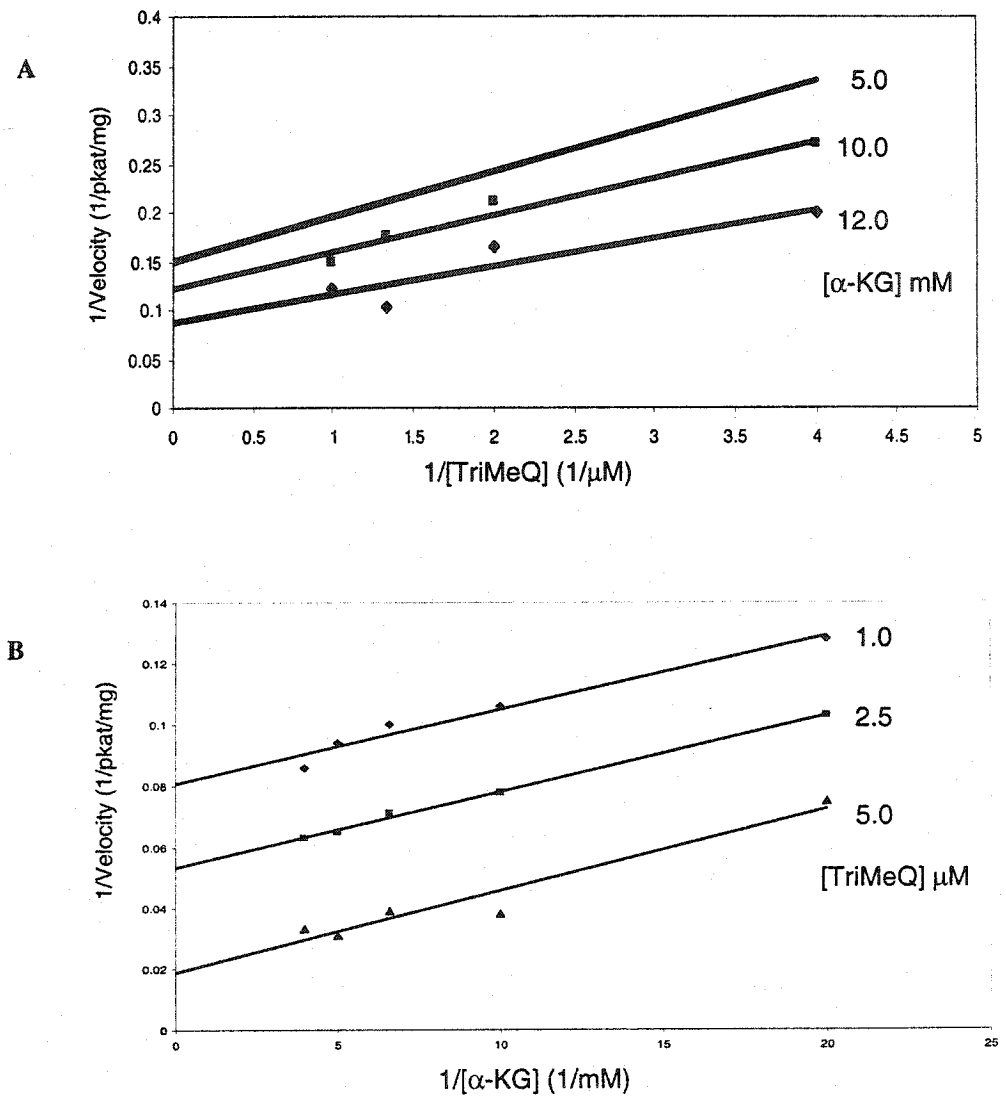


Figure 21: Double reciprocal plots of F6H catalyzing the conversion TriMeQ to TriMeQg

(A) The effect of varying concentrations of TriMeQ on the rate of F6H reaction in the presence of different, fixed concentrations of α-KG.

(B) The effect of varying concentrations of α-KG on the rate of F6H reaction in the presence of different, fixed concentrations of TriMeQ, under saturating oxygen conditions.

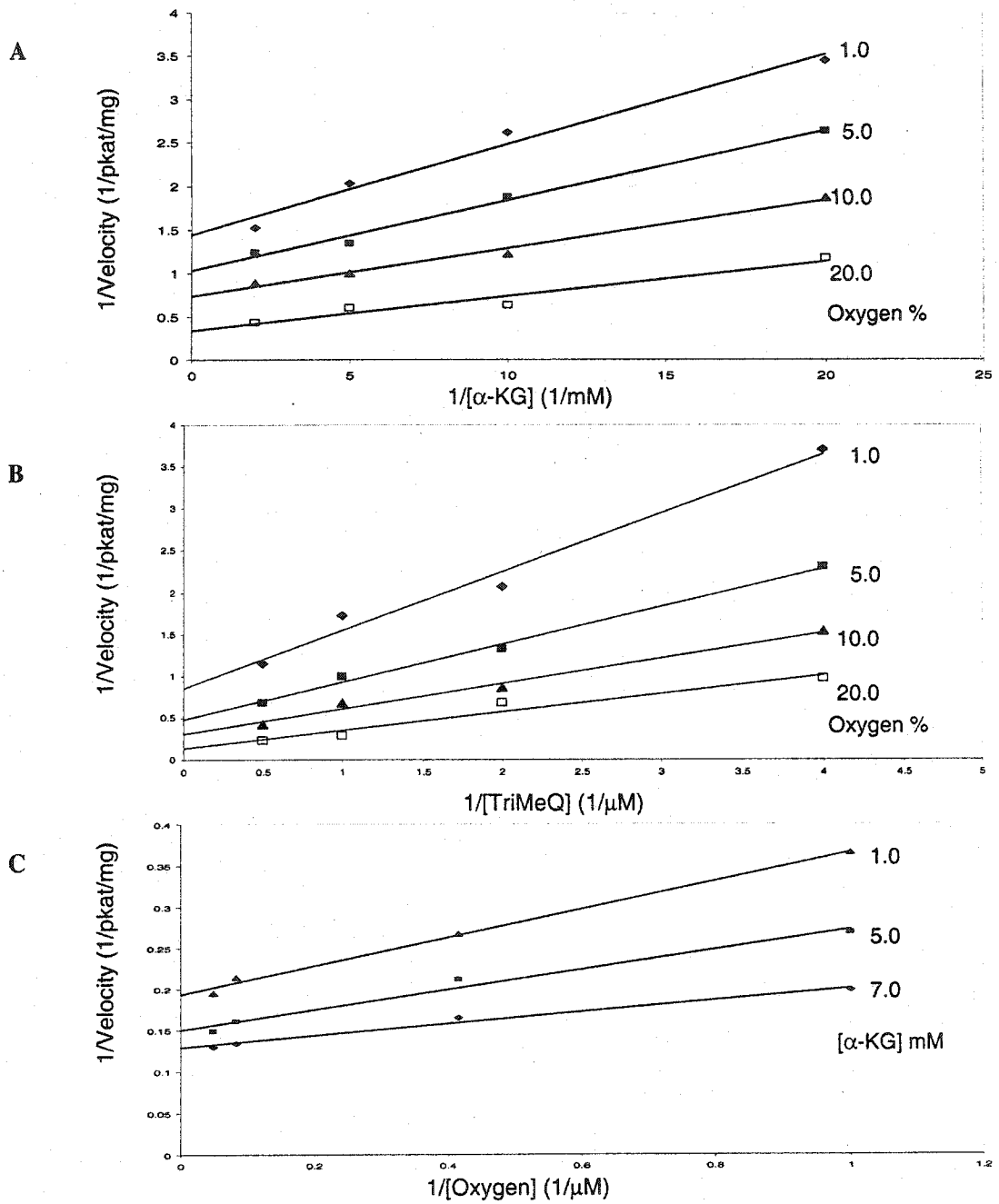


Figure 22: Double reciprocal plots of F6H catalyzing the conversion TriMeQ to TriMeQg

(A) Rate of F6H reaction in the presence of varying α-KG concentrations, at different, fixed concentrations of oxygen.
(B) Rate of F6H reaction in the presence of varying TriMeQ concentrations, at different, fixed concentrations of oxygen.
(C) Rate of F6H reaction in the presence of varying oxygen concentrations, at different, fixed concentrations of α-KG.

TABLE 5

Patterns of substrate interaction kinetics [a]

Varied substrate	Changing, fixed substrate	Constant substrate	Pattern	Apparent kinetic value [b]
TriMeQ	α-KG (a)	O_2	Converging	K_a = 60.0 μM
α-KG	O_2 (b)	TriMeQ	Converging	K_b = 5.26 μM
α-KG	TriMeQ (c)	O_2	Converging	K_c = 0.27 μM
O_2	α-KG	TriMeQ	Converging	
TriMeQ	O_2	α-KG	Converging	
O_2	TriMeQ	α-KG	Converging	
α-KG	TriMeQ	$O_{2\ (saturating)}$	Parallel	

[a] Abbreviations used: TriMeQ, 3,7,4´-trimethylquercetin; α-KG, α-ketoglutarate

[b] Replots from substrate interaction kinetics were used to determine the values of kinetic parameters.

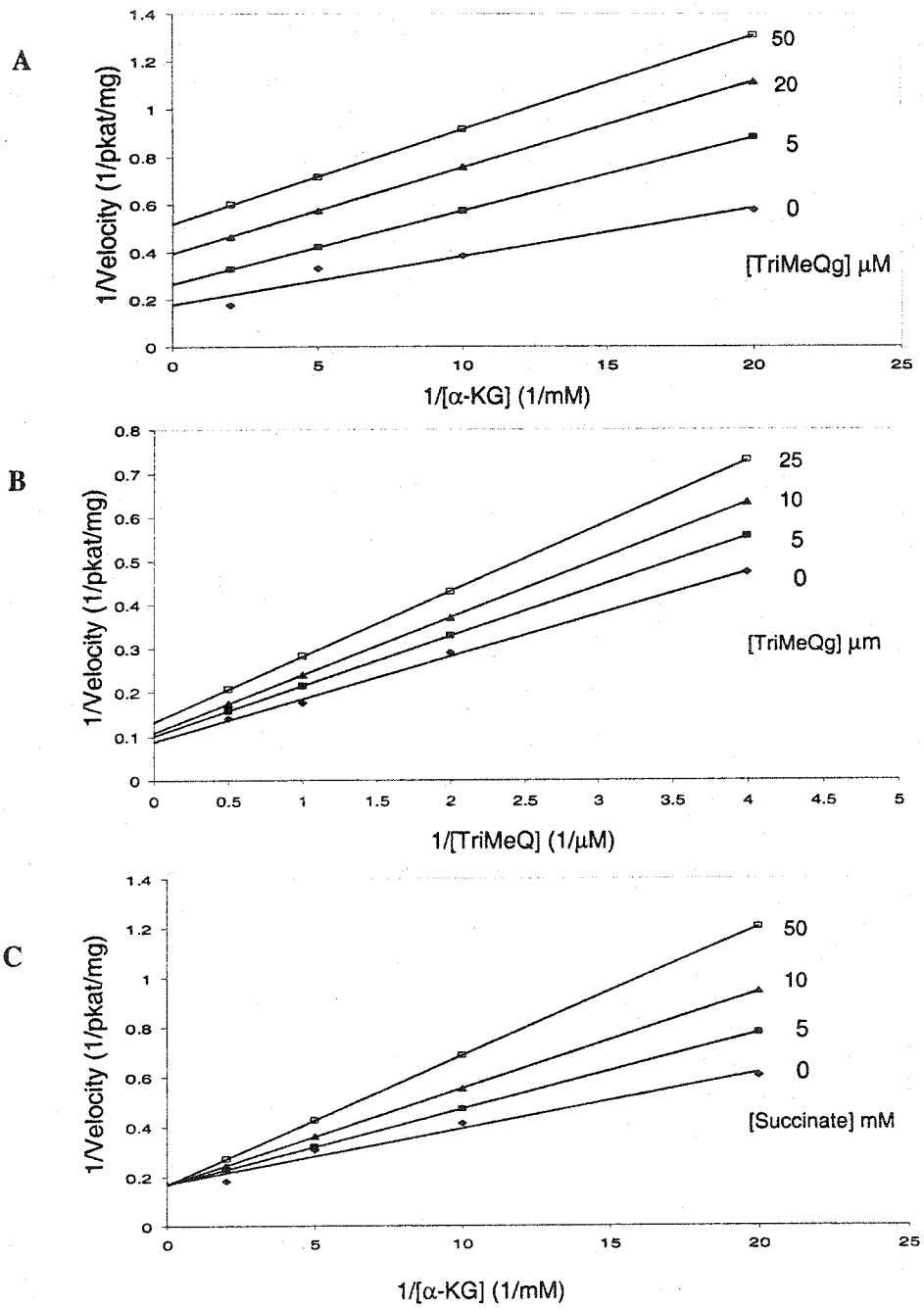


Figure 23: Product inhibition kinetics

(A) Effect of varying concentrations of α-KG on the reaction rate of F6H, in the presence of different fixed concentrations of TriMeQg.
(B) Effect of varying concentrations of TriMeQ on the reaction rate of F6H, in the presence of different fixed concentrations of TriMeQg.
(C) Effect of varying concentrations of α-KG on the reaction rate of F6H, in the presence of different fixed concentrations of succinate.

TABLE 6

Patterns of product inhibition kinetics [a]

Substrate	Product	Inhibition	Kinetic parameter [b]
α-KG	Succinate	Competitive	$K_{i\ \text{Succinate}} = 1.1$ mM
α-KG	TriMeQg	Non-Competitive	
TriMeQ	TriMeQg	Non-Competitive	$K_{i\ \text{TriMeQg}} = 118$ μM
TriMeQ	Succinate	Non-Competitive	
O_2	Succinate	Non-Competitive	
O_2	TriMeQg	Non-Competitive	

[a] Abbreviation used: TriMeQ, 3,7,4´-trimethylquercetin; TriMeQg, 3,7,4´-trimethyl-quercetagetin; α-KG, α-ketoglutarate.

[b] Replots from product inhibition kinetics were used to determine the values of the kinetic parameters.

was non-competitive with respect to TriMeQ. TriMeQg was found to be a non-competitive inhibitor of the hydroxylase reaction with respect to TriMeQ, α-KG or oxygen. These results suggest an ordered sequence of substrate binding and product release (Figure 24). A random mechanism was excluded, since this would have resulted in competitive patterns of product inhibition for the various combinations.

E.2 Cloning and Molecular Characterization of the Flavonol 6-Hydroxylase

E.2.1 Internal Peptide Identification

Several peptides were obtained from tryptic digestion of the ligand-affinity purified F6H protein, and identified by nano-electrospray tandem mass spectrometry after comparison with known peptide sequences in the Sequest database. One peptide in particular, Micro1 – DNGWILLHIGDSNGHR, exhibited a significant degree of similarity to other known flavonol αKDDs such as F3H, FLS and ANS, with six amino acids being strictly conserved among these sequences. Two other peptide sequences, Micro2 – KIVEACEDWG and Micro3 – TLMAGLACKLLGVL, exhibited limited similarity with this category of enzymes. However, given the isobaric nature of particular sets of amino acids, Ile and Leu for example, the identification of certain sequences was somewhat more ambiguous than others. Therefore, we chose to design degenerate primers for Micro1 and Micro2 to be used in the F6H cloning strategy.

E.2.2 Isolation of a Putative F6H cDNA Clone

Although three methods were employed in attempts to isolate a cDNA fragment encoding the F6H from a *C. americanum* bacteriophage expression library (Gauthier *et*

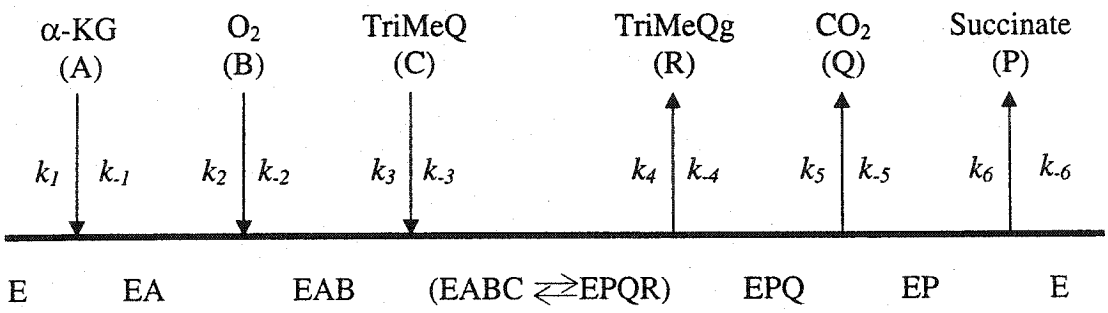


Figure 24: The proposed (ter ter) reaction mechanism for the F6H as derived from substrate interaction and product inhibition kinetic patterns

al., 1996), only two allowed the isolation of putative positive fragments. The immunological screening of the expression library using whole anti-D4H antisera was not successful since it did not produce any clones with a signal significantly higher than background. This was likely due to two separate factors. The first, may have resulted from the use of whole antisera, as opposed to an IgG-purified preparation; a shortcoming that could not be avoided, given the limited amount of available antiserum, and the number of membranes to be probed. Second, the antiserum produced a higher background signal with recombinant protein preparations as opposed to plant extracts. This is not unexpected since bacterial artifacts may have been present in the recombinant D4H preparation used for antibody production.

In order to isolate a putative F6H clone, a PCR-based strategy (Friedman *et al.*, 1990) was used to screen a *Chrysosplenium* cDNA bacteriophage library. The PCR-based screening strategy outlined in Figure 25, although less labor-intensive and time-consuming, yielded a significant number of false positives. These were limited by the use of oligonucleotide primers of high specificity in subsequent screens, optimized PCR cycling conditions and the inclusion of reagents that limit cross-reactivity, such as 5 % DMSO and 20 mM ammonium sulfate, in the PCR reactions.

In order to maximize the length of potential fragments, the initial screens utilized several primer combinations based on the internal peptide sequences obtained. Certain combinations did not amplify any fragments and/or produced false positives regardless of optimization. For the first round of screening, PCR primer sets consisting of either one degenerate peptide-specific primer (Micro1 or Micro2) and one vector primer (T3 and T7) or two peptide-specific primers were used.

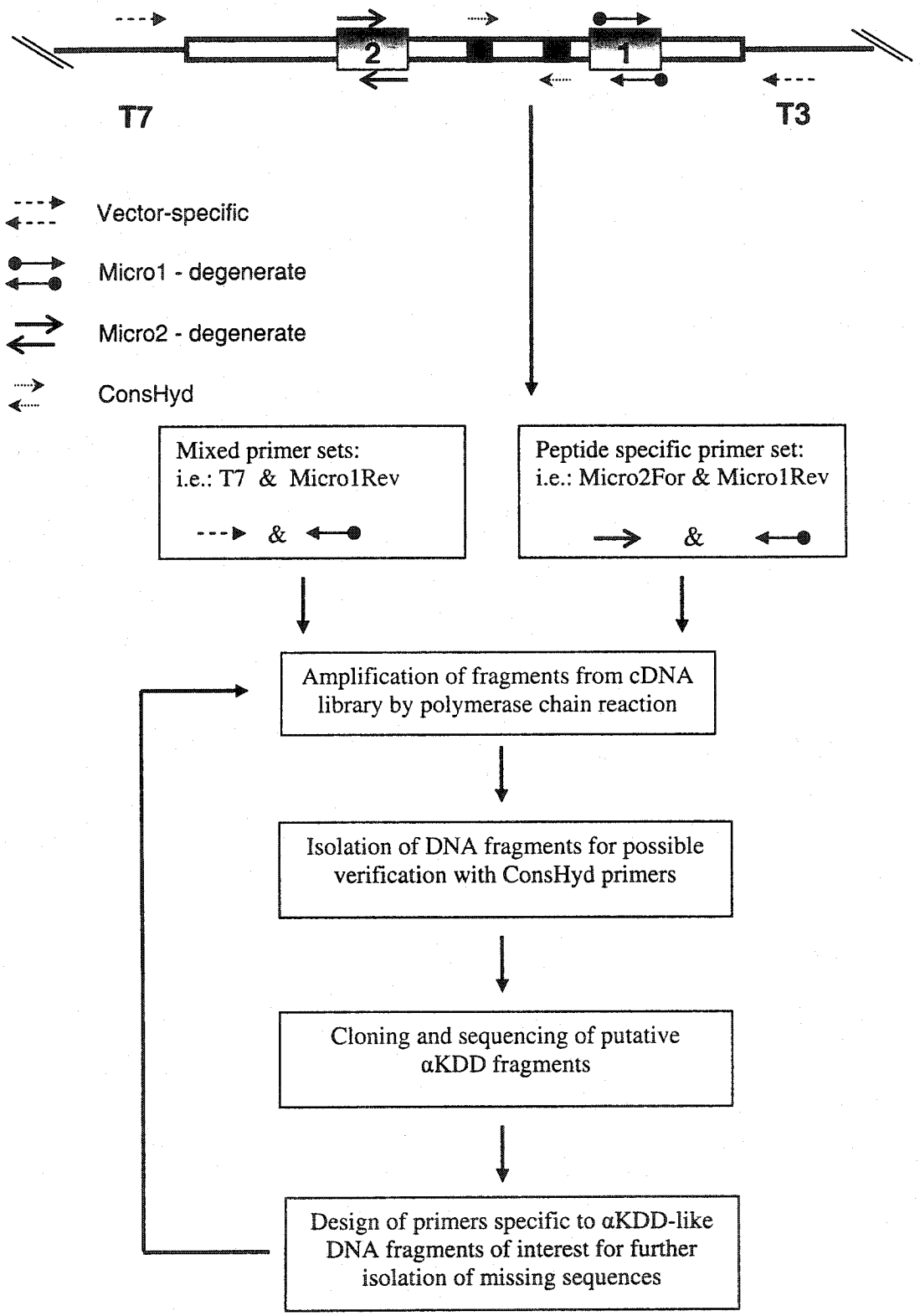


Figure 25: PCR-based screening strategy

To reduce the number of false positives, fragments obtained by PCR were verified, when possible, for the presence of conserved αKDD motifs by amplification using a set of primers designed to conserved regions (ConsHyd) if amplification was expected to occur across this conserved motif. Fragments amplified using microsequence-derived primers in combination with vector primers should contain this motif region. The conserved regions chosen for this verification step are located toward the 3′-end of the sequence, and are conserved among αKDD sequences. Any putatively positive fragments then were cloned and sequenced. The sequence information obtained was analyzed by BLAST and ClustalW alignment with other plant αKDDs. Among others, two fragments, identified as 1 and 2, were amplified using mixed primer sets; Fragment 1 (F6H·1; T7 and Micro1Rev) and Fragment 2 (ACCO·1; T7 and Micro2Rev), and both contained the peptide sequence to which the primer was designed, as well as other conserved αKDD motifs. However, Fragment 1 (602 bp) exhibited a higher degree of similarity with the flavonoid αKDD genes. In addition, a portion of the deduced amino acid sequence of Fragment 1 matched another fragment obtained from microsequencing (Micro3), with the exception of one amino acid residue. This may be attributed to the isobaric nature of certain amino acid pairs, making their initial identification from MS/MS analysis somewhat ambiguous. Fragment 2 (685 bp), on the other hand, shared significant similarity with known ACCO genes. Both sequences, however, were missing the 5′- and 3′-ends of the predicted full-length cDNA for F6H after this initial round of screening.

In order to isolate the remaining ends of putative F6H clones, new sets of primers were designed specifically to the ends of the Fragment 1 sequence (F6H·1For and Rev)

and the screening procedure was repeated with vector primers under stringent conditions. The specificity of these fragments then was verified by separate amplification reactions, using ConsHyd primers and nested primers (FNestFor and FNestRev) to Fragment 1. Subsequent screens yielded fragments that not only extended the F6H·1 sequence toward the C-terminal, but resulted in the isolation of other αKDD-related fragments as well, such as sequences sharing significant homology with F3H (>70 % identity at the amino acid level). Ultimately, PCR-screening allowed the amplification of a 1.245 kbp-long fragment (Figure 26) containing the 3′-end of this putative cDNA clone, including an in-frame stop codon (TAA), polyadenylation signal (ATATAA) and short poly-A tail. This fragment was termed *cF6H*. Nevertheless, a full-length open reading frame (ORF) was not obtained since this fragment lacked the translation start site. After repeated screenings using this strategy, no full-length fragment could be isolated.

The same cDNA library was screened with an oligonucleotide probe derived from the 602 bp-long F6H·1 sequence. Putatively positive clones were isolated and characterized, but again none yielded a full-length, putative F6H cDNA clone. On the other hand, several clones were isolated that exhibited similarity with related sequences, particularly *F3H* and *ACCO*, along with the *cF6H* fragment originally identified. None of *cF6H* clones isolated using this technique extended the 5′-terminal region. The ACCO-like clones (6 clones isolated) appeared to be full-length when compared with *ACCO* genes in the genbank database, while the F3H-like sequences (8 clones isolated) were likewise missing portions of their 5′-termini in a majority of the clones isolated. In addition, there seemed to be an abundance of the F3H-like clones as compared to the *cF6H*. The low abundance of *cF6H* clones may be attributed to sub-optimal F6H

caagagaagactcttaactctagattcgttgctagagatgaggattctctcgagagacct
Q E K T L N S R F V A R D E D S L E R P
aaggtttctgctatctacaacggatctttcgatgagatccctgttcttatctctcttgct
K V S A I Y N G S F D E I P V L I S L A
ggaatcgatatgactggagctggaactgatgctgctgctagaagatctgagatctgtaga
G I D M T G A G T D A A A R R S E I C R
aagatcgttgaggcttgtgaggattgggggatcttcggagagatcgatgatgaccatgga
K I V E A C E D W G I F G E I D D D H G
aagagagctgagatctgtgataagatcgttaaggcttgtgaggattggggagttttccaa
K R A E I C D K I V K A C E D W G V F Q
cctgatgagaagctggagtctgttatgtctgccgctaagaagggagatttcgttgttgat
P D E K L E S V M S A A K K G D F V V D
catggagttgatgctgaggttatctctcaatggactactttcgctaagcctacttctcat
H G V D A E V I S Q W T T F A K P T S H
actcaattcgagactgagactactagagatttccctaacaagcctgagggatggaaggct
T Q F E T E T T R D F P N K P E G W K A
actactgagcaatactctaggactcttatgggacttgcttgtaaacttcttggagttatt
T T E Q Y S R T L M G L A C K L L G V I
tctgaggctatgggacttgagaaggaggctcttactaaggcttgtgttgatatggatcaa
S E A M G L E K E A L T K A C V D M D Q
aaggttgttgttaactactaccctaagtgtcctcaacctgatcttactcttggacttaag
K V V V N Y Y P K C P Q P D L T L G L K
agacatactgatcctggaactatcactcttctacttcaagatcaagttggaggacttcaa
R H T D P G T I T L L L Q D Q V G G L Q
gctactagagatggaggaaagacttggatcactgttcaacctgtgaaggataacggatgg
A T R D G G K T W I T V Q P V K D N G W
atacttcttcatatcggagattctaacggacatagacatggacatttccttctaacgga
I L L H I G D S N G H R H G H F L S N G
agattcaagtctcatcaagcatacagatacagaagacctactaggggatctcctactttc
R F K S H Q A Y R Y R R P T R G S P T F
ggaactaaggtttctaactaccctccttgtcctgagcaatctcttgttagacctcctgct
G T K V S N Y P P C P E Q S L V R P P A
ggaagaccttacggaagagcgcttaacgctcttgatgctaagaagctggcttctgctaag
G R P Y G R A L N A L D A K K L A S A K
caacaacttgagtctgctgctatccttcttatctctgagcttgctgtagcttacatcatc
Q Q L E S A A I L L I S E L A V A Y I I
cttgctatccttccttcttctgagatcatcgctgaggaggggtacctctaaattctatct
L A I L P S S E I I A E E G Y L -
Taaccggctatc**atataa**tcatactacctagtacatggagcttgctgccttcttatctct
aattaatagactagcttacacgttttccttgtccaaaaaaaaaaa

Accession # - AY605048
Length – 1245nt
ORF – 1131 nt (376 a.a.)
Predicted *pI* – 5.1
Predicted Mr – 40.9 kDa

Boxes = Microsequences
Underline = Sequence termed F6H · 1
Circles = Catalytically important residues
Bold, underline = Potential glycosylation sites

Figure 26: Translated amino acid sequence for *cF6H*

expression conditions upon RNA extraction prior to library construction or a greater instability of the F6H RNA. At this point the *C. americanum* cDNA library was considered exhausted with respect to the coverage of total clones screened.

E.2.3 Alternative Approaches Used for the Isolation of *cF6H* ORF

RT-PCR was used in an attempt to isolate the F6H ORF. In this case, *C. americanum* RNA isolation proved difficult using the commercial kits or reagents; therefore, a modified manual protocol was employed that yielded RNA in sufficient quantities and of adequate purity. Initially, the RT-PCR system was tested for conformity with the results obtained from the library screen using two methods. First, RT-PCR using peptide-specific primers produced similarly-sized DNA fragments as those generated by the PCR-based screening technique, although the results were not consistently reproducible. Nevertheless, the fragments isolated were identical in sequence to the *cF6H*. On the other hand, PCR with one F6H-specific primer and oligo-dT often resulted in the amplification of non-specific fragments, regardless of the primer combination. It was unclear whether these discrepancies arose from differences in the isolated RNA or the reverse transcription reaction; however, measures were taken to ensure the reproducibility of the RT-PCR. On the other hand, there were differences in the quality of *C. americanum* maintained in the greenhouse, in that the tissue gradually exhibited signs of senescence, which may have resulted from pathogen attack. Attempts were made to restore the tissue to its original condition and to propagate uninfected tissue, yet none was successful for an extended period of time. For the reasons stated

above, RNA-based approaches were discontinued with regard to isolation of the 5′ F6H coding sequence.

Given the lack of fresh *C. americanum* tissue for use in 5′- rapid amplification of cDNA ends (RACE) experiments or RT-PCR at this point, inverse PCR was employed in an attempt to isolate the remaining 5′-end of the putative F6H sequence. Although this technique is relatively straight forward, it required significant troubleshooting with respect to the quality of genomic DNA and the conditions for monomeric ligation of digested DNA. Regardless of the restriction enzyme and the use of highly specific primers, no fragment was amplified that corresponded to the missing sequence. The DNA fragments that were isolated and sequenced showed no overlap with the *cF6H* sequence and likely resulted from cross-hybridization of the primers.

Subsequently, a modified GenomeWalker™ experiment was used to amplify the remaining ORF sequence. The main feature of this technique is the rapid isolation from genomic DNA of regions located upstream or downstream of a sequence-tagged site (cDNA sequence). This procedure involves the ligation of specific adaptors to blunt-ended genomic DNA in order to facilitate the amplification of sequences flanking specific regions of interest. Several attempts were made with this protocol to isolate any regions upstream of the F6H fragment and, regardless of the ligation conditions, gene-specific primers or *Taq* polymerase used, no *cF6H*-specific fragments could be amplified from the genomic DNA. The volume of the ligation reaction was increased in order to avoid intermolecular bond formation, however these attempts did not produce any significant results. A three-step thermal cycling protocol was used instead of the two-step protocol suggested by the manufacturer, where PCR annealing and extension

temperatures are identical. This modification could not be avoided since it was not possible to design adequate gene-specific primers (GSP1 and GSP2) with Tm ~70°C. In addition, the negative controls often gave unexpected banding patterns, further complicating the interpretation of the results observed in the test samples.

The DNA fragment termed *cF6H* translates into a 376 amino acid long sequence with a predicted molecular mass of 40.9 kDa. Given that the native protein exhibits a molecular mass of 43 to 45 kDa, following gel filtration or SDS-PAGE analysis respectively, it indicates that the cDNA fragment lacks ~10 to 15 N-terminal amino acid residues. However, the conserved regions and catalytically important residues are located in the C-terminal of αKDDs and are present in the translated *cF6H* sequence (encircled residues in Figure 26). In addition, previous studies have demonstrated that the N-terminal does not significantly contribute to enzyme activity for a dioxygenase involved in indole alkaloid biosynthesis (Vazquez-Flota *et al.*, 1997). Taking these factors into account, along with the exhaustion of the cDNA library and the unavailability of fresh tissue for the extraction of RNA, the *cF6H* fragment was cloned into prokaryotic and eukaryotic expression vectors to assess its enzyme activity.

E.2.4 Expression of *cF6H* cDNA in *E. coli*

The *cF6H* fragment was cloned into the *E. coli* expression vector pTrcHis. This construct allowed for the expression of a fusion protein with an N-terminally located $(His)_6$-tag under the control of the *trc* (*trp-lac*) promoter. Bacterial, recombinant cF6H (rF6Hb) production was induced with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) over a 4-hr period. The solubility of the fusion protein was assessed and F6H

activity was assayed. The level of expression of the fusion protein was monitored by Western analysis of cell lysates. Blots were probed with anti-$(His)_6$ antibodies, and detected either by chemiluminescence or by IRdye-800 detection. After IPTG induction, the antibody reacted with a 45 kDa protein band in the cell lysates carrying the pTrc-His-cF6H construct after IPTG induction (Figure 27A). The $(His)_6$-tag and amino acids resulting from the cloning process account for an additional ~ 3.5 kDa in the fusion construct. Purification of the recombinant protein was carried out using metal-ligand affinity resin (Ni-NTA) (Figure 27B) as well as size exclusion chromatography on Superose 12 (Figure 27C), and monitored by immunoblotting.

E.2.5 Determination of rF6Hb Protein Activity

The crude cell lysate did not exhibit significant F6H activity even after buffer exchange against buffer B (assay buffer; Table 7), which is a Tris-based assay as opposed to the phosphate-based buffer suggested by the manufacturer (Qiagen) for cell lysis and subsequent purification. The cofactor requirements for the bacterially expressed recombinant protein were nearly identical to those of the native plant protein. Fe^{2+} had the greatest effect on enzyme activity and enzyme reactivation in affinity-purified samples. This property was confirmed by the complete inhibition of enzyme activity upon incubation with 5 mM EDTA. Upon purification of the recombinant protein on Ni-NTA resin and buffer exchange with the Tris-HCl assay buffer, the protein exhibited some F6H activity (Table 7), albeit with two orders of magnitude lower than the Superose 12-purified plant protein (Table 1).

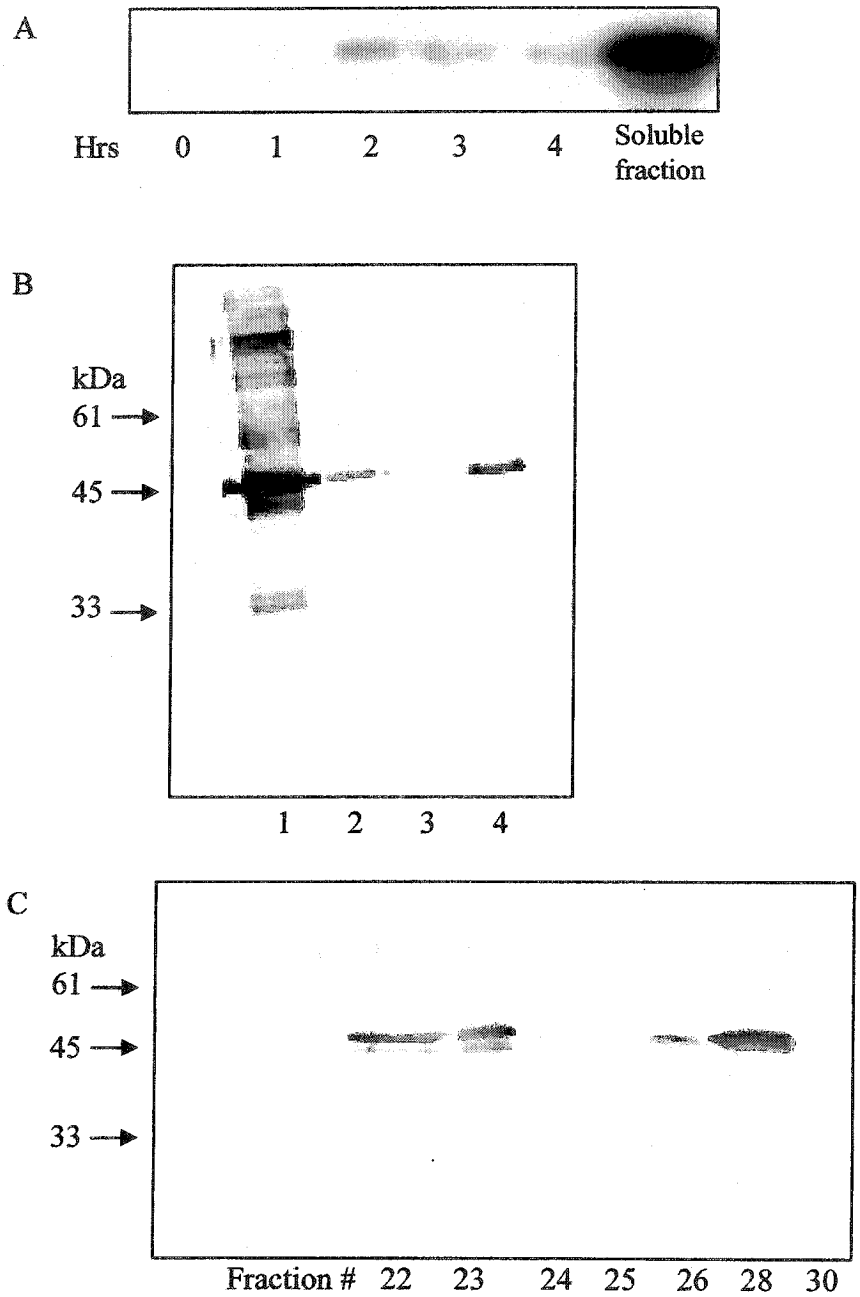


Figure 27: rF6Hb expression as detected by immunoblotting using anti-His antibody

(A) Time course induction of rF6Hb expression. Soluble fraction represents total cell lysate.
(B) Purification of rF6Hb on Ni-NTA; Lanes: **1**, flow-through (10 µg); **2**, wash (5 µg); **3**, Bound 1 (1 mL eluate, 2 µg); **4**, Bound 2 (2 mL eluate, 2 µg).
(C) Purification of rF6Hb eluted from Superose 12 (10 µg / fraction).

TABLE 7

Specific activity of rF6Hb protein

Step [a]	Specific Activity (pkat/mg)
Cell Lysate	0
Ni-NTA	0.091
Ni-NTA (30°C)	0.087
Ni-NTA + Superose 12	0.103
Superose 12	0.075
Cleavage of $(His)_6$-tag after Ni-NTA	0.044

[a] All assays were carried out using the direct enzyme assay with rF6Hb in the presence of assay buffer, αKDD cofactors and 5 μM 3,7,4′-trimethylquercetin.

The level of enzyme activity could not be significantly enhanced by size exclusion chromatography. The final activity obtained with the Ni-NTA purified enzyme preparation followed by gel filtration was 0.103 pkat/mg. Attempts were made to assay enzyme fractions collected at various time-points post induction, or of cultures grown at a lower temperature (30°C), but neither of these modifications resulted in any increase of enzyme activity.

E.2.6 Effect of Polyhistidine Tag on Enzyme Activity

To account for any inhibitory effect resulting from the presence of the $(His)_6$-tag, the fusion protein was incubated with enterokinase (EK) in order to cleave the tag from the recombinant protein. However, the enzyme activity of the cleaved protein was further reduced compared to the Ni-NTA-purified rF6Hb (Table 7), regardless of the stage at which cleavage was carried out. The loss of activity was possibly attributed to destabilization of the protein resulting from dialysis to remove inhibitors of the EK reaction that was followed by overnight incubation at elevated temperatures. Cleavage was similarly carried out overnight at 4°C, but the activity could not be completely restored. It is unlikely that the $(His)_6$-tag contributed to overall enzyme activity, since intact rF6Hb treated in the same manner as the preparation undergoing cleavage exhibited a similar reduction in enzyme activity.

E.2.7 Expression of *cF6H* cDNA in *P. pastoris*

The translated *cF6H* amino acid sequence possessed two potential glycosylation (O-GlcNAc) sites at Ser139 and Ser305 (Figure 26) that may contribute to catalytic

activity, perhaps by stabilizing a favorable tertiary conformation or through the modulation of interactions with other proteins. The effects of such post-translational modifications and protein folding were taken into account by expression of the recombinant protein in the eukaryotic expression system *P. pastoris*. This system, is relatively easy to manipulate and exhibits many of the advantages of eukaryotic expression including protein processing, folding and modification. *P. pastoris* can utilize methanol as its sole carbon source through the activity of a number of enzymes, one being alcohol oxidase (AOX1). The *AOX1* gene promoter is used to drive heterologous protein production in the presence of methanol and absence of glucose. In addition, the recombinant protein can be expressed in an intracellular or secreted fashion, depending on the presence of a signal peptide located N-terminally to the heterologous protein. These constructs also are amenable to metal-affinity ligand purification; however, the polyhistidine tag is located at the N-terminal for the intracellular expression construct and at the C-terminal for the secretory system.

Native F6H activity was shown to be dependent on pH; therefore, recombinant protein induction was carried out in a buffered culture medium as described previously. In addition, the chosen media (BMGY/BMMY) contained protein stabilizing factors, such as peptone and yeast extract that have been shown to reduce proteolysis, particularly of secreted proteins (Brierley *et al.*, 1994; Clare *et al.*, 1991). Recombinant protein production was induced with methanol over a 4-day period for both intracellular (rF6Hy) and secreted (rF6Hys) systems. Protein expression, monitored by Western blot analysis using anti-$(His)_6$ antibodies, was maximal 3 to 4 days post-induction in both cases

(Figures 28, 29) and protein eluted at a volume corresponding to ~40 kDa on Superose 12.

E.2.8 Enzyme Activity of rF6Hy/ rF6Hys Proteins

Direct enzyme activity assays carried out for the crude intracellular (rF6Hy) and secreted (rF6Hys) fractions in the assay buffer did not yield detectable F6H activity. Protein purification using Ni-affinity resin resulted in a significantly purified preparation with a specific activity of 0.17 pkat/mg for the intracellular construct, rF6Hy. On the other hand, the culture medium containing the secreted protein, rF6Hys, was subjected to ammonium sulfate concentration followed by Ni-NTA purification, yielding an enzyme with a specific activity of 0.11 pkat/mg. The specific activity of either recombinant protein construct could not be improved significantly by purification on size exclusion columns instead of/in addition to Ni-NTA purification (Table 8).

E.2.9 Substrate Preference of rF6Hb and rF6Hy Proteins

To confirm the substrate preference of the recombinant F6H, various quercetin derivatives were tested as potential substrates for the prokaryotic- and eukaryotic- expressed proteins, rF6Hb and rF6Hy. The substrate specificity was tested using the Ni-NTA-purified enzyme after buffer exchange with different flavonol substrates at different concentrations (1, 2.5, 5, 10 and 50 μM). The results (Table 9) indicate that 3,7,4´-TriMeQ is, in fact, the preferred substrate for both recombinant proteins, which did not accept quercetin, 3-methylquercetin, as is the case with the native F6H. On the other hand, 3,7-dimethylquercetin was accepted at 56% and 63% of the

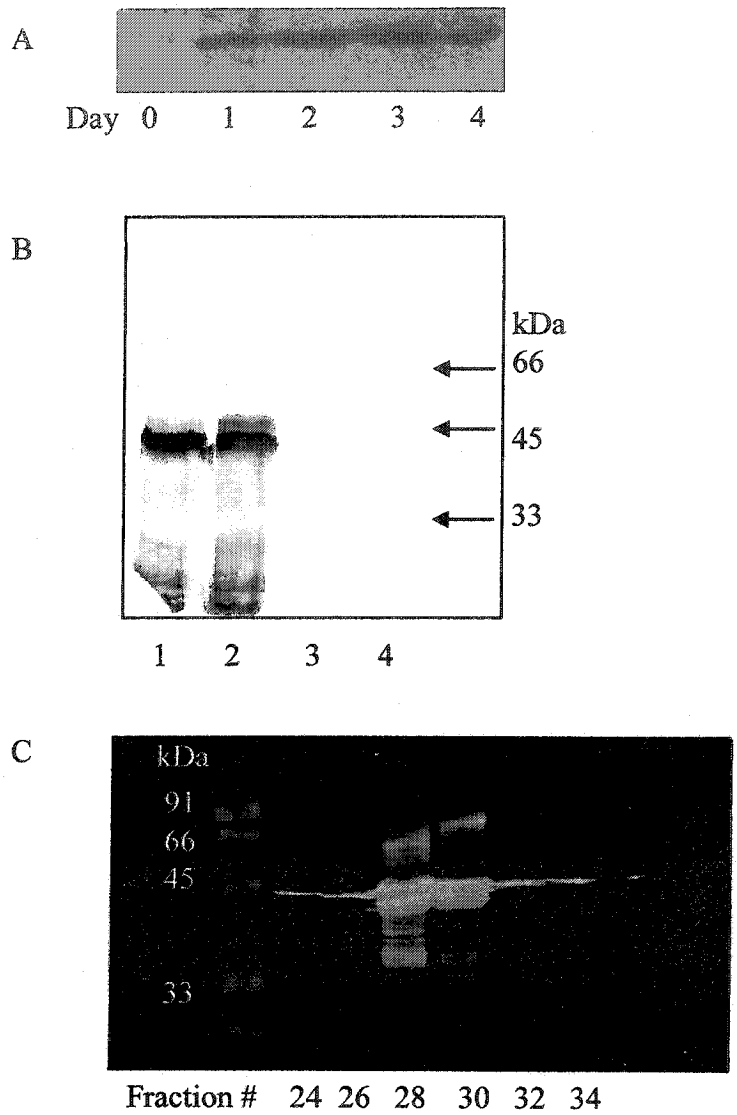


Figure 28: Expression of rF6Hy

(A) Time course induction of rF6Hy.
(B) Purification of rF6Hy on Ni-NTA. Lanes: **1**, Bound 2 (2 mL eluate; 5μg); **2**, Bound 1 (1 ml eluate; 5 μg); **3**, wash (10 μg); **4**, flow-through (10 μg);
(C) Purification of rF6Hy on Superose 12 (10 μg / fraction).

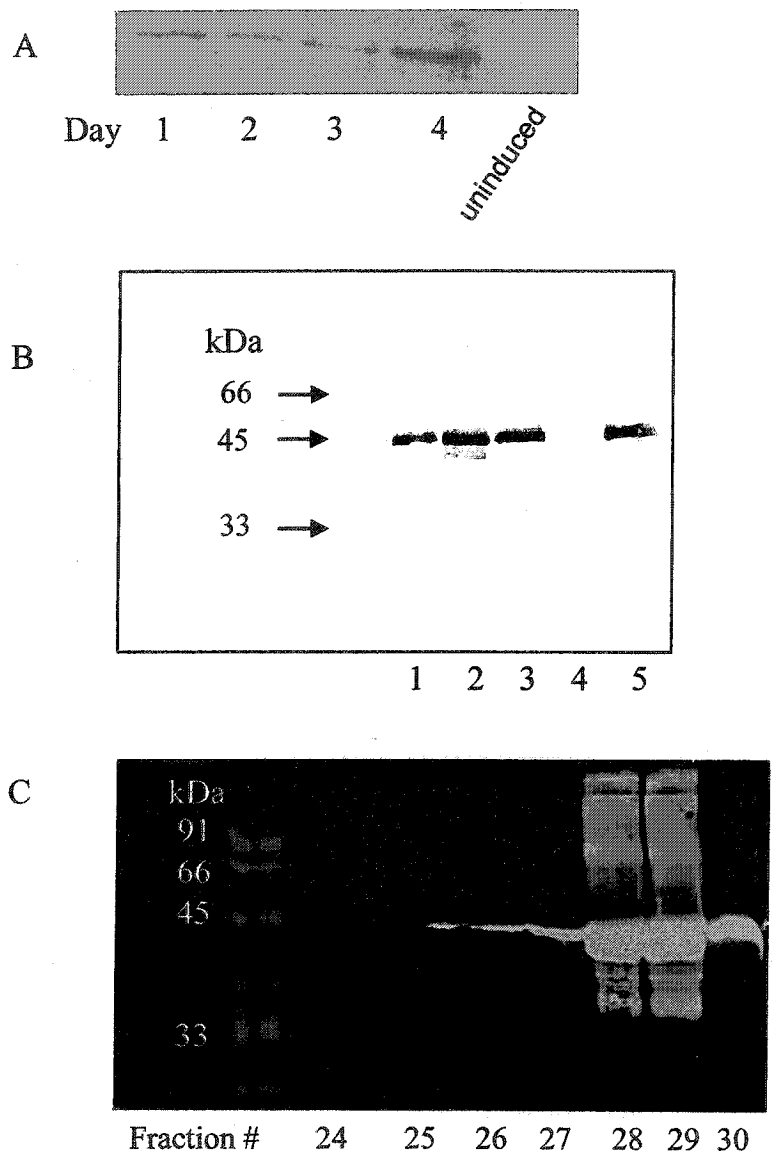


Figure 29: Expression of rF6Hys

(A) Time course induction of rF6Hys.
(B) Purification of rF6Hys on Ni-NTA. Lanes: **1**, Bound 3 (3 mL eluate; 2μg); **2**, Bound 2 (2 mL eluate; 2 μg); **3**, Bound 1 (1 mL eluate; 5 μg); **4**, wash (10μg); **5**, flow-through (10 μg);
(C) Purification of rF6Hys on Superose 12 (20 μg / fraction).

TABLE 8

Specific activity of intracellular (rF6Hy) and secreted (rF6Hys) proteins [a]

Step	Specific Activity (pkat/mg)
rF6Hy (intracellular)	
Cell Lysate	0
Ni-NTA	0.17
Ni-NTA + Superdex 75	0.21
Superdex 75	0.087
rF6Hys (secreted)	
$(NH_4)_2SO_4$ (35-70%)	0
$(NH_4)_2SO_4$ + Ni-NTA	0.11
$(NH_4)_2SO_4$ + Ni-NTA + Superdex 75	0.08
Superdex 75	0.04

[a] All assays were carried using the direct enzyme assay in the presence of assay buffer, αKDD cofactors and 5μM 3,7,4′-trimethylquercetin.

TABLE 9

Substrate specificity of the recombinant F6H

Both rF6Hb and rF6Hy were purified on Ni-NTA resin then Superose 12, and fractions were assayed with the indicated substrates, using the direct enzyme assay.

Substrate	Relative enzyme activity (%) rF6Hb [a]	rF6Hy [b]
3,7,4´-Trimethylquercetin	100	100
3,7-Dimethylquercetin	56	63
3,7,3´,4´-Tetramethylquercetin	11	21
3-Methylquercetin	0	4
Quercetin	0	0

[a] Estimated as 0.10 pkat/mg for 100% activity and 5 µM substrate (as optimal), however assays were also carried at 1, 2.5, 10 and 50 µM substrate.

[b] Estimated as 0.19 pkat/mg for 100% activity and 5 µM substrate (as optimal), however assays were also carried at 1, 2.5, 10 and 50 µM substrate.

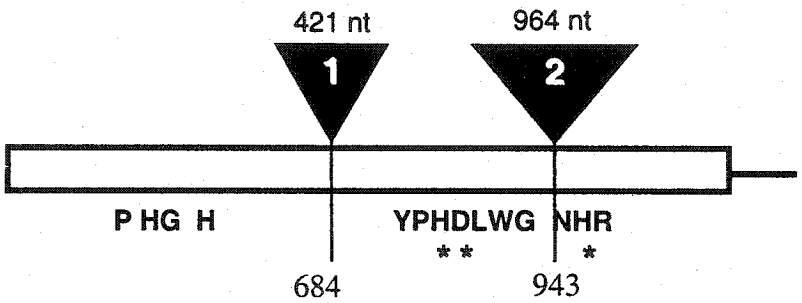
control activity by rF6Hb and rF6Hy, respectively. A similar set of alternative substrates as listed in Table 2 was also tested for the recombinant proteins, but no F6H activity could be detected regardless of the flavonoid concentration, aside from the flavonols rhamnetin (17 %) and kaempferol (5 %) at a substrate concentration of 50 μM.

E.2.10 Determination of Intron Number and Position

Genomic *C. americanum* DNA was used as template for the amplification reactions, with primers designed to the outermost regions of the *cF6H* sequence in order to isolate the genomic region coding for F6H. This produced a fragment of ~ 3.1 kbp (*gF6H*) identical in sequence to the *cF6H*, which contained two potential intronic sequences (Figure 30). The first intron, 421 bp long, is centrally located at position 684 of *cF6H*, whereas the second intron is significantly longer, 964 bp, and is located towards the 3′-end of the gene at position 943.

E.2.11 Southern Blot Analysis

Southern blots of genomic DNA isolated from *C. americanum* were probed with the *cF6H* fragment encoding the ORF. Single bands were detected in *Bam*HI-, *Eco*RI- and *Hin*dIII-digested genomic DNA (Figure 31). The internal sequence of *gF6H* does not contain any recognition sites for the above enzymes; however, one recognition site is located within the known sequence for *Kpn*I, *Nco*I, *Xho*I. Two major bands were detected in these lanes at estimated sizes ranging from 3.0-10.0 kbp (Figure 31). These results indicate that F6H is present as a single copy gene in *C. americanum*, particularly since potential tandem *F6H* repeats would have resulted in the observation of fragments



Intron 1 -

```
gccatttaca tgagttgagg tactcattgc gacctcttcg ttcgtcagtc
atcacaangc tttgaaagat tttcagctac cactctctcc tttatatatt
cattacacat ctcttctttc tatatctctc ttaatttagt cttttgtctt
cgtaattaca atggctccag gaactttgac tgagctagcc ggagagtcta
agctcaactc taaatttgtc agggacgaag atgaacggcc caaagtcgct
tacaatgtgt ttagcgacga aatcccggtg atctctctcg ccggtatcga
tgacgtcgat ggaaaaagag gagagatctg ccgtcagatc gttgaggctt
gtgagaattg gggcatcttc caagtggtcg atcacggcgt cgatactaac
ttagtggcgg atatgactcg t  421 nt
```

Intron 2-

```
attgctcctc ctgctactac gctgacatcc attgcgcatg agaaaaccct
acaacaaaaa ttcgtccgag acgaagacga gcgtccaaag gttgcctaca
acgatttcag caacgaaatt ccgatcatct cgcttgccgg gatcgatgag
gttgaaggcc gccgggccga gatttgcaag aagattgtgg acgcttgtga
ggactggggt attttccaga ttgttgatca tggagttgat gccgagctca
tatcggaaat gaccggtctc gccaaagagt tctttgattt gccatcggag
gagaagctcc gcttcgacat gtccggtggc aaaaagggtg gattcatcgt
gtccagtcat ttacagggag aagctgtgca agattggcgt gaaattgtga
cctacttttt atacccgatt cgccaccggg actactagag gtggccggac
aagccagagg catggaggga ggtgacgaag aagtacagcg acgagctgat
ggggctggca tgcaagctct tgggggtttt atcagaagca atggggttgg
atacagaggc attgacaaag gcatgtgtgg acatggacca aaaagtggtg
gtgaatttct atccgaagtg ccctcagccc gacctaactc ttggcctcaa
gcgccacacg gacccgggca caattaccct tttgcttcag gaccaagttg
gtggccttca ggctactagg gatgatggga agacatggat caccgttcaa
ccagtggaag gagcttttgt ggtcaatctc ggagatcatg gtcattttct
gagcaatggg aggttcaaga atgctgatca ccaagcagtg gtgaactcaa
acagcagcag gctgtccata gccacattcc agaacccagc tcaagatgca
atagtgtatc cactcagtgt gagggaggga gagaagccga ttcacgaggc
gccgatcacc taca  964 nt
```

Figure 30: Genomic organization of *gF6H* and sequences of introns 1 & 2

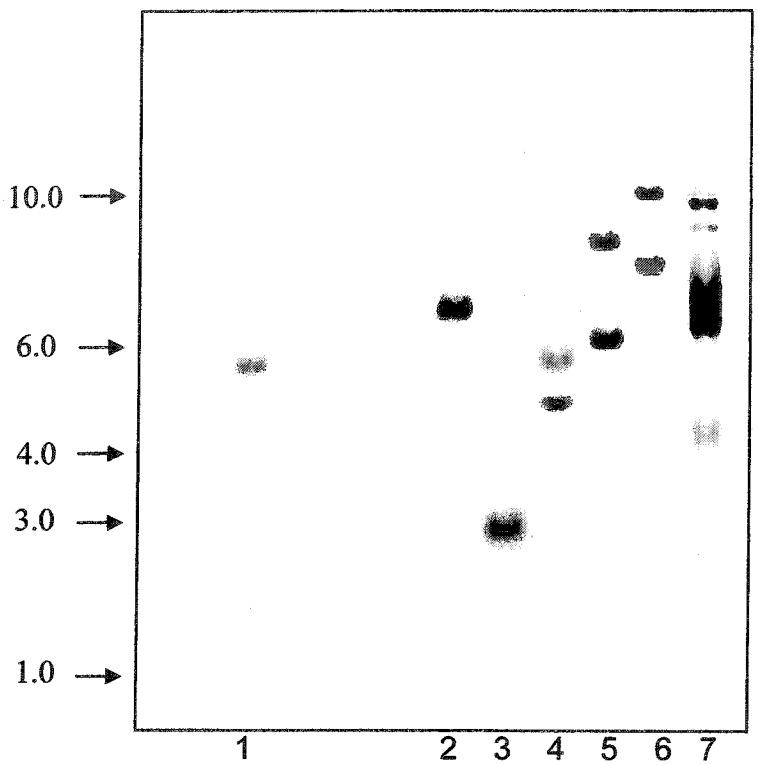


Figure 31: Southern blot of *C. americanum* genomic DNA

1, *Bam*HI
2, *Eco*RI
3, *Hind*III
4, *Kpn*I
5, *Nco*I
6, *Xho*I
7, Undigested vector containing probe sequence

larger than 6.2 kbp, which were not present in either lanes 1 and 3. Under conditions of reduced stringency, certain lanes exhibited more than the expected two bands, suggesting the existence of related sequences within the genome (data not shown).

E.2.12 Phylogenetic Analysis of *cF6H* and Related αKDD Sequences

cF6H from *C. americanum* exhibits highest amino acid similarity to a putative *F3H* from *Juglans nigra.* A distance method was employed to analyze the data and generate the phylogram since it is more suited to studying the relationship between a group of taxa based on the degree of sequence similarity as opposed to assessing the pathways of gene evolution through a parsimony approach. The latter analysis would not be particularly valid in this case since the nature of the relationships between the genes selected is thus far unclear. Phylogenetically, it is evident that homologous genes encoding biochemically characterized proteins with similar enzymatic activities in single pathways are clustered into related subgroups (Figure 32).

This clustering is likely the result of gene duplication and divergence. *cF6H* clusters with the *F3H* group of genes, although at the periphery of the *F3H* homologs from different species; thus suggesting an evolutionary relationship with this particular subgroup of biosynthetically related enzymes. This distinction is reinforced by the fact that the F6H reaction exhibits a strict position specificity and unifunctionality with respect to the reaction catalyzed, as is the case with F3H. However, recent evidence has demonstrated that FLS and ANS are dioxygenases with a multifunctional nature and broader substrate specificity (Turnbull *et al.*, 2004).

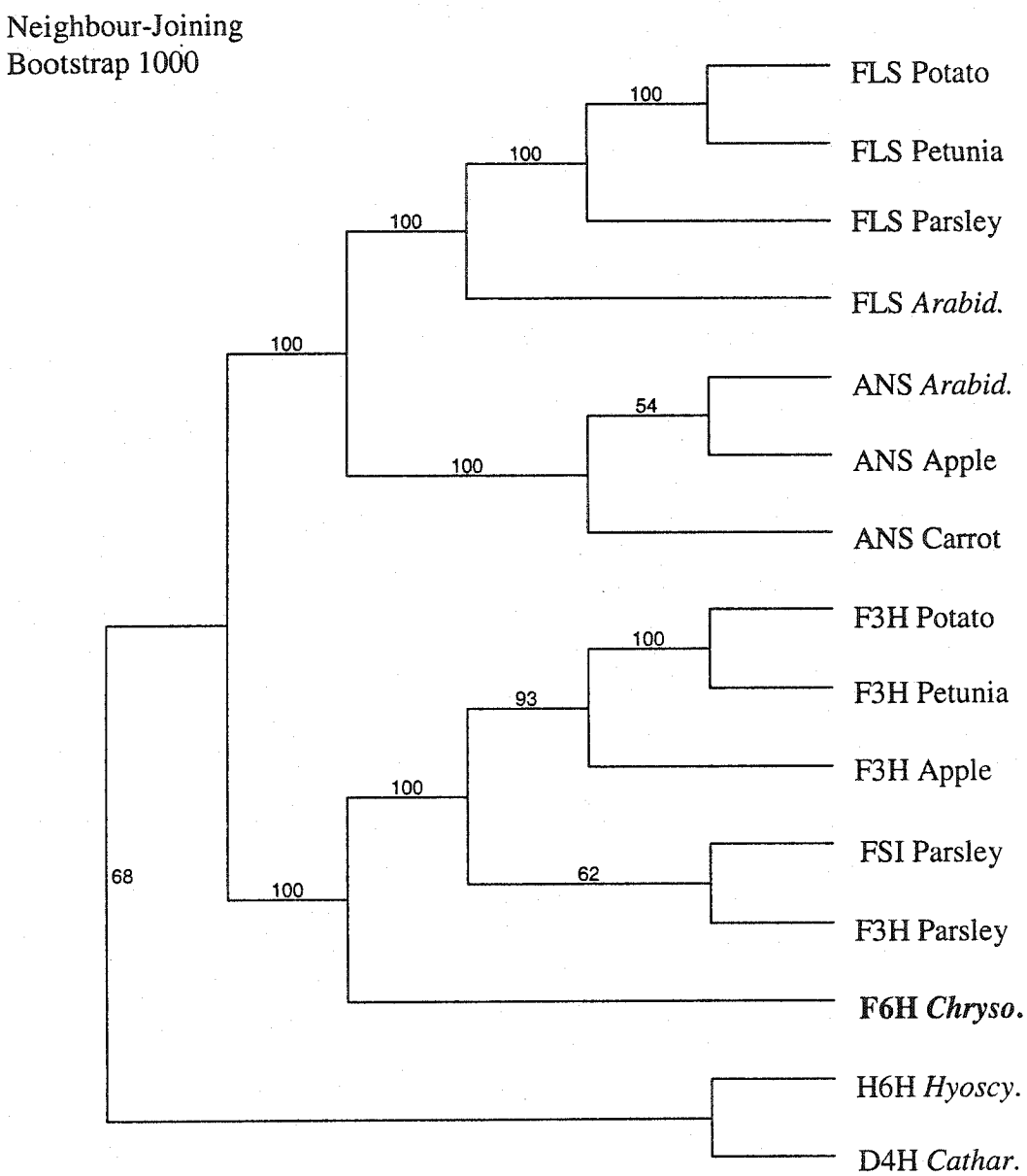


Figure 32: Phylogenetic analysis of *cF6H*

Accession numbers for above genes:
Potato FLS (X92178); Petunia FLS (Z22543); Parsley FLS (AY230249); *Arabidopsis* FLS (U84259); *Arabidopsis* ANS (U70478); Apple ANS (X71360); Carrot ANS (AF184273); Potato F3H (AY102035); Petunia F3H (X60512); Apple F3H (X69664); Parsley FSI (AY230247); Parsley F3H (AY230248); *Hyoscyamus* H6H (M62719); *Catharanthus* D4H (U71604).

E.3 Determination of 6-Hydroxylase Activity in *Chrysosplenium alternifolium*

C. alternifolium is another saxifragaceous species that is characterized by the presence of highly methylated derivatives of quercetagetin and 6-hydroxykaempferol, with the central flavonol intermediate being 3,6,7-trimethylquercetagetin (Jay and Voirin, 1976), although smaller amounts of di- and tetramethylquercetagetin also have been detected. Whereas polymethylated flavonol biosynthesis in *C. americanum* begins with the aglycone quercetin and 6-hydroxylation occurs after several methylation steps, that hydroxylation step in the homologous pathway in *C. alternifolium* seems to occur at an earlier stage, with the 6-*O*-methylation of dimethylquercetagetin. In order to assess the existence of an F6H-like activity in *C. alternifolium*, the hydroxylation reaction was investigated using different flavonol substrates (Table 10) and the F6H direct enzyme assay with an enzyme preparation purified according to the method developed for the F6H of *C. americanum* and comparison with authentic standards, when available. The results demonstrate that the highest level of hydroxylase activity occurred with 3,7-dimethylquercetin, although the specific activity of this reaction (0.75 pkat/mg; Superose-12 purified sample) was significantly lower than with a preparation from *C. americanum* of the same purification stage (8.53 pkat/mg). This may be due to an extraction and partial purification scheme that was not optimal for this tissue, or the use of a metabolite that is not the preferred substrate for the enzymatic reaction. The substrate preference of this enzyme indicates that 6-hydroxylation occurs at an earlier step in the pathway as compared to that of *C. americanum.*

Given that 6-hydroxylated intermediates occur in this tissue, as well as the detection of some hydroxylase activity, the genomic DNA of *C. alternifolium* was used

Table 10

6-Hydroxylation of quercetin derivatives in *Chrysosplenium alternifolium*

The ammonium sulfate (35-75%) fraction was assayed with the indicated substrates using the direct enzyme assay.

Substrate (10 μM) [a]	Relative Activity (%)
3,7-Dimethylquercetin	100 [b/c]
3,7,4′-Trimethylquercetin	37
3,7,3′,4′-Tetramethylquercetin	16
3-Monomethylquercetin	0
Quercetin	0

[a] Other substrate concentrations tested: 1, 2.5, 5, 20 μM. however, results did not vary significantly.

[b] Estimated as 0.75 pkat/mg for 100% activity

[c] The direct enzyme assay was used to determine F6H activity, and the identity of the reaction products was confirmed by HPLC analysis using authentic standards.

as a template for the isolation of potential *F6H* homologs using primers designed to the F6H coding sequence. These attempts did not result in the amplification of any novel hydroxylase sequences. In addition, when primers designed to the conserved regions of αKDDs, also present in F6H, were employed for PCR, sequences with significant homology to F3H, ACCO and gibberellin hydroxylases were amplified, but none of which was distinct enough to suggest the isolation of an F6H homolog.

F. DISCUSSION

A novel enzyme catalyzing the 6-hydroxylation of partially methylated flavonols has been purified to near homogeneity and characterized from *C. americanum* shoot tips, resulting in a 538-fold purification with a specific activity of 97.1 pkat/mg and a final recovery of 0.63 %. In contrast with the *Petunia* flavanone 3-hydroxylase, which was reported to undergo rapid proteolysis in crude extracts (Lukacin *et al.*, 2000), native F6H exhibits a single band on Western analysis throughout the purification process. Therefore, the low catalytic activity of the affinity-purified enzyme can not be attributed to proteolysis. Due to its low abundance and inherent instability, the purification scheme involved two separate affinity chromatography steps, along with other conventional chromatographic procedures. The most significant purification was obtained using the α-ketoglutarate affinity column, which resulted in one major protein band on denaturing gels, albeit with a decreased specific activity due to enzyme instability. In an attempt to prevent this inhibition while increasing the purity of the sample, an immunoaffinity column was prepared using the IgG-purified fraction of an anti-desacetoxyvindoline 4-hydroxylase antiserum. The use of a heterologous antibody raised against a plant dioxygenase of the same class (α-D4H) proved to be a useful tool in the purification and characterization of the F6H. The immunotitration of F6H activity was observed with protein preparations at different stages of purification. The probability that the F6H preparation contains the D4H protein, or that F6H may accept desacetoxyvindoline (DAV) as a substrate, was excluded since no new product was formed when an F6H preparation and DAV were incubated in the presence of α-ketoglutarate-dependent

dioxygenase cofactors (data not shown). Western blots of highly purified protein preparations produced only one band when probed with the anti-D4H antibody. Taken together, these results justified the use of α-D4H serum as an accessory tool for the purification of F6H from *C. americanum*. In addition, the immunoaffinity step of the purification scheme resulted in a sample of high purity in comparison to the Superose 12 fraction with regard to the number of bands observed on SDS gels (Table 1, Figure 12). Nevertheless, the ligand affinity column resulted in the enzyme preparation of highest purity with regard to elimination of extraneous protein.

The physicochemical properties of F6H are similar to those of other plant α-ketoglutarate-dependent dioxygenases with respect to molecular mass, pH optimum and *pI* value (Britsch and Grisebach, 1986; Hashimoto and Yamada, 1987; Britsch, 1990; De Carolis and De Luca, 1993). In addition, F6H behaves like flavanone 3-hydroxylase (Britsch and Grisebach, 1986) and flavone synthase I (Britsch, 1990) in that the F6H bound tenaciously on hydrophobic interaction chromatography (Phenyl-Sepharose) preventing its recovery. This property suggests the presence of significant hydrophobic regions within the protein structure. The cofactor requirements of *C. americanum* F6H are characteristic of other known plant α-ketoglutarate-dependent dioxygenases (Britsch, 1981; Britsch and Grisebach, 1986; Britsch, 1990; Saito *et al.*, 1999), particularly as to their dependence on ferrous ions for activity and reactivation. Optimal F6H activity was observed in the presence of ferrous ions, α-KG and ascorbate (Table 3). The stimulation of F6H activity observed upon the addition of catalase to the assay mixture, likely results from the removal of peroxides that affect dioxygenase activity and stability, as well as an elevated oxygen concentration generated from peroxide breakdown (Blanchard *et al.*,

1982). A similar effect also has been observed with the hyoscyamine 6-hydroxylase (Hashimoto and Yamada, 1987), flavanone 3-hydroxylase (Britsch, 1990) and desacetoxyvindoline 4-hydroxylase (De Carolis and De Luca, 1993), all of which are dioxygenases involved in plant secondary metabolism.

The F6H exhibits strict substrate specificity for position 6 of partially methylated flavonols with the highest preference for 3,7,4´-trimethylquercetin as substrate. The identity of the reaction product was thoroughly verified by TLC, HPLC, and MS analyses. The fact that F6H also accepts 3,7-dimethyl and 3,7,3´,4´- tetramethyl derivatives as substrates, but not the parent aglycone quercetin nor its monomethyl ether to a significant extent, indicates the requirement for a 7-methoxy group for catalytic activity. This implies its involvement in the biosynthesis of the polymethylated flavonols of *C. americanum.* It has been proposed that the cytosolic enzymes involved in the sequential methylation of these metabolites are organized on the surface of a multienzyme aggregate (Ibrahim, 1992). It is likely that the F6H is a component of this enzyme complement, as it introduces a hydroxyl group at position 6 of 3,7,4´-trimethylquercetin for subsequent *O*-methylation at this position. Immunolocalization of both enzymes, F6H and F6OMT, should provide further evidence in support of this view.

Although the physicochemical and kinetic properties of F6H compare well with other α-ketoglutarate-dependent dioxygenases, it exhibits a distinct substrate- and position specificity, since it hydroxylates position 6 of the aromatic ring A of a partially methylated flavonol. This is in contrast to other aromatic hydroxylations which are commonly catalyzed by cytochrome P450 monooxygenases, such as the flavonol 8-hydroxylase of *Tagetes patula* (Halbwirth, 1999) and the flavonoid 6-hydroxylase of

soybean involved in isoflavone biosynthesis (Latunde-Dada *et al.*, 2001). The fact that the F6H exhibits a preference for relatively hydrophobic substrates distinguishes it from other dioxygenases characterized to date (Britsch and Grisebach, 1986; Britsch, 1990; Holton *et al.*, 1993) which accept relatively polar flavonoid aglycones.

The fact that F6H exhibits a very low K_m for its flavonol substrate indicates a low concentration *in vivo* and a high specificity for its substrate. In addition, the K_m and catalytic efficiency of the F6OMT for 3,7,3´-trimethylquercetagetin are consistent with the lack of accumulation of the hydroxylated product in the intact plant (De Luca and Ibrahim, 1985b). Taken together, these results corroborate the channelled pathway proposed for the biosynthesis of polymethylated flavonols in *C. americanum* (Ibrahim, 1992). The K_m value of F6H for α-KG is similar to those reported for other plant dioxygenases involved in secondary metabolism (Britsch and Grisebach, 1986; Hashimoto and Yamada, 1987; De Carolis and De Luca, 1993). This suggests the requirement for a common binding site for the cosubstrate to chelate ferrous ions, whereas the binding site for the flavonol substrate is determined primarily by its molecular structure, as well as the microenvironment of the protein.

The kinetic data obtained suggest that the F6H follows the same kinetic patterns reported for other dioxygenases (Puistola *et al.*, 1980; De Carolis and De Luca, 1993; Britsch and Grisebach, 1980) Thus, an ordered, sequential ter ter reaction mechanism is predicted for the F6H, with α-ketoglutarate binding first, followed by oxygen then trimethylquercetin. Product release would be in the reverse order, *i.e.* the hydroxylated product, followed by CO_2 and succinate.

To obtain amino acid sequence information of the F6H, the purified protein was subjected to microcapillary RP-HPLC and nanoelectrospray tandem mass spectrometry that resulted in the identification of several peptides. An alignment of the Micro1 peptide sequence with seven known flavonoid dioxygenases, using the ClustalW program, revealed a significant degree of similarity with six amino acid residues being conserved among these sequences (data not shown). These included two flavanone 3-hydroxylase (Britsch *et al.*, 1992; Kim and Kim, 1998), three flavonol synthase (Holton *et al.*, 1993; Pelletier *et al.*, 1997; van Eldik *et al.*, 1997) and two anthocyanidin synthase (Davies, 1993; Saito *et al.*, 1999). On the other hand, a lower level of similarity was observed with the D4H (Vazquez-Flota *et al.*, 1997) and H6H (Matsuda *et al.*, 1991).

A *C. americanum* cDNA library was screened using degenerate oligonucleotide probes based on the peptide sequences derived from the purified plant protein. Regardless of the screening method a fragment encoding the full-length F6H ORF could not be isolated. *cF6H* encodes a near-full length clone, truncated by ~10 to 15 amino acids at the N-terminal, encoding a protein with a M_r ~41 kDa. The OFRs encoded by all the clones isolated were identical in sequence (data not shown), indicating that they represent the same protein. This is in agreement with the results obtained from Southern analysis suggesting that *F6H* is present as a single copy in the *C. americanum* genome. The clone isolated, *cF6H*, was confirmed to be an authentic fragment of the gene encoding the F6H, since the peptide sequences obtained from the native protein were recognized in the deduced amino acid sequence of the cDNA clone.

Analyses of αKDD genes (Figure 33) demonstrated that the conserved regions and catalytically important amino acids are located within the C-terminal half of the

```
CLUSTAL W (1.82) multiple sequence alignment

FSIParsley      ---MAPTTITALAKEKTLNLDFVR-DED--ERPKVA--YNQFSNEIP-IISLAGLD---- 47
F3HPetunia      IPRVTPSTLTALAEEKTLQTSFIR-DED--ERPKVA--YNQFSNEIP-IISLEGID---- 50
F6HChryso       ------------QEKTLNSRFVARDEDSLERPKVSAIYNGSFDEIPVLISLAGIDMTG- 46
FLSPetunia      --MKTAQGVSATLTMEVARVQAIASLSK--CMDTIPSEYIRSENEQPAATTLHGVVLQVP 56
                              :. .   :   ..:    .:.: *    :* *   :* *:

FSIParsley      --DDSDGRRPEICRKIVKACEDWGIFQVVD--HGIDSGLISEMTRLSREFFALPAEEK-- 101
F3HPetunia      --DET-GKRAEICDKIVKACEDWGVFQVVD--HGVDAEVISQMTTFAKEFFALPPEEK-- 103
F6HChryso       AGTDAAARRSEICRKIVEACEDWGIFGEIDDDHGKRAEICDKIVKACEDWGVFQPDEK-- 104
FLSPetunia      VIDLRDPDENKMVKLIADASKEWGIFQLIN--HGIPDEAIADLQKVGKEFFEHVPQEEKE 114
                       . ::   *..*.::**:*  ::..**       .:    .::    .:*:

FSIParsley      -LEYDTTGGKRGGFTISTVLQGDDAMDWREFVTYFSYPINARDYSR-WPKKPEGWRSTTE 159
F3HPetunia      -LRFDMSGGKKGGFIVSSHLQGEVVQDWREIVTYFSYPTRARDYSR-WPDKPEGWIAVTQ 161
F6HChryso       -LESVMSAAKKGDFVVDHGVDAEVISQWTTFAKPTSHTQFETETTRDFPNKPEGWKATTE 163
FLSPetunia      LIAKTPGSNDIEGYGTSLQKEVEGKKGWVDHLFHKIWPPSAVNYRY-WPKNPPSYREANE 173
                 :    . .  .:  .   : :    *         .    :   .:*.:* .:  ..:

FSIParsley      VYSEKLMVLGAKLLEVLSEAMGLEKGDLTKAC--VDMEQKVLINYYPTCPQPDLTLGVRR 217
F3HPetunia      KYSEKLMELACKLLDVLSEAMGLEKEALTKAC--VDMDQKVVVNFYPKCPEPDLTLGLKR 219
F6HChryso       QYSRTLMGLACKLLGVISEAMGLEKEALTKAC--VDMDQKVVVNYYPKCPQPDLTLGLKR 221
FLSPetunia      EYGKRMREVVDRIFKSLSLGLGLEGHEMIEAAGGDEIVYLLKINYYPPCPRPDLALGVVA 233
                 *.. :  :  :::  :* .:***   : :*.   ::   : :*:** **.***:**:

FSIParsley      HTDPGTITILLQDMVGGLQATRDGGKTWITVQPVE--GAFVVNLGD-----HGHYLSNGR 270
F3HPetunia      HTDPGTITLLLQDQVGGLQATKDNGKTWITVQPVE--GAFVVNLGD-----HGHFLSNGR 272
F6HChryso       HTDPGTITLLLQDQVGGLQATRDGGKTWITVQPVKDNGWILLHIGDSNGHRHGHFLSNGR 281
FLSPetunia      HTDMSYITILVPNEVQGLQVFKDG--HWYDVKYIP--NALIVHIGD-----QVEILSNGK 284
                *** . **:*: : * ***. :*... *  *: : ... :::::**:..  : . ****:

FSIParsley      FRNADHQAVVNSTSSRLSIATFQNPAQNAIVYPLKIREGEKAILDEAITYAEMYKKCMTK 330
F3HPetunia      FKNADHQAVVNSNSSRLSIATFQNPAPEAIVYPLKIREGEKSIMDEPITFAEMYRRKMSK 332
F6HChryso       FK--SHQAYRYRRPTRGS-PTFGTKVSN---YPP-CPEQSLVRPPAGRPYGRALNALDAK 334
FLSPetunia      YKSVYHRTTVNKDKTRMSWPVFLEPPSE--------HEVGPIPKLLSEANPPKFKTKKYK 336
                ::   *::      :* * ..*     :    .    *          .     .    *

FSIParsley      HIEVATRKKLAKEKRLQDEKA--KLEMKSKSADENLA----- 365
F3HPetunia      DLELARLKKQAKEQQLQAEVAAEKAKLESKPIEEILA----- 369
F6HChryso       KLASAKQQLESAAILLISELAVAYIILAILPSSEIIAEEGYL 376
FLSPetunia      DYVYCKLNKLPQ------------------------------ 348
                .   .  :  .       . :         . ..  :...
```

Figure 33: ClustalW alignment of *Chrysosplenium americanum* F6H with F3H genes from different organisms

*, strictly conserved regions

proteins. The amino acid residues involved in iron- (His222, Asp 224, His285) and α-KG-binding (Arg295, Ser297) as well as other conserved histidines are present within the *cF6H* sequence (Figures 26, 33). There is significantly more variation at the N-terminal, suggesting that this region may be involved in maintaining the protein's tertiary conformation, or in substrate binding and proper positioning within the active site. Given that various attempts to isolate the full-length F6H sequence were unsuccessful, along with previous reports of a similar truncated hydroxylase being catalytically active (Vazquez-Flota *et al.*, 1997) and the inherent variablility of the N-terminal, the truncated *cF6H* was expressed in heterologous systems to assess recombinant protein activity. In both prokaryote and eukaryote expression systems rF6H was soluble, exhibited similar physicochemical properties to the native protein and was not significantly more susceptible to proteolytic degradation. Nevertheless, rF6H exhibited a reduced specific activity in comparison to the native F6H, regardless of the expression system, degree of purification or the location of the polyhistidine tag. This suggests that the N-terminus may participate in substrate binding, either by directly contributing to the flavonol binding site, or by maintaining the appropriate conformation for substrate recognition and/or binding. Alternatively, the low enzyme activity may have resulted from the lack of a required cofactor or problems of complementation that are inherent to the native plant enzyme, although the highly purified plant preparations (α-KG-Sepharose stage) exhibited some F6H activity nonetheless. It may also have been possible that mutations were inadvertently introduced into the coding sequence during the PCR amplification process. However, this is highly unlikely since measures taken to limit this possibility

included the use of high fidelity DNA polymerase *Pfu* (Stratagene) and repeated sequencing of the different fragments and clones independently isolated.

A preliminary assessment of the apparent K_m values of rF6H for the flavonol substrate (63 μM) suggests that the binding affinity has been reduced 17-fold from the native protein (0.27 μM). In contrast, that for α-KG was only slightly different at 78 μM indicating that the binding site for the cosubstrate has not been substantially altered. It was not possible to conduct a detailed kinetic analysis of the recombinant protein, given the truncation of the N-terminal, reduced enzyme activity and the difficulties encountered in obtaining values within acceptable standard deviations.

Interestingly, as with the native plant protein, the recombinant F6H was functional as a monomer, but also was shown to associate *in vitro*. This may have resulted from the disruption during isolation of any weak interactions with other proteins, particularly those involved in polymethylated flavonol biosynthesis in the case of the native plant protein. Nevertheless, this association is disrupted upon incubation of the enzyme preparation with Fe^{2+} alone and with Fe^{2+} and α-KG (Lukacin and Britsch, 1997). A hydrophobic prediction plot of the translated *cF6H* (not shown) indicated the presence of a nonpolar region at the C-terminal end of the protein sequence. This region may possibly participate in dimerization or aggregation processes. Although the intracellular localization of F6H and its potential association with other flavonol biosynthetic enzymes remain to be elucidated, it is believed to be involved in the channelled pathway proposed for the biosynthesis of polymethylated flavonols in *C. americanum* (Ibrahim, 1992), analogous to the 'dynamic complex' suggested for the enzymes involved in the general pathway of flavonoid biosynthesis (Winkel-Shirley,

1999). Such a complex would allow for a more efficient catalysis of substrates, protection of rare or unstable compounds, rapid adjustments to changing concentration gradients and a more efficient regulation of the pathway as a whole.

The deduced amino acid sequence of *cF6H* exhibits similarity with other αKDDs, particularly at the C-terminal, and this conservation can be extended to the organization of the gene as a whole. *gF6H* possesses two introns at conserved locations for αKDDs. Prescott and John (1996) observed 4 potential intron positions that are conserved through the evolution of this class of enzymes. This suggests that αKDD sequences arose from a common ancestor possessing three introns to produce, through divergence, the variety observed today. The fact that the *F6H* is present as a single copy in this tissue has implications for it regulation, particularly if the other enzymes involved in this pathway are members of multigene families. In addition, its single copy nature may be significant, since there is a second hydroxylation that takes place at the 2′ position of the trimethylquercetin intermediate (Figure 2). Therefore, it is unlikely that another related αKDD enzyme is involved in the 2′-hydroxylation reaction. A comparable reaction was recently reported to be catalyzed by a cytochrome P450-dependent monooxygenase involving the 2′-hydroxylation of isoflavones in *Medicago truncatula* (Liu *et al.* 2003).

At the phylogenetic level, αKDDs comprise a superfamily with numerous subgroups whose families are defined by their shared motifs, such as those involved in flavonoid biosynthesis. These motifs, more often than not, comprise the active site of the enzyme or the binding domains for substrates and cofactors. Given that the entire complement of flavonoid biosynthetic genes, structural and regulatory, has yet to be cloned and characterized from a single plant species, any assumptions on the ancestry of

these genes based on the compilation of αKDDs originating from different organisms should be formulated with care. One problem lies in the determination of whether the sequences being compared are orthologous (common ancestor), paralogous (arose from a gene duplication event and subsequent divergence) or, the less common, xenologous (horizontal transfer). The topology of a set of taxa can be correctly ascertained from orthologous genes, but others will result in phylogenetic trees with incorrect topologies (Gray and Fitch, 1983). Nevertheless, assessing the level of similarity between homologous genes from different species provides insight as to whether a particular gene may share some evolutionary relatedness to a class of related enzymes.

The degree of similarity within groups (*i.e. F3Hs* from different species share ~ 70-80% similarity) is relatively higher than between groups. However, the level of similarity between *F3H* and *FLS* in different species ranges from 20-40%, although the values for the latter are significantly more variable. In addition, the genes encoding proteins in the same pathway of a given species, although not in the same subgroup, show expectedly higher identity than unrelated genes. The degree of similarity between genes encoding proteins catalyzing dioxygenase reactions of different substrates suggests a common reaction mechanism. However, it is difficult to predict the amino acids that are responsible for the evolution of a novel enzyme activity, since a gene can accumulate mutations that may result in amino acid changes that do not necessarily affect activity, whereas others are critical for a change of function. When aligned with related sequences, F6H exhibits a high degree of conservation of the motifs required for iron- and cosubstrate binding. Nevertheless, F6H is distinct from its most closely related genes in that certain amino acids present within its sequence are absent in those of F3H from

various species, as evidenced by the insertion of gaps for the proper alignment of the sequences in question, especially in certain regions of high conservation (Figure 33). Since the jellyroll structure is relatively tolerant of insertions, it is presumed therefore, that these differences would be the most likely to contribute to the difference in substrate preference, without being detrimental to the tertiary conformation of the active enzyme. Since F6H catalyzes a novel type of reaction for plant αKDDs, an aromatic hydroxylation, some of these alterations in the proteins primary structure would undoubtedly affect the reaction type. Further evidence of such distinctness is the fact that, although F6H clusters with the F3H group of flavonoid dioxygenases (Figure 32), it appears to have evolved in a paraphyletic fashion with respect to this group of genes. Therefore, it seems likely that F3H and F6H shared a common ancestor, which is reinforced by similarities in genomic organization, but significant divergence has since occurred resulting in the evolution of this novel activity. General flavonoid biosynthesis may have arisen from either a multifunctional dioxygenase or a F3H-like ancestor that gave rise to the later activities of FLS and ANS, or through both mechanisms depending on the species. In order for scientists to postulate a plausible evolutionary pathway for this class of dioxygenases, gene sequences and the corresponding *in vitro* activities of their encoded proteins must be compared from both primitive and higher-order plant species.

G. CLAIM TO ORIGINALITY AND CONTRIBUTION TO KNOWLEDGE

1) This work led to the discovery and characterization of a novel enzyme, flavonol 6-hydroxylase, involved in the biosynthesis and modification of an important group of plant secondary metabolites, the flavonols. This activity is significant since the products of hydroxylation reactions often serve as precursors for numerous substitution reactions.
2) The native enzyme was purified and the kinetic parameters determined for the hydroxylation reaction. This work contributes to the limited pool of information on plant ketoglutarate-dependent dioxygenases, an area of immense interest.
3) A near-full length *F6H* clone was isolated and characterized at the molecular level. This allowed the comparison of the F6H with other αKDDs, to assess similarities in genomic organization and evolutionary relatedness to other genes involved in plant secondary metabolism.
4) An active, heterologously-expressed *F6H* clone will be particularly useful in the production of polyclonal antibodies that can be used subsequently in experiments to determine the intracellular locations of enzymes involved in the biosynthesis of polymethylated flavonoids.
5) In addition, a clone with full enzyme activity, which can be isolated based on the *cF6H* sequence provided fresh tissue is available, can be used to assess the versatility of this enzyme with respect to substrate preference through site-directed mutagenesis, as well as to assess its potential applications in metabolic engineering.

H. REFERENCES

Anzellotti D and Ibrahim RK (2000) Novel flavonol 2-oxoglutarate-dependent-dioxygenase: affinity purification, characterization and kinetic properties. Arch Biochem Biophys **382:** 161-172

Balague C, Watson CF, Turner AJ, Rouge P, Picton S, Pech JC, Grierson D (1993) Isolation of a ripening and wound-induced cDNA from *Cucumis melo* L. encoding a protein with homology to the ethylene-forming enzyme. Eur J Biochem **212:** 27-34

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

Blanchard JS, Englard S, Kondo A (1982) γ-Butyrobetaine hydroxylase: a unique protective effect of catalase. Arch Biochem Biophys **219:** 327-334

Blume B, Grierson D (1997) Expression of ACC oxidase promoter-GUS fusions in tomato and *Nicotiana plumbaginifolia* regulated by developmental and environmental stimuli. Plant J **12:** 731-746

Bohm BA (1998) Introduction to Flavonoids. Harwood Academic Publishers, Amsterdam pp. 339-393

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem **72:** 248-254

Brierley RA, Davis GR, Holtz GC (1994) Production of insulin-like growth factor-1 in methylotrophic yeast cells. United States Patent 5,324,639

Britsch L (1990) Purification and characterization of flavone synthase I, a 2-oxoglutarate-dependent desaturase. Arch Biochem Biophys **282:** 152-160

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

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

Chiang HH, Hwang I, Goodman HM (1995) Isolation of the *Arabidopsis GA4* locus. Plant Cell **7:** 195-201

Chittum HS, Lane WS, Carlson BA, Roller PP, Lung FD, Lee BJ, Hatfield DL (1998) Rabbit β-globin is extended beyond its UGA stop codon by multiple suppressions and translational reading gaps. Biochemistry **37:** 10866-10870

Clare JJ, Romanos MA, Rayment FB, Rowedder JE, Smith MA, Payne MM, Sreekrishna K, Henwood CA (1991) Production of epidermal growth factor in yeast: high level secretion using *Pichia pastoris* strains containing multiple gene copies. Gene **105:** 205-212

Coles JP, Phillips AL, Croker SJ, Garcia-Lepe R, Lewis MJ, Hedden P (1999) Modification of gibberellin production and plant development in *Arabidopsis* by sense and antisense expression of gibberellin 20-oxidase genes. Plant J **17:** 547-556

Collins FW, De Luca, V, Ibrahim R.K., Voirin, B., Jay, M. (1981) Polymethylated flavonols of *Chrysosplenium americanum* I. Identification and enzymatic synthesis. Z. Naturforsch **36C:** 731-736

Croteau R, Kutchan, T., Lewis, N.G. (2000) Natural Products (Secondary Metabolites). *In* Biochemistry and Molecular Biology of Plants. Buchanan B, Gruissem, W., Jones, R., eds. American Society of Plant Physiologists, Rockville, pp 1250-1318

Cukovic D, Ehlting J, VanZiffle JA, Douglas CJ (2001) Structure and evolution of 4-coumarate:coenzyme A ligase (4CL) gene families. Biol Chem **382:** 645-654

Davies KM (1993) A *Malus* cDNA with homology to the *Antirrhinum, Candica* and *Zea A2* genes. Plant Physiol **103:** 1015

De Carolis E, De Luca V (1993) Purification, characterization, and kinetic analysis of a 2-oxoglutarate-dependent dioxygenase involved in vindoline biosynthesis from *Catharanthus roseus*. J Biol Chem **268:** 5504-5511

De Carolis E, De Luca V (1994) 2-oxoglutarate-dependent dioxygenase and related enzymes: biochemical characterization. Phytochem **36:** 1093-1107

De Luca V, Ibrahim RK (1985a) Enzymatic synthesis of polymethylated flavonols in *Chrysosplenium americanum*. I. Partial purification and some properties of S-adenosyl-L-methionine:flavonol 3-, 6-, 7-, and 4'-*O*-methyltransferases. Arch Biochem Biophys **238:** 596-605

De Luca V, Ibrahim RK (1985b) Enzymatic synthesis of polymethylated flavonols in *Chrysosplenium americanum*. II. Substrate interaction and product inhibition studies of flavonol 3-, 6-, and 4'-*O*-methyltransferases. Arch Biochem Biophys **238:** 606-618

Dixon RA, Steele CL (1999) Flavonoids and isoflavonoids - a gold mine for metabolic engineering. Trends Plant Sci **4:** 394-400

Dong JG, Fernandez-Maculet JC, Yang SF (1992) Purification and characterization of 1-aminocyclopropane-1-carboxylate oxidase from apple fruit. Proc Natl Acad Sci USA **89:** 9789-9793

Donnelly JK, Robinson DS (1995) Free radicals in foods. Free Radic Res **22**:147-76

Ecker JR (1995) The ethylene signal transduction pathway in plants. Science **268:** 667-675

Eichhorn E, van der Ploeg JR, Kertesz MA, Leisinger T (1997) Characterization of α-ketoglutarate-dependent taurine dioxygenase from *Escherichia coli.* J Biol Chem **272:** 23031-23036

Eng CM, Niehaus DJ, Enriquez AL, Burgert TS, Ludman MD, Desnick RJ (1994) Fabry disease: twenty-three mutations including sense and antisense CpG alterations and identification of a deletional hot-spot in the α-galactosidase A gene. Hum Mol Genet **3:** 1795-1799

Fernandez-Maculet JC, Dong JG, Yang SF (1993) Activation of 1-aminocyclo-propane-1-carboxylate oxidase by carbon dioxide. Biochem Biophys Res Commun **193:** 1168-1173

Forkmann G, Heller W (1999) Biosynthesis of flavonoids. *In* Comprehensive Natural Products Chemistry. Barton D, Nakanishi K, eds. Elsevier, Amsterdam, pp. 713-748

French CJ, Towers GHN (1992) Inhibition of infectivity of potato virus X by flavonoids. Phytochem **31:** 3017-3020

Friedman KD, Rosen NL, Newman PJ, Montgomery RR (1990) Screening of lambda-gt11 libraries. *In* PCR Protocols: A guide to methods and applications. Academic Press, San Diego, pp 253-257

Fukumori F, Hausinger RP (1993) Purification and characterization of 2,4-dichlorophenoxyacetate/α-ketoglutarate dioxygenase. J Biol Chem **268:** 24311-24317

Garcia I, Rodgers M, Lenne C, Rolland A, Sailland A, Matringe M (1997) Subcellular localization and purification of a p-hydroxyphenylpyruvate dioxygenase from cultured carrot cells and characterization of the corresponding cDNA. Biochem J **325:** 761-769

Gardner HW (1991) Recent investigations into the lipoxygenase pathway of plants. Biochim Biophys Acta **1084:** 221-239

Gauthier A, Gulick PJ, Ibrahim RK (1996) cDNA cloning and characterization of a 3'/5'-*O*-methyltransferase for partially methylated flavonols from *Chrysosplenium americanum*. Plant Mol Biol **32:** 1163-1169

Gray GS, Fitch WM (1983) Evolution of antibiotic resistance genes: the DNA sequence of a kanamycin resistance gene from *Staphylococcus aureus*. Mol Biol Evol **1:** 57-66

Halbwirth H, Wurst F, Forkmann G, Stich K (1999) Enzymatic A-ring hydroxylation of flavonols in position 8. *In* Abstracts of the annual meeting of the Phytochemical Society of North America, Montreal, Canada

Hashimoto T, Yamada Y (1987) Purification and characterization of hyoscyamine 6 β-hydroxylase from root cultures of *Hyoscyamus niger* L. Hydroxylase and epoxidase activities in the enzyme preparation. Eur J Biochem **164:** 277-285

Hedden P, Graebe JE, Beale MH, Gaskin P, MacMillan J (1984) The biosynthesis of 12α-hydroxylated gibberellins in a cell free system from *Cucurbita maxima* endosperm. Phytochem **23:** 569-574

Hegg EL, Que L, Jr. (1997) The 2-His-1-carboxylate facial triad--an emerging structural motif in mononuclear non-heme iron(II) enzymes. Eur J Biochem **250:** 625-629

Hollman PC, Katan MB (1998) Bioavailability and health effects of dietary flavonols in man. Arch Toxicol Suppl **20:** 237-248

Holton TA, Brugliera F, Tanaka Y (1993) Cloning and expression of flavonol synthase from *Petunia hybrida*. Plant J **4:** 1003-1010

Ibrahim RK (1992) Immunolocalization of Flavonoid Conjugates. *In*, Phenolic Metabolism in Plants. Stafford H, Ibrahim R eds. Plenum Press, New York, pp 25-61

Ibrahim RK, Anzellotti, D (2003) Enzymatic Basis of Flavonoid Diversity. *In* Integrative Phytochemistry: from ethnobotany to molecular ecology. JT Romeo ed, Plenum Press, New York, pp 1-36

Ibrahim RK, De Luca V, Khouri H, Latchinian L, Brisson L, Charest PM (1987) Enzymology and compartmentation of polymethylated flavonol glucosides in *Chrysosplenium americanum*. Phytochem **26:** 1237-1249

Jay M, Voirin, B (1976) Les flavonoides de deux especes du genre *Chrysosplenium*. Phytochem **15:** 517-519

John DC, Bulleid NJ (1994) Prolyl 4-hydroxylase: defective assembly of alpha-subunit mutants indicates that assembled α-subunits are intramolecularly disulfide bonded. Biochemistry **33:** 14018-14025

Khouri HE, Ishikura, N. and Ibrahim, R.K. (1986) FPLC purification and some properties of a partially methylated flavonol glucoside 2'/5'-*O*-methyltransferase from *Chrysosplenium americanum*. Phytochem **25:** 2475-2479

Khouri HE, Ibrahim RK (1987) Resolution of 5 position specific *O*-methyltransferases by fast protein liquid chromatofocusing. J Chromatog **407:** 291-297

Kim G, Kim Y (1998) Flavanone 3-Hydroxylase: cloning and sequencing from *Nicotiana tabacum*.Plant Physiol **116:** 1605

Kivirikko KI, Pihlajaniemi T (1998) Collagen hydroxylases and the protein disulfide isomerase subunit of prolyl 4-hydroxylases. Adv Enzymol Relat Areas Mol Biol **72:** 325-398

Ko K, Granell A, Bennett J, Cashmore AR (1990) Isolation and characterization of cDNAs from *Lycopersicon esculentum* and *Arabidopsis thaliana* encoding the 33 kDa protein of the photosystem II-associated oxygen-evolving complex. Plant Mol Biol **14:** 217-227

Koes RE, Spelt CE, van den Elzen PJ, Mol JN (1989) Cloning and molecular characterization of the chalcone synthase multigene family of *Petunia hybrida*. Gene **81:** 245-257

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

Lange T (1997) Cloning gibberellin dioxygenase genes from pumpkin endosperm by heterologous expression of enzyme activities in *Escherichia coli*. Proc Natl Acad Sci USA **94**: pp 6553-6558

Lange T, Hedden P, Graebe JE (1994) Expression cloning of a gibberellin 20-oxidase, a multifunctional enzyme involved in gibberellin biosynthesis. Proc Natl Acad Sci USA **91:** 8552-8556

Latchinian L, Ibrahim RK (1989) Characterization of a monoclonal antibody specific to a flavonol 2'-*O*-glucosyltransferase. Biochem Cell Biol **67:** 210-213

Latchinian L, Khouri, HE and Ibrahim RK (1987) Fast protein affinity chromatography of two flavonoid glucosyltransferases. J Chromatog **388:** 235-242

Latunde-Dada AO, Cabello-Hurtado F, Czittrich N, Didierjean L, Schopfer C, Hertkorn N, Werck-Reichhart D, Ebel J (2001) Flavonoid 6-hydroxylase from soybean (*Glycine max* L.), a novel plant 4-450 monooxygenase. J Biol Chem **276:** 1688-1695

Lincoln JE, Cordes S, Read E, Fischer RL (1987) Regulation of gene expression by ethylene during *Lycopersicon esculentum* (tomato) fruit development. Proc Natl Acad Sci USA **84:** 2793-2797

Liu JH, Lee-Tamon SH, Reid DM (1997) Differential and wound-inducible expression of 1-aminocylopropane-1-carboxylate oxidase genes in sunflower seedlings. Plant Mol Biol **34:** 923-933

Liu CJ, Huhman D, Sumner LW, Dixon RA (2003) Regiospecific hydroxylation of isoflavones by cytochrome p450 81E enzymes from *Medicago truncatula*. Plant J **36:** 471-484

Lukacin R, Britsch L (1997) Identification of strictly conserved histidine and arginine residues as part of the active site in *Petunia hybrida* flavanone 3β-hydroxylase. Eur J Biochem **249:** 748-757

Lukacin R, Groning I, Schiltz E, Britsch L, Matern U (2000) Purification of recombinant flavanone 3β-hydroxylase from *Petunia hybrida* and assignment of the primary site of proteolytic degradation. Arch Biochem Biophys **375:** 364-370

Maddison DR, Maddison, W.P. (2000) McClade, Version 4.02. Sinauer Associates, Sunderland, Massachusetts.

Martens S, Forkmann G, Matern U, Lukacin R (2001) Cloning of parsley flavone synthase I. Phytochem **58:** 43-46

Martin C, Prescott A, Mackay S, Bartlett J, Vrijlandt E (1991) Control of anthocyanin biosynthesis in flowers of *Antirrhinum majus*. Plant J **1:** 37-49

Martin DN, Proebsting WM, Hedden P (1999) The *SLENDER* gene of pea encodes a gibberellin 2-oxidase. Plant Physiol **121:** 775-781

Matsuda J, Okabe S, Hashimoto T, Yamada Y (1991) Molecular cloning of hyoscyamine 6β-hydroxylase, a 2-oxoglutarate-dependent dioxygenase, from cultured roots of *Hyoscyamus niger*. J Biol Chem **266:** 9460-9464

Menssen A, Hohmann S, Martin W, Schnable PS, Peterson PA, Saedler H, 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

Middleton E, Kandaswami C. (1994) The impact of plant flavonoids on mammalian biology: implications for immunity, inflammation and cancer. *In* The Flavonoids: Recent Advances in Research Since 1986. JB Harborne, ed. Chapman and Hall, London, pp 619-652

Ming LJ, Que L, Jr., Kriauciunas A, Frolik CA, Chen VJ (1991) NMR studies of the active site of isopenicillin N synthase, a non-heme iron(II) enzyme. Biochemistry **30:** 11653-11659

Myllyharju J, Kivirikko KI (1997) Characterization of the iron- and 2-oxoglutarate-binding sites of human prolyl 4-hydroxylase. EMBO J **16:** 1173-1180

Myllyla R, Gunzler V, Kivirikko KI, Kaska DD (1992) Modification of vertebrate and algal prolyl 4-hydroxylases and vertebrate lysyl hydroxylase by diethyl pyrocarbonate. Evidence for histidine residues in the catalytic site of 2-oxoglutarate-coupled dioxygenases. Biochem J **286:** 923-927

Myllyla R, Majamaa K, Gunzler V, Hanauske-Abel HM, Kivirikko KI (1984) Ascorbate is consumed stoichiometrically in the uncoupled reactions catalyzed by prolyl 4-hydroxylase and lysyl hydroxylase. J Biol Chem **259:** 5403-5405

Nakanishi H, Okumura N, Umehara Y, Nishizawa NK, Chino M, Mori S (1993) Expression of a gene specific for iron deficiency (*Ids3*) in the roots of *Hordeum vulgare*. Plant Cell Physiol **34:** 401-410

Nozaki M (1979) Oxygenases and Dioxygenases. *In* Topics in Current Chemistry Boschke FL, ed. Vol 78. Springer-Verlag, New York, pp 147-186

Okumura N, Nishizawa NK, Umehara Y, Ohata T, Nakanishi H, Yamaguchi T, Chino M, Mori S (1994) A dioxygenase gene (*Ids2*) expressed under iron deficiency conditions in the roots of *Hordeum vulgare*. Plant Mol Biol **25:** 705-719

Osborne DJ (1989) Biochemical and Physiological Aspects of Ethylene Production in Lower and Higher Plants. Kluver, Dordrecht

Pelletier MK, Murrell JR, Shirley BW (1997) Characterization of flavonol synthase and leucoanthocyanidin dioxygenase genes in *Arabidopsis*. Further evidence for differential regulation of "early" and "late" genes. Plant Physiol **113:** 1437-1445

Phillips AL, Ward DA, Uknes S, Appleford NE, Lange T, Huttly AK, Gaskin P, Graebe JE, Hedden P (1995) Isolation and expression of three gibberellin 20-oxidase cDNA clones from *Arabidopsis*. Plant Physiol **108:** 1049-1057

Prescott AG (1993) A dilemma of dioxygenases (or where biochemistry and molecular biology fail to meet). J Exp Botany **44:** 849-861

Prescott AG (2000) Two-oxoacid dependent dioxygenases: inefficient enzymes or evolutionary driving force? *In* Evolution of Metabolic Pathways. Romeo JT, *et al.*, eds, Elsevier Science Ltd., pp 249-284

Prescott AG, John P (1996) Dioxygenases: Molecular structure and role in plant metabolism. Annu Rev Plant Physiol Plant Molec Biol **47:** 245-271

Pueppke SG (1996) The genetic and biochemical basis for nodulation of legumes by Rhizobia. Critic Rev Biotechnol **16:** 1-51

Puistola U, Turpeenniemi-Hujanen TM, Myllyla R, Kivirikko KI (1980) Studies on the lysyl hydroxylase reaction. I. Initial velocity kinetics and related aspects. Biochim Biophys Acta **611:** 40-50

Rausher MD, Miller RE, Tiffin P (1999) Patterns of evolutionary rate variation among genes of the anthocyanin biosynthetic pathway. Mol Biol Evol **16:** 266-274

Rice-Evans CA, Miller NJ (1996) Antioxidant activities of flavonoids as bioactive components of food. Biochem Soc Trans **24:** 790-795

Roach PL, Clifton IJ, Fulop V, Harlos K, Barton GJ, Hajdu J, Andersson I, Schofield CJ, Baldwin JE (1995) Crystal structure of isopenicillin N synthase is the first from a new structural family of enzymes. Nature **375:** 700-704

Ruetschi U, Nordin I, Odelhog B, Jornvall H, Lindstedt S (1993) γ-Butyrobetaine hydroxylase. Structural characterization of the *Pseudomonas* enzyme. Eur J Biochem **213:** 1075-1080

Saito K, Kobayashi M, Gong Z, Tanaka Y, Yamazaki M (1999) Direct evidence for anthocyanidin synthase as a 2-oxoglutarate-dependent oxygenase: molecular cloning and functional expression of cDNA from a red forma of *Perilla frutescens*. Plant J **17:** 181-189

Saito N, Harborne JB (1992) Correlation between anthocyanin type, pollinator and flower color in the Labiatae. Phytochem **31:** 3009-3015

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol **4:** 406-425

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, Second Edn. Cold Spring Harbor Laboratory, Cold Spring Harbor

Sambrook J, Russell DW (2001) Molecular Cloning: a laboratory manual. Edn 3. Cold Spring Harbor Press, Cold Spring Harbor

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA **74:** 5463-5467

Schofield CJ, Baldwin, J.E., Byford, M.F., Clifton, I., Hadju, J., Hensgens, C., Roach, P. (1997) Proteins of the penicillin biosynthesis pathway. Curr Opin Struct Biol **7:** 857-864

Segel I (1975) Enzyme Kinetics: behaviour and analysis of rapid equilibrium and steady-state enzyme systems. John Wiley and Sons, New York

Seidman JD (1994) Screening of recombinant DNA libraries. *In* Current Protocols on Molecular Biology. John Wiley and Sons, New York, pp 6.0.3-6.3.6

Shorrosh BS, Dixon RA (1991) Molecular cloning of a putative plant endomembrane protein resembling vertebrate protein disulfide-isomerase and a

phosphatidylinositol-specific phospholipase C. Proc Natl Acad Sci USA **88:** 10941-10945

Sparvoli F, Martin C, Scienza A, Gavazzi G, 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

Stafford H (1991) Flavonoid evolution: An enzymatic approach. Plant Physiol **96:** 680-685

St-Pierre B, Vazquez-Flota FA, De Luca V (1999) Multicellular compartmentation of *Catharanthus roseus* alkaloid biosynthesis predicts intercellular translocation of a pathway intermediate. Plant Cell **11:** 887-900

Suzuki K, Yun DJ, Chen XY, Yamada Y, Hashimoto T (1999) An *Atropa belladonna* hyoscyamine 6β-hydroxylase gene is differentially expressed in the root pericycle and anthers. Plant Mol Biol **40:** 141-152

Swofford DL (2000) PAUP. Sinauer Associates, Sunderland, Massachusetts.

Taylor LP, Jorgensen R (1992) Conditional male fertility in chalcone synthase-deficient petunia. J Heredity **83:** 11-17

Thomas SG, Phillips AL, Hedden P (1999) Molecular cloning and functional expression of gibberellin 2-oxidases, multifunctional enzymes involved in gibberellin deactivation. Proc Natl Acad Sci USA **96:** 4698-4703

Thomson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W- improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. Nuc Acids Res **22:** 4673-4680.

Turnbull JJ, Nakajima J, Welford RWD, Yamazaki M, Saito K, Schofield CJ (2004) Mechanistic studies on three 2-oxoglutarate-dependent oxygenases of flavonoid biosynthesis. J Biol Chem **279:** 1206-1216

van Eldik GJ, Reijnen WH, Ruiter RK, van Herpen MM, Schrauwen JA, Wullems GJ (1997) Regulation of flavonol biosynthesis during anther and pistil development, and during pollen tube growth in *Solanum tuberosum*. Plant J **11:** 105-113

van Tunen AJ, Koes RE, Spelt CE, van der Krol AR, Stuitje AR, Mol JN (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

Vazquez-Flota F, De Carolis E, Alarco AM, De Luca V (1997) Molecular cloning and characterization of desacetoxyvindoline-4-hydroxylase, a 2-oxoglutarate dependent-dioxygenase involved in the biosynthesis of vindoline in *Catharanthus roseus* (L.) G. Don. Plant Mol Biol **34:** 935-948

Vazquez-Flota FA, De Luca V (1998) Jasmonate modulates development- and light-regulated alkaloid biosynthesis in *Catharanthus roseus*. Phytochem **49:** 395-402

Weiss D, van der Luit AH, Kroon JT, Mol JN, 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

Winkel-Shirley B (1999) Evidence for enzyme complexes in the phenylpropanoid and flavonoid pathways. Plant Physiol **107:** 142-149

Wisman E, Hartmann U, Sagasser M, Baumann E, Palme K, Hahlbrock K, Saedler H, Weisshaar B (1998) Knock-out mutants from an *En-1* mutagenized *Arabidopsis thaliana* population generate phenylpropanoid biosynthesis phenotypes. Proc Natl Acad Sci USA **95:** 12432-12437

Wojtaszek P, Smith CG, Bolwell GP (1999) Ultrastructural localisation and further biochemical characterisation of prolyl 4-hydroxylase from *Phaseolus vulgaris*: comparative analysis. Int J Biochem Cell Biol **31:** 463-477