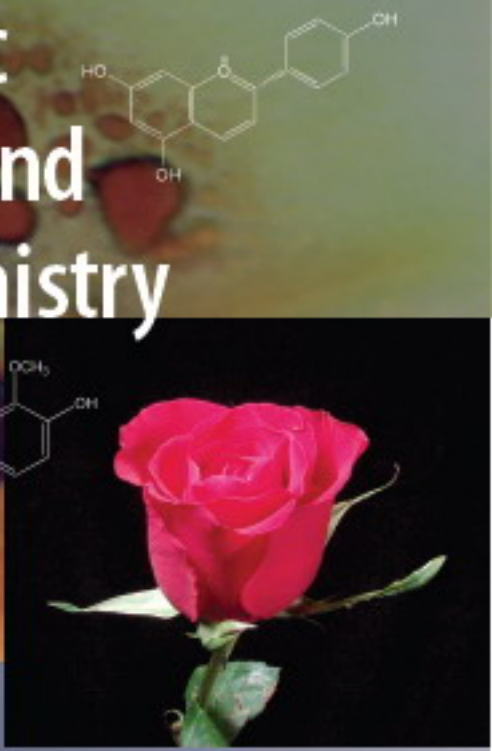
Vermerris · Nicholson

Phenolic Compound Biochemistry

Wilfred Vermerris
Ralph Nicholson

Phenolic Compound Biochemistry



Springer

PHENOLIC COMPOUND BIOCHEMISTRY

PHENOLIC COMPOUND BIOCHEMISTRY

By

WILFRED VERMERRIS
University of Florida, Gainesville, FL, U.S.A.

and

RALPH NICHOLSON
Purdue University, West Lafayette, IN, U.S.A.

Springer

A C.I.P. Catalogue record for this book is available from the Library of Congress.

ISBN-10 1-4020-5163-8 (HB)
ISBN-13 978-1-4020-5163-0 (HB)
ISBN-10 1-4020-5164-6 (e-book)
ISBN-13 978-1-4020-5164-7 (e-book)

Published by Springer,
P.O. Box 17, 3300 AA Dordrecht, The Netherlands.

www.springer.com

Printed on acid-free paper

CONTENTS

CHAPTER 2
Chemical Properties of Phenolic Compounds

CHAPTER 3
Biosynthesis of Phenolic Compounds

CHAPTER 4
Isolation and Identification of Phenolic Compounds

PREFACE

Phenolic compounds represent a large group of molecules with a variety of functions in plant growth, development, and defense. Phenolic compounds include signaling molecules, pigments and flavors that can attract or repel, as well as compounds that can protect the plant against insects, fungi, bacteria, and viruses. Most phenolic compounds are present as esters or glycosides rather than as free compounds. Tannins and lignin are phenolic polymers. Tannins are used comercially as dyes and astringents, and lignin accounts for structural ridgidity of cells and tissues and is essential to vascular development. From this brief overview it is apparent that phenolic compounds make up a large and fascinating family.

Our intention is to provide the reader with an overview of this family of compounds that will show their diversity and provide a basis for continued investigations. The target audience is intended to include phytopathologists, botanists, agronomists, entomologists, and people with a general interest in plant biochemistry. This book will also be a valuable resource when used as a textbook in a course on phenolic compounds, aimed at advanced undergraduate or beginning graduate students in the life sciences. While writing this book we assumed the reader would have basic knowledge of organic chemistry, biochemistry of DNA, RNA, proteins and lipids, and cell physiology. At the end of the chapters we have provided some additional references for further reading, either to obtain background information, or to obtain more details.

The focus of this book is centered on structure, nomenclature and occurrence of phenolic compounds (Chapter 1), and their chemical properties (Chapter 2). Chapter 3 describes the biosynthetic pathways leading to the major classes of phenolics. This chapter presents an up-to-date overview of the genetic approaches that have been used to elucidate these pathways. Chapter 4 presents an overview of methods for the isolation and identification of plant phenolic compounds. Given that much of the recent

advances in the identification of phenolic compounds have been made possible through the development of mass spectrometry, we have dedicated an entire chapter to the use of mass spectrometry in the identification of phenolic compounds (Chapter 5). This chapter was written by Dr. Karl V. Wood of the Mass Spectrometry Facility in the Department of Chemistry at Purdue University. We are grateful for his contribution to this book. Chapter 6 highlights the role of phenolic compounds in plant defense. We have included a number of examples, including phytoanticipins, phytoalexins and lignin synthesized in response to pathogen attack, to complement the more chemical nature of the other chapters, and to illustrate the important role phenolic compounds play in plant survival. The final chapter is an overview of some of the positive effects phenolics can have on human health.

We would like to thank Dr. Karl Wood for helpful comments on the manuscript, Drs. William Bourdoncle, Surinder Chopra, Hans Jung and Michael McMullen for providing details on specific topics, Ana Saballos for helpful discussions, Shelly Dunk and Flora Mauch for assistance with the references, and Zuzana Bernhardt and Ineke Ravesloot at Springer Publishers for their encouragement and flexibility.

This book would not have been possible without our supporters at home, Heather (R.L.N.) and Lauren and Deirdre (W.V.), and we are grateful for their patience during the writing of this book, especially when this took place during evenings and weekends.

West Lafayette, IN
Spring 2006

W. Vermerris
Associate Professor of Agronomy,
University of Florida Genetics Institute,
Gainesville, FL

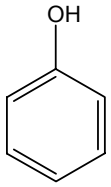
R. L. Nicholson
Professor of Botany and Plant Pathology,
Purdue University, West Lafayette, IN

Chapter 1

FAMILIES OF PHENOLIC COMPOUNDS AND MEANS OF CLASSIFICATION

1. DEFINITIONS

What are phenolic compounds? They are compounds that have one or more hydroxyl groups attached directly to an aromatic ring. Phenol (**1.1**) is the structure upon which the entire group is based. The aromatic ring in this case is, of course, benzene.



(1.1)

The phenols are in many ways similar to alcohols of aliphatic structures where the hydroxyl group is attached to a chain of carbons. The phenolic hydroxyl group, however, is influenced by the presence of the aromatic ring. Because of the aromatic ring, the hydrogen of the phenolic hydroxyl is labile, which makes phenols weak acids. We will deal with their chemical properties separately in Chapter 2.

Polyphenols are compounds that have more than one phenolic hydroxyl group attached to one or more benzene rings. The term is somewhat misleading since it tends to make people think of polymers of individual phenol molecules. Of course such polymers exist. Phenolic compounds are characteristic of plants and as a group they are usually found as esters or glycosides rather than as free compounds. It is important to realize this if you want to extract phenols from plant tissues.

2. CLASSIFICATION

The term phenolics covers a very large and diverse group of chemical compounds. These compounds can be classified in a number of ways. Harborne and Simmonds (1964) classified these compounds into groups based on the number of carbons in the molecule (Table 1.1).

Table 1-1. Classification of phenolic compounds

Structure	Class
C_6	simple phenolics
C_6 - C_1	phenolic acids and related compounds
C_6 - C_2	acetophenones and phenylacetic acids
C_6 - C_3	cinnamic acids, cinnamyl aldehydes, cinnamyl alcohols
C_6 - C_3	coumarins, isocoumarins, and chromones
C_{15}	chalcones, aurones, dihydrochalcones
C_{15}	flavans
C_{15}	flavones
C_{15}	flavanones
C_{15}	flavanonols
C_{15}	anthocyanidins
C_{15}	anthocyanins
C_{30}	biflavonyls
C_6-C_1-C_6, C_6-C_2-C_6	benzophenones, xanthones, stilbenes
C_6, C_{10}, C_{14}	quinones
C_{18}	betacyanins
Lignans, neolignans	dimers or oligomers
Lignin	polymers
Tannins	oligomers or polymers
Phlobaphenes	polymers

An alternative classification has been used by Swain and Bate-Smith (1962). They grouped the phenols in “common” and “less common” categories. Ribéreau-Gayon (1972) grouped the phenols into three families as follows:

1. Widely distributed phenols – ubiquitous to all plants, or of importance in a specific plant
2. Phenols that are less widely distributed – limited number of compounds known
3. Phenolic constituents present as polymers.

In this chapter each of the classes will be discussed, and structures of commonly found or important compounds will be presented. The compounds will be presented in the order listed in Table 1-1. The main focus will be on phenolic compounds that are widely distributed.

3. CLASSES OF PHENOLIC COMPOUNDS

3.1 Simple phenolics

Simple phenolics are substituted phenols. The *ortho*, *meta* and *para* nomenclature refers to a 1,2-, 1,3- and 1,4-substitution pattern of the benzene ring, respectively, where in this case one of the functional groups is the hydroxyl group. With three functional groups, the substitution pattern can be 1,3,5, which, when all three substituents are identical, is designated as a *meta*-tri-substitution pattern, whereas the 1,2,6, substitution pattern is indicated by the prefix '*vic*'(Figure 1-1).

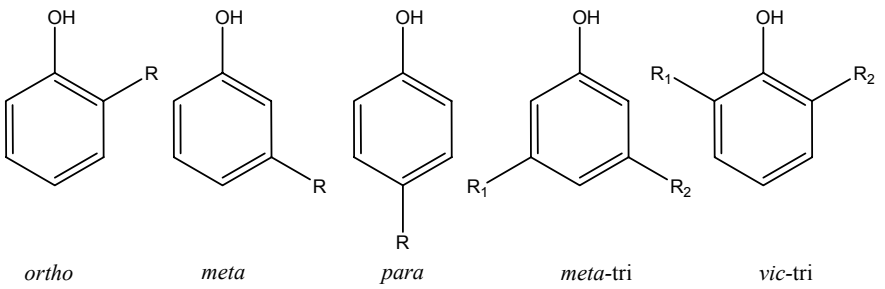
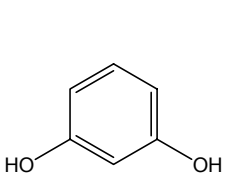
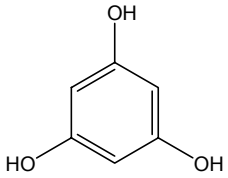


Figure 1-1. Nomenclature for substitution patterns of phenolic compounds. R, R_1 and R_2 are generic substituents.

Examples include resorcinol (1,3-dihydroxybenzene; **1.2**), a *meta*-dihydroxylated simple phenolic, and phloroglucinol (1,3,5-trihydroxybenzene; **1.3**), a *meta*-trihydroxylated simple phenolic.



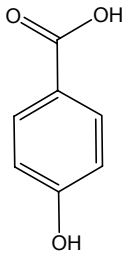
(1.2)



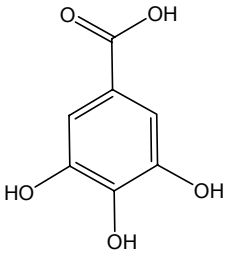
(1.3)

3.2 Phenolic acids and aldehydes

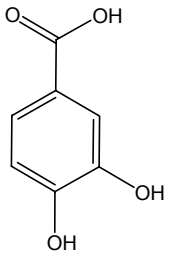
Hydroxy-benzoic acids are characterized by the presence of a carboxyl group substituted on a phenol. Examples include *p*-hydroxybenzoic acid (**1.4**), gallic acid (**1.5**), protocathechuic acid (**1.6**), salicylic acid (**1.7**) and vanillic acid (**1.8**). Related are hydroxybenzoic aldehydes, such as vanillin (**1.9**), which have an aldehyde group in stead of a carboxyl group.



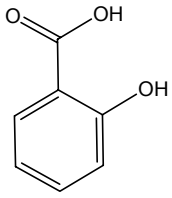
(1.4)



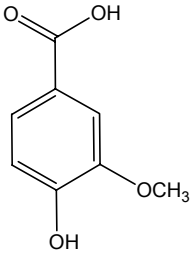
(1.5)



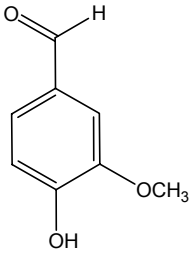
(1.6)



(1.7)



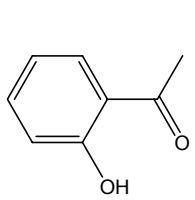
(1.8)



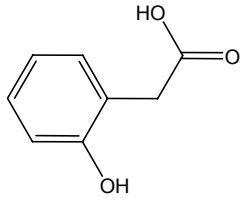
(1.9)

3.3 Acetophenones and phenylacetic acids

Phenones are C_6-C_2 compounds that are rarely found in nature. Examples include 2-hydroxyacetophenone (**1.10**) and 2-hydroxyphenyl acetic acid (**1.11**).



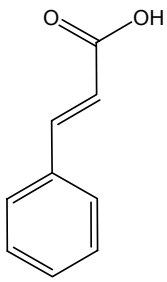
(1.10)



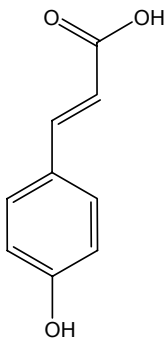
(1.11)

3.4 Cinnamic acids

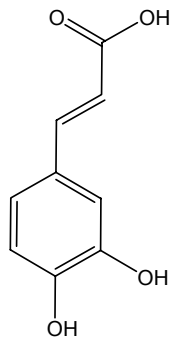
There are six common cinnamic acids, which have a C_6 - C_3 skeleton. All plants probably contain at least three of them. Shown below are cinnamic, acid (**1.12**), *p*-coumaric acid (**1.13**), caffeic acid (**1.14**), ferulic acid (**1.15**), 5-hydroxyferulic acid (**1.16**), and sinapic acid (**1.17**).



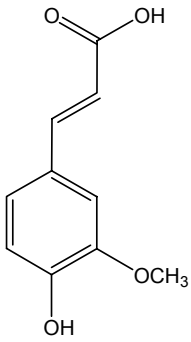
(1.12)



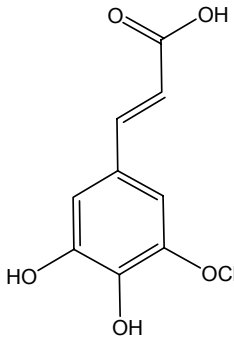
(1.13)



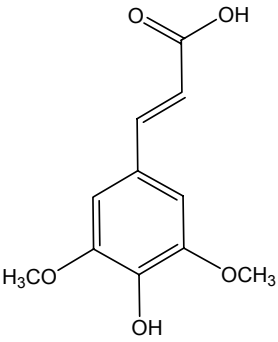
(1.14)



(1.15)

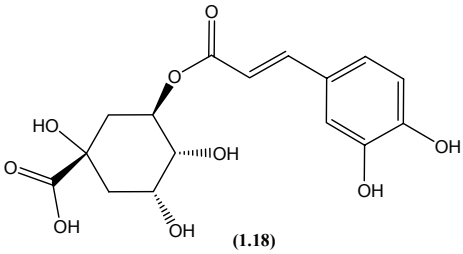


(1.16)



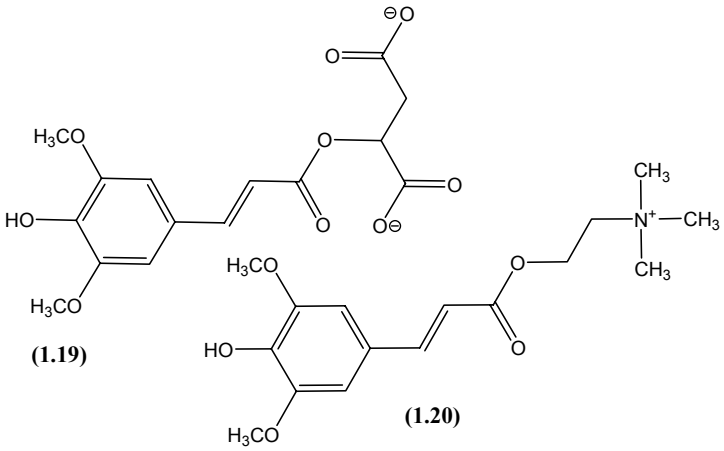
(1.17)

Cinnamic acids are commonly found in plants as esters of quinic acid, shikimic acid, and tartaric acid. For example, chlorogenic acid (**1.18**) is an ester of caffeic acid and quinic acid.



(1.18)

Cinnamic esters are also found as sugar esters, or as esters of a variety of other organic acids. For example, sinapoyl esters represent a class of UV-absorbing compounds in the family of the *Brassicaceae*. Examples include sinapoyl malate (**1.19**) present in leaves, and sinapoyl choline (**1.20**) present in roots (Ruegger and Chapple, 2001).



(1.19)

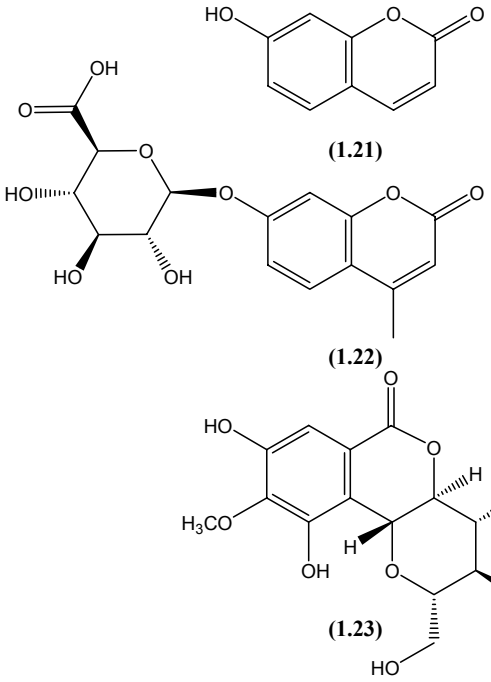
(1.20)

3.5 Coumarins

Coumarins also have a C_6-C_3 skeleton, but they possess an oxygen heterocycle as part of the C_3-unit. There are numerous coumarins, many of which play a role in disease and pest resistance, as well as UV-tolerance. The coumarin umbelliferone (**1.21**) is popular in enzyme assays. Umbelliferone esters can be used as a substrate for non-specific esterase enzyme assays and in fluorescent immunoassays (Jacks and Kircher, 1967). In order to quantify the enzyme activity of the popular reporter gene β-glucuronidase (GUS), plant extracts can be incubated with 4-methylumbelliferyl β-D-glucuronide (4-MUG; **1.22**), which upon hydrolysis

by GUS produces the fluorescent compound 4-methylumbelliferone (4-MU), along with glucuronic acid (Gallagher et al., 1992).

Isocoumarins, such as bergenin (**1.23**) have a structure similar to coumarins, but the position of the oxygen and carbonyl groups within the oxygen heterocycle are reversed. Isocoumarins also play a role in defense responses. For example, bergenin has been shown to inhibit the growth of the powdery mildew on pea (Prithiviraj et al., 1997).

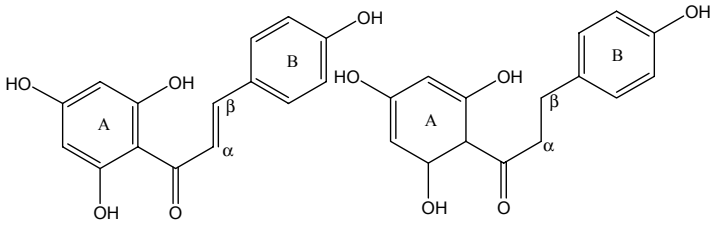


3.6 Flavonoids

Flavonoids are C_{15} compounds all of which have the structure C_6-C_3-C_6. Flavonoids may be grouped into three big classes based on their general structure. In each case, two benzene rings are linked together by a group of three carbons. It is the arrangement of the C_3 group that determines how the compounds are classified.

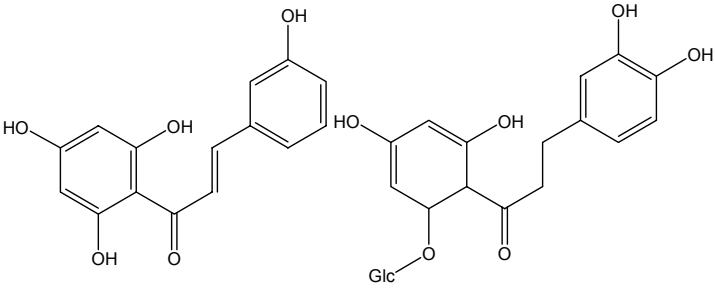
3.6.1 Chalcones

Chalcones (**1.24**) and dihydrochalocones (**1.25**) have a linear C_3-chain connecting the two rings. The C_3-chain of chalcones contains a double bond, whereas the C_3-chain of dihydrochalcones is saturated.



(1.24) **(1.25)**

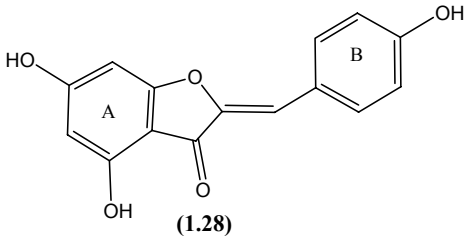
Chalcones, such as butein (**1.26**), are yellow pigments in flowers. An example of a dihydrochalcone is phloridzin (phloretin-2′-O-D-glucoside) (**1.27**), a compound found in apple leaves, and which has been reported to have anti-tumor activity (Nelson and Falk, 1993).



(1.26) **(1.27)**

3.6.2 Aurones

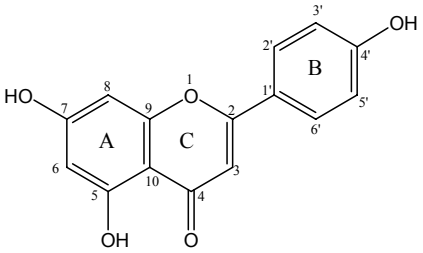
Aurones (**1.28**) are formed by cyclization of chalcones, whereby the *meta*-hydroxyl group reacts with the α-carbon to form a five-member heterocycle. Aurones are also yellow pigments present in flowers.



(1.28)

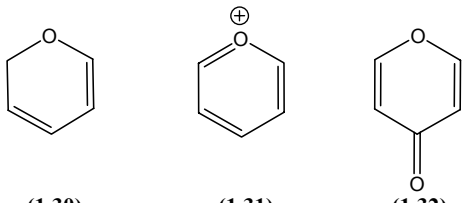
3.6.3 Flavonoids

Typical flavonoids, such as flavanone (**1.29**), have a six-member heterocycle. Flavonoids have an A-, B-, and C-ring, and are typically depicted with the A-ring on the left-hand side. The A-ring originates from the condensation of three malonyl-CoA molecules, and the B-ring originates from *p*-coumaroyl-CoA. These origins explain why the A-ring of most flavonoids is either *meta*-dihydroxylated or *meta*-trihydroxylated.



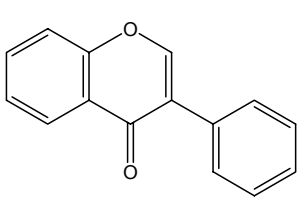
(1.29)

In typical flavonoids one of the *meta*-hydroxyl groups of the A-ring contributes the oxygen to the six atom-heterocycle. The six member oxygen heterocycle of typical flavonoids may be a pyran (**1.30**), pyrylium (**1.31**), or pyrone ring (**1.32**). The B-ring is typically mono-hydroxylated, *ortho*-dihydroxylated, or *vic*-trihydroxylated. The B-ring may also have methyl-ethers as substituents.

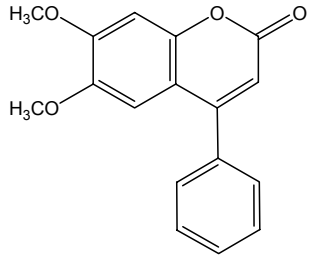


(1.30) **(1.31)** **(1.32)**

Isoflavones, isoflavanones and neoflavonoids are also members of the flavonoid group. They all have the C_6-C_3-C_6 structure but the B-ring is in a different position on the oxygen heterocycle. Examples are isoflavone (**1.33**) and the neoflavonoid dalbergin (**1.34**).



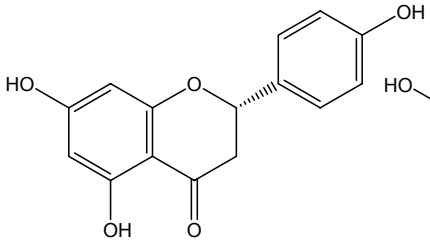
(1.33)



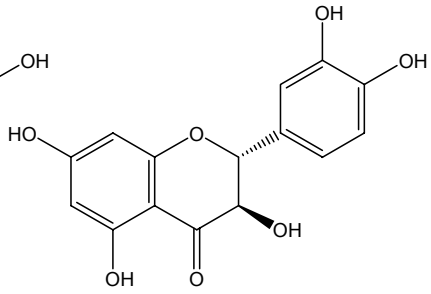
(1.34)

3.6.3.1 Flavanones

The heterocycle of flavanones also contains a ketone group, but there is no unsaturated carbon-carbon bond. The A- and B-ring can be substituted analogous to the flavones, as in naringenin (**1.35**).



(1.35)



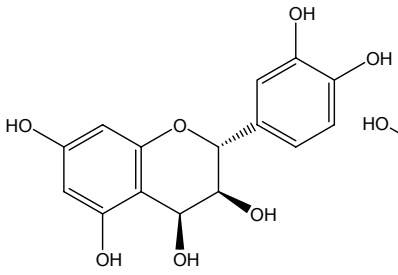
(1.36)

3.6.3.2 Flavanonols

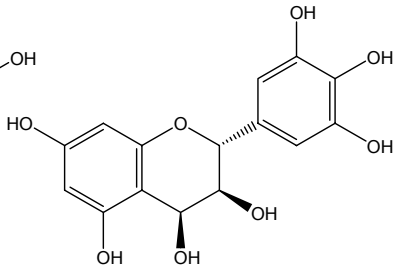
Flavanonols are also known as dihydroflavonols and often occur in association with tannins in heartwood. An example is taxifolin (**1.36**), also known as dihydroquercitin.

3.6.3.3 Leucoanthocyanidins

Leucoanthocyanidins are also referred to as flavan-3,4-*cis*-diols. They are synthesized from flavanonols via a reduction of the ketone moiety on C4. Examples are leucocyanidin (**1.37**) and leucodelphinidin (**1.38**). These compounds are often present in wood and play a role in the formation of condensed tannins.

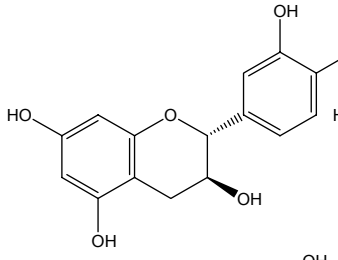


(1.37)

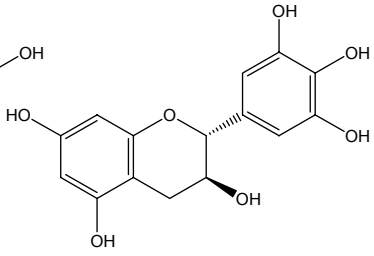


(1.38)

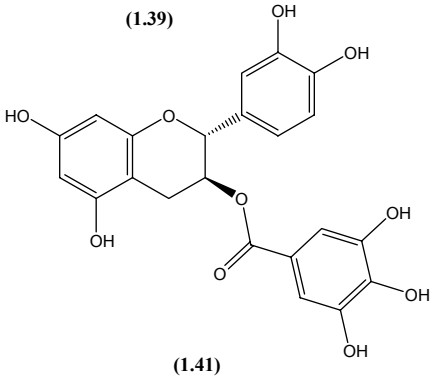
Because of their completely saturated heterocycle, leucoanthocyanidins, together with flavan-3-ols are referred to as flavans. Examples of flavan-3-ols are catechin (**1.39**) and gallocatechin (**1.40**). The 'gallo' in the latter compound refers to the *vic*-tri-hydroxy substitution pattern on the B-ring. Unlike most other flavonoids, the flavans are present as free aglycones or as polymers of aglycones, *i.e.* they are not glycosylated.



(1.39)



(1.40)

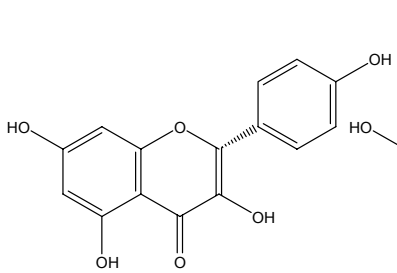


(1.41)

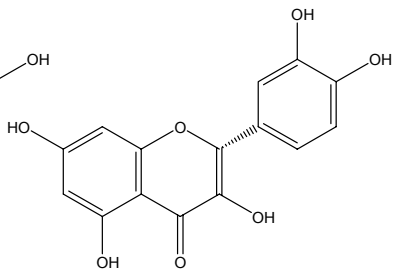
Catechins (**1.41**) can also be found as gallic acid esters that are esterified at the 3′ hydroxyl group. Note the difference between the gallic acid ester of catechin (**1.41**) and gallocatechin (**1.40**).

3.6.3.4 Flavones

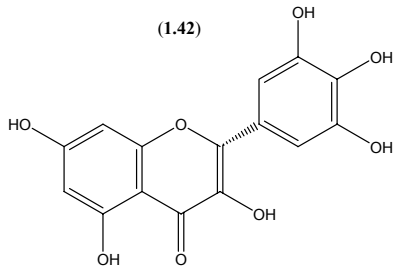
The heterocycle of flavones contains a ketone group, and has an unsaturated carbon-carbon bond. Flavones are common in angiosperms. The most widely distributed flavones in nature are kaemferol (5,7,4′ hydroxyflavone; **1.42**), quercetin (5,7,3′,4′ hydroxyflavone; **1.43**), and myricetin (5,7,3′,4′,5′ hydroxyflavone; **1.44**).



(1.42)



(1.43)

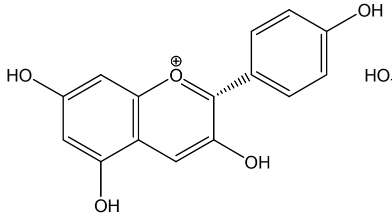


(1.44)

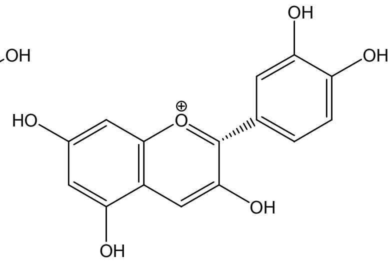
3.6.3.5 Anthocyanidins and deoxyanthocyandins

The heterocycle of anthocyanidins is a pyrilium kation. Anthocyanidins are typically not found as free aglycones, with the excepton of the following widely distributed, colored compounds: Pelargonidin (orange-red; **1.45**), cyanidin (red; **1.46**), peonidin (rose-red; **1.47**), delphinidin (blue-violet; **1.48**), petunidin (blue-purple; **1.49**), and malvidin (purple; **1.50**). A convenient mnemonic is: PCP–DPM. The most common anthocyanidin is cyanidin. These compounds are present in the vacuoles of colored plant tissues such as leaves or flower petals. The color of the pigment depends on

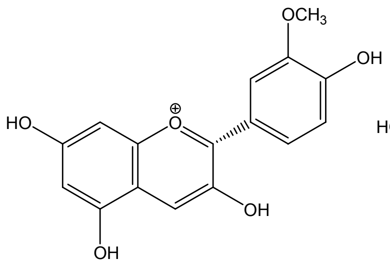
the pH, metal ions present, and the combination of substituted sugars and acylesters. Different colors can also result from the presence of combinations of several anthocyanidins (Figure 1.2).



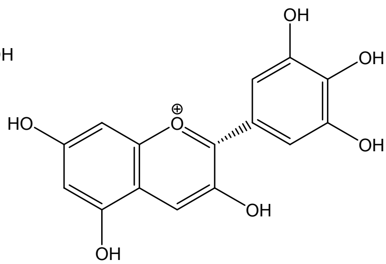
(1.45)



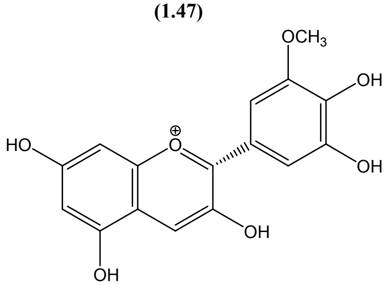
(1.46)



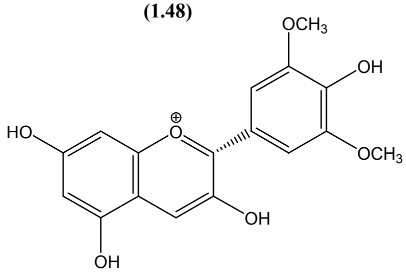
(1.47)



(1.48)



(1.49)



(1.50)

Note that each of the six common anthocyanidins has the basic structure of the flavylium cation (2-phenyl benzopyrilium; **1.31**).

Other anthocyanidins exist, and can be categorized into two groups:

1. Those where either the C5 or C7 position is substituted with a methoxyl group
2. The deoxyanthocyanidins, which do not contain a hydroxyl group at the C3 position

There are five deoxyanthocyanidins: Apigeninidin (**1.51**), luteolinidin (**1.52**), 7-methoxyapigeninidin (**1.53**), 5-methoxy-luteolinidin (**1.54**), and the caffeic acid ester of arabinosyl 5-*O*-apigeninidin (**1.55**).

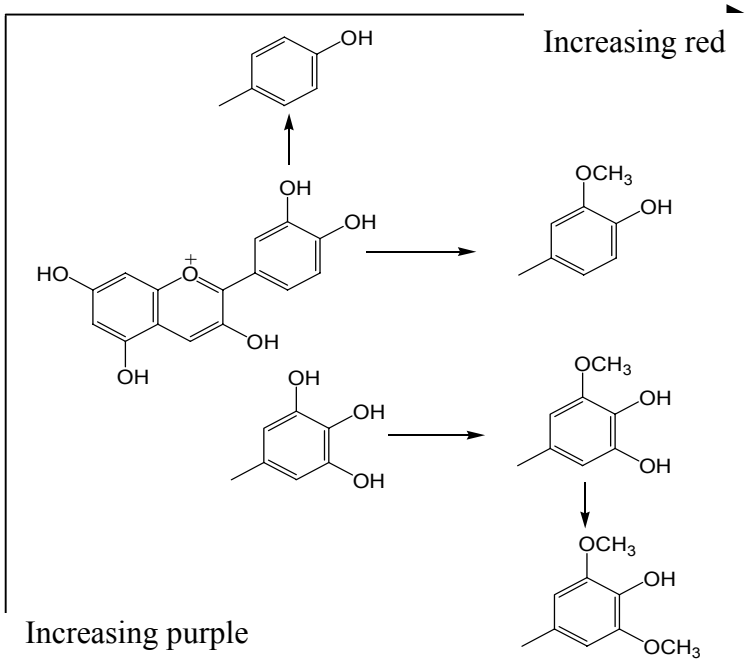
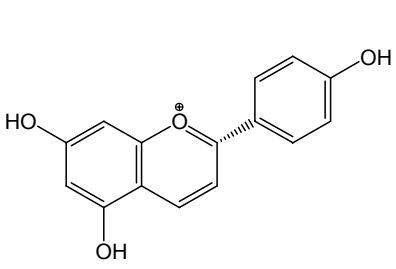
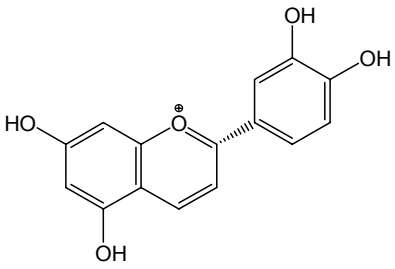


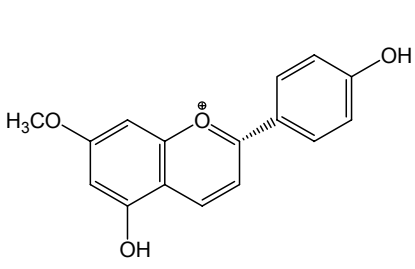
Figure 1-2. Impact of ring substitution on the color of anthocyanidins.



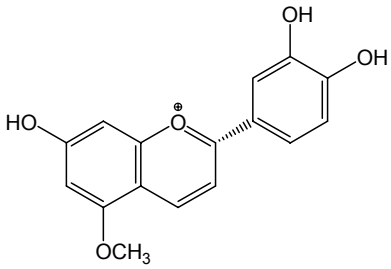
(1.51)



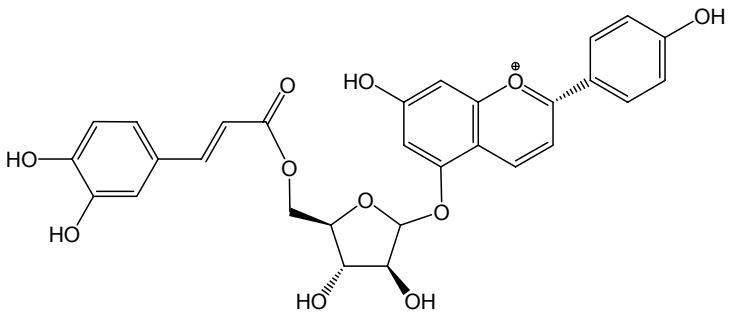
(1.52)



(1.53)



(1.54)



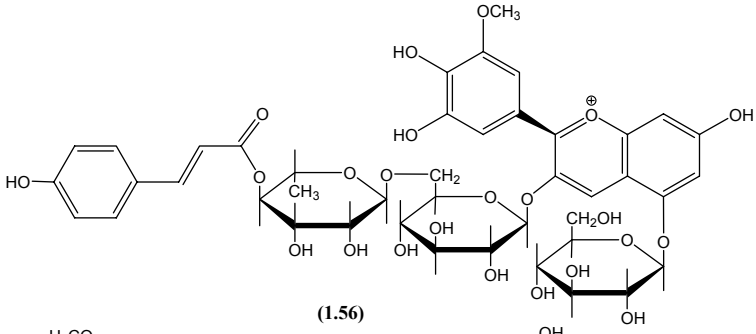
(1.55)

3.6.3.6 Anthocyanins

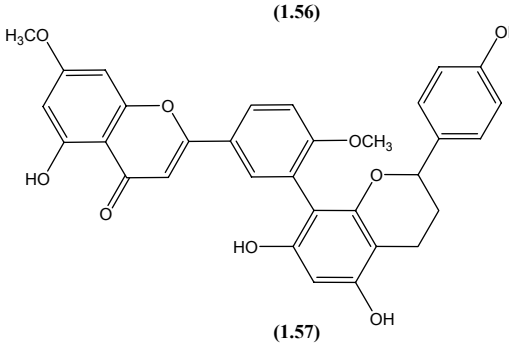
Anthocyanins are water-soluble glycosides of anthocyanidins. The most common glycoside is the 3-glycoside. If a second sugar is present, it is almost always at the 5-hydroxyl position, and almost always a glucose residue. Such compounds are called 3,5-dimonosides. In addition, there are a few rare 3,7-substitutions. While glucose is the most common sugar, substitutions of other sugars, such as arabinose, are sometimes observed. Anthocyanins can also be acylated. In this case an organic acid – typically *p*-coumaric acid (**1.11**), caffeic acid (**1.12**), or ferulic acid (**1.13**) – is esterified to the sugar. An example is petanin (3-[6-*O*-(4-*O*-*E*-*p*-coumaroyl-*O*-α-rhamnopyranosyl)-β-glucopyranoside]-5-*O*-β-glucopyranoside; **1.56**), a compound found in *Solanaceae*.

3.7 Biflavonyls

Biflavonyls have a C_{30} skeleton. They are dimers of flavones such as apigenin or methylated derivatives and are found in gymnosperms. Few compounds are known. The most familiar is ginkgetin (**1.57**) from *Ginkgo biloba* (fossil tree or Japanese silver apricot).



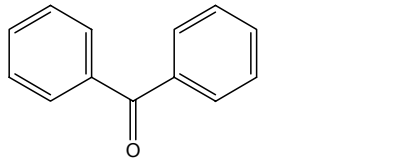
(1.56)



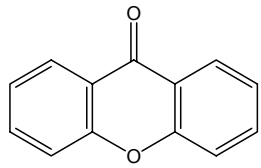
(1.57)

3.8 Benzophenones, xanthones and stilbenes

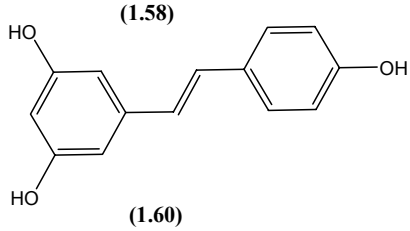
Benzophenones and xanthones have a C_6-C_1-C_6 structure, whereas stilbenes have a C_6-C_2-C_6 structure. Xanthones are yellow pigments in flowers, and stilbenes are associated with heartwood of trees. Shown are the structures of benzophenone, (**1.58**), xanthone (**1.59**), and the stilbenes resveratrol (**1.60**) and pinosylvin (**1.61**).



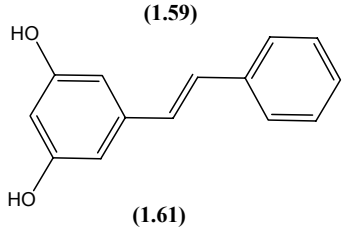
(1.58)



(1.59)



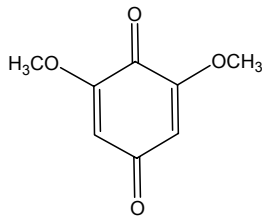
(1.60)



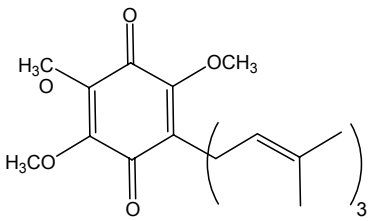
(1.61)

3.9 Benzoquinones, anthraquinones and naphthaquinones

Benzoquinones, such as 2,6-dimethoxybenzoquinone (**1.62**), are present in root exudates of maize and stimulate parasitic plants to form haustoria (Matvienko et al., 2001). Ubiquinones, such as ubiquinone(*3*) (**1.63**), where (*3*) indicates the number of isoprenoid sidechains, is also known as Coenzyme Q and has a role in electron transport in the mitochondria.

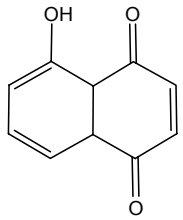


(1.62)

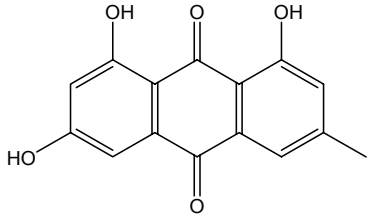


(1.63)

Naphthaquinones are rare. Among the naphthaquinones juglone (**1.64**) is relatively common. It is found in walnuts. Anthraquinone is the most widely distributed of the quinones in higher plants and fungi. There are numerous compounds. The anthtraquinone emodin (**1.65**) occurs as a rhamnoside in rhubarb roots.



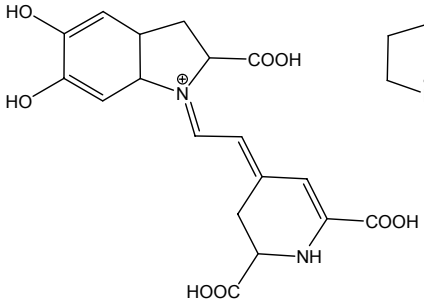
(1.64)



(1.65)

3.10 Betacyanins

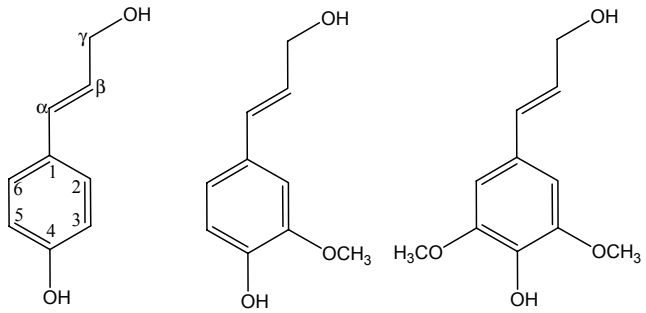
Betacyanins are red pigments and account for the red color of beets (*Beta vulgaris*). They are unique compounds to the Centrospermae. They have absorption spectra that resemble anthocyanins, but they contain nitrogen. An example is betanidin (**1.66**). Betacyanins are normally found as glycosides. Betaxanthins are chemically related to betacyanins, but they are not phenols. An example is indicaxanthin (**1.67**). These compounds are yellow pigments, and are also unique to the Centrospermae.



(**1.66**) (**1.67**)

3.11 Lignans

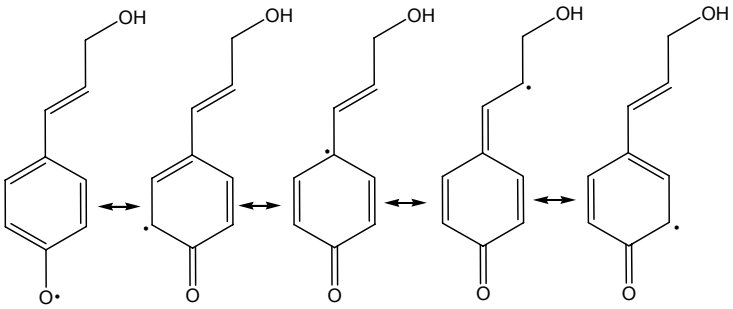
Lignans are dimers or oligomers that result from the coupling of monolignols – *p*-coumaryl alcohol (**1.68**), coniferyl alcohol (**1.69**), and sinapyl alcohol (**1.70**), with coniferyl alcohol being the most common monolignol used in lignan biosynthesis. Lignans are present in ferns, gymnosperms and angiosperms. They are localized in woody stems and in seeds and play a role as insect deterrents. Some of these compounds have medicinal properties.



(1.68) **(1.69)** **(1.70)**

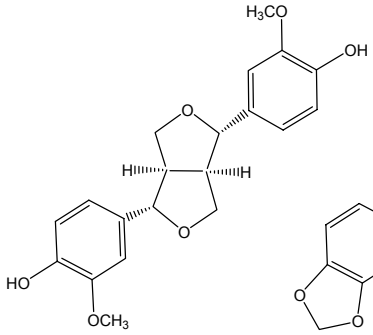
Lignan biosynthesis results from the reaction of monolignol radicals. The monolignol radicals (**1.71**), in this example derived from *p*-coumaryl alcohol (**1.68**), are generated enzymatically by activated cell-wall bound peroxidases (see also Chapter 2, Section 1.8.2.3), which eliminate the proton on the *para*-hydroxyl-group of the phenol. The radical electron can be delocalized along the phenol ring, but also along the propane tail, so that the

carbon at the 1, 3 and 5 positions of the ring, as well as the β-carbon of the propane tail, become reactive.

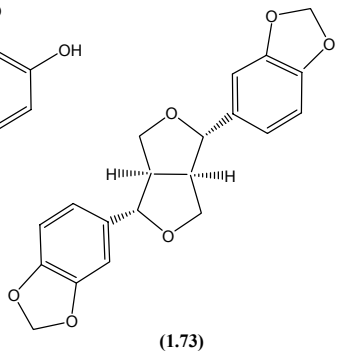


(1.71)

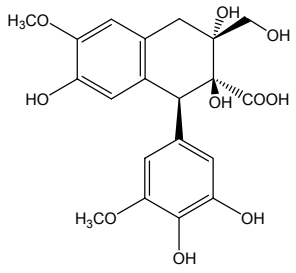
The term lignan typically refers to dimers of monolignols that are linked via an 8-8′ (β-β′) bond, whereas the term neolignan refers to dimers and oligomers that contain bonds other than the 8-8′ bond. Most lignans are optically active, and typically only one enantiomer is found in a given species. Examples of lignans include (+)-pinoresinol (**1.72**), (+)-sesamin (**1.73**), and (–)-plicatic acid (**1.74**).



(1.72)



(1.73)



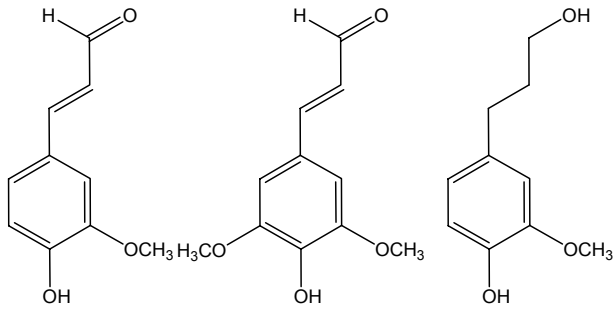
(1.74)

The stereo-selective formation of certain lignans has been shown to be mediated by 'dirigent' proteins. These proteins hold the monolignols in a specific orientation, but have no catalytic activity (Davin et al., 1997; see also Chapter 3).

3.12 Lignin

Lignin is a phenolic polymer. It is the second most abundant bio-polymer on Earth (after cellulose), and plays an important role in providing structural support to plants. Its hydrophobicity also facilitates water transport through the vascular tissue. Finally, the chemical complexity and apparent lack of regularity in its structure make lignin extremely suitable as a physical barrier against insects and fungi.

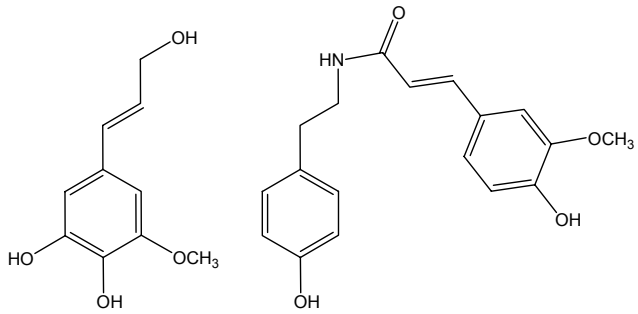
Like lignans, lignin is synthesized primarily from three monolignol precursors: *p*-coumaryl alcohol (**1.68**), coniferyl alcohol (**1.69**), and sinapyl alcohol (**1.70**). Additional compounds are incorporated into the lignin, but typically in small quantities. Some of these compounds include: coniferaldehyde (**1.75**; Pillonel et al., 1991; Halpin et al., 1994; Ralph et al., 2001), sinapaldehyde (**1.76**; Pillonel et al., 1991), dihydroconiferyl alcohol (**1.77**; Ralph et al., 1997), 5-hydroxyconiferyl alcohol (**1.78**; Lapierre et al., 1988; Ralph et al., 2001; Marita et al., 2003), tyramine ferulate (**1.79**; Ralph et al., 1998), *p*-hydroxy-3-methoxybenzaldehyde (**1.80**; Kim et al., 2003), *p*-hydroxybenzoate (**1.81**; Landucci et al., 1992), *p*-coumarate (**1.13**; Lu and Ralph, 1999) and acetate (Ralph, 1996).



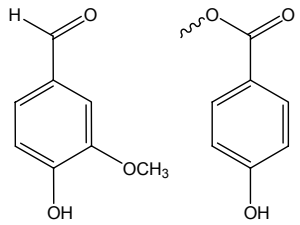
(1.75) **(1.76)** **(1.77)**

The latter three compounds are esterified to the γ-carbon of the monolignols. These compounds are found in higher quantities in certain mutants or genetically engineered plants in which the expression of specific lignin biosynthetic genes has been altered (Sederoff et al., 1999; Boerjan

et al., 2003; Ralph et al., 2004). The presence of these compounds in lignin has prompted a broader definition of lignin, based more on the function than on a narrowly defined chemical composition (Brunow et al., 1999).



(1.78) **(1.79)**

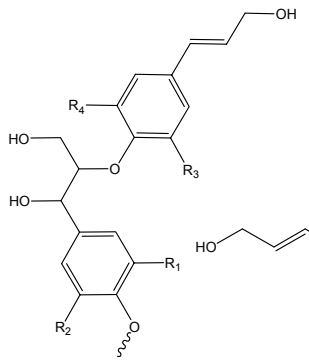


(1.80) **(1.81)**

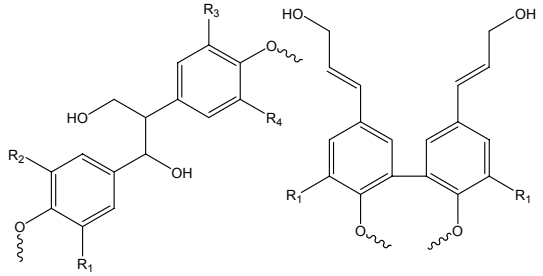
Lignin is formed through a radical-mediated polymerization process, but lignin is not optically active (Ralph et al., 1999), and the structure of lignin is believed to be under chemical control, rather than under the control of dirigent proteins or enzymes (Hatfield and Vermerris, 2001). The lignin polymer enlarges as additional monolignol radicals react with reactive sites on the polymer.

After polymerization the different lignin subunits are referred to as *p*-hyrdoxyphenyl (H), guaiacyl (G), and syringyl (S) residues, depending on whether they originated from *p*-coumaryl alcohol, coniferyl alcohol, or sinapyl alcohol, respectively.

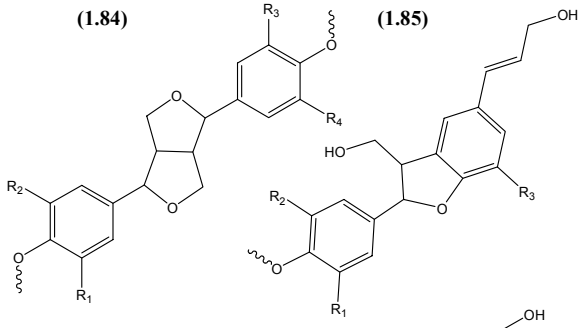
Different kinds of interunit linkages can be formed depending on the position of the delocalized radical electron at the time two radicals are coupled. The most common interunit linkage in lignin is the β-*O*-4 linkage (**1.82**). Other coupling modes include: 5-*O*-4′ (**1.83**), β-1 (**1.84**), 5-5′ (**1.85**), β-β′ (**1.86**), β-5 (**1.87**), and the dibenzodioxocin linkage (**1.88**; Brunow et al., 1998). The interunit linkages involving the β-carbon are favored. The 5-5′



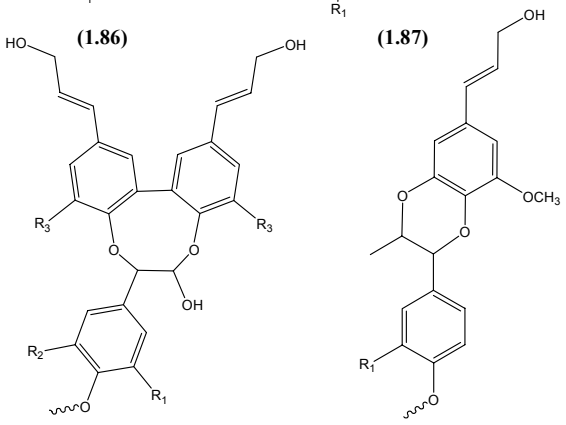
(1.82) **(1.83)**



(1.84) **(1.85)**



(1.86) **(1.87)**



(1.88) **(1.89)**

and the 5-*O*-4′ linkages are present in only small amounts, and tend to originate from preformed oligomers, rather than from the addition of new monolignol radicals to the growing ligninpolymer (Ralph et al., 2004). In plants that accumulate substantial amounts of 5-hydoxyconiferyl alcohol (**1.78**) as a result of reduced activity of the enzyme caffeic acid *O*-methyltransferase (see Chapter 3), the benzodioxane linkage (**1.89**) has been identified. This is a novel linkage between two subunits involving a β-*O*-4′ *and* an α-*O*-5′ bond (Ralph et al., 2001; Marita et al., 2003). The substituents on the phenol ring in structures (**1.82**) through (**1.89**) are indicated with an R. In the case of H-residues both substituents are hydrogens, in the case of G-residues C3 contains a methoxyl group, and in the case of syringyl residues both C3 and C5 contain methoxyl groups.

3.13 Tannins

Tannins comprise a group of compounds with a wide diversity in structure that share their ability to bind and precipitate proteins. The name tannins refers to the process of tanning animal skin to form leather. This process has been known since prehistoric times, when animal hides were treated with animal fat and brain tissue. Chemically this resulted in the cross-linking of the collagen chains in the hide. Throughout much of history the tanning process was performed with tannins derived from plants, until minerals such as aluminum and chromium replaced the use of plant tannins during the last century. As part of Japanese and Chinese natural medicine tannins have been used as anti-inflammatory and antiseptic compounds. They have also been used to treat a wide array of illnesses, including diarrhea and tumors in the stomach or duodenum (Khanbabaee and Van Ree, 2001). Another application of tannins is in wine and beer production, where they are used to precipitate proteins.

Tannins are abundant in many different plant species, in particular oak (*Quercus* spp.), chestnut (*Castanea* spp.), staghorn sumac (*Rhus typhina*), and fringe cups (*Tellima grandiflora*). Tannins can be present in the leaves, bark, and fruits, and are thought to protect the plant against infection and herbivory.

Tannins can be classified in three groups: condensed tannins, hydrolysable tannins, and complex tannins (Khanbabaee and Van Ree, 2001). These groups can then be further divided, as shown in Figure 1-3.

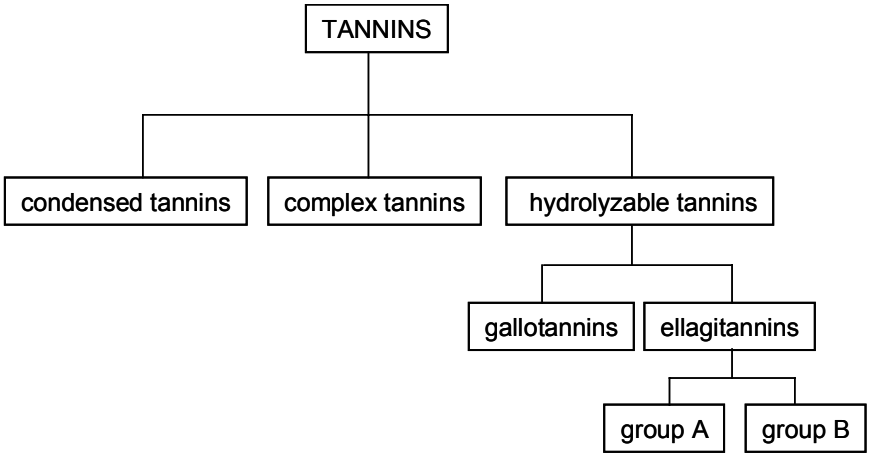
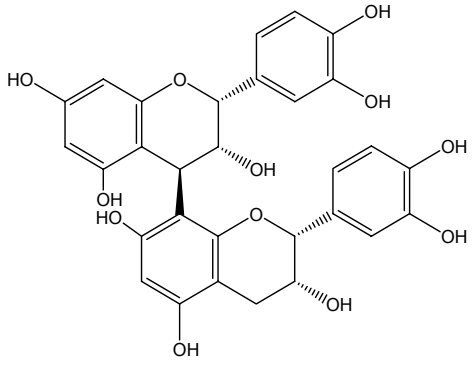


Figure 1-3. Classification of tannins.

3.13.1 Condensed tannins

Condensed tannins are also referred to as proanthocyanidins. They are oligomeric or polymeric flavonoids consisting of flavan-3-ol (catechin) units. Hydrolysis under harsh conditions, such as heating in acid, yields anthocyanidins. An example of a condensed tannin is procyanidin B_2 (epicatechin-(4β→8′)-epicatechin; **1.90**). In this case the interflavanyl linkage is between C4 of the 'lower' unit, and C8 of the 'upper' unit. The linkage can also be between C4 of one unit and C6 of the second unit.



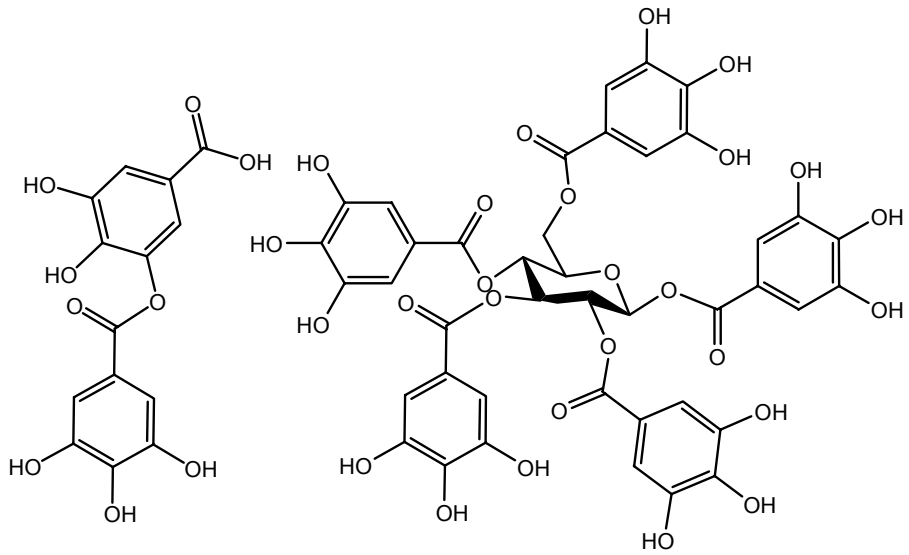
(**1.90**)

Polymers are formed through the action of acids or enzymes. Polymers made up of between two and ten residues are called flavolans. Polymers

made up of more than 50 catechin units have been identified (Khanbabaee and Van Ree, 2001).The degree of polymerization affects the ability to precipitate proteins. This is of importance in wine making, where a high level of condensed tannins, especially in red wines, can result in the dry feeling on the inside of the mouth.

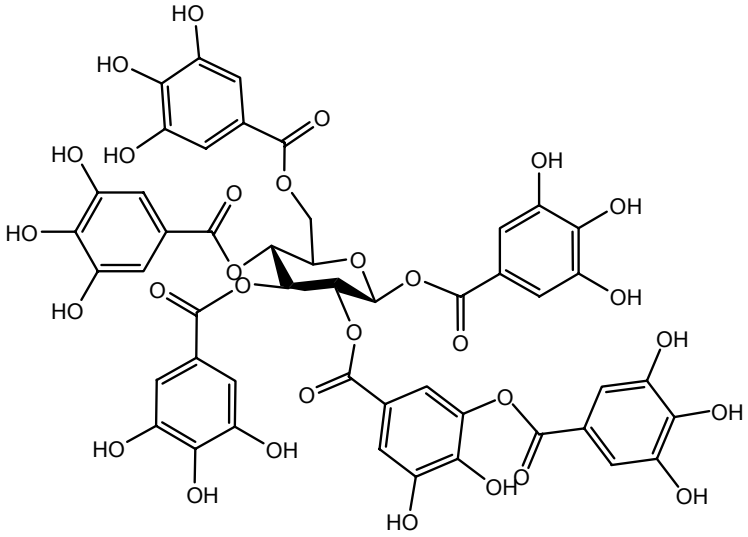
3.13.2 Gallotannins

Gallotannins are hydrolysable tannins with a polyol core (referring to a compound with multiple hydroxyl groups) substituted with 10-12 gallic acid residues. Gallotannins contain the characteristic *meta*-depside bonds (**1.91**) between gallic acid residues. This bond is more labile than an aliphatic ester bond, and can be methanolyzed with a weak acid in methanol. In contrast, methanolysis of an aliphatic ester bond requires methanol with a strong mineral acid and heat.



(1.91) **(1.92)**

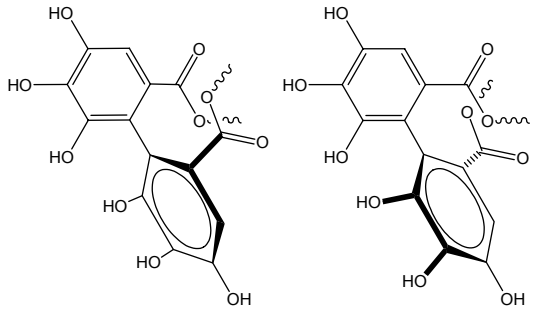
The most commonly found polyol is D-glucose, although some gallotannins contain catechin and triterpenoid units as the core polyol. Gallotannins with a D-glucose core are synthesized from 1,2,3,4, 6-pentagalloylglucose (1,2,3,4,6-penta-*O*-galloyl-β-D-glucopyranose; **1.92**; see also Chapter 3). An example of a gallotannin is the hexagalloylated compound 2-*O*-digalloyl-1,3,4,6-tetra-*O*-galloyl-β-D-glucopyranose (**1.93**), where the additional gallic acid residue is located on C2 of the glucopyranose ring.



(1.93)

3.13.3 Ellagitannins

Ellagitannins are also hydrolysable tannins derived from pentagalloylglucose (**1.92**), but unlike gallotannins, they contain additional C-C bonds between adjacent galloyl moieties in the pentagalloylglucose molecule. This C-C linkage is formed through oxidative coupling between the two adjacent galloyl residues, and results in the formation of a hexahydroxydiphenoyl (HHDP) unit, which can have either the *S*- (**1.94)** or the *R*-configuration (**1.95)**.

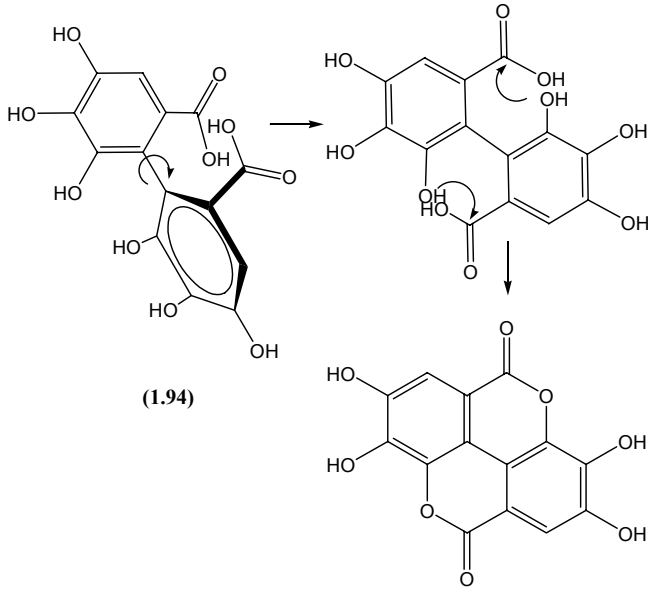


(1.94) **(1.95)**

The chirality is the result of the limited free rotation around the axis of the C-C bond due to the two ester bonds between the galloyl residues and the

polyol (indicated by the wavy bonds in **1.94** and **1.95**), combined with the presence of the *ortho*-substituents that create steric hindrance.

The name ellagitannins is derived from ellagic acid (**1.96**), which is formed spontaneously from hexahydroxydiphenic acid (**1.94/1.95**) in aqueous solution *via* an *intra*-molecular esterification reaction.



(1.94)

(1.96)

With the glucopyranose molecule in the $^4C_1$ conformation (a chair conformation with C4 above the plan, as shown in (**1.92**)), the most common linkages are between galloyl residues at the 2- and 3-positions of the glucopyranose ring, and/or between those at the 4- and 6-positions. These are referred to as *Group A* ellagitannins. In addition, ellagitannins with the less common 3,4-linkage, such as identified in the compound cercidinin A from the bark of *Cercidiphyllum japonicum* are included in this group (Tanaka et al., 2001). *Group B* ellagitannins have a glucopyranose molecule in the energetically less favorable $^1C_4$ chair or boat conformations. In this case 1,6-, 1,3-, 2,4- or 3,6-C-C linkages can be formed between galloyl residues. Figure 1-4 displays the possible different configurations. It should be noted that it is possible to have combinations of different linkages, such as, for example, 3,6- *and* 2,4-linkages.

The configuration of the HHDP esters (*R* or *S*; as shown in Figure 1-4) varies depending on the position of the HHDP unit, and is in the

energetically most favorable configuration. This is dictated by the stereochemistry of the sugar molecule (Haslam and Cai, 1993). Aside from the glucopyranose ring, there are also ellagitannins consisting of an open-chain polyol. The open chain ellagitannins identified so far all contain a 2,3-linked HHDP unit (Haslam and Cai, 1993).

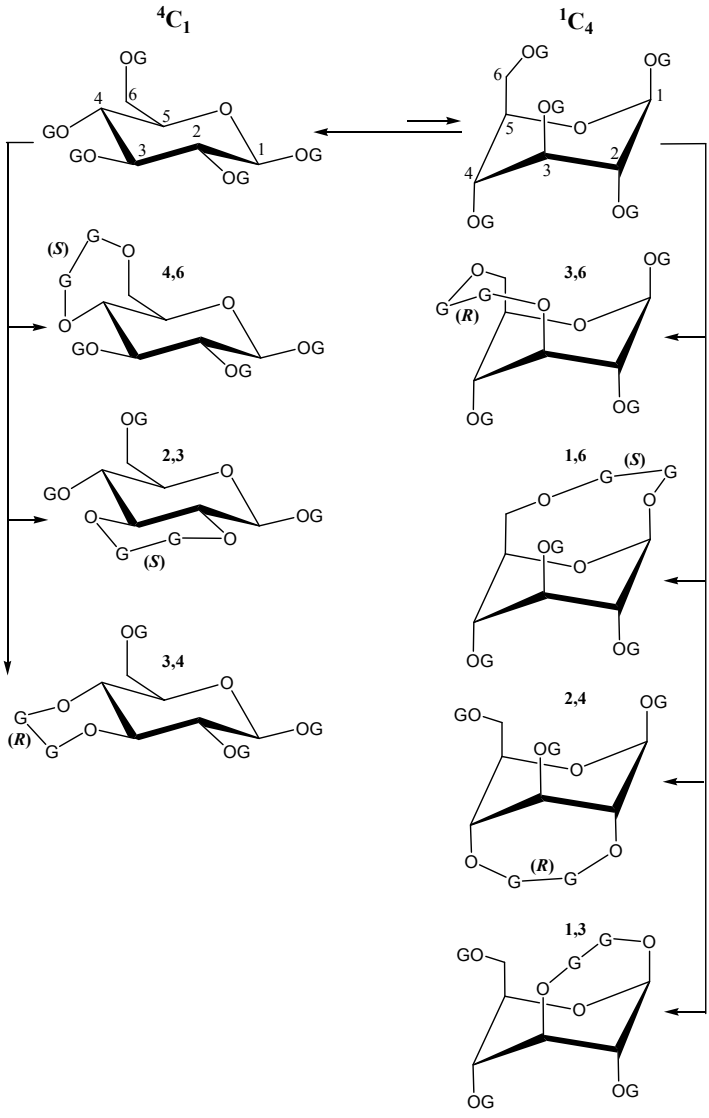
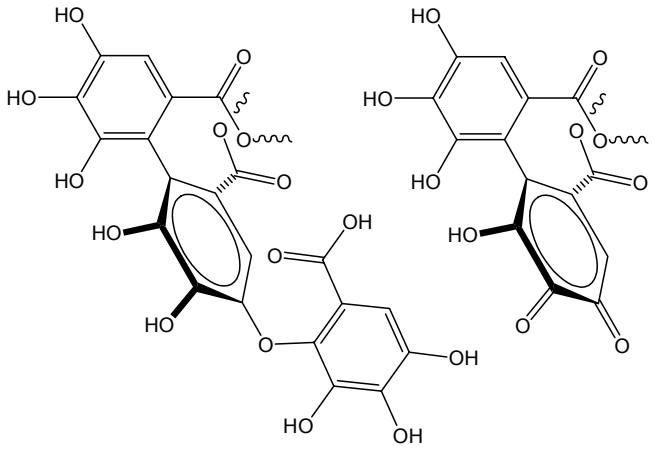


Figure 1-4. Possible linkages between adjacent galloyl residues in D-glucose-based ellagitannins. Ellagitannins from Group A are shown on the left, from Group B on the right. The (*S*) and (*R*) refer to the conformation of the HHDP-units.

Further modification of the HHDP unit is possible. Adjacent galloyl residues can undergo further oxidative coupling with participation of the aromatic hydroxyl groups. The valoneoyl unit (**1.97**), for example, arises from the linkage between a galloyl residue with an HHDP-unit. Furthermore, the *meta*- and *para*-hydroxyl groups on one of the galloyl-moieties can be oxidized, resulting in a dehydro-HHDP unit (**1.98**).



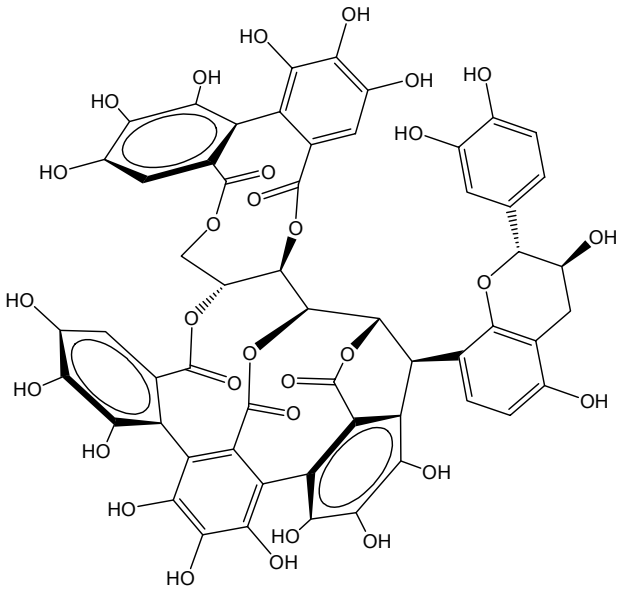
(1.97) **(1.98)**

3.13.4 Complex tannins

Complex tannins are defined as tannins in which a catechin unit (**1.39**) is bound glycosidically to either a gallotannin or an ellagitannin unit. As the name implies, the structure of these compounds can be very complex. An example is Acutissimin A (**1.99**). This is a flavogallonyl unit bound glucosidically to C1, with an additional three hydrolyzable ester bonds to a D-glucose-derived open-chain polyol.

This complex tannin is formed during the aging process of red wine, whereby the catechin unit originates from the grapes, and the ellagitannin, in this case vescalagin, originates from the oak barrels.

Acutissimin A has been shown to be a powerful inhibitor of DNA topo-isomerase II, an enzyme required for the division of cancer cells, and a target for chemotherapeutic drugs (Quideau et al., 2003). Based on these findings, however, it is an overstatement to consider red wine a cancer preventative. Red wine contains other compounds with medicinal activity which will be discussed in more detail in Chapter 7.



(1.99)

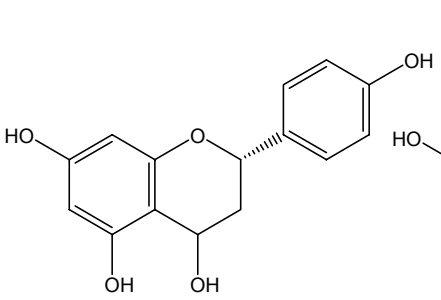
3.14 Phlobaphenes

Phlobaphenes are phenolic polymers that can be present in floral organs of maize (*Zea mays* L.), including the pericarp (the hard, outermost layer of the kernel, derived from the ovary wall), the cob, the husks (the leaves covering the ear), the tassel glumes, the cob pith, and the tassel pith. Accumulation of phlobaphenes results in a red pigmentation (Styles and Ceska, 1989). Certain lines of sorghum (*Sorghum bicolor* L. (Moench)) also produce phlobaphenes (Boddu et al., 2005).

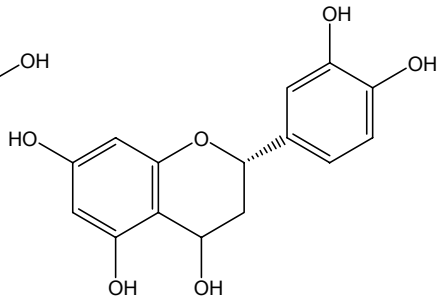
The structure of phlobaphenes is poorly understood. These compounds are believed to be polymers of flavan-4-ols, notably apiferol (**1.100**) and luteoferol (**1.101**) (Shirley-Winkel, 2001). Both of these monomers are derived from naringenin (**1.35**).

The polymerization is thought to be under chemical, rather than enzymatic control, and give rise to a polymer (**1.102**) in which the monomers are linked via a 4-8′ linkage. The C-C bonds between the flavan-4-ol monomers would be difficult to break, which could help explain the difficulties with the structural elucidation of phlobaphenes. The reaction mechanism whereby the hydroxyl group on C-4 is eliminated and a C-C

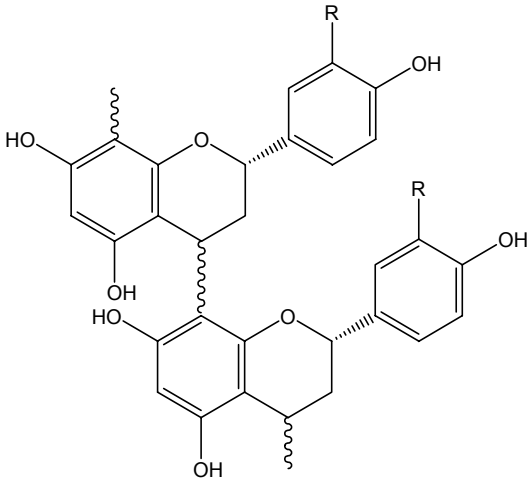
bond between two monomers is formed, is however, not easily imagined, and it is possible that the structure of phlobaphenes is different than is currently postulated.



(1.100)



(1.101)



(1.102)

4. REFERENCES

Boddu, J., Svabek, C., Ibraheem, F., Jones, D., and Chopra, S., 2005, Characterization of a deletion allele of a sorghum Myb gene *yellow seed1* showing loss of 3-deoxyflavonoids, *Plant Sci.* **169**: 542-552.

Boerjan, W., Ralph, J., and Baucher, M., 2003, Lignin biosynthesis. *Ann. Rev. Plant Biol*. **54**: 519-546.

Brunow, G., Lundquist, K. and Gellerstedt, G., 1999, Lignin. *In*: Sjöström, E. and Alén R., (eds.), *Analytical Methods in Wood Chemistry, Pulping, and Papermaking*, Springer-Verlag, Germany, pp. 77–124.

Gallagher, S. R., 1992, *GUS Protocols: Using the GUS Gene as a Reporter of Gene Expression*, Academic Press, San Diego.

Halpin, C., Knight, M. E., Foxon, G. A., Campbell, M. M., Boudet, A. M., Boon, J. J., Chabbert, B., Tollier, M.-T., and Schuch, W., 1994, Manipulation of lignin quality by downregulation of cinnamyl alcohol dehydrogenase, *Plant J.* **6**: 339-350.

Haslam, E., and Cai, Y., 1994, Plant polyphenols (vegetable tannins) : gallic acid metabolism, *Nat. Prod. Rep*. **11**: 41-65.

Hatfield, R. and Vermerris, W., 2001, Lignin formation in plants: the dilemma of linkage specificity, *Plant Physiol*. **126**: 1351-1357.

Harborne, J. B. and Simmonds, N. W., 1964, *Biochemistry of Phenolic Compounds*, Academic Press, London, pp. 101.

Jacks, T. J., and Kircher, H. W., 1967, Fluorometric assay for hydrolytic activity of lipase using fatty acyl esters of 4-methylumbelliferone *Anal. Biochem*. **21**: 279-285.

Khanbabaee, K., and Van Ree, T., 2001, Tannins: Classification and definition, *Nat. Prod. Rep*. **18**: 641-649.

Kim, H., Ralph, J., Lu, F., Ralph, S. A., Boudet, A-M., MacKay, J. J., Sederoff, R. R., Ito, T., Kawai, S., Ohashi, H. and Higuchi, T. (2003) NMR analysis of lignins in CAD-deficient plants. Part 1. Incorporation of hydroxycinn-amaldehydes and hydroxybenzaldehydes into lignins, *Org. Biomol. Chem*. **1**: 158-281.

Landucci, L. L., Deka, G. C. and Roy, D. N., 1992, A 13C NMR study of milled wood lignins from hybrid *Salix* clones, *Holzforsch*. **46**: 505-511.

Lapierre, C., Tollier, M. T., and Monties, B., 1988, Mise en évidence d'un nouveau type d'unité constitutive dans les lignines d'un mutant de maïs *bm3*, *C. R. Acad. Sci. Paris, Ser. III* **307**: 723-728.

Lu, F., and Ralph, J., 1999, Detection and determination of *p*-coumaroylated units in lignin. *J. Agric. Food Chem.* **47**: 1988-1992.

Marita, J., Vermerris, W., Ralph, J., and Hatfield, R. D., 2003, Variation in the cell wall composition of maize *brown midrib* mutants, *J. Agric. Food Chem.* **51**: 1313-1321.

Matvienko, M., Torres, M. J., and Yoder, J. I., 2001, Transcriptional responses in the hemiparasitic plant *Triphysaria versicolor* to host plant signals, *Plant Physiol*. **127**: 272-282.

Nelson, J. A., and Falk, R. E., 1993, The efficacy of phloridzin and phloretin on tumor cell growth, *Anticancer Res.* **13**: 2287-2292.

Niemetz, R., and Gross, G. G., 2005, Enzymology of gallotannin and ellagitannin biosynthesis, *Phytochem.* **66**: 2001-2011.

Pillonel, C., Mulder, M. M., Boon, J. J., Forster, B., and Binder, A., 1991, Involvement of cinnamyl-alcohol dehydrogenase in the control of lignin formation in *Sorghum bicolor* L. Moench, *Planta* **185**: 538-544.

Prithiviraj, B., Singh, U. P., Manickam, M., Srivastava, J. S., and Ray A. B., 1997, Antifungal activity of bergenin, a constituent of *Flueggea microcarpa*, *Plant Path*. **46**: 224-228.

Quideau, S., Jourdes, M., Saucier, C., Glories, Y., Pardon, P., Baudry, C., 2003, DNA Topoisomerase inhibitor acutissimin A and other flavano-ellagitannins in red wine, *Angew. Chem*. **115**: 6092-6099.

Ribéreau-Gayon, P., 1972, *Plant Phenolics*, Oliver and Boyd, Edinburgh, pp. 254.

Ralph, J., 1996, An unusual lignin from Kenaf, *J. Nat. Prod.* **59**: 341-342.

Ralph, J., Hatfield, R. D., Piquemal, J., Yahiaoui, N., Pean, M., Lapierre, C., Boudet, A. M., 1998, NMR characterization of altered lignins extracted from tobacco plants down-regulated for lignification enzymes cinnamyl alcohol dehydrogenase and cinnamyl-CoA reductase, *Proc. Natl. Acad. Sci. USA* **95**: 12803-12808.

Ralph, J., Lapierre, C., Marita, J. M., Kim, H., Lu, F., Hatfield, R. D., Ralph, S., Chapple, C., Franke, R., Hemm, M. R., Van Doorsselaere, J., Sederoff, R. R., O'Malley, D. M., Scott, J. T., Mackay, J. J., Yahiaoui, N., Boudet, A. M., Pean, M., Pilate, G., Jouanin, L., and Boerjan, W., 2001, Elucidation of new structures in lignins of CAD- and COMT-deficient plants by NMR. *Phytochem.* **57**: 993-1003.

Ralph, J., Lundquist, K., Brunow, G., Lu, F., Kim, H., Schatz, P. F., Marita, J. M., Hatfield, R. D., Ralph, S. A., Christensen, J. H., Boerjan, W., 2004, Lignins: Natural polymers from oxidative coupling of 4-hydroxyphenylpropanoids, *Phytochem. Revs*. **3**: 29-60.

Ralph, J., MacKay, J. J., Hatfield, R. D., O'Malley, D. M., Whetten, R. W., and Sederoff, R. R., 1997, Abnormal lignin in a loblolly pine mutant, *Science* **277**: 235-238.

Ruegger, M., and Chapple, C., 2001, Mutations that reduce sinapoylmalate accumulation in *Arabidopsis thaliana* define loci with diverse roles in phenylpropanoid metabolism, *Genetics* **159**: 1741-1749.

Sederoff, R. R., Mackay, J. J., Ralph, J., and Hatfield, R. D., 1999, Unexpected variation in lignin, *Curr. Opin. Plant Biol.* **2**: 145-152.

Styles, E. D., and Ceska, O., 1989, Pericarp flavonoids in genetic strains of

Zea mays, *Maydica*, **34**: 227-237.
Swain, T. and Bate-Smith, E. C., 1962, Flavonoid compounds, in: *Comparative Biochemistry Vol. III.,* M. Florkin and H. S. Mason, eds., Academic Press, New York, NY, pp. 755-809.
Tanaka, T., Nonaka, G., Ishimatsu, M., Nishioka, I., and Kouno, I., 2001, Revised structure of cercidinin A, a novel ellagitannin having (*R*)-hexahydroxydiphenoyl esters at the 3,4-positions of glucopyranose, *Chem. Pharm. Bull.* **49**: 486-487.
Winkel-Shirley, B., 2001, Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology, *Plant Physiol*. **126**: 485-493.

Suggested titles for further reading

Grotewold, E., 2005, *The Science of Flavonoids*, Springer, New York, NY, pp. 274.

Chapter 2

CHEMICAL PROPERTIES OF PHENOLIC COMPOUNDS

1.1 The benzene ring

The atomic number of carbon is six, which means that the carbon atom has six protons and six electrons. Electrons are present in orbitals around the atom's nucleus. An orbital is defined as the space around the nucleus where an electron with certain energy is most likely to be found. There are typically several different orbitals around a nucleus, representing different energy levels and different spatial distributions.

The electron configuration of carbon is $1s^2\ 2s^2\ 2p^2$, meaning that the 1s, 2s and 2p orbitals each contain two electrons. Covalent bonds between atoms result from the sharing of one or more pairs of electrons. This is the result of overlapping *atomic* orbitals that form *molecular* orbitals. Carbon is able to make four covalent bonds, even though based on its electron configuration only two unpaired electrons – those in the 2p-orbitals - are available. The ability to make four instead of two covalent bonds results from *hybridization*, whereby one of the 2s electrons is promoted to the 2p orbital (electron configuration: $1s^2\ 2s^1\ 2p^3$), followed by the formation of four equivalent sp^3 orbitals with slightly higher energy than the 2s orbital. The promotion of one of the 2s-electrons to the 2p-orbital can also result in the formation of three sp^2 and one 2p orbitals (sp^2 hybridization), or in two sp orbitals and two 2p orbitals (sp hybridization). The angle between these different orbitals is always such that the distance between the orbitals is maximal. Thus, the four sp^3 orbitals make an angle of 109.5°. This can be visualized as a regular tetrahedron, with the C-atom in the center and the

orbitals pointing into the corners. The three sp^2 orbitals make an angle of 120° with each other, and lie in a plane. The two sp orbitals make an angle of 180° with each other.

If the bond formed between two adjacent atoms is symmetric along the axis between the two nuclei, we refer to the bond as a σ-bond. In contrast, the bond that is formed between two electrons in the p-orbitals is called a π-bond. These molecular orbitals have a very different shape. The single bond between two carbon atoms (-C-C-) is formed between electrons in the sp^3-orbitals (σ-bond) of two adjacent carbon atoms (Figure 1a). A double bond between two C-atoms (-C=C-) consists of a σ-bond formed between electrons in the sp^2 orbitals of two adjacent C-atoms, and a π-bond formed between two 2p-orbitals (Figure 2-1).

Figure 2-1. Formation of a σ and a π-bond between two sp^2 hybridized carbon atoms. The p-orbital has two lobes, one above and one below the C-atom.

A single bond has a length of 1.54 Å (1Å = 1Ångstrom = 10^{-10} m). A double bond is 1.34 Å. In *conjugated* molecules there is an alternation between single and double bonds, such that the π-electrons can be shared by all C-atoms that are part of the conjugated system. These electrons are referred to as delocalized electrons, because they are less confined to the axis of the bond. A conjugated molecule is considered *aromatic* if it contains a cyclic π-system with $(4n+2)$ π-electrons (n = 1, 2, 3, …). Hence, benzene is the simplest aromatic compound (n = 1). The π-electrons of benzene are present in a molecular orbital that lies above and below the plane formed by the C-atoms (Figure 2-1). The bond between the C-atoms is 1.39 Å, which lies in between the length of the single and double bond.

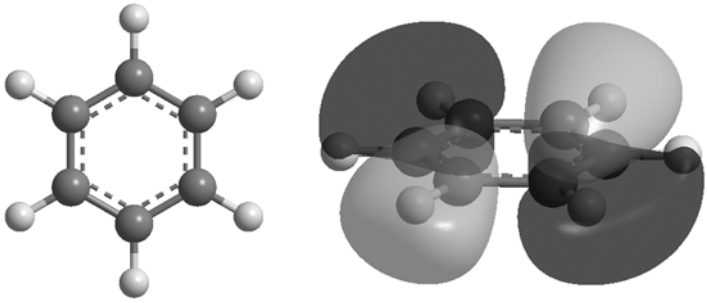
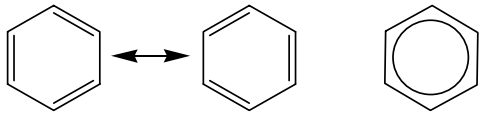


Figure 2-2. Spatial representation (ball-and-stick model) of benzene, with C-atoms in grey and H-atoms in white. The dotted lines between the C-atoms represent the delocalized electrons. The image on the right shows the surface area of the highest occupied molecular orbital (HOMO). Note how the π-electrons are above and below the benzene ring.

The benzene ring is usually depicted as one of two *mesomeric* structures (**2.1**). The double arrow indicates that the true structure of the molecule lies somewhere in between the two drawn structures. It is therefore more accurate to use structure **2.2**, since the six C-C bonds of the ring are identical, with the π-electrons delocalized over the entire ring. The configuration shown in **2.2** is, however, less convenient for drawing reaction mechanisms.

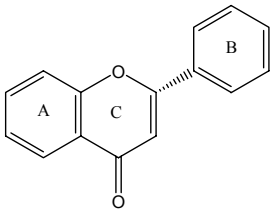


(2.1) **(2.2)**

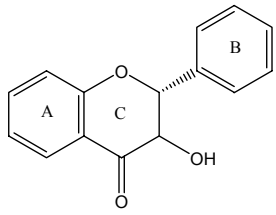
The delocalization of the π-electrons is energetically favorable, and this affects the reactivity of aromatic compounds: There is a tendency towards restoring aromaticity. This is why aromatic compounds, in contrast to regular alkenes (linear chains of carbon atoms containing at least one double bond), do not easily undergo *addition* reactions, whereby a double bond is replaced by two single bonds. Aromatic compounds show a preference for *substitution* reactions, which means that atoms are replaced.

1.2 Planar versus non-planar bonds

The three-dimensional structure of molecules is important, because it affects both chemical reactivity and biological activity. The number of bonds the carbon atom is involved in is important in this respect, because it affects whether the molecule is planar (*i.e.* flat) or non-planar. Since the angle between the sp^3 orbitals is 109.5°, alkanes (a chain of carbon atoms connected via single bonds) cannot lie in a flat plane. A double bond, however, confines the neighboring single bonds to a plane. Hence, because of the conjugated structure, the benzene ring is planar.



(2.3)



(2.4)

The spatial structure of flavone (**2.3**) and flavanonol (**2.4**) is very different. The presence of the double bond in the C-ring of flavone results in the A- and C-rings being planar, whereas the C-ring of flavanonol is not planar.

1.3 The acidic nature of the phenolic hydroxyl group

Since phenol is benzene with a hydroxyl group, the reactivity of phenol and phenolic compounds is in many ways dictated by the chemical properties of the benzene ring. The first property to consider is acidity. A compound is considered an acid when it can release a proton (H^+) while in solution. The acid constant K_a of a compound defines to what extent the proton is released. Strong acids will completely dissociate, whereas weak acids (HA) are at equilibrium with their dissociated state:

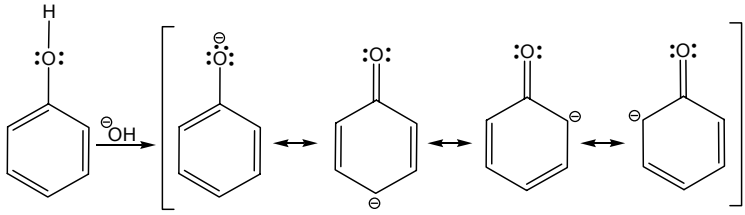
$$HA \rightleftharpoons A^- + H^+$$

The K_a is defined as:

$$K_a = \frac{[H^+]\,[A^-]}{[HA]}$$

The pK_a is the negative ^{10}log of K_a, which tends to result in more convenient numbers.

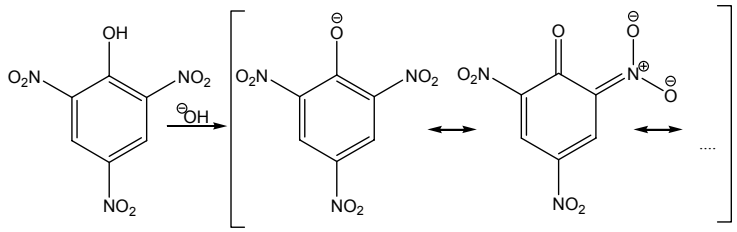
Phenolic compounds are, in general, weak acids. Compared to the hydroxyl group of unsubstituted aliphatic alcohols, however, the phenolic OH-group is more acidic. The reason for this is that the anion formed after abstracting the proton from the hydroxyl group is relatively stable because of the existence of several mesomeric structures. The anion is referred to as the phenolate anion. Hence, phenol (**2.5**) is a weak acid, with a pK_a value of 10. This places phenol in between carboxylic acids (pK_a = 4-5) and aliphatic alcohols (pK_a = 16-19).



(2.5)

1.3.1 The effects of substituents on acidity

The actual pK_a of phenolic compounds depends upon the overall structure of the molecule and the nature of the substituents on the aromatic ring(s). The compound 2,4,6-trinitrophenol (**2.6**) has a pK_a value 0.71, equivalent to that of a strong acid. This low pK_a-value is the result of the electron-pulling NO_2 substituents. The presence of these groups stabilizes the anion that results after the proton is removed.



(2.6)

1.3.2 Use of acidity in separations

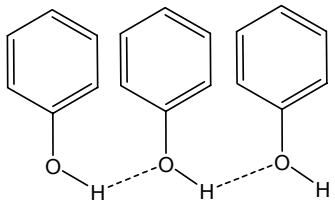
The pK_a is important since it affects the way certain phenolic compounds are extracted. If we consider having a mixture of phenols that range from weak acids to strong acids, the addition of sodium carbonate (Na_2CO_3) or sodium bicarbonate ($NaHCO_3$) to the mixture will allow separation of these phenols. The weak base picks up the H^+ from the strong acids or the more acidic phenols. This results in formation of phenolate salts of the phenols that are soluble in water.

The less acidic phenols are not neutralized or do not lose a H^+ and tend to remain as "free phenols" and will not be as soluble in water. These compounds can then be extracted with organic solvents.

1.4 Hydrogen bonding and the phenolic hydroxyl group

The hydrogen bond is an electrostatic interaction between a hydrogen atom bound to an electronegative atom such as oxygen, fluorine or nitrogen, and the free electrons of other atoms. An actual covalent bond is not possible, because that would result in the presence of more than two electrons in the orbital around the hydrogen atom.

The proton of the hydroxyl group of phenol is an ideal candidate for hydrogen bonding. Shown below (**2.7**) are three phenol molecules, with the hydrogen bonds indicated by dotted lines.

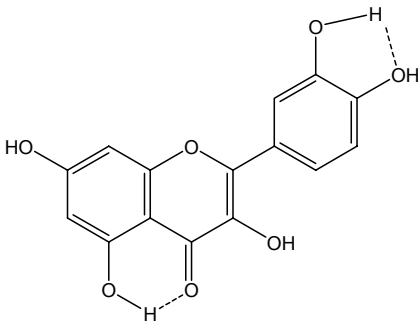


(2.7)

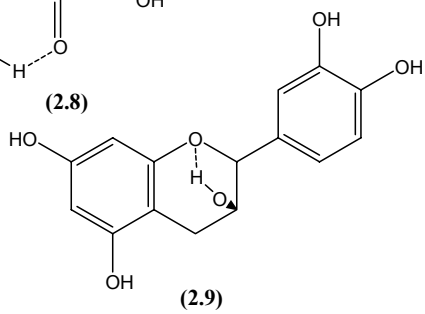
The hydrogen bond is weaker than a shared covalent bond, and the distance between the hydrogen nucleus and the oxygen nucleus is approximately 1.5 times larger than that of the covalent bond. The presence of hydrogen bonds raises the melting and boiling points of compounds, because more energy is required to break intermolecular bonds. The presence of hydrogen bonds can alter the UV and IR spectra of a given compound.

1.4.1 Intra- and inter-molecular hydrogen bonds

Phenolic compounds may form both inter- and intra-molecular hydrogen bonds, referring to bonds that are formed between or within molecules, respectively. Intra-molecular hydrogen bonds are common between adjacent hydroxyl groups (*ortho*-substitution), or groups in the *ortho*-position relative to a carbonyl group. An example is quercetin (**2.8**). The B-ring of flavonoids is more stable with respect to hydroxyl groups. The hydroxyl groups on the B-ring are placed in either the *ortho*- or the tri-*vic* configuration, both of which allow hydrogen bonding. The hydroxyl groups on the A-ring are typically in the *meta*-position, which precludes hydrogen bonds from forming. Catechin (**2.9**) may form a hydrogen bond between the hydroxyl group and the oxygen of the heterocycle.



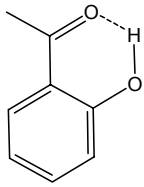
(2.8)



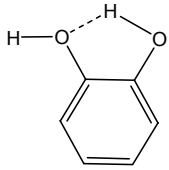
(2.9)

1.4.2 Stability of the hydrogen bond ring

*Intra*molecular hydrogen bonding is considered to reduce the reactivity of the phenolic hydroxyl group. Thus, it reduces solubility in alcohol, and may reduce the ability to form esters and ethers. Note that the hydrogen bond results in the formation of a ring. This ring also has a level of stability. The six-member rings are more stable and stronger than the five-member rings. Compare for example *O*-hydroxyacetophenone (six-member ring; **2.10**) to catechol (a five-member ring; **2.11**).

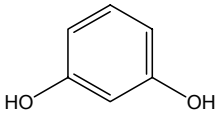


(2.10)

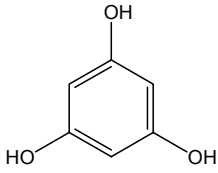


(2.11)

*Inter*molecular hydrogen bonds raise melting points and solubility. Table 2.1 lists physical properties of phenol (**2.5**), resorcinol (**2.12**) and phloroglucinol (**2.13**) that are influenced by hydrogen bonds.



(2.12)



(2.13)

Table 2-1. Impact of substituents on solubility

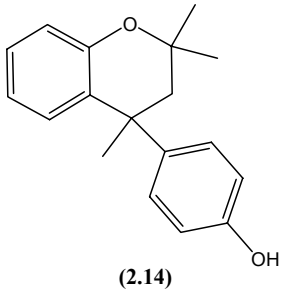
	phenol	resorcinol	phloroglucinol
Melting point (°C)	41	118	218
Solubility in ethanol	very good	1 g/0.9 ml	1 g/12 ml
Solubility in water	1 g/15 ml	1 g/0.9 ml	1 g/100 ml

Intermolecular hydrogen bonds make it difficult to purify phenolic compounds from mixtures, because of the interactions between different molecules, including the solvent.

Hydrogen bond formation in phenolic compounds can be summarized with the following general rules:

1. Unless they are sterically hindered, all phenolic compounds take part in hydrogen bonding.
2. Intramolecular hydrogen bonds are less stable than intermolecular hydrogen bonds. The formation of intramolecular hydrogen bonds diminishes reactivity, whereas the formation of intermolecular hydrogen bonds can complicate purification.

3. Phenolic compounds that form intermolecular hydrogen bonds are typically solid at room temperature.
4. Ring structures can be formed as a result of hydrogen bonds. Six-member rings are more stable than five-member rings.
5. Some phenolic compounds can form flat hexagonal structures with the aromatic rings facing outward, and linked by hydrogen bonds. The internal space thus formed contains solvent. Such compounds are called inclusion compounds or clathrates. An example is dianin (**2.14**), which can form clathrates in more than 50 different solvents.



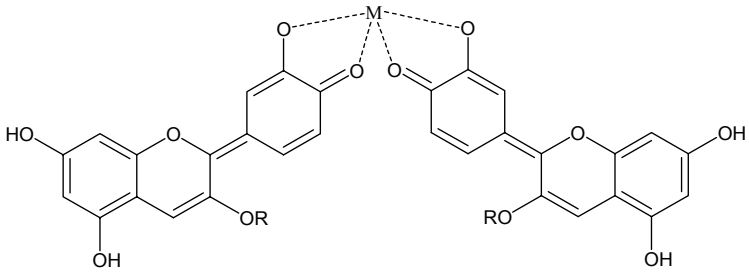
(2.14)

1.5 Metal complexes

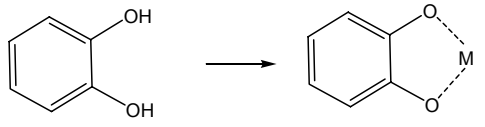
Metal complexes of phenols are important in nature and useful in the laboratory. The metals involved usually include iron, aluminum and magnesium. In nature the flavonoids account for most red, blue, and violet - and to some extent yellow - colors. The majority of yellow colors are the result of the presence of carotenoids and aurones.

The precise color of anthocyanins depends on the substitution pattern, the pH (red in acid, blue in base), but also on the formation of complexes formed with iron, aluminum and magnesium ions. Such complexes can be rather large, and chemically diverse. An example is the blue pigment protocyanin (**2.15**) from cornflower. This is a cyanidin 3,5-diglucoside complex of two molecules of cyanidin linked through their *o*-diphenol groups with one Al^{3+} or Fe^{3+} ion. Cyanocentaurin is even more complex: four molecules of cyanidin 3,5-diglucoside, iron, and three molecules of biflavone glycoside.

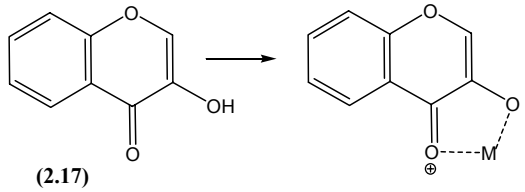
Several structures are capable of forming metal complexes: *o*-dihydroxy phenols (**2.16**), 3-hydroxychromones (**2.17**), 5-hydroxychromones (**2.18**), and *o*-hydroxycarbonyls (**2.19**).



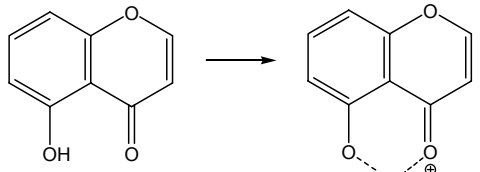
(2.15)



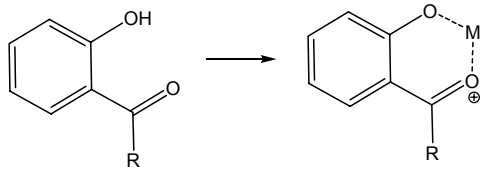
(2.16)



(2.17)



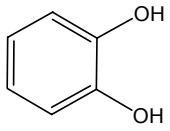
(2.18)



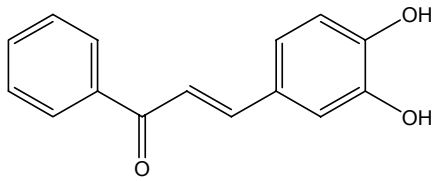
(2.19)

The overall structure of the molecule determines the reactivity of the molecule with the metal, and the presence of the metal ion will impact the chemical properties of the complex. The degree to which the chemical

properties are altered as a result of complex formation depends on the structure of the phenolic compound. For example, aluminum chloride has less effect on the absorption spectrum of catechol (**2.20**) than on that of 3,4-dihydroxychalcone (**2.21**).



(2.20)

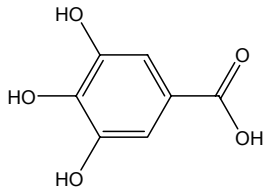


(2.21)

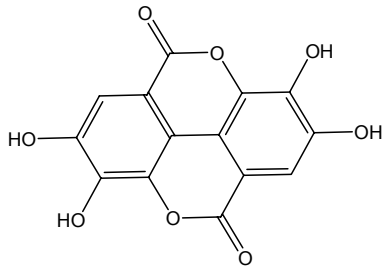
Metal complexes are used for compound identification. They can shift or change absorption spectra, change the R_f of compounds in thin layer chromatography, and change visual colors used in chromatography.

1.6 Esterification

Esters (RCOOR) are formed by reaction of a carboxylic acid with the hydroxyl group of an alcohol. The hydroxyl group of phenolic compounds can participate in ester formation. Esters of two phenols are not particularly common in nature. The most familiar is the diester of gallic acid (**2.22**), which is ellagic acid (**2.23**), along with other gallotannin compounds.

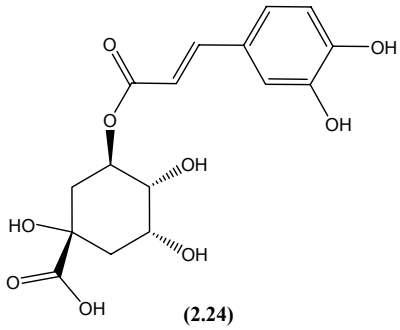


(2.22)



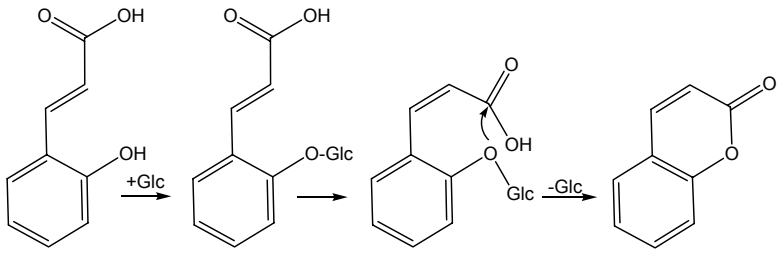
(2.23)

Generally the esters of phenols found naturally are compounds where the phenol contributes the carboxyl group, and another compound contributes the alcoholic hydroxyl group. The hydroxycinnamic acids do not seem to undergo intermolecular condensation, but esters with quinic acid and other acids do occur. For example, chlorogenic acid is an ester of caffeic acid and quinic acid (3-caffeoylquinic acid; **2.24**).



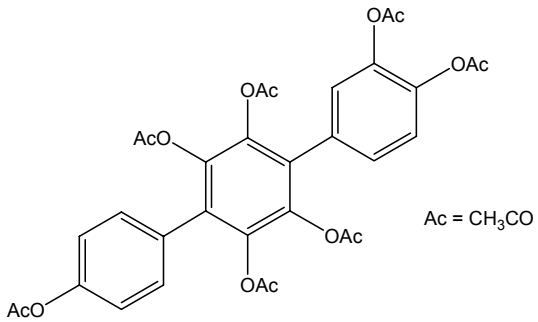
(2.24)

o-Hydroxycinnamic acids undergo intramolecular esterification to yield lactones that are called coumarins (see also Chapter 1, Section 3.5). Shown below is the formation of coumarin (**2.28**) from coumaric acid (**2.25**; by definition coumaric acid has the hydroxyl group in the *ortho*–position), which involves glycoslyation (**2.26**; see also Section 1.7), isomerization from the *trans*- to the *cis*-form (**2.27**), and intramolecular esterification.



(2.25) **(2.26)** **(2.27)** **(2.28)**

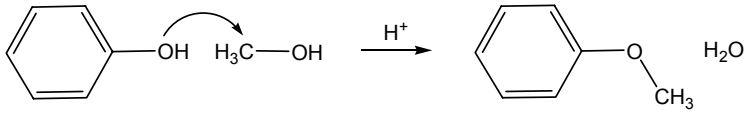
Of interest to mycologists, the basidiomycete *Polyporus leucomelas* produces the phenolic ester protoleucomelone (**2.29**).



(2.29)

1.7 Ethers and glycosides

Ethers (R-O-R) are frequently found as natural products in nature. The most common ether is that of methanol and the phenolic hydroxyl group. The methyl ether (methoxyl group) is very stable and therefore not reactive. Shown below is the formation of methoxybenzene (**2.30**) from phenol and methanol.



(2.5) **(2.30)**

Glycosides – formed between a sugar molecule and an alcohol – are in some sense similar to ethers. Glycosides are formed between the sugar molecule in a ring conformation (pyranose or furanose form) and an alcohol. The example below shows D-glucose (**2.31**), in equilibrium with β-D-glucopyranose (**2.32**).

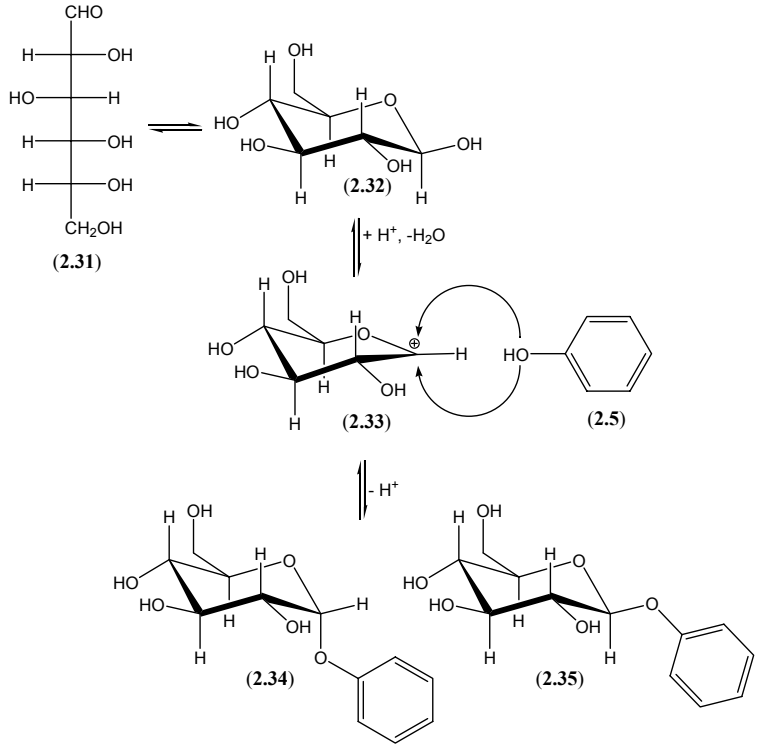


Figure 2-3. Formation of phenolic glycosides.

In presence of acid, cation **2.33** reacts with phenol (**2.5**) to result in a mixture of the α- and β-glucoside (**2.34** and **2.35**, respectively). Chemically this reaction involves the formation of an acetal (**2.34** and **2.35**) from a semi-acetal (the pyranose **2.32**; the furanose could also react). Unlike typical ether bonds, the glycoside bond is susceptible to acid hydrolysis.

1.8 Oxidation of the phenolic hydroxyl group

Oxidation of phenols is one of the most important aspects of these compounds to the biologist. Oxidation of phenolic compounds can result in the browning of tissues. Well-known examples are the browning of fruits after they have been cut. Oxidation can also result in the formation of metabolites that are toxic to animals and plants, and that can account for spoilage of foods in processing. On the other hand, toxic compounds formed from the oxidation of phenolics can inhibit pathogenic microorganisms. Certain phenols are used as retardants or antioxidants to prevent the oxidation of fatty acids.

1.8.1 Auto-oxidation of phenolic compounds

Auto-oxidation refers to the formation of cross-linked structures as a result of exposure to light and oxygen. Under the influence of light, oxygen can abstract a proton, thereby generating a radical. This is particularly likely to occur if the proton is adjacent to a double bond, because the radical electron can be delocalized, thus lowering the energy.

Given their aromatic nature, phenolic compounds are easily auto-oxidized. The radical that is generated can subsequently react with other radicals to form a dimer. Since the radical electron is delocalized, several structures can be formed depending on the precise location of the radical electrons at the time of the reaction.

Figure 2-4 shows how radicals of catechol (**2.11**) can react to form mixtures of tetrahydroxy-biphenyls (**2.36**) and quinines (**2.37**). Another example, shown in Figure 2-5, shows the formation of dimers of *p*-cresol (**2.38**).

A variety of complex compounds can arise through these mechanisms, including biflavonyls and bianthraquinones. An example of the latter is the compound iridoskyrim (**2.39**) formed by the fungus *Penicillium islandicum*.

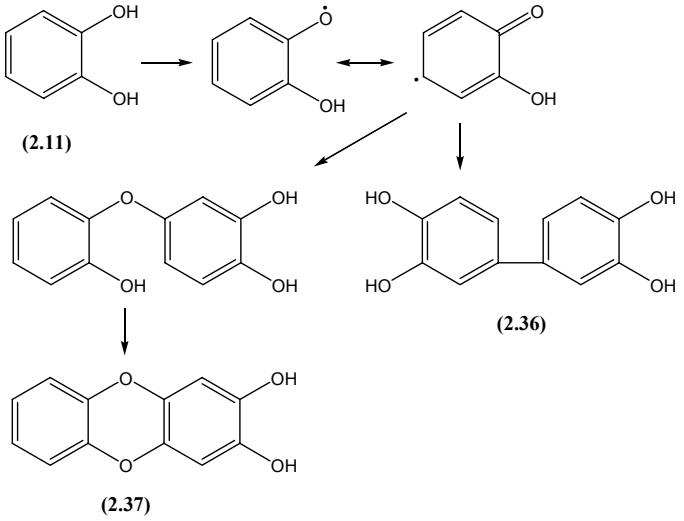


Figure 2-4. Auto-oxidation of catechol can result in the formation of different dimers.

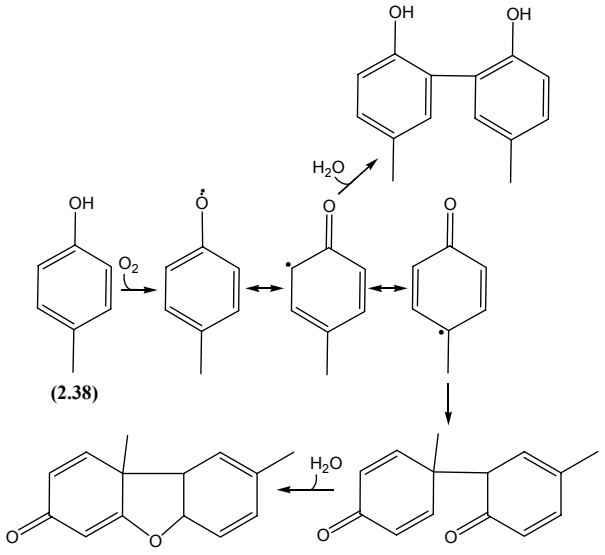
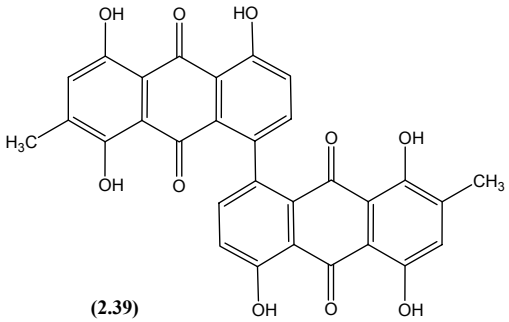


Figure 2-5. Auto-oxidation of *p*-cresol can result in the formation of different dimers.



(2.39)

1.8.2 Enzymatic oxidation of the phenolic hydroxyl group

An alternative mechanism for the oxidation of phenolic compounds is enzyme-catalyzed oxidation. Several classes of enzymes can catalyze this reaction. According to the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (NC-IUBMB), these enzymes are part of the E.C. 1 class of oxidoreductases (see the Internet web site: http://www.chem.qmul.ac.uk/iubmb/enzyme/EC1). The three main classes of enzymes that catalyze the oxidation of phenolic compounds are the oxidoreductases that use oxygen as electron acceptor (E.C. 1.10.3), the peroxidases (E.C. 1.11.1), and monophenol monooxygenase (E.C. 1.14.18.1).

1.8.2.1 E.C. 1.10.3

This class includes enzymes that use diphenols or related compounds as electron donors and oxygen as the acceptor, thereby forming the oxidized donor and water. Members include catechol oxidase (E.C. 1.10.3.1), laccase (E.C. 1.10.3.2), and *o*-aminophenol oxidase (E.C. 1.10.3.4). Laccase is also known as *p*-diphenoloxidase, whereas catechol oxidase is also known as diphenoloxidase, phenoloxidase, polyphenoloxidase, *o*-diphenolase, phenolase and tyrosinase. Many of these names are also used in reference to a different enzyme, monophenol monooxygenase (E.C. 1.14.18.1). This enzyme will be discussed further in Section 1.8.2.2.

The phenol-oxidizing enzyme tyrosinase has two types of activity: (***1***) phenol *o*-hydroxylase (cresolase) activity, whereby a monophenol is converted into an *o*-diphenol *via* the incorporation of oxygen, and (***2***) cathecholase activity, whereby the diphenol is oxidized. The two reactions are illustrated in Figure 2-6, in the conversion of tyrosine (**2.40**) to L-DOPA (3,4-dihydroxyphenylalanine; (**2.41**), dopaquinone (**2.42**), and indole-5,6-quinone carboxylate (**2.43**), which is further converted to the brown pigment

melanin *via* enzyme-mediated oxidation (reviewed by Sánchez-Ferrer et al. (1995)). Melanin is the major determinant of skin color in humans (Sturm, 1998), and is formed when the cut surfaces of fruits, such as apples, bananas and avocados, are exposed to air.

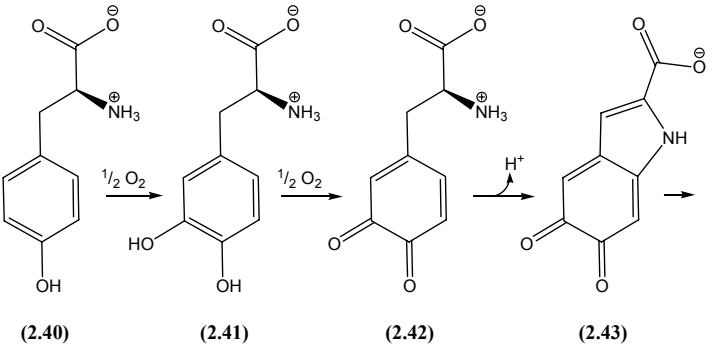


Figure 2-6. Tyrosinase-catalyzed oxidation of tyrosine results in precursors of melanin.

Laccase catalyzes the oxidation of *p*-diphenols to *p*-quinones. Shown in Figure 2-7 is the oxidation of 1,4-dihydroxybenzene (**2.44**) to *p*-quinone (**2.45**).

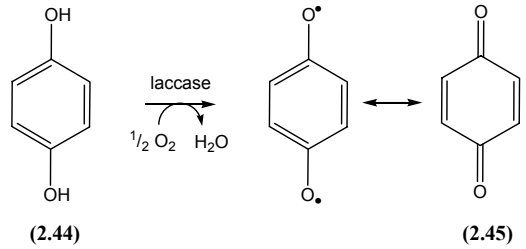
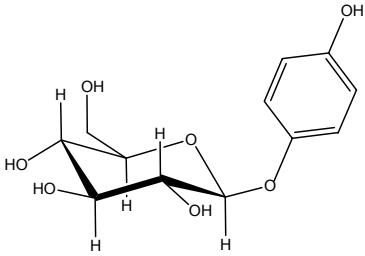
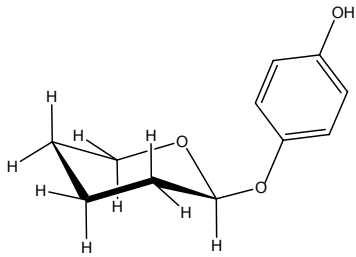


Figure 2-7. The oxidation of 1,4-dihydroxybenzene to *p*-quinone.

This enzyme exhibits no hydroxylase activity and is involved in the final synthesis of many naturally occurring *p*-quinones, *e.g.* the naphthaquinone juglone in walnut (**1.58**) and the benzoquinone arbutin (hydroquinone-β-D-glucopyranoside; **2.46**). Arbutin is a plant cryo-protectant that stabilizes membranes (Hincha et al., 1999). This compound has medicinal properties and has, for example, been used to treat urinary tract infections in humans. It is also used to lighten skin color, because it inhibits tyrosinase and hence the formation of melanin. The derivative deoxyarbutin (**2.47;** note the difference in the sugar molecule) was recently reported to be considerably more effective as a skin-lightening compound (Boissy et al., 2005).



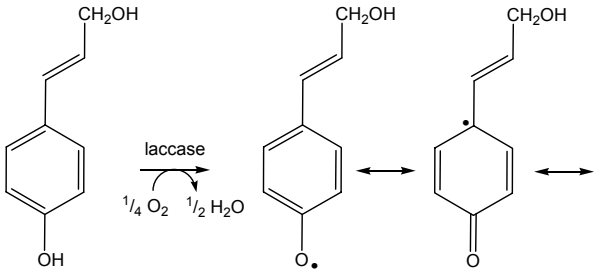
(2.46)



(2.47)

While initially controversial, there is evidence that laccases play a role in the polymerization of the cell wall polymer lignin (see Chapter 1, section 3.12), which occurs *via* the oxidative coupling of monolignol radicals with reactive (oxidized) sites on the lignin polymer. The evidence for their involvement comes from a number of studies in which laccases were localized to lignifying tissues in woody species through the use of histochemical stains. Furthermore, the expression of the *laccase* genes was shown to be specific for lignifying tissues. In addition, when laccases purified from these tissues were mixed with monolignols under aerobic conditions, a dehydrogenation polymer (DHP) with lignin-like characteristics was formed (Sterjiades et al., 1992; Driouich et al., 1992; Boa et al., 1993; Ranocha et al., 1999). Down-regulation of laccase genes in poplar through the introduction of antisense constructs did not, however, impact lignin content nor subunit composition, but did have an effect on cell wall structure (Ranocha et al., 2002).

Figure 2-8 shows laccase-mediated generation of radicals of coniferyl alcohol (**2.48**), with the stoichiometry of the reaction adjusted for coniferyl alcohol. The radical electron is delocalized, enabling the formation of various interunit linkages, as discussed in Chapter 1.



(2.48)

Figure 2-8. Laccase-catalyzed formation of coniferyl alcohol radicals.

1.8.2.2 E.C. 1.14.18.1

The E.C. 1.14 class of monooxygenases contains enzymes acting on paired donors, with the incorporation or reduction of molecular oxygen. Monophenol monooxygenase (E.C. 1.14.18.1) catalyzes the same reactions as catechol oxidase (E.C. 1.10.3.1; see Section 1.8.2.1) if only 1,2-benzenediols are available as substrate. In this case one of the monophenols acts as a donor for the oxidation of the other monophenol, and one atom of oxygen is incorporated. Common names like tyrosinase and phenolase can refer to both catechol oxidase and monophenol monooxygenase, but they can be distinguished based on the fact that E.C. 1.14 uses paired donors. In addition, the E.C. number is generally indicated in the text to further clarify this.

1.8.2.3 E.C. 1.11.1.

The E.C. 1.11.1 subclass contains the peroxidases, which use hydrogen peroxide (H_2O_2) as electron acceptor to oxidize the donor, thereby forming the oxidized donor and water. Members include horseradish peroxidase (E.C. 1.11.1.7; also known as guaiacol peroxidase and scopoletin peroxidase), manganese peroxidase (E.C. 1.11.1.13) and diarylpropane peroxidase (E.C. 1.11.1.14). All three classes are hemoproteins. Horseradish peroxidase and related peroxidases are involved in the oxidative coupling of lignans, lignin, and tannins. Mechanistically, hydrogen peroxide oxidizes the active site of the peroxidase enzyme, and upon binding of the substrate in the active site, the substrate becomes oxidized and the enzyme returns to its reduced state.

Peroxidases are encoded by large multi-gene families, which has complicated the study of individual peroxidase enzymes (*cf.* Christensen et al., 1998). Manganese and diarylpropane peroxidases are used by white rot fungi (basiodiomycetes) to degrade lignin *via* oxidation.

The origin of the H_2O_2 that peroxidases use is not entirely clear. The best studied peroxidases are the ones involved in cell wall lignification, and several mechanisms that describe the generation of H_2O_2 have been identified in different plant species, as will be discussed below.

Ogawa et al. (1997) investigated the formation of H_2O_2 and the superoxide radical ($O_2^{\cdot-}$) in spinach (*Spinacia oleracea*) hypocotyls with the use of histochemical stains. Nitroblue tetrazolium (NBT) is used to detect $O_2^{\cdot-}$ radicals. The colored reaction product formazan was only detected in the vascular tissue of developing spinach hypocotyls if CuZn-superoxide dismutase (CuZn-SOD; E.C. 1.15.1.1) was inhibited by DDC

(*N,N*-diethyldithiocarbamate), suggesting that CuZn-SOD effectively catalyzes the elimination of the $O_2^{\cdot-}$ radicals. The mechanism for the elimination of these radicals is through dismutation of superoxide into oxygen (O_2) *via* oxidation, and H_2O_2 *via* reduction. Imidazole and DPI (diphenyleneiodonium), inhibitors of NAD(P)H-oxidase (E.C. 1.6.99.6), were shown to suppress the formation of formazan, indicating the involvement of NAD(P)H-oxidase in superoxide radical formation. NADPH is nicotinamide dinucleotide phosphate, a compound that is used as an electron donor throughout cellular metabolism. Based on these results they proposed a mechanism for the generation of H_2O_2, shown below in Figure 2-9, that involves the concerted action of a membrane-bound NADPH oxidase and a CuZn-superoxide dismutase The H_2O_2 that is generated by the combined action of these two enzymes is then used to activate a peroxidase that oxidizes monolignols.

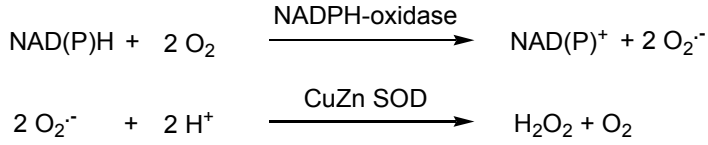


Figure 2-9. Formation of H_2O_2 through the concerted action of NAD(P)H oxidase and CuZn-superoxide dismutase, as proposed by Ogawa et al. (1997)**.**

There are cases where the peroxidase is not activated by H_2O_2, but by a reaction product instead. Ferrer et al. (1990) described the oxidation of the auxin indole-3-acetic acid (IAA; **2.49**) and molecular oxygen by a cell wall peroxidase that was able to oxidize coniferyl alcohol (**2.48**) in the absence of H_2O_2.

Auxins are plant hormones involved in a number of developmental processes in plants, including embryo development, leaf formation, and apical dominance (reviewed by Leyser (2005), and Woodward and Bartel (2005)). IAA is transported through the extracellular space, and would thus be readily available as a reductor to cell-wall bound peroxidases. The cell-wall bound peroxidases in this study were isolated from lupin (*Lupinus alba*), and the oxidation of coniferyl alcohol at the expense of IAA was monitored spectrophotometrically in the UV range of the spectrum. The IAA was converted to oxindoles. Previous studies showed that 3-methylene 2-oxindole (**2.54**) is the predominant oxindole formed (Ricard and Job, 1974). Ferrer et al. (1990) showed that the oxidation of coniferyl alcohol was dependent on the concentration of IAA, whereby high concentrations inhibited the reaction. This makes sense physiologically, since lignification is associated with a terminal developmental process, whereas high levels of

auxin are correlated with growth and differentiation. Folkes et al. (2002) proposed a reaction mechanism for the peroxidase-mediated oxidation of IAA without the involvement of H_2O_2, which is shown in Figure 2-10.

Oxidation of IAA (**2.49**) results in cation **2.50**, which undergoes decarboxylation and results in the skatolyl radical (**2.51**). This compound reacts with molecular oxygen to form peroxyl radical **2.52**. With IAA or another cellular reductor, the hydroperoxide **2.53** is formed. It is this compound that activates the peroxidase, and thus allows the oxidation of other substrates, such as coniferyl alcohol. Among the degradation products of **2.53**, 3-methylene 2-oxindole (**2.54**) is the most abundant.

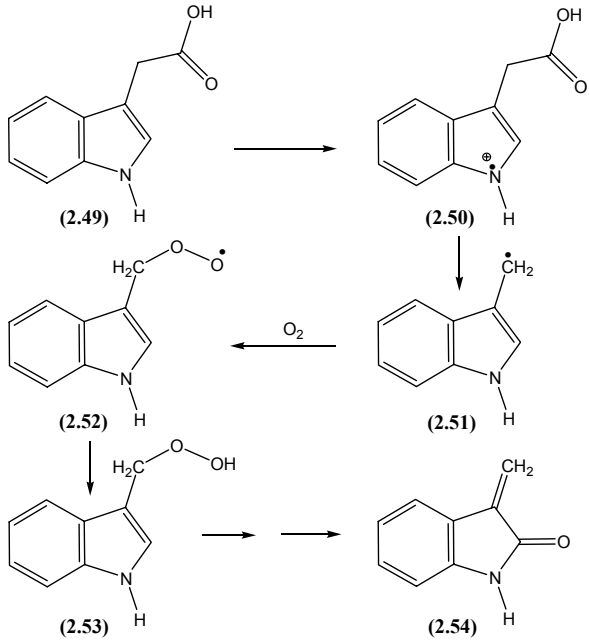
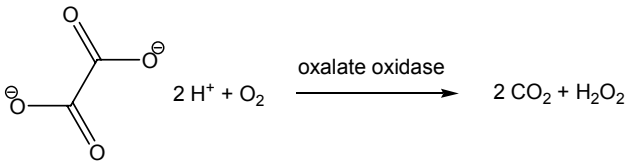


Figure 2-10. Oxidation of IAA.

An alternative mechanism for the generation of H_2O_2 was described by Caliskan and Cuming (1998), who studied the wheat protein germin. This protein is synthesized *de novo* when wheat embryos germinate. It was shown to be highly resistant to proteolytic degradation, and to have oxalate oxidase (E.C.. 1.2.3.4) activity. This enzyme catalyzes the oxidation of oxalate (**2.55**), and the formation of H_2O_2 as shown in Figure 2-11. Given that in 9-day old seedlings both the oxalate oxidase mRNA and protein were localized to the vascular tissue, the authors speculated that wheat germin

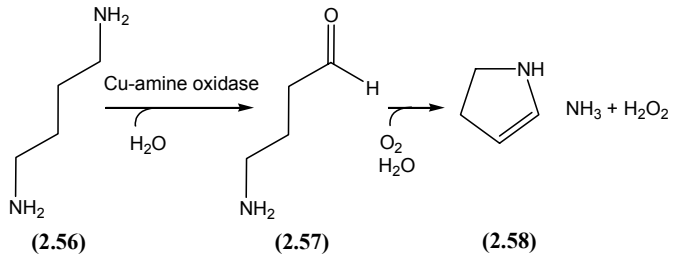
plays a role in providing H_2O_2 in those tissues of the seedling where the cell wall needs to be cross-linked to restrict cell growth.



(2.55)

Figure 2-11. Formation of H_2O_2 via oxidation of oxalate by oxalate oxidase.

A fourth possibility is the generation of H_2O_2 *via* oxidation of putrescine (butane-1,4-diamine; **2.56**). This reaction is catalyzed by copper amine oxidase (E.C. 1.4.3.6). Copper amine oxidases are homodimers in which each unit contains a copper ion and a 1,3,5-trihydroxyphenylalanine quinine co-factor. In plants copper amine oxidases generally oxidize putrescine to 4-aminobutanal (**2.57**). This latter compound undergoes spontaneous cyclization to Δ^1 pyrroline (**2.58**), ammonia, and H_2O_2, as shown in Figure 2-12 (Medda et al., 1995).



(2.56) **(2.57)** **(2.58)**

Figure 2-12. Formation of H_2O_2 via oxidation of putrescine by copper amine oxidase.

The copper amine oxidase in the model plant *Arabidopsis thaliana* is encoded by the *ATAO1* gene (Møller and McPherson, 1999). *In situ* hybridizations and analyses of transgenic plants expressing a reporter gene under control of the *ATAO1* promoter revealed expression of the *ATAO1* gene in the root cap and the vascular tissue. This would make it feasible that H_2O_2 generated *via* this mechanism could be used by a peroxidase involved in lignification and cross-linking of cell wall proteins in Arabidopsis.

Ros-Barceló et al. (2002) analyzed lignifying xylem of *Zinnia elegans*. Based on the inhibition of H_2O_2 production as a result of treatment with

imidazole, an involvement of NADPH-oxidase (E.C. 1.6.3.1) was hypothesized. This enzyme catalyzes the formation of H_2O_2 from the oxidation of NADPH with molecular oxygen, as shown in Figure 2-13.

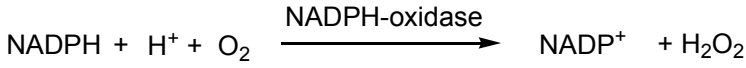


Figure 2-13. Formation of H_2O_2 via oxidation of NADPH by NADPH oxidase.

Önnerud et al. (2004) recently proposed a mechanism involving a redox shuttle for the oxidative coupling of coniferyl alcohol (**2.48**), as it may occur during lignification. Given that lignin is very compact, the authors speculated that enzymes such as peroxidases may be too large to be effective. They investigated whether manganese (II) oxalate (**2.55**; see Figure 2-11) could function as a redox shuttle. Figure 2-14 depicts how Mn(II) oxalate is reduced to Mn(III) oxalate by a membrane or cell-wall bound manganese peroxidase (E.C. 1.11.1.13). The Mn(III) diffuses into the cell wall, oxidizes monolignols and the lignin polymer, and returns to the manganese peroxidase to get oxidized again.

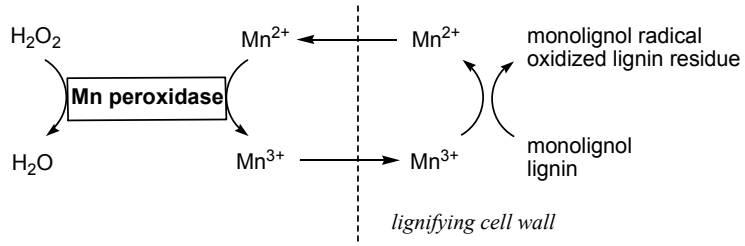


Figure 2-14. Oxidation of monolignols via a manganese oxalate redox shuttle as proposed by Önnerud et al. (2004).

Based on this overview, it is clear that there are many different mechanisms by which H_2O_2 can be generated, and it is possible that even more mechanisms exist. Further research is needed to determine to what extent these mechanisms are unique to particular plant species, tissues, metabolic processes, or developmental stages. Given that the availability and concentration of enzyme substrates is likely to fluctuate as a function of both the developmental stage and the environmental conditions, the availability of multiple mechanisms to generate H_2O_2 offers a high degree of flexibility.

1.9 Reactive oxygen species and antioxidants

Radicals are molecules with a free (unpaired) electron that are highly reactive. Radicals are formed in all living organisms during oxidation reactions that occur as part of normal metabolism. Under certain circumstances such as environmental stress, wounding, and pathogen attack, the concentration of free radicals is increased beyond the normal levels. Radicals can do considerable damage to living organisms when left unchecked. This results in part from their reactivity, particularly towards DNA and membranes (lipids and proteins), and in part from the chain reactions they can initiate. Chain reactions occur when a radical reacts with another molecule, abstracts an electron, and thereby creates a new radical that can react with other molecules.

Reactive oxygen species (ROS) are molecules that contain an oxygen atom and that are highly reactive as a result of the presence of a free radical, or a configuration of the oxygen atom whereby there are more electrons than usual. Examples of the first class include the hydroxyl radical ($^{\cdot}OH$), and the superoxide radical ($O_2^{\cdot -}$), whereas the peroxide (O_2^{2-}) and hypochlorite ($ClO^{\cdot -}$) ions belong to the second class. Hydrogen peroxide (H_2O_2) is also considered as a ROS because of its reactivity (Halliwell, 1991), as we have seen in Section 1.8.2.3. The hydroxyl radical is the most reactive. In fact, it is considered the most reactive radical known, with an ability to react with a very wide range of (bio-)molecules. It can be produced *via* the Fenton reaction, first described in 1894 (Figure 2-15).

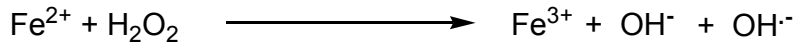


Figure 2-15. Production of the hydroxyl radical via the Fenton reaction.

Because of the reactivity of the hydroxyl radical, and the fact that the ingredients are inexpensive, the Fenton reaction is used on a commercial scale to treat waste water. The Fenton reaction can also occur with copper as the transition metal. Given that Fe^{2+}, Cu^{+}, and H_2O_2 are abundantly present in biological systems, hydroxyl radicals can be generated *via* the Fenton reaction *in vivo*. Reviews by Schützendübel and Polle (2002) and by Valko et al. (2005) describe the impact of ROS in plants and humans, respectively.

A particularly damaging reaction is the reaction between the hydroxyl radical and unsaturated fatty acid side chains of phospholipids in the cell membrane, a reaction referred to as lipid peroxidation (Figure 2-17).

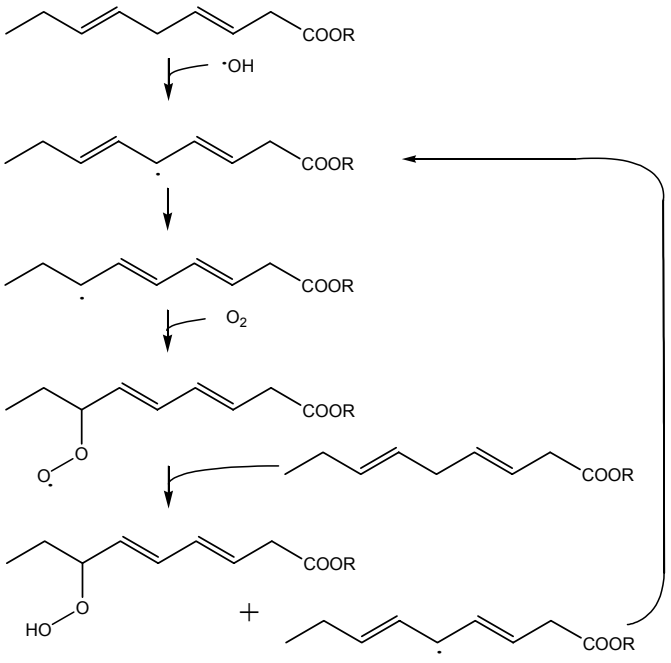


Figure 2-17. Lipid peroxidation. A hydroxyl radical abstracts a hydrogen from a fatty acid or lipid molecule. After rearrangement to a conjugated structure, the radical reacts with oxygen to form a peroxyl radical. The newly formed peroxyl radical can initiate a chain reaction whereby new peroxyl radicals are formed.

The hydroxyl radical will abstract a hydrogen atom from the fatty acid, creating a fatty acid radical with the free electron on a carbon atom in the chain. The radical will typically undergo a rearrangement resulting in a more stable conjugated structure. This newly generated radical can crosslink with a nearby fatty acid radical. Alternatively, the fatty acid radical can react with molecular oxygen to produce a peroxyl radical. This then sets in motion a chain reaction, whereby many new peroxyl radicals are generated. As a consequence of lipid peroxidation, the fluidity of the membrane can be affected, and membrane proteins, especially receptors, can become part of the radical reactions, affecting their function. Ultimately, the membrane can collapse.

More recently, the impact of excess iron on carcinogenesis has been studied. Toyokuni (2002) reported that in kidney cells of rats an overload of iron can result in carcinogenesis because of Fenton-reaction induced damage to a tumor suppressor gene. This is currently an active area of research.

Living organisms have developed various ways to deal with ROS. One mechanism is enzymatic inactivation. The enzyme superoxide dismutase, (E.C. 1.15.1.1) catalyzes the dismutation of superoxide into oxygen (O_2) *via* oxidation, and H_2O_2 *via* reduction (see also Section 1.8.2.3). The H_2O_2, which is reactive itself, is removed through the action of the enzymes catalase (E.C. 1. 11.1.6) and glutathione peroxidase (E.C. 1.11.1.9). Catalase catalyzes the conversion of H_2O_2 to water and oxygen, whereas glutathione peroxidase catalyzes the formation of oxidized glutathione (G-S-S-G) from reduced glutathione (G-SH), at the expense of H_2O_2 (Halliwell, 1991).

The other commonly used mechanism to inactivate ROS is through the use of antioxidants. Antioxidants can react with the radical, but rather than turning into another reactive molecule, these compounds are relatively stable in the presence of the radical electron. As a consequence, they scavenge the radical electrons, quench the chain reaction, and avoid further damage. The relative stability of antioxidants containing a radical electron is generally the result of the presence of conjugated bonds, so that the radical electron can be delocalized. As a consequence, aromatic compounds in general, and phenolic compounds in particular are very effective antioxidants. Examples of delocalized radical electrons in phenolic compounds are given in structures **2.11** and **2.38**. The antioxidant properties of phenolics will be discussed in more detail in Chapter 7.

2. REFERENCES

Bao, W., O'Malley, D. M., Whetten, R., and Sederoff, R.R., 1993, A laccase associated with lignification in loblolly pine xylem. *Science* **260**: 672-674.

Boissy, R. E., Visscher, M., deLong, M. A., 2005, Deoxyarbutin: a novel reversible tyrosinase inhibitor with effective *in vivo* skin lightening potency, *Exp. Dermatol*. **14**: 601-608.

Christensen, J. H., Bauw, G., Gjesing Welinder, K., Van Montagu, M., and Boerjan, W., 1998, Purification and characterization of peroxidases correlated with lignification in poplar xylem, *Plant Physiol*. **118**: 125-135.

Driouich, A., Laine, A. C., Vian, B., and Faye, L., 1992, Characterization and localization of laccase forms in stem and cell cultures of sycamore, *Plant J*. **2**: 13-24.

Ferrer, M. A., Pedreño, M. A., Muñoz, R., and Ros Barceló, A., 1990, Oxidation of coniferyl alcohol by cell wall peroxidases at the expense of indole-3-acetic acid and O_2, *FEBS Lett*. **276**: 127-130.

Folkes, L. K., Rossiter, S., and Wardman, P., 2002, Reactivity toward thiols and cytotoxicity of 3-methylene-2-oxindoles, cytotoxins from indole-3-acetic acids, on activation by peroxidases, *Chem. Res. Toxicol*. **15**: 877-882.

Halliwell, B., 1991, Reactive oxygen species in living systems, *Am. J. Med.* **91**: 14S

Hincha, D. K., Oliver, A. E., and Crowe, J. H., 1999, Lipid composition determines the effects of arbutin on the stability of membranes, *Biophys. J.* **77**: 224-234.

Leyser, O., 2005, Auxin distribution and plant pattern formation: how many angels can dance on the point of PIN?, *Cell* **121**: 819-822.

Medda, R., Padiglia, A., Pedersen, J. Z., Rotilio, G., Finazzi Agró, and Floris, G, 1995, The reaction mechanism of copper amine oxidase: detection of intermediates by the use of substrates and inhibitors, *Biochem*. **34**: 16375-16381.

Møller, S. G., and McPherson, M. J., 1998, Developmental expression and biochemical analysis of the *Arabidopsis atao1* gene encoding an H_2O_2-generating diamine oxidase, *Plant J.* **13**: 781-791.

Ogawa, K., Kanematsu, S., and Asada, K., 1997, Generation of superoxide anion and localization of CuZn-superoxide dismutase in the vascular tissue of spinach hypocotyls: their association with lignification, *Plant Cell Physiol*. **38**: 1118-1126.

Önnerud, H., Zhang, L., Gellerstedt, G., and Henriksson, G., 2004, Polymerization of monolignols by redox shuttle–mediated enzymatic

oxidation: a new model in lignin biosynthesis I, *Plant Cell* **14**: 1952-1963.

Ranocha P., Chabannes, M., Chamayou, S., Danoun, S., Jauneau, A., Boudet, A.-M., and Goffner, D., 2002, Laccase down-regulation causes alterations in phenolic metabolism and cell wall structure in poplar, *Plant Physiol.* **129**: 145-155.

Ranocha, P., McDougall, G., Hawkins, S., Sterjiades, R., Borderies, G., Stewart, D.,, Cabanes-Macheteau, M., Boudet, A.-M., and Goffner, D., 1999, Biochemical characterization, molecular cloning and expression of laccases – a divergent gene family – in poplar, *Eur. J. Biochem*. **259**: 485-495.

Ricard, J. and Job, D., 1974, Reaction mechanisms of indole-3-acetate degradation by peroxidases. A stopped-flow and low-temperature spectroscopic study, *Eur. J. Biochem.* **44**: 359-374.

Ros-Barceló, A., Pomar, F., López-Serrano, M., Martínez, P., Pedreño, M. A., 2002, Developmental regulation of the H_2O_2-producing system and of a basic peroxidase isoenzyme in the *Zinnia elegans* lignifying xylem, *Plant Physiol. Biochem*. **40**: 325-332.

Sánchez-Ferrer, A., Rodríguez-López, J. N., García-Cánovas, F., and García-Carmona, F., 1995, Tyrosinase: a comprehensive review of its mechanism, *Biochim. Biophys. Acta* **1247**: 1-11.

Schützendübel, A., and Polle, A., 2002, Plant responses to abiotic stresses: heavy metal-induced oxidative stress and protection by mycorrhization, *J. Exp. Bot.* **53**: 1351-1365.

Sterjiades, R., Dean, J. F. D., and Eriksson, K.-E. L., 1992, Laccase from sycamore maple (*Acer pseudoplatanus*) polymerizes monolignols, *Plant Physiol.* **99**: 1162-1168.

Sturm, R. A., 1998, Human pigmentation genes and their response to UV radiation, *Mutat. Res. Rev*. **422**: 69-76.

Toyokuni, S., 2002, Iron and carcinogenesis: from Fenton reaction to target genes, *Redox Rep*. **7**: 189-197.

Valko, M., Morris, H., and Cronin, M. T., 2005, Metals, toxicity and oxidative stress, *Curr. Med. Chem*. **12**: 1161-1208.

Woodward, A. W., and Bartel, B., 2005, Auxin: regulation, action, and interaction, *Ann. Bot. (Lond.),* 2005, **95**: 707-735.

For further reading:

Denisov, E., and Afanas'ev, I. B., 2005, *Oxidation and Antioxidants in Organic Chemistry and Biology*, CRC Press, Boca Raton, pp. 1024.

Smirnoff, N., *Antioxidants and Reactive Oxygen Species in Plants*, Blackwell Publishing, Oxford, pp. 302.

Chapter 3

BIOSYNTHESIS OF PHENOLIC COMPOUNDS

1. INTRODUCTION

This chapter provides an overview of the biosynthetic pathways leading to the major classes of phenolic compounds as outlined in Chapter 1. This overview is intended to show the origin of the different families of phenolic compounds, many of which share precursors. In addition to an overview of established pathways, newly elucidated steps in the biosynthesis of several classes of compounds will be highlighted.

Detailed studies of biosynthetic pathways have recently become much more feasible due to the availability of mutants in which genes that affect phenolic metabolism are defective. Genes consist of coding regions and regulatory sequences. The coding region is transcribed into messenger RNA (mRNA) by (generally) RNA polymerase II and stipulates the amino acid sequence of the protein encoded by the gene. The regulatory sequences are located at the upstream (5′) and sometimes also the downstream (3′) end of the coding region and control the spatio-temporal expression of the gene. 'Spatio' refers to cells, tissues or organs where a gene is expressed, whereas 'temporal' refers to the developmental stage at which a gene is expressed. Mutations are genetic alterations in the DNA that may affect gene expression. This can be the result of changes in regulatory elements of the gene, so that the gene is no longer expressed, expressed at lower levels, or has a different spatio-temporal pattern of expression. Alternatively, mutations can affect the part of the gene that encodes the protein. In this case truncated proteins are synthesized, or proteins in which a critical amino

acid has been substituted, affecting substrate binding ability, substrate specificity, or the ability to catalyze the substrate conversion. In these cases the protein can no longer function in its normal way, which oftentimes leads to an altered phenotype, such as altered growth characteristics, plant architecture, or color.

Below follows a brief overview of commonly used strategies for protein isolation, gene cloning, and protein characterization to describe the principles of using biochemistry and genetics for the elucidation of biosynthetic pathways. Next, biochemical pathways leading to the different classes of phenolic compounds will be presented. Each compound will be referred to with one number (*e.g.* **3.12**), which will be used throughout the chapter, even if the compound is an intermediate of several different pathways and is thus included in several sections of this chapter. Enzymes in this chapter are referred to by their name and E.C. number whenever possible. The E.C. numbers were designated by the Enzyme Commission of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (IUBMB). This information can be accessed via the Internet at *www.expasy.org*.

2. PROTEIN ISOLATION AND PURIFICATION

The traditional methodology for the isolation of proteins involves biochemical separation techniques, whereby the protein of interest is isolated from all the other proteins based on its unique physical and chemical properties. This includes molecular weight (and hence size), shape (for example, globular versus rod), hydrophobicity, and net electric charge.

In the case of enzymes involved in biochemical pathways, the isolation is often based on activity assays. The nature of the activity assay depends on the enzymatic reaction and can involve, for example, the detection of a product on a thin-layer chromatography (TLC) plate (see Chapter 4, Section 1.2.1), the appearance or disappearance of a specific absorbance in a spectrophotometric assay, or a coupled assay involving the oxidation or reduction of a co-factor such as nicotinamide dinucleotide (NAD(H)), which can be measured by changes in fluorescence.

The isolation procedure starts with the preparation of a cell extract in which the enzyme activity can be detected. This typically involves grinding the tissue in an extraction buffer so that the cell contents, including the proteins, become accessible. Protease inhibitors, such as phenylmethyl

sulphonylfluoride (PMSF), are added to the extraction buffer to avoid proteolytic degradation of the enzyme. The first fractionation is generally a centrifugation step, whereby the enzyme is precipitated if it is bound to the cell wall or the cell membrane, and otherwise ends up in the supernatant. The supernatant will also contain membrane fractions derived from the organelles (mitochondria, chloroplasts, endoplasmic reticulum, Golgi complex), which can be precipitated by ultracentrifugation if desired.

Soluble proteins can be separated from each other based on variation in their solubility in high-salt solutions. Most proteins will precipitate if the salt concentration exceeds a certain level, a process referred to as 'salting out'. The salt concentration at which proteins precipitate varies depending on the size, shape, and the proportion and distribution of polar, apolar and charged amino acids. A successful strategy to separate proteins from each other, therefore, is to add a salt, followed by centrifugation to remove the precipitated proteins. Additional salt can be added to the supernatant, so that different protein fractions representing different solubility values can be generated. The precipitated proteins can be redissolved by diluting the precipitate in a low-salt buffer. Ammonium sulfate is commonly used to precipitate proteins, because most proteins will precipitate in a saturated (4 Molar) solution of this salt. Furthermore, little heat is generated when ammonium sulfate is dissolved, thus preventing heat-induced denaturation of the proteins. The ammonium sulfate can be added as a powder, or as a saturated solution. Enzyme activity assays are performed on the different fractions that are generated, so that the amount of ammonium sulfate that needs to be added in order to remove a subset of the proteins in the extract can be determined. Alternatively, other salts, polyethylene glycol, apolar solvents, protamine sulfate, or trifluoroacetic acid can be used to precipitate proteins. After this initial enrichment step, chromatography is typically used to further purify the enzyme. The main types of chromatography will be discussed below.

Ion exchange chromatography relies on variation in charge between different proteins. In this case the chromatography column is filled with a resin harboring fixed charged groups, and counter ions of the opposite charge. The protein mixture is loaded onto the column, and proteins with a charge opposite of the charge of the resin will replace the counter ions and adsorb to the column, whereas proteins with the same charge as the resin and uncharged proteins can be removed by flushing the column with a low-salt buffer. An elution buffer of increasing ionic strength (increasing salt concentration) or changing pH is pumped through the column. Depending on the strength of the electrostatic interaction between the protein and the

matrix, the proteins elute from the column at different ionic strengths or pH-values of the elution buffer. The column can be recharged afterwards.

Hydrophobic interaction chromatography relies on hydrophobic interactions between apolar amino acid residues in the proteins and a resin containing hydrophobic groups, such as *n*-octyl or phenyl groups. After the protein mixture is applied to the column, an elution buffer with decreasing ionic strength is used. Hydrophilic proteins will elute first, whereas hydrophobic proteins elute last.

Bio-affinity chromatography is based on adsorption of the target enzyme to a resin to which the substrate of the target enzyme has been covalently linked. An example is the purification of a cellulose-degrading enzyme through the use of a column containing a cellulose-based resin. Proteins that have no affinity for the substrate are removed from the column with a wash buffer, and the target enzyme is eluted through the application of a buffer containing unbound competitor molecules, or through the use of a buffer with a high ionic strength or a pH that reduces the affinity of the enzyme for the resin-bound substrate.

Gel filtration or size exclusion chromatography relies on separation of proteins by size through a matrix made of small beads in a chromatography column. The beads contain pores of a fixed size. Large proteins will move faster through the column than small proteins, because the large proteins move in between the beads of the matrix, whereas the small proteins move through the pores. The size of the beads and the size of the pores within the beads determine the fractionation range, *i.e.* the range in molecular weight that can be effectively separated on the column. The dimensions of the column and the flow rate of the elution buffer determine the resolution.

While there are no set protocols for the specific order in which the different chromatographic separations are carried out, ion exchange and hydrophobic interaction chromatography typically precede gel filtration and bio-affinity chromatography. The latter can involve expensive resins and is often performed after various contaminants have been removed during earlier purification steps.

High-performance liquid chromatography (HPLC) and fast protein liquid chromatography (FPLC) rely on the same separation principles as the traditional chromatography columns, but tend to be much faster because of high flow rates that are possible due to the uniform bead size and the mechanical strength of the beads. See also Chapter 4, section 1.2.2.

Regardless of the chromatography method that is used, fractions containing subsets of the proteins in the sample are collected at the bottom of the column. Each of the fractions is assayed for enzyme activity. In addition, the complexity of the fraction is evaluated by separating the proteins in a small sample of the fraction using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). This method separates proteins in mixtures based on their size.

The purification is complete when the protein of interest has been purified to (apparent) homogeneity, meaning that a fraction has been obtained in which the protein of interest is the only protein, or in which any remaining contaminants are below the level of detection.

3. GENE CLONING STRATEGIES

The availability of mutants with altered phenolic content or composition enables cloning of the mutated gene, and subsequently cloning of the wild-type (normal) version of the gene. It is important to keep in mind that if a plant lacks a certain phenolic compound as a result of a mutation, the wild-type version of the gene – referred to as the wild-type allele – plays a role in the biosynthesis of this compound. Mutations occur spontaneously as a result of errors during DNA replication, but can also be induced by the use of alkylating agents such as ethyl methanesulfonate (EMS) and diethylsulfate (DES), radiation with fast neutrons, and insertional elements (see Section 3.1). The gene name is written in italics, and the protein encoded by the gene is written in regular font style. Depending on the species, the whole name or just the first letter of the wild-type allele is capitalized, whereas the mutant allele is referred to in lower case letters.

The sequence of the gene can be used to deduce the amino acid sequence of the protein encoded by the gene. The DNA and amino acid sequences can then be used to identify similar sequences in the large sequence databases such as GenBank (www.ncbi.nlm.nh.gov) or SWISSPROT (www.expasy.org). The chemical data obtained form the mutant combined with the sequence data from the gene that is defective in the mutant can then provide information on the function of the gene in the biosynthesis of a certain class of phenolic compounds.

Several methods to clone genes will be discussed in the next three sections. The ability to clone genes was revolutionized by the development

of the polymerase chain reaction (PCR). This is an *in vitro* method for the amplification of specific regions of DNA or cDNA. cDNA (complementary DNA) is a DNA copy of an mRNA molecule and which is synthesized *in vitro* by a retroviral reverse transcriptase. cDNA is more stable than mRNA, and more compact than the corresponding genomic DNA region, because cDNA does not contain introns.

PCR was developed in 1984 by Dr. Kary Mullis and co-workers (Saiki et al., 1985), who subsequently won (half of) the 1993 Nobel Prize in Chemistry for his work. The breakthrough that made PCR possible was the use of a heat-stable DNA polymerase for the synthesis of DNA. The polymerase is typically isolated from a thermophilic archaebacterium, such as *Thermus aquaticus, Thermococcus gorgonarius,* or *Thermus ubiquitus*. The name of the polymerase reflects its origin. In the examples above, the polymerase is referred to as *Taq*, *Tgo* and *Tub* polymerase, respectively. These polymerases are primer- and template-dependent, *i.e.* they require a nucleotide with a free 3′ OH group that is part of a short double stranded DNA structure, to which they will add a new nucleotide that complements the nucleotide on the template strand. The primers are designed by the researcher and are typically oligonucleotides ('oligo's') of 18-25 residues long. This length results in sequences that in most cases will define a unique site in the genome of the species of interest. Exceptions may be oligo's that bind to repetitive sequences or that bind to a conserved sequence in a gene that is part of a multigene family.

PCR involves three steps: **1**) denaturation, performed at 94°C, to melt (separate) the two DNA strands of the template, **2**) primer annealing, performed at 45-70°C depending on the length and GC-content of the primer, and **3**) extension, performed at 72°C (or a specific temperature recommended by the manufacturer of the enzyme), during which the DNA delineated by the primers is being synthesized by the polymerase. This process is repeated 20-40 times, and during each cycle each template strand is being duplicated. Consequently, there is an exponential amplification of the target DNA, so that PCR with pico- or nanogram quantities of template DNA will result in enough product to perform further manipulations (cloning, sequencing, transfection). Given the expense of the enzyme, PCR is performed in small volumes (10-50 µl) in a thermal cycler, which is a machine specifically designed for this process. PCR conditions typically need to be optimized empirically in order to obtain highly specific products and a reasonable yield. Varying the annealing temperature of the primers, the total number of cycles, and the concentrations of primers, nucleotides, and the cofactor Mg^{2+} will generally allow the identification of suitable

amplification conditions. PCR has become a routine technique in plant, animal and microbial biology as well as medicine, and is used for cloning of genes and cDNA's, as well as for genotyping using so-called molecular markers (Innis et al., 1999).

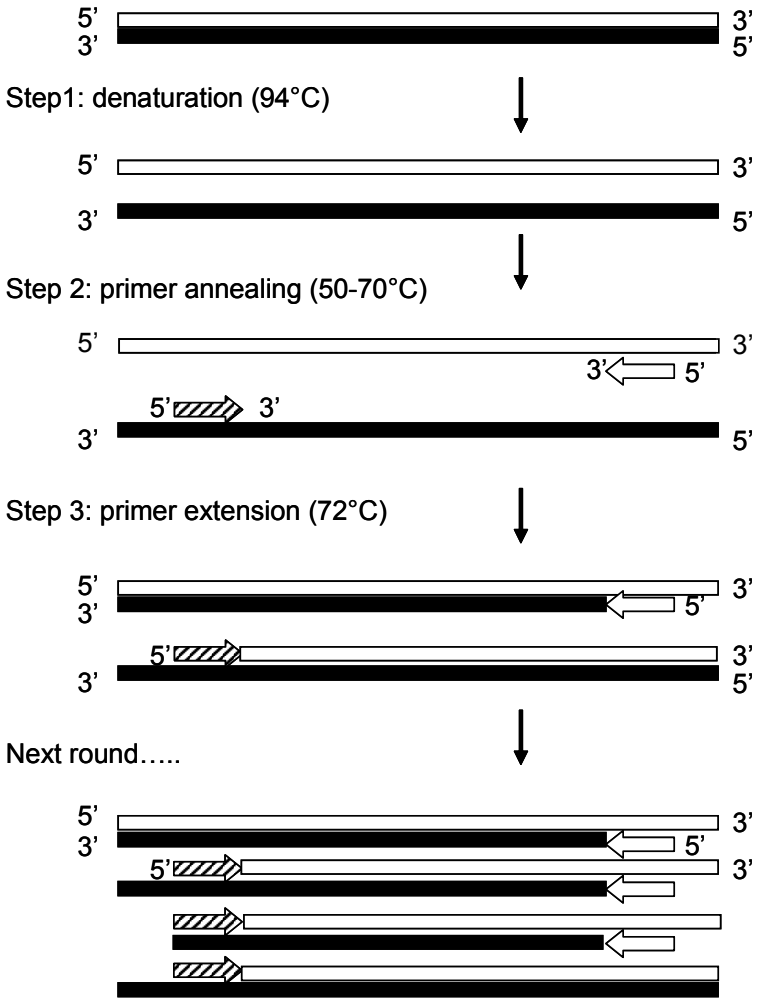


Figure 3-1. Schematic representation of the polymerase chain reaction. The two strands of the template DNA are represented by the white and black bars. The upper (forward) and lower (reverse) primers are indicated by a hatched and solid white arrow, respectively. The 5′ and 3′ refer to the corresponding hydroxyl groups on the ribose residue in the DNA backbone, and indicate the directionality of the DNA. DNA polymerases synthesize DNA in the 5′ to 3′ direction.

3.1 Insertional mutagenesis

Insertional mutagenesis is based on the insertion of a known sequence, referred to as a *tag*, in the gene of interest. Depending on whether the tag is a transposable element or T-DNA, the method of cloning genes is referred to as *transposon tagging* or *T-DNA tagging* (Walbot, 1992). Transposable elements, or transposons, are mobile genetic elements that were first discovered in maize (*Zea mays* L.) by the late Dr. Barbara McClintock (McClintock, 1947), who won the 1983 Nobel Prize in Medicine/Physiology for her pioneering research. The three major transposable element systems in maize are *Activator* (*Ac*)/*Dissociation* (*Ds*), *Supressor mutator* (*Spm*)/*defective Suppressor mutator* (*dSpm*) and Robertson's *Mutator* (*MuDR*)/*Mu*. The *Spm*/*dSpm* system is also known as *Enhancer* (*En*)/*Inhibitor* (*I*). All of these transposon systems are two-element systems: the first element (*Ac*, *Spm* (*En*), *MuDR*) is the autonomous element which encodes the transposase enzyme necessary for transposition, and the second element is the non-autonomous element, which can only transpose in the presence of an active autonomous element. In addition to these transposable element systems, the maize genome contains many other transposable elements, as well as retrotransposons. The latter class of mobile elements replicate through an RNA intermediate. The other species in which transposable elements are commonly used for cloning purposes is snapdragon (*Antirrhinum majus*). The elements in this species are called *Tam,* where *Tam1* is the autonomous element and *Tam2* is the non-autonomous element. Several additional *Tam* elements have been identified (reviewed by Schwarz-Sommer et al., 2003). The *Ac* and *Spm* (*En*) systems have also been used in species that do normally not harbor transposable elements themselves, including the model plant *Arabidopsis thaliana*, tomato (*Lycospersicon esculentum*) and tobacco (*Nicotiana tabacum*).

T-DNA is DNA transferred by the soil borne pathogen *Agrobacterium tumefaciens*. This bacterium transfers T-DNA, which is harbored on the bacterium's tumor-inducing (Ti) plasmid, to cells of the host plant via wound sites. The T-DNA then integrates in the genome of the plant, in a more or less random fashion. If the T-DNA inserts into a gene, the gene will likely loose its normal function. The principle of *Agrobacterium*-mediated transformation of plants has recently been reviewed (Gelvin, 2000; 2003).

There are two strategies to use insertional mutagenesis for cloning purposes. One is the *direct tagging* approach. In this case a mutant, for example a spontaneous or chemically induced mutant, has already been identified, but no information on the nature of the mutated gene is available.

The mutant is then crossed with a wild-type line carrying an active transposable element, and the F_1 progeny is screened for the presence of a plant with the mutant phenotype. Such a mutant will only be identified if the wild-type allele of the gene of interest has been mutated, presumably as the result of an insertion.

The second approach that can be taken is the *random tagging* method. This method is based on the principle that the insertion elements can insert in any gene, so that all of the genes controlling the trait of interest can be uncovered, as long as the mutation is not lethal. Since most mutations are recessive, meaning that both the maternal and paternal copy of the gene need to be defective in order to see a mutant phenotype, the screening is typically performed using F_2 families in which mutations will segregate. Given the relative low mutation rate (1:10,000 to 1:1,000,000), large populations of plants need to be screened. Therefore, an efficient method of screening has to be available in order to identify mutants of interest.

Once an insertional mutant with the desirable phenotype has been obtained, the cloning strategy is similar regardless whether transposons or T-DNA were used. In the direct tagging strategy the mutant is crossed with a wild-type plant to produce F_1 progeny. Since most mutations are recessive, the F_1 progeny is self pollinated to produce F_2 progeny. The plants of the F_2 population are scored for the presence of the mutation. DNA is isolated from these plants, as well as from the wild-type siblings. In the case of the random tagging strategy, DNA is isolated from the mutant and its wild-type siblings from the same family. Under both scenarios the isolated DNA is then used to identify the insertion element that is likely the cause of the mutation. There are several methods available to do this, including the use of the traditional Southern blot hybridized with a radioactive or chemically labeled probe that will hybridize to the insertion element (Federoff et al, 1984; Tan et al., 1997), methods based on the polymerase chain reaction (PCR) (Liu et al., 1995; Frey et al., 1997; Ribot, 1998), or plasmid rescue (Behringer and Medford, 1992; Meyer et al., 1996). All of these methods will ultimately result in the isolation of the DNA adjacent to the insertion element, which is typically the gene of interest. Sequencing of this DNA followed by homology searches in the large public sequence databases can then provide information on the identity and function of the gene of interest. The function will ultimately need to be tested with additional experiments.

3.2 Map-based cloning

Map-based cloning strategies work best with species with completely sequence genomes. At this time plantgenomes that have been sequenced include the model plant *Arabidopsis thaliana* (The Arabidopsis Genome Initiative, 2000), rice (*Oryza sativa*; International Genome Sequencing Project, 2005) and the poplar tree (*Populus trichocarpa*; Joint Genome Institute, http://genome.jgi-psf.org/Poptr1/Poptr1.home.html). In this case the first step after identifying a mutant of interest is placing the mutation on the genetic map. This is done by crossing the mutant with a wild-type plant that has a different genetic origin. This will generate F_1 progeny, which, for the same reason as described in Section 3.1, is self pollinated to produce an F_2 population. The mutants and wild-type plants are identified in this F_2 population, DNA is isolated, followed by the determination of the genotype of these plants at a large number of genetic loci across the genome.

Genotyping is generally performed with so-called molecular markers, which detect genetic variation across the genome. Mapping the mutation involves identifying the molecular marker(s) that are associated with the presence of the mutation. If the mutation is closely linked to a particular marker, the mutant plants will show predominantly the marker allele from the mutant parent, whereas the wild-type plants will show the marker allele from the wild-type parent. Once the precise map location has been determined, the gene sequence can be obtained based on the available genome sequence. Oftentimes several genes are located in the region where the mutation maps, but typically some of these genes can be discarded based on their known or deduced function. Introduction of a wild-type copy of the gene into a mutant plant via transformation and showing that the transformed plants have a wild-type phenotype is proof that the right gene has been identified. An example of the map-based cloning of a gene involved in the biosynthesis of a phenolic compound can be found in Franke et al. (2002a).

3.3 The candidate-gene approach

The candidate-gene approach (Pflieger et al., 2001) became possible after the establishment of large DNA and protein databases. When using this approach, a mutant of interest is characterized chemically, and based on the knowledge of the pathway most likely affected by the mutation, a candidate gene is proposed. Thus, the candidate gene is defined as the gene that, if it were defective, would cause the observed mutant phenotype. Once such a candidate gene has been proposed, the sequence databases can be searched in an attempt to identify DNA or protein sequences from the candidate gene.

If the DNA sequence is available in a database, PCR can be used to obtain the gene from the mutant. After sequencing the gene, sequence comparison between the gene from the mutant and the sequence in the database (or better yet, the sequence obtained from the wild-type progenitor of the mutant) can reveal whether or not the candidate gene is the gene that is responsible for the mutation. If no mutations are identified during the sequence comparison, additional sequence, especially from the upstream regulatory elements, needs to be obtained in order to exclude the candidate gene from further consideration. If there is no evidence for mutations, a new candidate gene needs to be considered. This approach will also work if the sequence of the candidate gene is available from a different but related species, because the degree of sequence homology is typically high between related species.

The efficiency of the candidate-gene approach can be improved considerably if comparisons between the mutant and a wild-type (progenitor) show differences in gene expression (observed via northern blots or reverse-transcription PCR), differences in enzyme activity, or differences in the amount of protein (observed via western blotting) (Bout and Vermerris, 2003). If differences are observed prior to the cloning of the candidate gene, the evidence in support of the candidate gene being mutated and responsible for the mutant phenotype is much stronger. It is, however, possible that the mutation has no impact on gene expression levels, and that the mutation results in reduced enzyme activity. Enzyme assays performed with crude protein extracts are not always specific for one enzyme. Therefore, even if the enzyme encoded by the candidate gene is less active, a different protein with similar activity may obscure differences in enzyme activity between the mutant and the wild-type control.

3.4 QTL mapping

Quantitative traits such as, for example, weight, height and yield, show a continuum in values because they tend to be controlled by multiple independent genetic loci. In addition, they are influenced by the environment. In the case of plant height, for example, rainfall and soil mineral concentrations will play a role. The identification of genes controlling quantitative traits is more complex than the identification of genes controlling traits that are inherited in a qualitative or discrete manner, such as, for example, flower color. A *quantitative trait locus* (QTL; this is also the abbreviation for the plural, quantitative trait loci), defined as a genetic locus delineated by two molecular markers on a genetic map of the species of interest and affecting a quantitative trait of interest, can be identified in an F_2 population generated from two parental lines that differ as

much as possible from each other with respect to the trait of interest. By evaluating a large F_2 population (at least several hundred individuals) one ensures that so many recombination events between the two parental genomes are represented that the two parental genomes have essentially been shuffled. The individual plants in the F_2 population are evaluated for the trait of interest, and their genotypes are established with the use of genetic and/or molecular markers. One of several possible statistical analyses is performed to identify associations between the trait of interest and specific alleles at various genetic loci. In the simplest scenario, all individuals with high values for the trait share a particular allele at a given locus, and the individuals with low values for the trait share the other allele at that locus.

The F_2 population should be evaluated in several locations and/or years to separate the genetic and environmental effects on the trait. The QTL is then mapped to a region of the chromosome. Oftentimes, several QTL are identified for a given quantitative trait, each representing a portion of the variance for the trait. A relatively dense marker map will help narrow down any chromosomal regions of interest as much as possible. The identification of a QTL will often lead to the identification of a candidate gene at that position. When the genome sequence is available, one or more candidate genes can be identified relatively easily and evaluated further. When the genome sequence is not (yet) available, map based cloning using, for example, bacterial artificial chromosomes (BACs) can be considered, as long as the map interval is small enough and sufficient resources are available. Further background on the theory and application of this approach, referred to as QTL mapping, can be found in the review by Doerge (2002).

4. ISOLATION AND CHARACTERIZATION OF RECOMBINANT PROTEINS

Knowing the sequence of the gene allows the isolation of the cDNA. Unlike the genomic DNA, cDNA's do not contain introns, which are stretches of non-coding sequences interspersed between the exonic (coding) DNA regions. The introns are removed after transcription through a process called splicing. Intron-exon boundaries and splicing mechanisms vary between different species, so that introduction of foreign genomic sequences into a particular species will not automatically result in the synthesis of a functional protein. Introduction and expression of cDNA sequences, however, generally does result in the synthesis of functional proteins. Such proteins are called recombinant proteins.

Recombinant protein can be obtained in large amounts via over-expression of the cDNA of interest in, for example, the bacterium *Escherichia coli*, the yeasts *Saccharomyces cerevisiae* and *Pichia pastoris*, or the Sf9 insect cell line of *Spodoptera frugiperda*. This is typically achieved by the introduction of a plasmid vector containing the cDNA. He plasmid is a circular DNA replicon that is maintained because it confers a selective advantage, such as an antibiotic resistance. The use of an appropriate promoter results in high expression levels of the cDNA. As a consequence, the abundance of recombinant proteins in these systems is typically much higher than that of the native protein in a crude plant extract. The higher abundance facilitates the purification of the protein to apparent homogeneity. Purification of the recombinant protein can be achieved through traditional biochemical separation techniques, including various types of chromatography (see Section 2). Many of the expression systems, however, are based on so called tags at the N- or C-terminus of the protein. These tags allow binding of the recombinant protein to a specially formulated resin. The crude cell extract containing the recombinant protein is loaded on a column containing the resin, and the recombinant protein becomes bound to the resin, typically *via* an electrostatic interaction. Undesirable proteins are subsequently removed by washing the column with a buffer that does not disrupt the binding of the recombinant protein. The recombinant protein is eventually eluted with a buffer that disrupts the interaction between the recombinant protein and the resin.

An example of the use of tags to purify recombinant proteins is the addition of six histidine (His) residues to a recombinant protein. The His-residues are encoded by the DNA of the expression vector and will be added to the protein during translation. The purification of the recombinant protein is achieved through immobilized metal affinity chromatography (IMAC). Since His can chelate transition metal ions such as Ni^{2+}, Zn^{2+}, and Cu^{2+}, the His-tag will allow the recombinant protein to bind to a resin containing Ni^{2+} ions, such as Ni^{2+}-Sepharose. Use of Ni^{2+} appears to be more successful than the other ions. The recombinant protein is eluted from the column by washing with imidazole, which competes for the Ni^{2+} binding sites. Alternatively, a low-pH buffer can be used to elute the recombinant protein. The low pH disrupts the electrostatic interactions between the recombinant protein and the Ni^{2+} on the column, but has the risk of denaturing the protein of interest. In general the presence of the six His-residues has no impact on protein function, but the tag can be removed by endoproteolytic cleavage.

An alternative method to purify recombinant proteins, similar in principle to the His-tag described above, is the fusion with glutathione S-transferase

(GST). This is an enzyme that has high affinity for glutathione. The crude cell extract containing the recombinant GST-fusion protein is loaded on a column of a resin containing glutathione (such as glutathione Sepharose). Undesirable proteins are removed by washing, and the recombinant GST-fusion protein is eluted by washing the resin with reduced gluthathione, which competes with the recombinant protein. The enzyme thrombin is used to specifically remove the GST-tag from the purified recombinant protein. Plasmid vectors to synthesize recombinant proteins with His-tags or GST-fusion proteins, as well as various other types of tags are available from various biochemical supply companies.

Data from *in vitro* activity assays with these purified recombinant proteins can typically be interpreted much more easily than data obtained from experiments with crude or partially purified protein extracts, because (1) there will be no competing proteins with similar activity present in the assay, and (2) there will no enzymes present that convert the product generated by the enzyme of interest, and hence reduce the effective product concentration. A potential downside of the use of recombinant protein over crude extracts is the fact that critical co-factors that will ensure proper activity may not be present in the purified protein fraction. If that is the case, the researcher will have to empirically determine which co-factor and at what concentration needs to be included in the assay. Another consideration is that the native protein may have undergone post-translational processing, such as acetylation, glycosylation, myristoylation, etc. These modifications may not occur or may not occur properly when the protein is expressed in bacterial, fungal or insect cells. Assuming that these potential problems do not occur or can be dealt with, the availability of pure recombinant protein will enable the determination of substrate specificity, as well as kinetic experiments in which the rate of conversion is measured as a function of time and/or substrate concentration.

5. CARBOHYDRATE CATABOLISM

All plant carbohydrates are derived from photosynthesis, the process during which CO_2 from the atmosphere is fixed and converted to carbohydrates with energy from light. The carbohydrates that are generated during photosynthesis form the building blocks for all other carbon-based compounds in the cell, including the phenolic compounds. The precursors for plant phenolic compounds are derived from two catabolic processes in the plant cell: glycolysis and the pentose phosphate pathway. For detailed

background on carbohydrate metabolism, please refer to Dennis and Blakeley (2000) or Plaxton and McManus (2006).

5.1 Glycolysis

Glycolysis, also known as the Embden-Meyerhof-Parnas pathway, is the catabolic process during which carbohydrates generated during photosynthesis are broken down to pyruvate, and ultimately CO_2 (Figure 3-1). This process fulfils two fundamental roles: It oxidizes hexoses to generate ATP, reductant, and pyruvate, and it produces building blocks for anabolism. Glycolysis in plants has been reviewed extensively by Plaxton (1996).

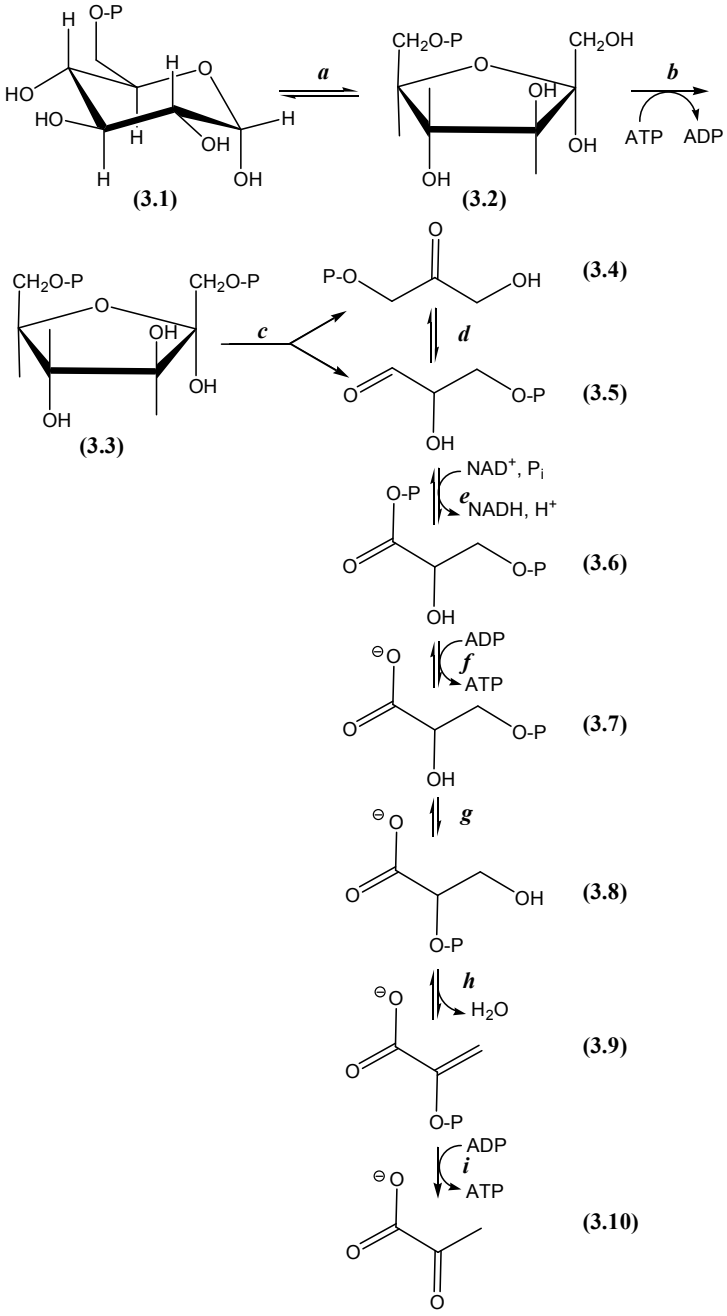
Glycolysis starts with the conversion of glucose-6-phosphate (**3.1**) to fructose-6-phosphate (**3.2**) by the enzyme hexose phosphate isomerase. Glucose-6-phosphate (**3.1**) can be generated from glucose via hexokinase, or from glucose-1-phosphate resulting from hydrolysis of sucrose or starch. In the latter case, phosphoglucomutase converts glucose-1-phosphate to glucose-6-phosphate. The pentose phosphate pathway (see Section 5.2) generates fructose-6-phosphate (**3.2**), which can enter glycolysis at this point.

Fructose-6-phosphate (**3.2**) is converted to fructose-1,6-bisphosphate (**3.3**) by phosphofructokinase. This compound is subsequently broken down by fructose bisphosphate aldolase to dihydroxyacetone-phosphate (DHAP; **3.4**) and glyceraldehyde-3-phosphate (GA3P; **3.5**). DHAP is converted to GA3P by triose-phosphate isomerase. In the subsequent reaction GA3P is converted to glycerate-1,3-bisphosphate (**3.6**) by GAP-dehydrogenase. Glycerate-1,3-bisphosphate is converted by glycerate-3-phosphate kinase to glycerate-3-phosphate (**3.7**), which is subsequently converted to glycerate-2-phosphate (**3.8**) by glycerate phosphate mutase. An enolase, phosphopyruvate hydratase, converts glycerate-2-phosphate to phospho*enol*pyruvate (PEP; **3.9**). As a final step, PEP is converted to pyruvate (**3.10**) by pyruvate kinase. Alternatively, PEP can enter the shikimate pathway (see Section 6).

5.2 The pentose phosphate pathway

The plant can also use the pentose phosphate pathway to break down glucose, but the main purpose of this pathway is the generation of reducing power in the form of NADPH (nicotinamide adenine dinucleotide

phosphate). In addition, this pathway provides sugar intermediates that can serve as building blocks for aromatic amino acids and nucleic acids.



The pentose phosphate pathway can be divided in two phases: an oxidative phase during which glucose-6-phosphate is converted to ribulose-5-phosphate, and a non-oxidative phase constituting of a series of reversible reactions in which two pentose-phosphate residues are converted to a series of sugar-phosphate molecules of differing lengths (Figure 3-2).

The oxidative part of the pentose phosphate pathway starts with the oxidation of glucose-6-phosphate (**3.1**) to gluconolactone-6-phosphate (**3.11**) by glucose-6-phosphate 1-dehydrogenase with the reduction of $NADP^+$ to NADPH. Gluconolactone-6-phosphate (**3.11**) is converted to gluconate-6-phosphate (**3.12**) by gluconate-6-phospate lactonase. Irreversible oxidative decarboxylation of (**3.12**) by gluconate-6-phosphate reductase results in ribulose-5-phosphate (**3.13**), with the generation of another NADPH molecule.

Ribulose-5-phosphate (**3.13**) can be converted to ribose-5-phosphate (**3.14**) and xylulose 5-phosphate (**3.15**), by the enzymes ribose-5-phosphate isomerase and ribulose 5-phosphate 3-epimerase, respectively. The two pentose-phosphate molecules, **3.14** and **3.15**, are converted to a C_3 and a C_7 sugar-phosphate, glyceraldehyde 3-phosphate (**3.4**) and sedoheptulose-7-phosphate (**3.16**), respectively, *via* the action of a transketolase.

Transketolases are characterized by their ability to transfer a two-carbon unit from a ketose to an aldehyde. The C_3 and C_7 sugar-phosphates can subsequently be converted to a C_4 and a C_6 sugar-phosphate, erythrose 4-phosphate (**3.17**) and fructose 6-phosphate (**3.2**), respectively. This reaction is catalyzed by a transaldolase, which transfers a three-carbon glyceraldehyde unit from an aldose to a ketose. Erythrose-4-phosphate (**3.17**) can be used in the shikimate pathway (see Section 6). A second transketolase reaction can generate a second fructose-6-phosphate (**3.2**) and glyceraldehyde-3-phosphate (**3.4**) residue from erythrose-4-phosphate (**3.17**) and xylulose-5-phosphate (**3.15**). Hexose-phosphate isomerase converts the

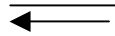
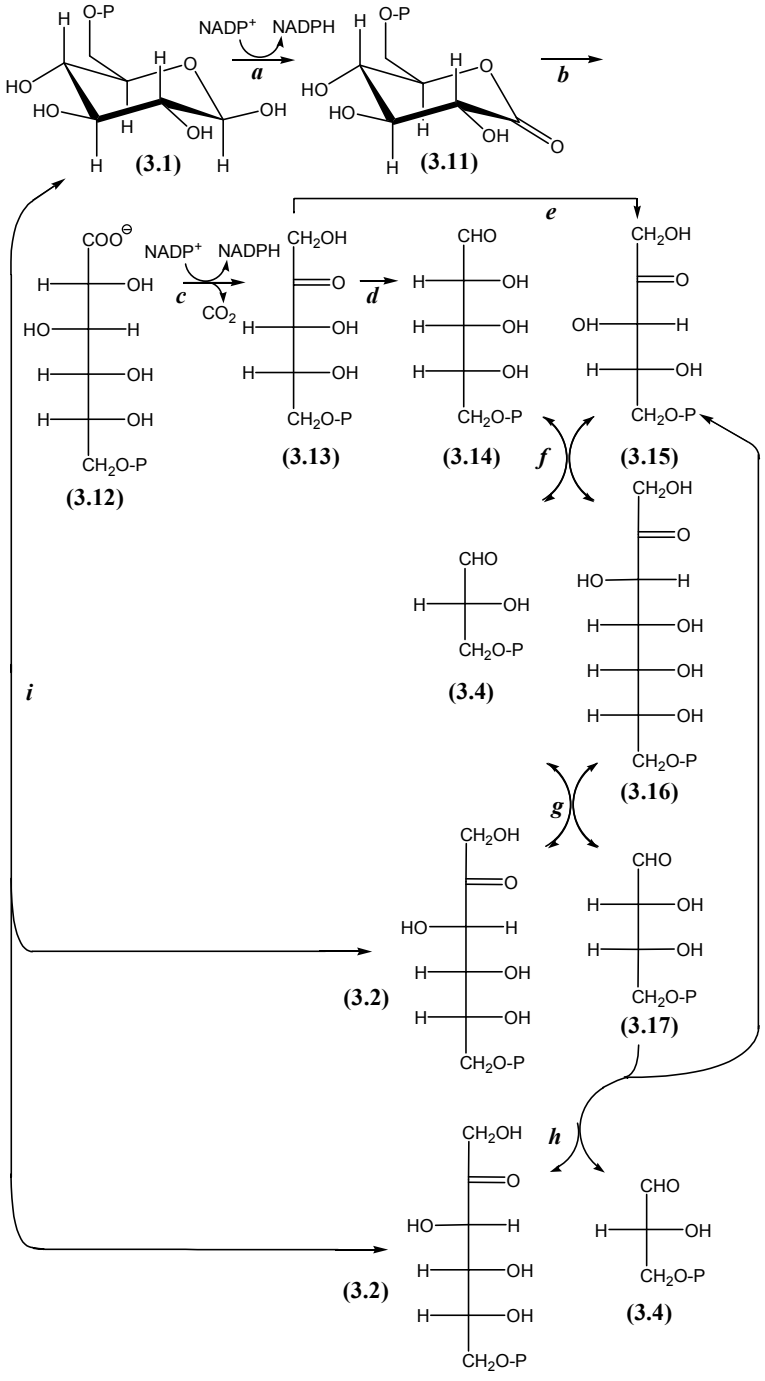


Figure 3-2. Glycolysis. The enzymes involved in this pathway are: **(*a*)** hexose phosphate isomerase (E.C. 5.3.1.9), **(*b*)** phosphofructokinase (E.C. 2.7.1.1), **(*c*)** fructose bisphosphate aldolase (E.C. 4.1.2.13), **(*d*)** triose-phosphate isomerase (E.C. 5.3.1.1), **(*e*)** GAP-dehydrogenase (E.C. 1.2.1.12), **(*f*)** glycerate-3-phosphate kinase (E.C. 2.7.2.3), **(*g*)** glycerate phosphate mutase (E.C. 5.4.2.1), **(*h*)** enolase (phosphopyruvate hydratase; E.C. 4.2.1.11), and **(*i*)** pyruvate kinase (E.C. 2.7.1.40).


NADP+
NADPH
a
b
(3.1)
(3.11)
e
c
CO2
d
(3.12)
(3.13)
(3.14)
f
(3.15)
(3.4)
(3.16)
g
(3.17)
i
(3.2)
h
(3.2)
(3.4)

two fructose-6-phosphate molecules to glucose-6-phosphate (**3.1**), which can enter the pentose phosphate pathway again to generate additional NADPH.

So in summary, three glucose-6-phosphate (**3.1**) molecules (3 x C_6) are oxidized to three ribulose-5-phosphate (**3.13**) residues (3 x C_5) and three molecules of CO_2 (3 x C_1) under generation of six molecules of NADPH. The three ribulose-5-phosphate residues are then converted to one glyceraldehyde-3-phosphate (**3.14**) molecule (1 x C_3) and two fructose-6-phosphate (**3.2**) molecules (2 x C_6). Fructose-6-phosphate can be converted to glucose-6-phosphate and reenter the oxidative part of the pentose phosphate pathway. Fructose-6-phosphate and glyceraldehydes can also serve as intermediates in glycolysis (Section 5.1), which offers the cell considerable flexibility in terms of its metabolic flux.

The availability of the Arabidopsis genome sequence revealed that there are multiple genes encoding the different enzymes in the oxidative pentose phosphate pathway (reviewed by Kruger and von Schaewen, 2003). Studies across a range of species indicate that genes encoding individual isozymes may be differentially expressed in different tissues, at different developmental stages, and in response to different growth conditions, especially those that alter demand for NADPH or intermediates of the oxidative pentose phosphate pathway for biosynthesis. The ability to use specific isoforms that allow optimal performance under certain conditions offers the plant a greater degree of metabolic flexibility.

6. THE SHIKIMATE PATHWAY

The shikimate pathway results in the biosynthesis of chorismate, which can subsequently serve as a recursor for the biosynthesis of the aromatic amino acids tryptophan, phenylalanine and tyrosine. The biochemistry of

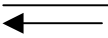


Figure 3-3. The pentose phosphate pathway. The enzymes involved in this pathway are: **(*a*)** glucose-6-phosphate dehydrogenase (E.C. 1.1.1.49), **(*b*)** gluconate-6-phospate lactonase (E.C. 3.1.1.31), **(*c*)** gluconate-6-phosphate reductase (E.C. 1.1.1.44), **(*d*)** ribose-5-phosphate isomerase (E.C. 5.3.1.6), **(*e*)** ribulose5-phosphate 3-epimerase (E.C. 5.1.3.1.), **(*f*)** transketolase (E.C. 2.2.1.1), **(*g*)** transaldolase (E.C. 2.2.1.2), **(*h*)** transketolase, and **(*i*)** hexose-phosphate isomerase (E.C. 5.3.1.9).

the shikimate pathway has been extensively reviewed by Weaver and Herrmann (1997) and Hermann and Weaver (1999).

The shikimate pathway is common to both plants and microorganisms (Figure 3-3). Shikimate is synthesized from the substrates phospho*enol*pyruvate (**3.9**) and erythrose 4-phosphate (**3.17**). These two precursors are derived from glycolysis and the pentose phosphate pathway, respectively, and are condensed to 3-deoxy-D-*arabino*-heptulosonate 7-phosphate (DAHP; **3.18**) by the enzyme DAHP synthase. The subsequent steps result in the formation of 3-dehydroquinate (**3.19**) by the enzyme 3-dehydroquinate synthase, 3-dehydroshikimate (**3.20**) by the enzyme 3-dehydroquinate dehydratase, and finally shikimate (**3.21**) by the enzyme shikimate dehydrogenase.

Shikimate is further converted to shikimate 3-phosphate (**3.22**) by shikimate kinase, and subsequently to 5-*enol*pyruvylshikimate 3-phosphate (EPSP; **3.23**) by 5-*enol*pyruvylshikimate 3-phosphate synthase. EPSP is then converted to chorismate (**3.24**) by chorismate synthase.

Chorismate is at a branch point for the biosynthesis of aromatic amino acids: tryptophan on the one hand, and phenylalanine (**3.27**) and tyrosine (**3.28**) on the other hand. While this is strictly speaking no longer part of the shikimate pathway, the biosynthesis of phenylalanine and tyrosine is included in Figure 3-3, because they are the precursors of the important class of phenolic compounds, the phenylpropanoids, as well as several other classes of phenolic compounds. This requires the conversion of chorismate to prephenate (**3.25**), catalyzed by chorismate mutase, and arogenate (**3.26**), catalyzed by prephenate aminotransferase. The enzyme arogenate dehydratase converts arogenate to phenylalanine (**3.27**), whereas the enzyme arogenate dehydrogenase generates tyrosine (**3.28**).

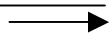
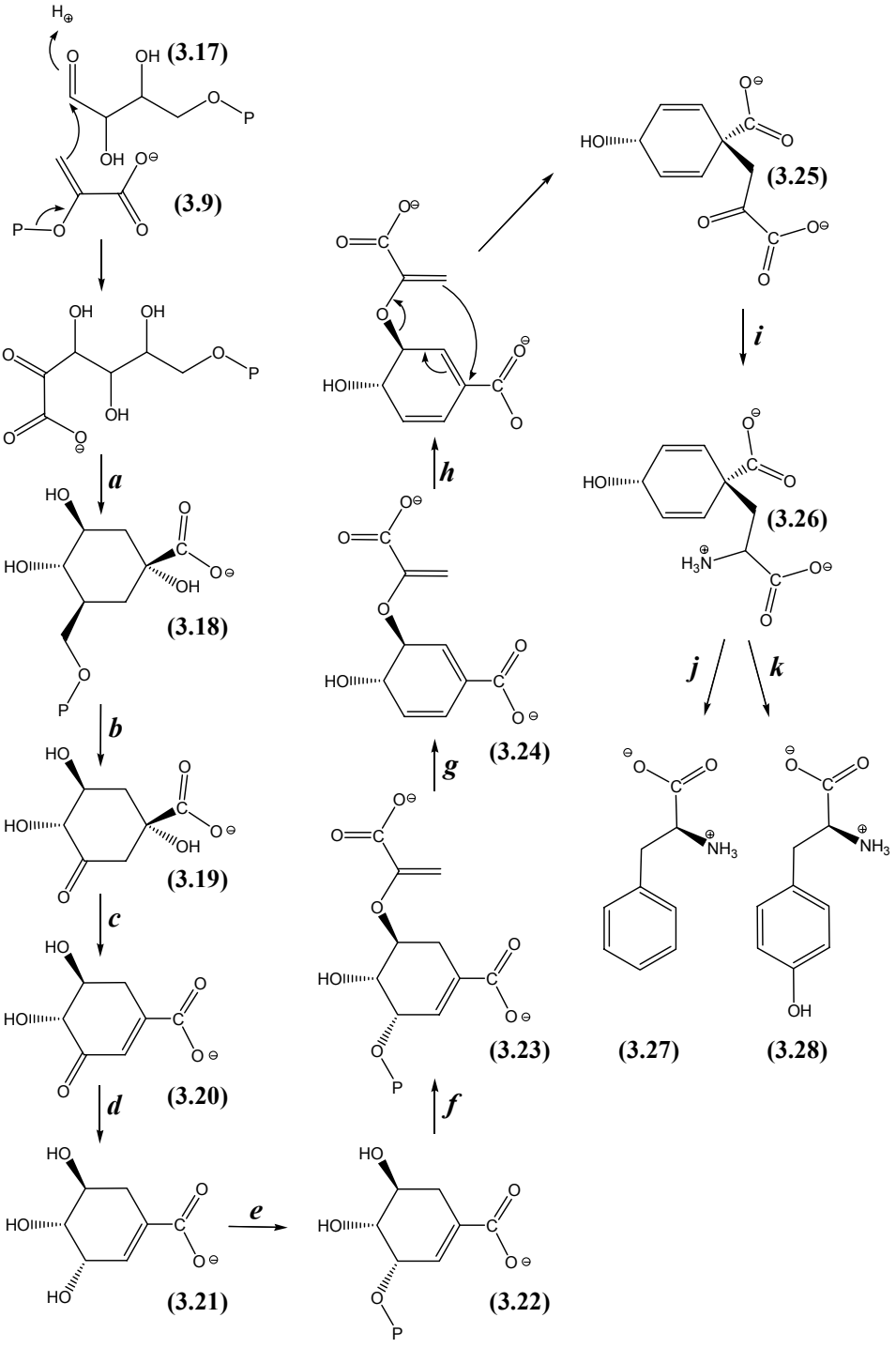


Figure 3-3. The shikimate pathway. The enzymes involved in this pathway are: **(*a*)** DAHP synthase (E.C. 2.5.1.54), **(*b*)** 3-dehydroquinate synthase (E.C. 4.2.3.4), **(*c*)** 3-dehydroquinate dehydratase (E.C. 4.2.1.10), **(*d*)** shikimate dehydrogenase (E.C. 1.1.1.25), **(*e*)** shikimate kinase (E.C 2.7.1.71), **(*f*)** 5-*enol*pyruvylshikimate 3-phosphate synthase (E.C. 2.5.1.19), **(*g*)** chorismate synthase (E.C. 4.2.3.5), **(*h*)** chorismate mutase (E.C. 5.4.99.5), **(*i*)** prephenate aminotransferase (E.C. 2.6.1.78 and E.C. 2.6.1.79), **(*j*)** arogenate dehydratase (E.C. 4.2.1.91), and **(*k*)** arogenate dehydrogenase (E.C. 1.3.1.43, E.C. 1.3.1.78, E.C. 1.3.1.79).


(3.17)
(3.9)
(3.18)
(3.19)
(3.20)
(3.21)
(3.22)
(3.23)
(3.24)
(3.25)
(3.26)
(3.27)
(3.28)
a
b
c
d
e
f
g
h
i
j
k

7. THE GENERAL PHENYLPROPANOID PATHWAY

The general phenylpropanoid pathway (Figure 3-4), as the name implies, generates a substrate common to a number of phenylpropanoid compounds, including flavonoids, monolignols, hydroxycinnamic acids, sinapoyl esters, coumarins and stilbenes. The general phenylpropanoid pathway starts with phenylalanine (**3.27**) generated via the shikimate pathway (see Section 6). Deamination of phenylalanine is catalyzed by the enzyme phenylalanine ammonia lyase (PAL) and results in cinnamic acid (**3.29**). Cinnamic acid is subsequently hydroxylated by cinnamic acid 4-hydroxylase (C4H) to give *p*-coumaric acid (**3.30**). In graminaceous species, such as maize, this compound can also result from the deamination of tyrosine (**3.28**). *In vitro* assays with recombinant enzyme demonstrated that the catalytic activity towards tyrosine resides in the same enzyme, in other words, in grasses PAL has activity against both phenylalanine and tyrosine (Roesler et al., 1997). *p*-Coumaric acid (**3.30**) is converted to *p*-coumaroyl Coenzyme A (**3.31**) by the enzyme 4-coumaric acid:CoA ligase (4CL). The general phenylpropanoid pathway ends with *p*-coumaroyl Coenzyme A (**3.31**), as further reactions lead to the biosynthesis of specific classes of compounds.

Older text books and articles typically describe the general phenylpropanoid pathway as the pathway leading to the full set of hydroxycinnamic acids (*p*-coumaric acid (**3.30**), caffeic acid (**3.32**), ferulic acid (**3.33**), 5-hydroxyferulic acid (**3.34**) and sinapic acid (**3.35**), as well as their corresponding CoA-esters (**3.31, 3.36-3.39**). This is indicated by the grey structures in Figure 3-4. Recent advances in the cloning and characterization of genes encoding enzymes involved in phenylpropanoid metabolism have made it possible to determine substrate specificity and catalytic activity of a number of enzymes. This provided evidence

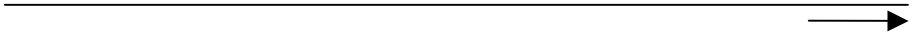
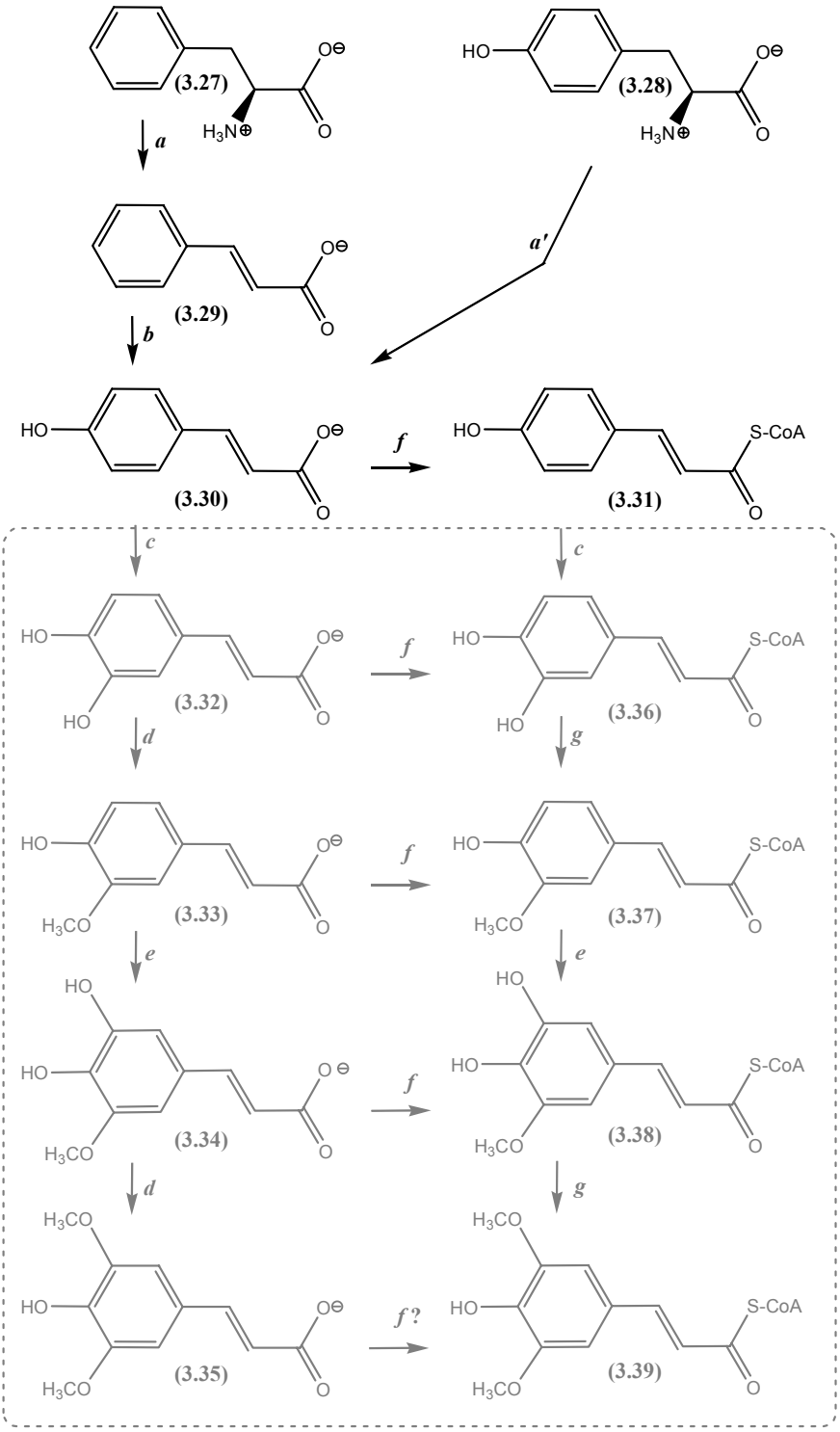


Figure 3-4. The general phenylpropanoid pathway. The enzymes involved in this pathway are: **(*a*)** phenylalanine ammonia lyase (PAL; E.C. 4.3.1.5), **(*b*)** cinnamic acid 4-hydroxylase (C4H; E.C. 1.14.13.11), and **(*f*)** 4-coumaric acid:CoA ligase (4CL; E.C. 6.2.1.12). **(*a'*)** depicts tyrosine ammonia lyase activity in PAL of graminaceous species. The grey structures in the box represent an older version of the phenylpropanoid pathway in which the ring substitution reactions were thought to occur at the level of the hydroxycinnamic acids and/or hydroxycinnamoyl esters. The enzymes involved in these conversions are (*c*) coumarate 3-hydroxylase (C3H; E.C. 1.14.14.1), (*d*) caffeate *O*-methyltransferase (COMT; EC 2.1.1.68), (*e*) ferulate 5-hydroxylase (F5H; EC 1.14.13), and (*g*) caffeoyl-CoA *O*-methyltransferase (CCoA-OMT; EC 2.1.1.104). These enzymes are discussed in more detail in Section 10.


(3.27)
(3.28)
(3.29)
(3.30)
(3.31)
(3.32)
(3.33)
(3.34)
(3.35)
(3.36)
(3.37)
(3.38)
(3.39)
a
a'
b
c
d
e
f
g
f?
HO
H3N⊕
O⊖
O
S-CoA
H3CO

against the hydroxylation and methylation reactions occurring at the level of the hydroxycinnamic acids. Rather, these reactions appear to occur at the level of the hydroxycinnamoyl esters, aldehydes, and alcohols (reviewed by Humphreys and Chapple, 2002). The hydroxycinnamic acids are thought to be generated from cinnamaldehydes, as will be discussed in Section 12. As a consequence, the current model of the general phenylpropanoid pathway is more streamlined, as shown in black in Figure 3-4. There is some evidence, however, that in certain plant species, specifically poplar (Meyermans et al., 2000), variations of the general phenylpropanoid pathway exist that are different from the general model depicted here.

8. BIOSYNTHESIS OF PHENOLIC ACIDS

Phenolic acids are generally not abundant in most plants. There are a few exceptions: gallic acid (**1.5**) and salicylic acid (SA; **1.8**). Gallic acid is a precursor for the ellagitannins and gallotannins, which will be discussed in more detail in section 15 of this chapter. Salicylic acid is an important defense compound because it mediates systemic acquired resistance (SAR), a resistance mechanism whereby SA is used as a signaling molecule to relay information on pathogen attack to other parts of the plant. Upon receiving the SA signal, a general defense response is activated that includes the biosynthesis of pathogenesis-related (PR) proteins. This will be discussed in more detail in Chapter 6, Section 2.4.

8.1 Salicylic acid

There are two possible biosynthetic routes to SA, one plant-specific, and the other more similar to the one found in bacteria (Figure 3-5). The 'plant-specific' biosynthetic route to SA starts with phenylalanine (**3.27**), which, as part of the general phenylpropanoid pathway (see Section 7), is converted to cinnamic acid (**3.29**) by the enzyme phenylalanine ammonia lyase (PAL). Cinnamic acid is converted to benzoic acid (**3.40**), probably through a process similar to β–oxidation of fatty acids. The hydroxylation of C2, catalyzed by the enzyme benzoic acid 2-hydroxylase (BA2H), results in SA (**3.41**). Characterization of purified tobacco BA2H showed that this enzyme is a cytochrome P450-dependent oxygenase (Léon et al., 1995). Alternatively, BA2H may hydroxylate cinnamic acid (**3.29**) on C2 to produce coumaric acid (**3.42**), which, after oxidation of the propane side chain, results in SA (**3.41**).

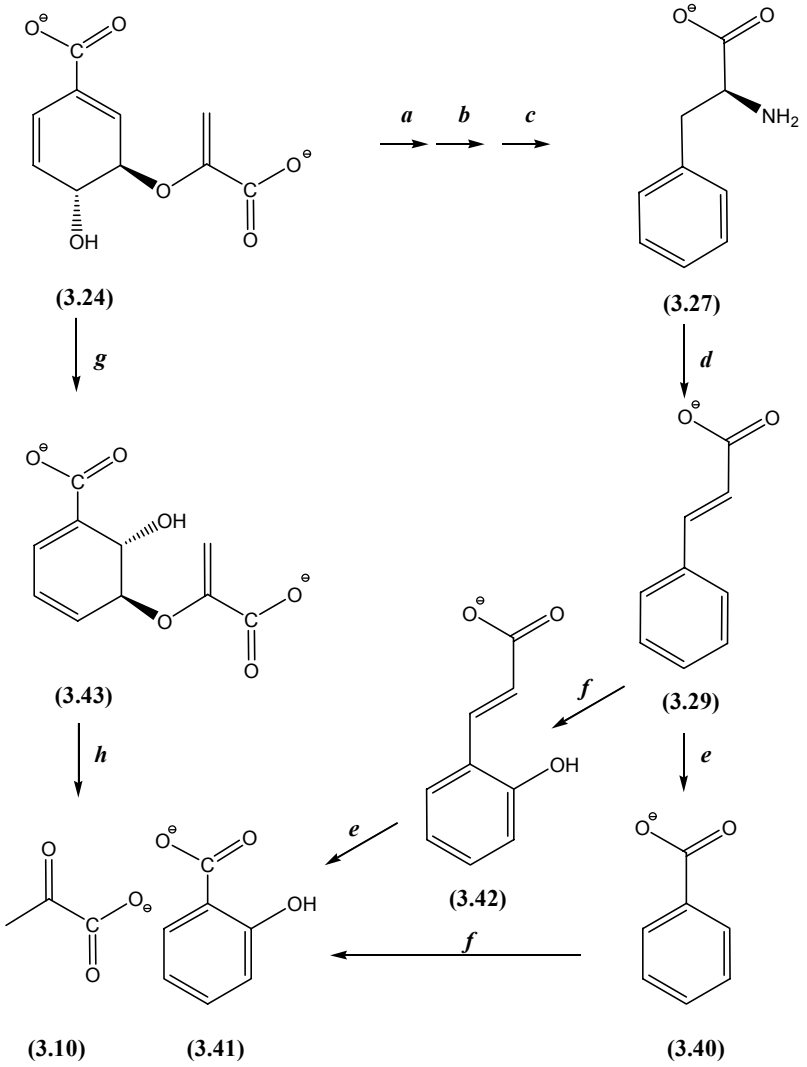


Figure 3-5. Biosynthesis of salicylic acid. The enzymes involved in this pathway are: **(*a*)** chorismate mutase (E.C. 5.4.99.5), (***b***) prephenate aminotransferase (E.C. 2.6.1.78 and E.C. 2.6.1.79), (***c***) arogenate dehydratase (E.C. 4.2.1.91), **(*d*)** phenylalanine ammonia lyase (E.C. 4.3.1.5), **(*e*)** presumed β–oxidation by a yet to be identified enzyme, **(*f*)** benzoic acid 2-hydroxylase, **(*g*)** isochorismate synthase (E. C. 5.4.4.2), and (***h***) a putative plant pyruvate lyase.

An alternative biosynthetic pathway towards SA was hypothesized to exist after experiments with labeled benzaldehyde, benzyl alcohol, and phenylalanine resulted in lower incorporation of the label in SA than

expected (Ribnicky et al., 1998). Sequence analysis of the Arabidopsis genome revealed two genes with homology to the bacterial gene encoding isochorismate synthase, an enzyme involved in SA production in bacteria such as *Pseudomonas aeruginosa*. One of these genes, *ICS1* was shown to be up-regulated when Arabidopsis leaves were infected with bacterial pathogens (Wildermuth et al., 2001). In addition, two mutants that were lacking a functional copy of the *ICS1* gene produced less SA, exhibited a reduced pathogenesis-related gene expression, and were more susceptible to pathogens. The ICS1 protein contains a putative plastid transit sequence and cleavage site that would be consistent with biosynthesis of SA from the plastid-synthesized pool of chorismate (**3.24**). In bacteria isochorismate (**3.43**) is converted to SA and pyruvate (**3.10**) by the enzyme pyruvate lyase. It is yet unclear whether plants would use a similar enzyme. Wildermuth et al. (2001) speculated that the SA synthesized by isochorismate synthase is crucial for SAR, whereas SA synthesized from phenylalanine may be important in mediating cell necrosis in response to certain pathogens or fungal elicitors.

8.2 Gallic acid

The biosynthesis of gallic acid (**3.47**) has been under investigation for more than 50 years. Different biosynthetic routes have been proposed, as depicted in Figure 3-6: (***1***) direct biosynthesis from an intermediate of the shikimate pathway, (***2***) biosynthesis *via* phenylalanine (**3.27**), cinnamic acid (**3.29**), *p*-coumaric acid (**3.30**), caffeic acid (**3.32**), and 3,4, 5-trihydroxycinnamic acid (**3.44**), or (***3***) biosynthesis via caffeic acid (**3.32**) and protocatechuic acid (**3.45**). The possibility that different pathways co-existed in different species or even within one species was also considered.

Werner et al. (1997) used ^{13}C-labeled glucose in feeding experiments with the fungus *Phycomyces blakesleeanus* and in leaves of sumac (*Rhus typhina*). After incubation with the labeled glucose, gallic acid and aromatic amino acids were isolated and subjected to NMR analyses. The NMR data indicated that in both the fungus and the plant the carbon atoms in gallic acid were biosynthetically similar to those in shikimate (**3.21**), and different from those in phenylalanine (**3.27**) and tyrosine (**3.28**). Based on these data 5-dehydroshikimate (**3.46**) was proposed as the most likely precursor of gallic acid. Ossipov et al. (2003) indeed isolated an enzyme from birch (*Betula pubescence*) that was able to reduce 5-dehydroshikimate to gallic acid.

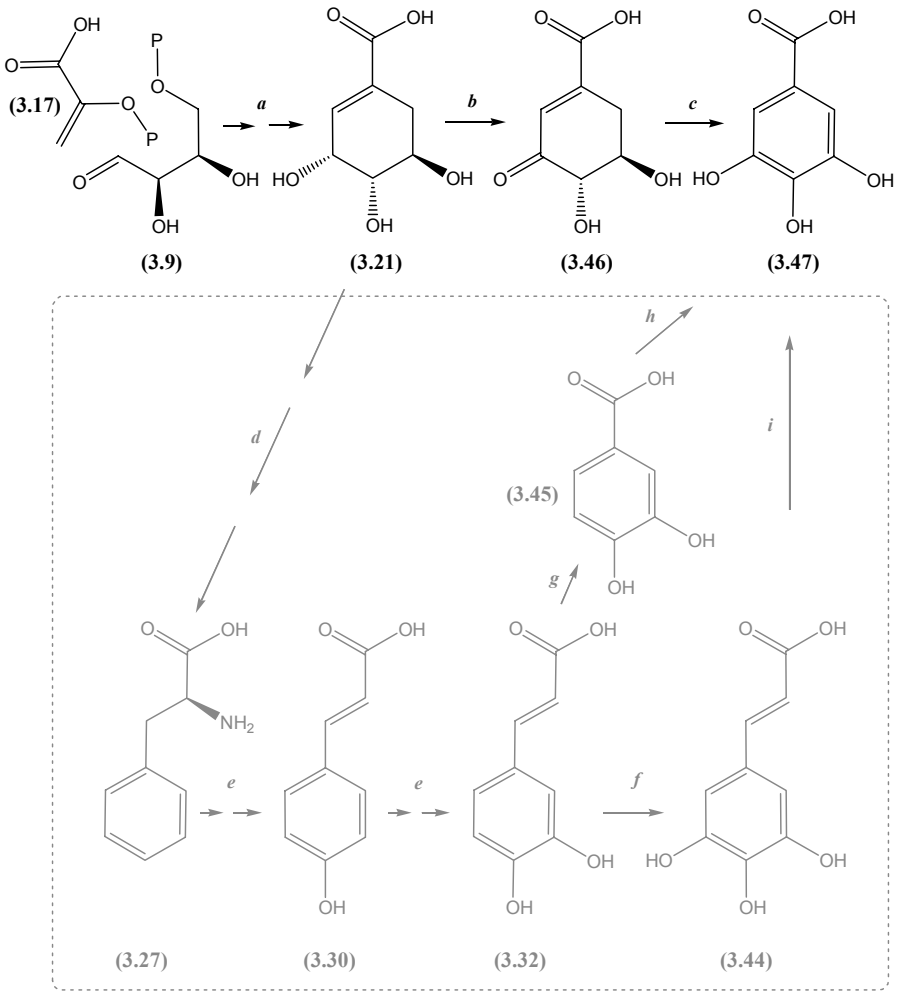


Figure 3-6. Biosynthesis of gallic acid in sumac (*Rhus typhina*). The enzymes involved in this pathway are: (***a***) enzymes involved in the shikimate pathway (see Figure 3-3), (***b***) proposed shikimate dehydrogenase, (***c***) proposed 5-dehydroshikimate dehydrogenase identified in birch by Ossipov et al. (2003), (***d***) enzymes involved in the biosynthesis of phenylalanine (see Figure 3-3), (***e***) enzymes involved in the general phenylpropanoid pathway (see Figure 3-4), (***f***) proposed caffeic acid 5-hydroxylase, (***g***) and (***i***) oxidizing enzymes that reduce the propane side chain from C3 to C1, (***h***) proposed protocatechuic acid 5-hydroxylase. The structures in black represent the pathway based on the experiments by Werner et al. (2004), which is responsible for > 90% of the gallic acid synthesis. The grey structures in the box represent alternative pathways that had been hypothesized to exist, but that do not contribute (significantly) to the biosynthesis of gallic acid.

Further evidence in favor of 5-dehydroshikimate (**3.46**) as a precursor of gallic acid came from a recent study by Werner et al. (2004) in which the ratio of oxygen isotopes ($^{16}O/^{18}O$) in gallic acid was measured. The ^{16}O isotope of oxygen is by far the most common (99.8 atom%), with the ^{17}O and ^{18}O isotopes representing 0.04 and 0.2 atom%, respectively. The isotope abundance (δ) in a biological compound is typically expressed as the relative difference of the isotope ratio of the compound to that of an international standard, and is expressed in ‰. In the case of ^{18}O, ocean water is used as the standard. Interestingly, the isotope abundance of ^{18}O varies among the main sources of oxygen in plant-based compounds. The $\delta^{18}O$ values of CO_2 and O_2 are approximately +40 and +24‰, whereas the value for groundwater varies between –10 and +2‰. In this case positive numbers are indicative of $^{18}O/^{16}O$ ratios higher than ocean water. Hence, $\delta^{18}O$ of a particular compound varies depending on its biosynthetic origin (Schmidt et al., 2001).

If gallic acid was synthesized from an intermediate of the shikimate pathway, the three phenolic oxygen atoms would originate from the carbohydrate erythrose 4-phosphate (**3.17**). In the alternative route *via* phenylalanine, the phenolic oxygen atoms would have to be introduced by a monooxygenase using atmospheric oxygen. Since carbohydrates and molecular oxygen have different $\delta^{18}O$ values, the $\delta^{18}O$ value for gallic acid will reflect the origin of its oxygen atoms. Werner et al. (2004) measured $\delta^{18}O$ of gallic acid and water isolated from sumac leaves using an isotope ratio mass spectrometer and calculated that the biosynthesis of gallic acid from 5-dehydroshikimate (**3.46**) was consistent with the experimental value, whereas biosynthesis from phenylalanine (**3.27**) was not.

9. BIOSYNTHESIS OF FLAVONOIDS AND CONDENSED TANNINS

9.1 Structural genes and enzymes

The identification and isolation of genes involved in flavonoid biosynthesis has benefited from the fact that many of the flavonoids are colored compounds. Mutant phenotypes are therefore often easily identifiable based on variation in color. In Arabidopsis many of the genes involved in flavonoid bioysnthesis have been uncovered based on the change in seed coat (testa) color. Wild-type Arabidopsis seeds have a brown color, and mutations in flavonoid biosynthetic genes result in yellow or pale brown color because the underlying cotyledons are visible. These mutants are referred to as

transparent testa (*tt*) mutants. A total of 21 of these mutants have been identified, resulting from either chemical mutagenesis with EMS, or ionizing radiation (X-ray or fast neutrons). This includes 19 *tt* mutants, and two *transparent testa glabra* (*ttg*) mutants, which have pale seeds but also lack trichomes (leaf hairs) (reviewed by Winkel-Shirley, 2001). The absence of flavonoids in the seed coat reduces seed dormancy, and some of the *tt* mutants were actually identified based on their reduced dormancy, as opposed to the seed coat color.

Maize mutants with altered flavonoid metabolism can also be identified based on variation in color, either of the seeds, the vegetative parts of the plant, or the floral structures (anthers and silks). Petunia (*Petunia hybrida*) and snapdragon (*Antirrhinum majus*) have also been widely used as model species for the elucidation of flavonoid biosynthesis (reviewed by Winkel-Shirley, 2001). In the description of genes involved in flavonoid biosynthesis presented in this section, the emphasis will be on maize and Arabidopsis.

Flavonoid biosynthesis (Figure 3-7) is initiated from the condensation of *p*-coumaroyl-CoA (**3.31**) with three molecules malonyl-CoA (**3.48**), which is catalyzed by the enzyme chalcone synthase (CHS), and gives rise to 4,2′,4′,6′ tetrahydroxychalcone (**3.49**). This compound can undergo a number of reactions that give rise to the different classes of compounds described in Section 3.6 of Chapter 1.

CHS in maize is encoded by the *Colorless2* (*C2*) gene. Mutations in this gene result in yellow kernels as a result of a colorless aleurone (Reddy and Coe, 1962). The aleurone is the cell layer under the pericarp (the hard outer cell layer of the maize kernel). The *C2* gene was cloned *via* transposon tagging with the *Spm* transposon by Wienand et al. (1986). Niesback-Klösgen et al. (1987) further characterized the gene. In Arabidopsis CHS is encoded by the *TT4* locus (Feinbaum and Ausubel, 1988).

The product generated by CHS, 4,2′,4′,6′ tetrahydroxychalcone (**3.49**) is the substrate for aurones, yellow pigments common in the petals of flowers, that contain a five-member ring and additional hydroxyl groups on the B-ring. An example is aureusidin (4,6,3′,4′-tetrahydroxyaurone; **3.50**). Aureusidin synthase, the enzyme responsible for the formation of aureusidin (**3.50**) was isolated from 32 kg of yellow snapdragon buds *via* a series of biochemical separations (Nakayama et al., 2000). Oligonucleotide primers based on partial amino acid sequence obtained from the isolated protein enabled the isolation of the corresponding cDNA clone. The cDNA was shown to encode a 64 kDa protein with similarity to polyphenoloxidases

from various plant species. The mature aureusidin synthase can be produced after cleavage of a 10 kDa N-terminal transit peptide, and a 15 kDa C-terminal peptide of unknown function. The mature aureusidin synthase is a 39 kDa copper-containing glycoprotein that catalyzes both the hydroxy lation of the B-ring and the oxidative cyclization of the 5-member ring characteristic for aurones. The compound 3,4,2′,4′,6′ pentahydroxychalcone was shown to be a better substrate for the production of aureusidin (**3.50**) and the aurone bracteatin (4,6,3′,4′,5′-pentahydroxyaurone). Given the similarity of aureusidin synthase to polyphenol oxidases (PPO's), Nakayama et al. (2001) investigated the specificity of the enzyme, and concluded it was a highly specific PPO with substrate specificity for chalcones with a 4-mono or 3,4-dihydroxy substitution pattern.

The formation of a six-member ring from 4,2′,4′,6′ tetrahydroxychalcone (**3.49**), catalyzed by chalcone isomerase (CHI), results in the flavanone naringenin (**3.51**). In Arabidopsis CHI is encoded by the *TT5* locus (Shirley et al., 1992) Naringenin is subsequently converted by flavanone 3-hydroxylase (F3H) to yield the flavanonol dihydrokaempferol (**3.52**). This compound can be converted by flavone synthase (FLS) to the flavone kaempferol (**3.53**). Alternatively, the B-ring of dihydrokaempferol can be substituted with additional hydroxyl groups by flavonoid 3′-hydroxylase (F3′H) or flavonoid 3′,5′-hydroxylase (F3′5′H) to produce dihydroquercetin (**3.54**) and dihydromyricetin (**3.55**), respectively. These latter two compounds can subsequently be converted by FLS to their corresponding flavones, quercetin (**3.56**) and myricetin (**3.57**). F3H and F3′H are encoded by the Arabidopsis *TT6* and *TT7* genes, respectively (Pelletier and Shirley, 1996; Schoenbohm et al., 2000).

The flavanonols can also give rise to anthocyanins. This involves a multi-step process whereby they are first reduced to leucoanthocyanidins – leucopelargonidin (**3.58**), leucocyanidin (**3.59**) and leucodelphinidin (**3.60**) – by the enzyme dihydroflavonol 4-reductase (DFR) (Figure 3-7), followed by dehydration and glycosylation (Figure 3-8). The enzyme DFR is encoded by the maize *Anthocyaninless1* (*A1*) gene, based on the colorless aleurone layer in the seed, and lack of pigmentation of the green tissues of the plant in the *a1* mutant. This gene was cloned by O'Reilly et al. (1985) using transposon tagging, and its sequence was analyzed by Schwarz-Sommer et al. (1987). In Arabidopsis DFR is encoded by the *TT3* locus (Shirley et al., 1992).

Anthocyanidin synthase (ANS) is the enzyme that dehydrates the leucoanthocyanidins (Figure 3-8). While this enzyme had been postulated to exist, and cDNA's encoding the putative ANS had been identified (*e.g. A2* in

maize; Messen et al., 1990), *in vitro* ANS activity was only recently demonstrated in extracts from the beefsteak plant (*Perilla frutescens* (L.) Britt; Saito et al., 1999).

ANS is a 2-oxoglutarate-dependent oxygenase that is thought to abstract a hydrogen radical from C2 of leucoanthocyanidin (**3.61**) to yield the radical (**3.62**) (Figure 3-8). Following a second hydrogen abstraction at C3, the 2-flaven-3,4-diol (**3.63**) is formed. The reaction can also occur in the reverse order, *i.e.* abstraction of the hydrogen at C3 followed by the one on C2. The colorless 2-flavene-3,4-diol is hydrated to 3-flavene-2,3-diol (**3.64**), which under acid conditions can give rise to anthocyanidin (**3.65**). The glycosylation of anthocyanidins results in the formation of anthocyanins (**3.66**) and is catalyzed by UDP-glucose:flavonoid 3-*O*-glucosyltransferase, also referred to as anthocyanidin 3-glycosyl transferase (3GT). In maize this enzyme is encoded by the *Bronze1* (*Bz1*) gene (Dooner and Nelson, 1977; Larson and Coe, 1977), and mutations in this gene result in bronze aleurone in the seed and brown vegetative tissues. The *Bz1* gene was cloned *via* transposon-tagging with *Ac* by Federoff et al. (1984). The maize *bronze2* (*bz2*) mutant looks very similar to the *bz1* mutant. The *Bz2* gene encodes a glutathione S-transferase required for tagging anthocyanins, synthesized as a result of Bz1 activity, with glutathione. This modification appears necessary for transfer of anthocyanins into the vacuole. The *Bz2* gene was cloned via transposon-tagging with *Mu* and *Ds* by McLaughlin and Walbot (1987) and Theres et al. (1987), respectively.

Condensed tannins (**3.68**) arise from polymerization of flavonoids. Polymerization starts with the condensation of a 2,3-*cis*-flavonol residue (**3.67**) onto a 2,3-*trans*-flavonol 'starter' residue, after which additional 2,3-*cis*-flavonol residues polymerize. The biosynthesis of the monomers of condensed tannins was poorly understood until the discovery of the Arabidopsis *banyuls* mutant. This mutant, named after a French wine, displays transparent testa as a result of accumulation of red anthocyanins, and loss of condensed tannins in the seed coat (Albert et al., 1997). The *BANYULS* gene was cloned and shown to encode an enzyme with similarity to DFR (Devic et al., 1999). Thus, it was initially proposed that the *BANYULS* gene encoded leucoanthocyanidin reductase (LAR), which reduces leucoanthocyanidins to the 2,3-*trans*-flavonol 'starter' residue. This was subsequently disproved, as the product of the *BANYULS* gene did not show any activity towards leucoanthocyanidins, but instead was shown to use anthocyanidins as a substrate (Xie et al., 2003). The resulting products are 2,3-*cis*-flavonols (**3.67**). The biosynthetic origin of the 2,3-*trans*-flavonols remains to be elucidated, but several uncharacterized DFR-like genes

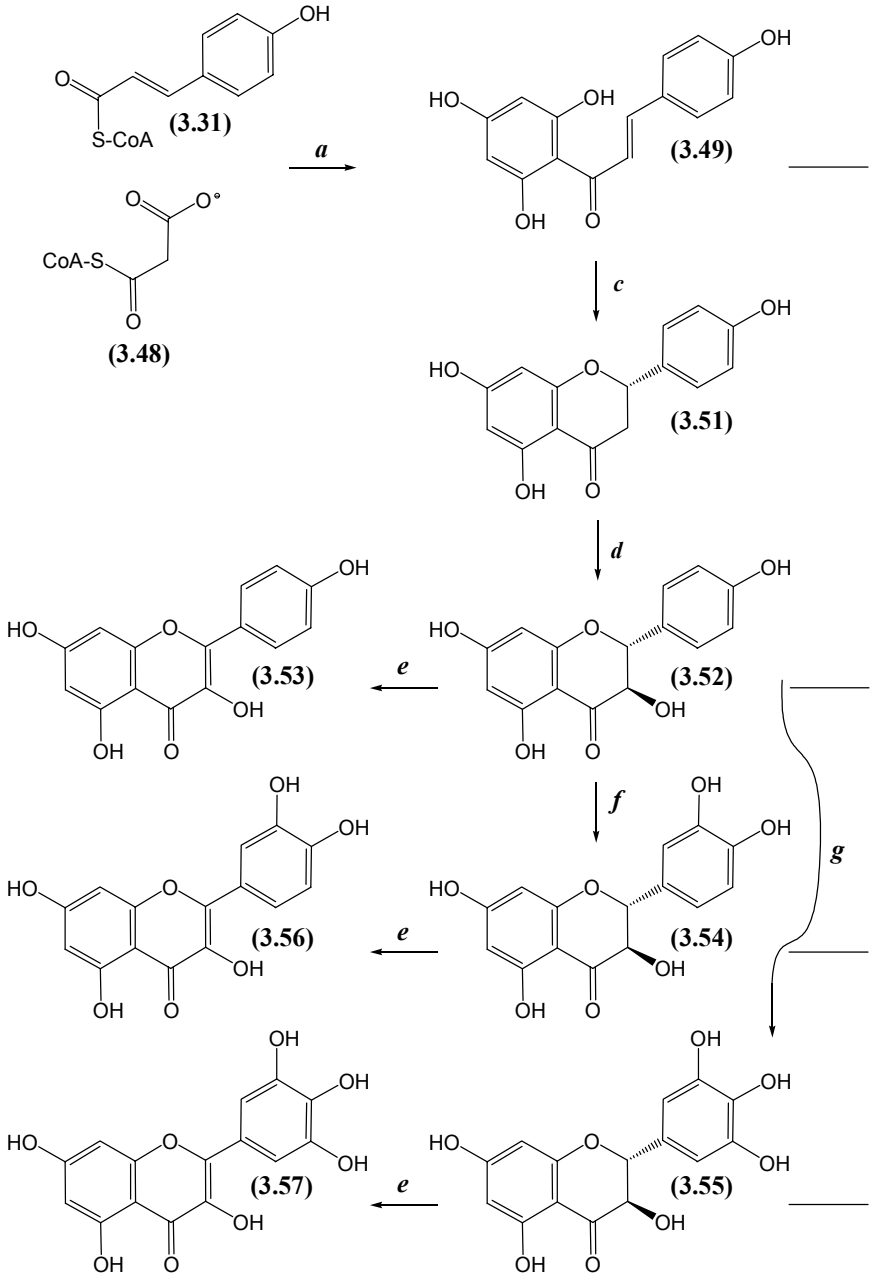
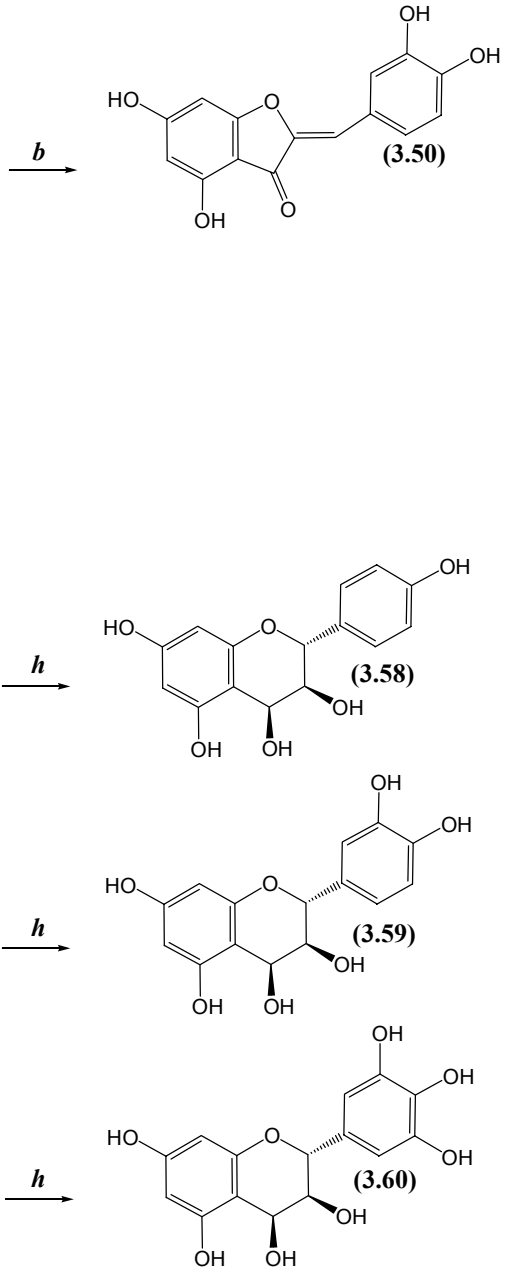


Figure 3-7. Flavonoid biosynthesis (this page and next page). The enzymes involved in this pathway are: **(*a*)** chalcone synthase (E.C. 2.3.1.73), **(*b*)** aureusidin synthase (E.C. 1.21.3.6), **(*c*)** chalcone isomerase (E.C. 5.5.1.6), **(*d*)** flavanone 3-hydroxylase (E.C. 1.14.11.9), **(*e*)** flavone synthase (E.C. 1.14.11.22), **(*f*)** flavonoid 3′-hydroxylase (E.C. 1.14.13.21),



(*g*) flavonoid 3′,5′-hydroxylase (E.C. 1.14.13.88), and **(*h*)** dihydroflavonol 4-reductase (E.C. 1.1.1.219).

have been identified in the Arabidopsis genome, one or more of which may encode LAR.

A detailed analysis of the Arabidopsis *tt12* mutant revealed that the vacuole in the endothelial cells (the innermost layer of the testa) accumulated lower levels of condensed tannins, and that instead the cytosol contained higher levels of these compounds. The *TT12* gene was cloned using T-DNA tagging. Sequence analysis revealed that the TT12 protein shows similarity to a multidrug secondary transporter of the MATE (multidrug and toxic compound extrusion) family. Based on the phenotype and the sequence similarity, the most likely role of the TT12 protein is that it functions as a vacuolar transporter for the precursors of condensed tannins (leucocyanidin and catechin), as it is very unlikely that the polymeric procyanidin itself may be handled by a transporter. Consequently, the pale brown color of the *tt12* seeds may be the result of the accumulation of precursors of condensed tannins in the cytoplasm (Debeaujon et al., 2001).

9.2 Regulatory genes

Efforts to elucidate the genetic basis of biosynthetic pathways tend to result in the identification of structural genes, *i.e.* genes encoding enzymes that catalyze the conversion of intermediates in a specific pathway. This is in part due to the relatively high expression of the structural genes, which results in abundant representation in expressed sequence tag (EST) libraries, or high probability of identification when cDNA subtraction methods or other methods that rely on differential expression are used.

In order to fully understand the biochemical pathway, it is also important to identify the regulatory mechanisms that control both the timing (developmental stage, environmental cues) and the location (cells, tissues or organs) of the biosynthesis of a particular compound or class of compounds. The regulation of biosynthetic pathways is mediated by regulatory genes that encode transcription factors and repressors that can enhance and inhibit the expression of the structural genes, and that are under control of internal (*e.g.* developmental) and/or external (*e.g.* environmental) cues. The identification of regulatory genes is generally much more difficult, in part because the expression of regulatory genes tends to be much lower than the expression of the structural genes.

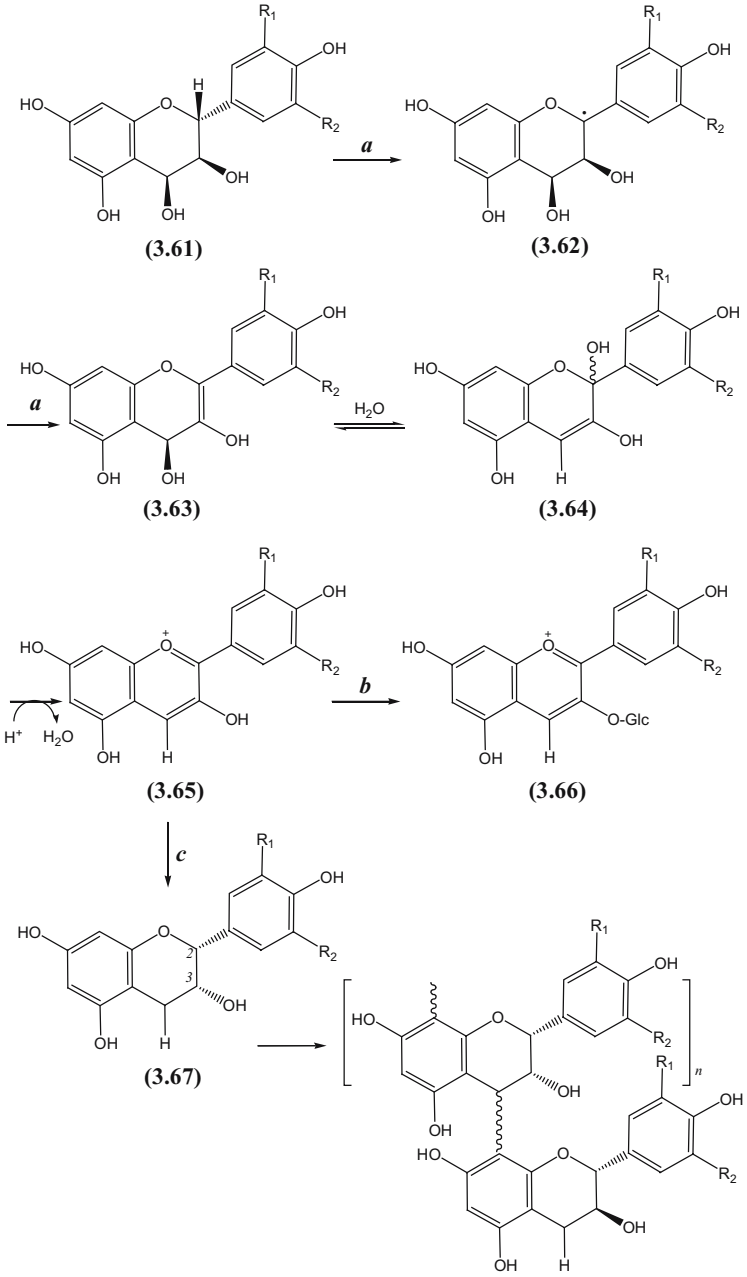


Figure 3-8. Biosynthesis of anthocyanins and condensed tannins. The enzymes involved in this pathway are: **(*a*)** anthocyanidin synthase (E.C. 1.14.11.19), **(*b*)** anthocyanin 3-glycosyl transferase, and **(*c*)** BANYULS.

Furthermore, there is considerable sequence similarity between transcription factors, even though they may operate in different biosynthetic pathways. Hence, based on sequence similarity it is often difficult to identify the target genes of a given transcription factor.

Interestingly, regulatory genes involved in the biosynthesis of anthocyanins were identified relatively early on. This was possible because of a number of mutants that accumulated high levels of anthocyanins, or that accumulated anthocyanins in tissues where they were normally not found, in combination with the easily scorable phenotype of these mutants.

The *Red color* (*R*) and *Booster* (*B*) genes are regulatory genes that control the tissue-specific deposition of anthocyanins in maize. These two genes were shown to independently activate the same target gene *Bz1* (Dooner and Nelson, 1977). In most maize lines the *R* gene is actually represented by a small family of homologous genes that map closely together and that are thought to have arisen through gene duplication and divergence. A functional *R* gene is required for the pigmentation of all plant tissues. More than 50 naturally occurring *R* alleles have been identified, which differ from each other in the spatial and temporal patterns of anthocyanin accumulation. The particular pattern of pigmentation displayed by a given plant is the result of the combined expression of all *R* family members that it contains. The standard *R* locus is responsible for pigmentation of the aleurone, anthers, and coleoptile. This phenotype is due to the expression of two tightly linked members of the *R* gene family, *S* and *P*. The *S* gene controls pigmentation of the aleurone of the kernel, whereas the *P* gene controls pigmentation of the anthers and coleoptile of the plant. Another member of the *R* gene family is the *Lc* gene, which conditions the pigmentation of the leaf midrib, the ligule and auricle (both of these tissue are at the boundary of the leaf blade and leaf sheath), several tissues in the male inflorescence (glume, lemma, palea), and the seed pericarp. The *R-nj* gene was cloned by transposon tagging (Dellaporta et al., 1988) and used to isolate genomic and cDNA clones of the *Lc* member of the *R* gene family. The *Lc* gene was shown to encode a transcriptional activator of the *myc* class. This transcription factor is required for the accumulation of transcripts of the *chalcone synthase* (*C2*) and *dihydroflavonol4-reductase* (*A1*) genes in the anthocyanin biosynthetic pathway (Ludwig et al., 1989). The *R-r* gene complex was characterized in detail at the molecular level by Robbins et al. (1991). It was shown to contain three complete *R* transcription units *P*, *S1*, *S2*, and one incomplete unit, *Q*. The *P* gene controls plant pigmentation, whereas the *S* genes control the pigmentation of the aleurone layer in the seed.

The *B* gene was cloned by Chandler et al. (1989) based on postulated sequence homology to the *R1* gene. A probe derived from the cloned *R1* gene was used to screen Southern blots obtained from a population of recombinant inbred lines in which two *B* alleles (*b* and *B*) were segregating. In addition to a strong hybridization signal corresponding to the *R1* gene itself, a weak hybridization signal was detected, the size of which varied depending on whether the *b* or the *B* allele was present. The *B-Peru* allele, conferring a blotched pigmentation pattern to the leaf sheaths, was subsequently cloned from a sub-genomic library. The coding sequence and the direction of transcription were determined. Chandler et al. (1989) also showed a direct correlation between the expression level of the *B* gene and the expression of the target genes *A1* and *Bz1*. This was based on expression analyses of maize plants homozygous for the *b* null allele, the weakly expressed *B-Peru* allele, and the highly expressed *B-I* allele. The *B-Peru* gene and the corresponding cDNA were subsequently sequenced (Radicella et al., 1991) and shown to encode a protein with similarity to the *myc* transcription factors. Comparison of the *B-Peru* allele with the sequence of the *B-I* allele indicated that the variation in expression pattern was the result of sequence variation in the promoter and the 5′ part of the gene (Radicella et al., 1992).

The maize *C1* (*Colorless1*) gene regulates the expression of the structural genes *C2*, *A1*, *Bz1*, *Bz2*, and *A2* in the seed. Presence of a dominant *C1* allele, in combination with functional copies of the abovementioned target genes results in seeds with colored (purple) aleurone. The *C1* gene was cloned independently by Cone et al. (1986) and Paz-Arez et al. (1986) using a transposon-tagging strategy, and shown to encode a transcription factor of the *myb* class (Paz-Ares et al., 1987). Aside from a number of mutant *c1* alleles that reduce or abolish the expression of the gene, a dominant mutant allele *C1-I* was identified. The protein encoded by the *C1-I* allele lacks 21 amino acid residues at the carboxy terminus of the protein and contains a mutation that results in an amino acid substitution (Paz-Ares et al., 1990). As a result of these two changes, the mutant C1-I protein can bind to the promoter regions of its target genes, but not activate transcription, and hence acts as a dominant inhibitor of the functional C1 protein. The similarity between the R and B proteins, both at the protein sequence level, and in molecular complementation studies, prompted Cone et al. (1993a) to investigate whether a similar relationship existed between the *C1* and *Purple plant* (*Pl*) genes. The *C1* gene determines the color of the aleurone, whereas the *Pl* gene determines the color of the vegetative and reproductive parts of the plant. Evidence in support of this hypothesis came from the existence of alleles for both genes that were only expressed after exposure to light. Cone et al. (1993a)

were thus able to clone the maize *Pl* gene with a DNA fragment from the *C1* locus as a probe to screen genomic and cDNA libraries. Based on sequence comparison of the *Pl* and *C1* cDNA's, more than 90% of the amino acids at the amino and carboxyl terminal domains that are important for the regulatory function of the C1 protein were shown to be identical. The difference between the light-dependent expression of the *pl* allele and the essentially constitutive expression of the *Pl* allele was subsequently shown to reside in the promoter (Cone et al., 1993b). The *pl* allele was expressed at very low levels, but was not a null allele. Rather, Cone et al. (1993b) proposed a threshold model, whereby the amount or concentration of the Pl protein had to be above a minimum in order for the structural anthocyanin biosynthetic genes to be activated.

In summary, anthocyanin biosynthesis in maize requires a combination of *R1* and *C1* or *B* and *Pl*, in addition to functional structural genes. The *R1* and *C1* combination is necessary for anthocyanin biosynthesis in the seeds, whereas the *B* and *Pl* combination stipulates anthocyanin synthesis in the vegetative parts of the plant. A range of modifications to this general model, such as tissue-specific deposition or light-dependent deposition, exists as a consequence of the availability of numerous mutant alleles.

In Arabidopsis the *TT2*, *TT8, TTG1* and *TTG2* genes have been shown to be regulatory genes controlling the biosynthesis of flavonoids. Walker et al. (1999) cloned the *TTG1* gene using map-based cloning and showed that this gene encodes a protein with four 'WD40' repeats. WD40 proteins have diverse roles in intracellular signaling, including control of the cell cycle and vesicular trafficking. The *TTG1* protein is thought to act as a signaling molecule that activates transcription factors of the *myc* class. Nesi et al. (2000) cloned the *TT8* gene using T-DNA tagging. Sequence analysis revealed that this gene encodes a protein containing a basic helix-loop-helix at its C-terminus, with similarity to the maize *R* protein and other transcription factors of the *myc* class. Based on expression data obtained with quantitative PCR, Nesi et al. (2000) concluded that the TT8 protein is important for the expression of the structural genes *DFR* and *BAN*. They also provided data that *TTG1* and *TT2* are important regulators of these two genes.

The cloning of the *TT2* gene by T-DNA tagging was reported by Nesi et al. (2001). This gene encodes an R2R3 *myb* protein that shows similarity to the maize C1 protein. In order to test whether the TT2 protein functions as a transcriptional activator, Nesi et al. (2001) were able to show that this protein is localized in the nucleus, that the spatio-temporal expression

pattern is consistent with the production of condensed tannins in the seed coat, and that over-expression of *TT2* in the presence of a functional TT8 protein results in ectopic expression (*i.e.* throughout the plant) of the *BAN* gene. Hence, TT2 and TT8 appear to work together to regulate the expression of *BAN*.

The *TTG2* gene encodes a WRKY transcription factor (Johnson et al., 2002). This class of transcription factors is characterized by the presence of a WRKYGQK amino acid sequence (one-letter amino acid code) near the N-terminal region and a conserved $C\text{-}X_{4\text{-}5}\text{-}C\text{-}X_{22\text{-}23}\text{-}H\text{-}X_{1}\text{-}H$ sequence that resembles zinc finger motifs. Together, these motifs are referred to as the WRKY domain. WRKY transcription factors were initially implicated in the response to wounding or pathogen attack, but based on the phenotype of the *ttg2* mutant, which includes developmental defects to the trichomes, it is clear that they can also function in plant development. Based on the lack of condensed tannins in the endothelial cells, the *TTG2* gene is thought to play a role in the regulation of structural genes, possibly *BAN*, involved in the synthesis of condensed tannins.

The *TT1* gene was cloned from a *transparent testa* mutant that was allelic to the original *tt1* mutant. The new mutant originated from the insertion of the maize *En* transposon (see Section 3.2) that had been introduced in Arabidopsis *via* transformation. The insertion of the *En* element enabled the cloning of the *TT1* gene (Sagasser et al., 2002). The deduced amino acid sequence of the TT1 protein revealed the presence of two zinc fingers, one near the N-terminus and one near the C-terminus of the protein, and an additional two zinc fingers in the C-terminal part of the protein. The presence of zinc fingers is indicative of a role as transcription factor. The TT1 sequence, however, revealed very limited homology with known proteins. TT1 and a small number of other plant proteins were classified as a novel group of transcriptional activators, the WIP subfamily of zinc finger proteins, where WIP refers to the first three conserved amino acids. The *TT1* gene was shown to be expressed in the endothelial cells of the seed coat using a reporter gene construct with the *GUS* gene (see Chapter 1, Section 3.5). The expression of *BAN* was reduced, but not completely eliminated in the *tt1* mutant, suggesting that TT1 is not as specific as TT8 in its regulation of *BAN* expression. TT1 may instead play a more general role in the differentiation of the endothelial tissue.

The maize *P* gene has been implicated in the regulation of phlobaphene synthesis. Phlobaphenes are red pigments that accumulate in the pericarp of the maize kernel, as well as various other parts of the plant, including the

cob, husks, and tassel glumes. Phlobaphenes arise from the polymerization of flavan-4-ols, although the exact structure is not known (see Chapter 1, Section 3.14). Alleles of the *P* gene are typically referred to by a two-letter suffix that reflects their expression in the pericarp ad the cob. For example, the *P-rr* allele results in red pericarp and red cobs, whereas the *P-rw* allele results in red pericarp and white cobs. The *P* gene was isolated using a *P-vv* allele that resulted in variegated patterns of phlobaphene deposition in both pericarp and cob as a result of an Ac insertion in a *P-rr* allele (Lechelt et al., 1989). The *P* gene encodes two different transcripts that have the 5′ exon and first intron in common, but that differ from each other at the 3′ end (Grotewold et al., 1991). The proteins encoded by the *P* gene contain a DNA binding domain resembling that of the *myb*-class of transcriptional activators, and this domain of the P protein is similar to that of the R and Pl proteins that regulate anthocyanin biosynthesis. The functional P protein regulates the expression of the *A1* (*DFR*) and the *C2* (*CHS*) genes (Grotewold et al., 1991; 1994).

10. MONOLIGNOL BIOSYNTHESIS

The monolignols are the building blocks of lignans (Section 11) and lignin (Section 12), whereas some of the intermediates of monolignol biosynthesis serve as precursors for hydroxycinnamic acids (Section 13) and sinapoyl esters (Section 14). Monolignols are synthesized from *p*-coumaroyl-CoA (**3.31**) generated *via* the shikimate and general phenylpropanoid pathways (see Sections 6 and 7). As part of monolignol biosynthesis *p*-coumaroyl-CoA (**3.31**) can undergo two types of modifications: reduction of the carboxyl group on the propane side chain to an alcohol, and substitution of the phenyl ring (Figure 3-9). The two predominant monolignols are coniferyl alcohol (**3.79**) and sinapyl alcohol (**3.81**). *p*-Coumaryl alcohol (**3.70**) and 5-hydroxyconiferyl alcohol (**3.80**) are generally much less abundant, and are found only in trace amounts in some species or tissues.

The reduction of *p*-coumaroyl-CoA (**3.31**) to *p*-coumaryl aldehyde (**3.69**) is catalyzed by the enzyme cinnamoyl-CoA : NADP oxidoreductase (CCR). This enzyme was initially purified from soybean cultures (Wegenmayer et al., 1976), and was later on efficiently isolated from lignifying cambium of eucalyps (*Eucalyptus gunnii*) (Goffner et al., 1994). A *CCR* cDNA was identified in a cDNA library that was screened with oligonucleotiede derived from the peptide sequence of the CCR protein. CCR is considered the first enzyme committed towards the biosynthesis of monolignols and shows

homology to the flavonoid biosynthetic gene flavonol 4-reductase (Lacombe et al., 1997).

The substitution of the phenyl ring necessary for the biosynthesis of coniferyl alcohol (**3.79**) and sinapyl alcohol (**3.81**) begins with the hydroxylation of C3. This is a conversion that requires the formation of the ester of *p*-coumaroyl-CoA with D-quinate (**3.73**) or shikimate (**3.74**) catalyzed by the enzyme hydroxycinnamoyl-CoA shikimate/quinate hydroxy-cinnamoyl transferase (HCT; Hoffmann et al., 2003). The hydroxylation of this ester intermediate is catalyzed by the enzyme *p*-coumaroyl-CoA 3′-hydroxylase (C3′H; Schoch et al., 2001; Franke et al., 2002a,b). The resulting shikimate or quinate ester (**3.75**; **3.76**) is subsequently hydrolyzed by the same HCT, resulting in caffeoyl-CoA (**3.36**).

The enzyme responsible for the hydroxylation of C3 was extremely difficult to identify. It had been postulated to be a phenol oxidase, dioxygenase, or a cytochrome P450 monooxygenase, but biochemical approaches aimed at isolating a protein displaying activity toward *p*-coumaric acid (**3.30**) were unsuccessful. With the availability of the *Arabidopsis thaliana* genome sequence, Schoch et al. (2001) performed a phylogenetic analysis of the genes encoding cytochrome P450 enzymes (for a review on this class of enzymes, see Chapple (1998)). This analysis resulted in the identification of CYP98A3 as a putative C3H. Expression analyses confirmed that the gene encoding CYP98A3 was expressed in tissues that would be expected to synthesize caffeic acid derivatives, including lignified tissues, and wounded tissues that produced chlorogenic acid (**1.18**). After cloning of the corresponding cDNA and expression of the recombinant enzyme in yeast, substrate specificity could be investigated. The enzyme did not show activity towards *p*-coumaric acid (**3.30**), *p*-coumaroyl-CoA (**3.31**), the *p*-coumaroyl glucose ester, nor the *p*-coumaroyl 4-glucoside. Based on experiments by Heller and Kühnl (1985) and Kühnl et al. (1987) with parsley cell suspension cultures, Schoch et al. (2001) were able to show activity of the recombinant CYP98A3 enzyme towards the shikimate and D-quinate esters of *p*-coumaroyl-CoA (**3.71; 3.72**), which resulted in the shikimate and D-quinate esters of *p*-caffeoyl CoA (**3.73; 3.74**), respectively. When the numbering of carbon atoms in the *p*-coumaroyl-CoA esters is taken into consideration, the carbon at the 3′ position – not the 3 position – is hydroxylated. Consequently, the enzyme is now referred to as C3′H. Franke et al. (2002a) showed that the Arabidopisis *reduced epidermal fluorescence8* (*ref8*) mutant was unable to synthesize caffeic acid (**3.32**) as a result of a defective copy of the *C3′H* gene.

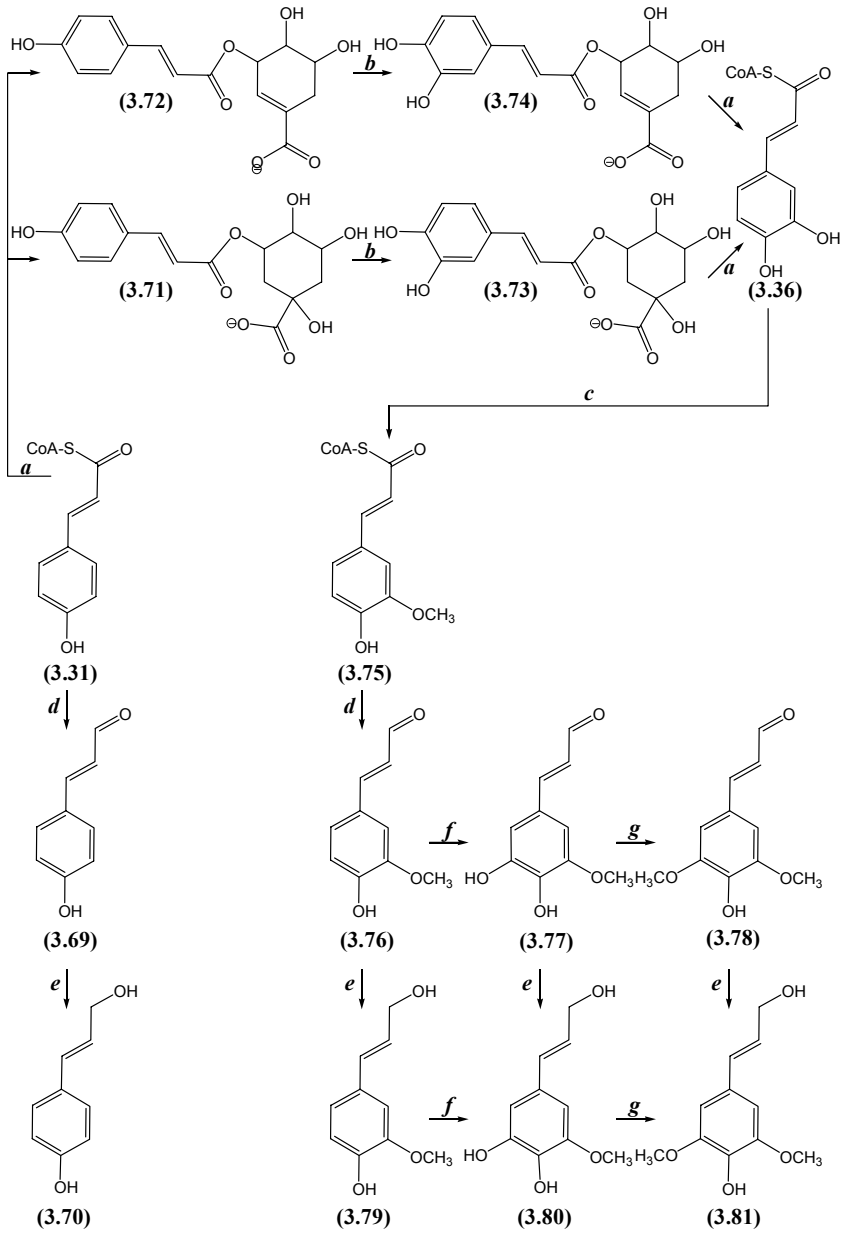


Figure 3-9. Biosynthesis of monolignols. The enzymes involved in this pathway are: **(*a*)** hydroxycinnamoyl-CoA shikimate/quinate hydroxy-cinnamoyl transferase, **(*b*)** *p*-coumaroyl-CoA 3′-hydroxylase (E.C. 1.14.14.1), **(*c*)** caffeoyl-CoA *O*-methyltransferase (E.C. 2.1.1.104), **(*d*)** cinnamoyl-CoA reductase (E.C. 1.2.1.44) **(*e*)** cinnamyl alcohol dehydrogenase (E.C. 1.1.1.195), **(*f*)** coniferyl aldehyde/coniferyl alcohol 5-hydroxylase (E.C. 1.14.13), **(*g*)** coniferaldehyde/coniferyl alcohol *O*-methyltransferase (E.C. 2.1.1.68).

HCT, the enzyme responsible for the formation of the D-quinate and shikimate esters of *p*-coumaroyl CoA (**3.71**; **3.72**), was identified by Hoffmann et al. (2003) in stem extracts of tobacco. Separation of the proteins with HPLC resulted in a fraction containing HCT activity. The protein was partially sequenced, and degenerate primers were synthesized to amplify the corresponding cDNA. Purification of the recombinant protein from *E. coli* enabled more detailed studies on substrate specificity and catalytic properties. These studies showed that HCT was able to catalyze the esterification reaction of D-quinate and shikimate with both *p*-coumaroyl-CoA (**3.31**) and caffeoyl-CoA (**3.36**). The enzyme could also catalyze the reverse reaction, *i.e.* the hydrolysis of the ester, thus producing caffeoyl CoA and either D-quinate or shikimate. The role of this enzyme in phenylpropanoid metabolism was further demonstrated via down-regulation of the *HCT* gene in Arabidopsis and tobacco (*Nicotiana benthamiana*). The transgenic plants were dwarfed, accumulated caffeoylquinate esters in their leaves, and showed different lignin subunit composition (Hoffmann et al., 2004).

Caffeoyl-CoA (**3.36**) is methylated to feruoyl-CoA (**3.75**) by the enzyme caffeoyl-CoA *O*-methyltransferase (CCoA-OMT). CCoA-OMT had been implicated in disease responses based on its induction in carrot (*Daucus carota*) cell suspension cultures that were treated with elicitors (Kühnl et al., 1989). A more general role of CCoA-OMT in phenylpropanoid metabolism was proposed after Ye et al. (1994) showed that the *CCoA-OMT* gene was up-regulated during the *in vitro* development of lignified tracheary elements derived from *Zinnia elegans* mesophyl cells. Feruoyl-CoA (**3.75**) is subsequently reduced to coniferaldehyde (**3.76**) by CCR, analogous to the reduction of *p*-coumaroyl CoA (**3.31**) to *p*-coumaryl aldehyde (**3.69**).

Coniferaldehyde (**3.76**) can undergo several fates, some of which can ultimately lead to the same end product. It can be reduced to coniferyl alcohol (**3.79**) by the enzyme cinnamyl alcohol dehydrogenase (CAD). Alternatively, the enzyme coniferyl aldehyde/coniferyl alcohol 5-hydroxylase (C5H), also known by its less accurate name ferulic acid 5-hydroxylase (F5H; Humphreys et al., 1999) can catalyze the hydroxylation of C5 to result in 5-hydroxyconiferyl aldehyde (**3.77**). C5H is also able to form 5-hydroxyconiferyl alcohol (**3.80**) from coniferyl alcohol (**3.79**). This enzyme was initially identified as F5H, after analysis of the Arabidopsis *ferulic acid hydroxylase1* (*fah1*) mutant, which was isolated in a mutant screen based on reduced levels of the UV-fluorescent sinapoyl esters (Section 13; Chapple et al., 1992). The *FAH1* gene was cloned using a T-DNA tagged mutant allele (Meyer et al., 1996), which revealed that the

gene encoded a cytochrome P450 monooxygenase with homology to flavonoid 3′, 5′ hydroxylases. Substrate specificity of recombinant F5H was evaluated by Humphreys et al. (1999) and Osakabe et al. (1999). Their analyses revealed that F5H had much higher activity towards coniferaldehyde (**3.76**) and coniferyl alcohol (**3.79**) than against ferulic acid (**3.33)**.

Methylation of 5-hydroxyconiferyl aldehyde (**3.77**) and 5-hydroxy coniferyl alcohol (**3.80**) by the enzyme 5-hydroxyconiferaldehyde/ 5-hydroxyconiferyl alcohol *O*-methyltransferase results in sinapaldehyde (**3.78**) and sinapyl alcohol (**3.81**), respectively. The enzyme catalyzing this step is known by the historic but inaccurate name caffeic acid *O*-methyl transferase (COMT). So COMT is now thought to be responsible for the methylation of the hydroxyl group on C5, whereas CCoA-OMT is responsible for methylation of the hydroxyl group on C3. This explains why mutations in the *COMT* gene, such as in the maize *brown midrib3* mutant (Vignols et al., 1995) and the sorghum *brown midrib26* mutant (Bout and Vermerris, 2003), result in reductions in lignin units derived from sinapyl alcohol, and not in lignin subunits derived from coniferyl alcohol (see also Section 12).

As described above, sinapyl alcohol (**3.81**) can be synthezed *via* methylation of 5-hydroxyconiferyl alcohol (**3.80**) by COMT. An alternative route is *via* the reduction of sinapaldehyde (**3.78**) by CAD or, in the case of aspen (*Populus tremuloides*) and several other angiosperm trees, sinapyl alcohol dehydrogenase (SAD; Li et al, 2001). *SAD* cDNA's were identified as a distinct class of hybridizing fragments during the screening of an aspen cDNA library derived from lignifying xylem tissue with a probe derived from an aspen *CAD* cDNA. Analysis of the substrate specificity of the recombinant protein generated by expression of a *SAD* cDNA in *E.coli* indicated that SAD had a 60-fold higher affinity for sinapaldehyde than coniferaldehyde.

CAD is encoded by a multigene family in Arabidopsis (Raes et al., 2003; Goujon et al., 2003) and rice (*Oryza sativa*; Tobias and Chow, 2005), and probably in many other species. Mutants of Arabidopsis in which the genes encoding two distinct isoforms of CAD, CAD-C and CAD-D, were down-regulated as a result of T-DNA insertions were analyzed by Sibout et al. (2003). The reduction in CAD-C activity resulted in minor changes in lignin composition, whereas reduction CAD-D activity resulted in a 45% and 24% reduction in lignin residues derived from sinapyl alcohol in stem and root tissue, respectively. Taken together with the fact that both isoforms display activity towards both coniferaldehyde and sinapaldehyde, these data

suggest that the CAD enzymes in Arabidopsis do not display the same substrate *specificity* for either sinapaldehyde or coniferaldehyde as was observed in aspen. Rather, in Arabidopsis, and possibly many other species, the biosynthesis of coniferyl alcohol and sinapyl alcohol appears to be catalyzed by a combination of isoforms, some of which have a *preference* towards one of the substrates. The combination of isoforms varies depending on the developmental stage and the tissue (Sibout et al., 2003).

11. LIGNAN BIOSYNTHESIS

Lignans are synthesized from the oxidative coupling of monolignol radicals (see Section 10 in this chapter, and Chapter 1, Section 3.11). The monolignol radicals are generated through the action of laccases (E.C. 1.10.3.2) or peroxidases (E.C. 1.11.1.7; see also Chapter 2, section 1.8). Unlike lignin, however, lignans are optically active, and typically only one enantiomer is present in a given species. This means that the coupling between the monolignols is under regio-chemical control, whereby both the coupling site and the orientation of the two monomers are controlled. Analysis of *Forsythia intermedia* stem extracts, which accumulate high levels of the lignan (+)-pinoresinol (**3.82**) resulted in the purification of a protein that has no catalytic activity, but that is able to stipulate the formation of (+)-pinoresinol from two coniferyl alcohol (**3.79**) radicals (Davin et al., 1997). This protein was referred to as 'dirigent protein'. It is thought to hold the two coniferyl alcohol radicals in a specific configuration during bond formation so that both the position of the bond (8-8′) and the conformation are controlled (Figure 3-10). Once the bond is formed, intramolecular cyclization results in (+)-pinoresinol (**3.82**). This particular dirigent protein was only able to recognize coniferyl alcohol radicals and none of the other monolignol radicals. Since the initial discovery of this protein from *Forsythia intermedia*, homology searches in sequence databases have revealed the existence of additional genes encoding putative dirigent proteins, from a variety of species (Gang et al., 1999; Davin and Lewis, 2000). Recently, evidence for the existence of dirigent proteins able to direct the formation of 8-*O*-4′ linkages between coniferyl alcohol and sinapyl alcohol was reported (Lourith et al., 2005).

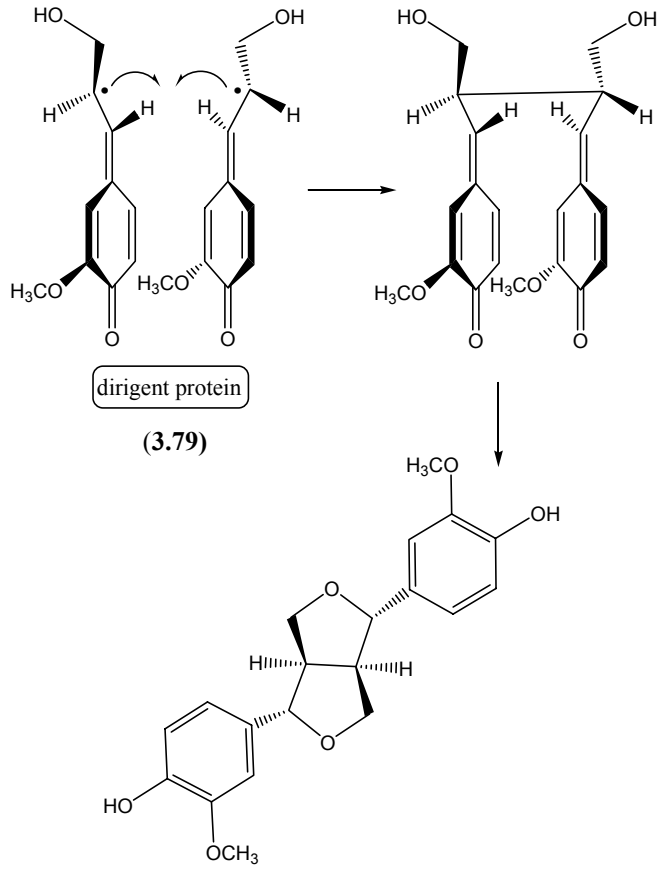


Figure 3-10. Dirigent-protein mediated formation of (+)-pinoresinol from two coniferyl alcohol radicals. The formation of the ring results from intramolecular cyclization.

Lignans represent an extremely diverse group of compounds. This is the result of both structural diversity and stereo-selective biosynthesis. One particular plant species generally makes only one enantiomer of a particular compound. The other enantiomer may be synthesized by a different species. As a consequence, it is virtually impossible to summarize the biosynthesis of lignans in general. Instead, the focus here will be on the biosynthesis of the lignan podophyllotoxin in a number of different plant species, as an illustration of the different biosynthetic routes that can be used to synthesize the same compound.

Podophyllotoxin (**3.86**) is used as an ectopic antiviral agent in the treatment of venereal warts. It is too cytotoxic to be ingested, but after chemical modifications podophyllotoxin can be used as a powerful anti-cancer

drug in the treatment of small cell lung cancer, acute leukemia, and non-Hodgkin's lymphoma (Canel et al., 2000). Podophyllotoxin is isolated from the plant *Podophyllum peltatum*, commonly known as Mayapple (also referred to as Devils's apple or hog apple), or from the related species *Podophyllum hexandrum* (Himalayan mayapple). In order to be able to produce podophyllotoxin without depleting the natural populations of these wild species, attempts have been made to produce podophyllotoxin in tissue culture. In this context it is important to know the biosynthetic route to podophyllotoxin production, so that lines that produce high levels of the relevant biosynthetic enzymes can be selected. Broomhead et al. (1991) investigated the biosynthesis of podophyllotoxin in *Podophyllum hexandrum* using radioactive feeding experiments in whole plants. They proposed the biosynthetic route displayed in Figure 3-11. This shows how two coniferyl alcohol (**3.79**) radicals are coupled in a stereo-selective manner to form matairesinol (**3.83**). Subsequent hydroxylation and methylation of C5´ and lactone ring formation at C3 and C4 gives yatein (**3.84**). Ring closure results in the formation of deoxypodophyllotoxin (**3.85**), which is hydroxylated to yield podophyllotoxin (**3.86**).

In addition to podophyllotoxin itself, closely related compounds, such as 5-methoxypodophyllotoxin (**3.89**), can be used as a precursor for anti-cancer drugs. *Linum* species (flax, linseed) can produce podophyllotoxin or substituted podophyllotoxins in tissue culture with yields of up to 0.35% of dry weight. The biosynthesis of 5-methoxypodophyllotoxin in *Linum flavum* was investigated by Xia et al. (2000), whereas Seidel et al. (2002) investigated podophyllotoxin (**3.86**) production in cell cultures of *Linum album*.

Xia et al (2000) proposed a biosynthetic route starting coniferyl alcohol (**3.79**) and subsequent formation of (+)-pinoresinol (**3.82**), as shown in Figure 3-12. The enzyme pinoresinol/lariciresionol reductase converts this compound to (+)-lariciresinol (**3.87**), and then to (–)-secoisolariciresionol (**3.88**). The enzyme secoisolariciresinol dehydrogenase converts (**3.88**) into (–)-matairesinol (**3.83**). The conversion from (–)-matairesinol (**3.83**) to podophyllotoxin (**3.86**) is likely to be similar to the route shown in Figure 3-11.

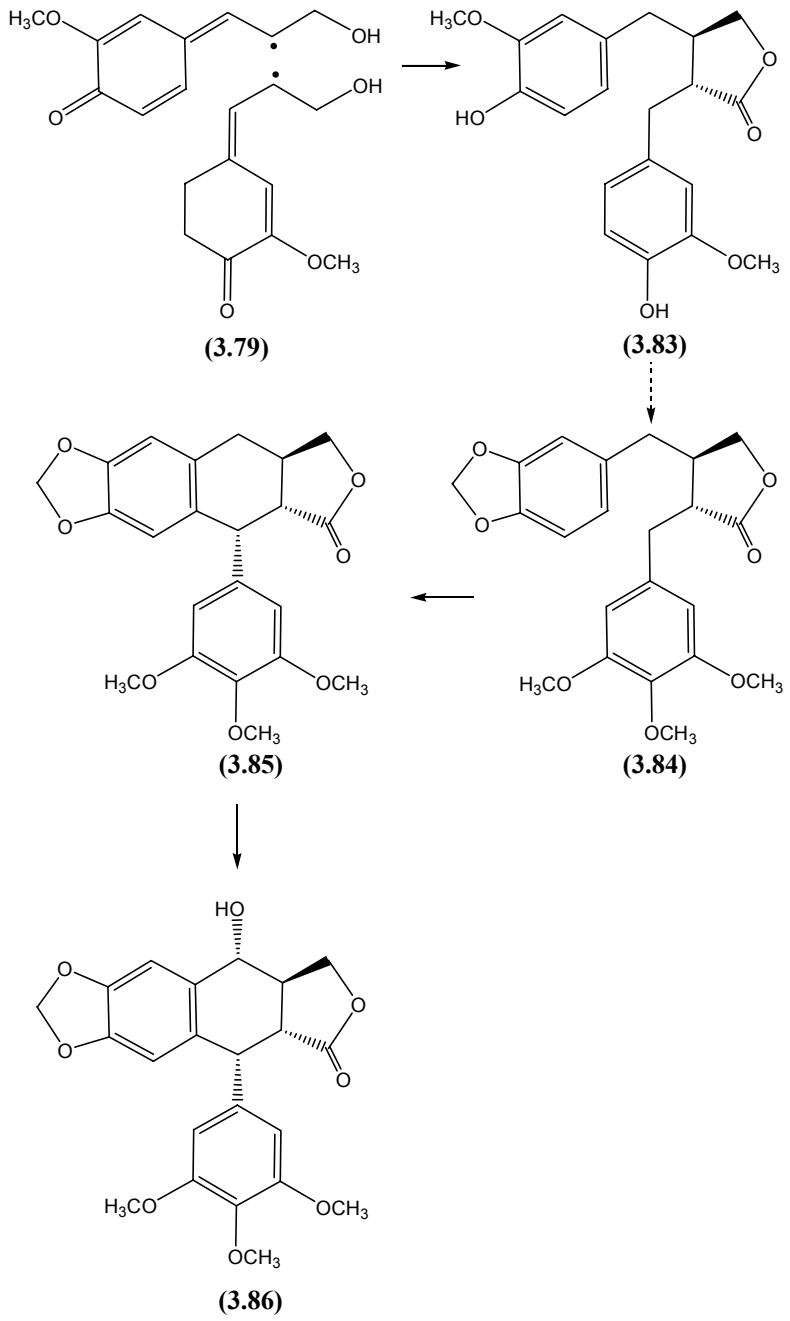


Figure 3-11. Biosynthesis of podophyllotoxinin in *Podophyllum hexandrum*, according to Broomhead et al. (1991). The dotted arrow refers to a series of reactions that have not yet been fully elucidated: formation of a methylene dioxybridge, and hydroxylation and methylation reactions.

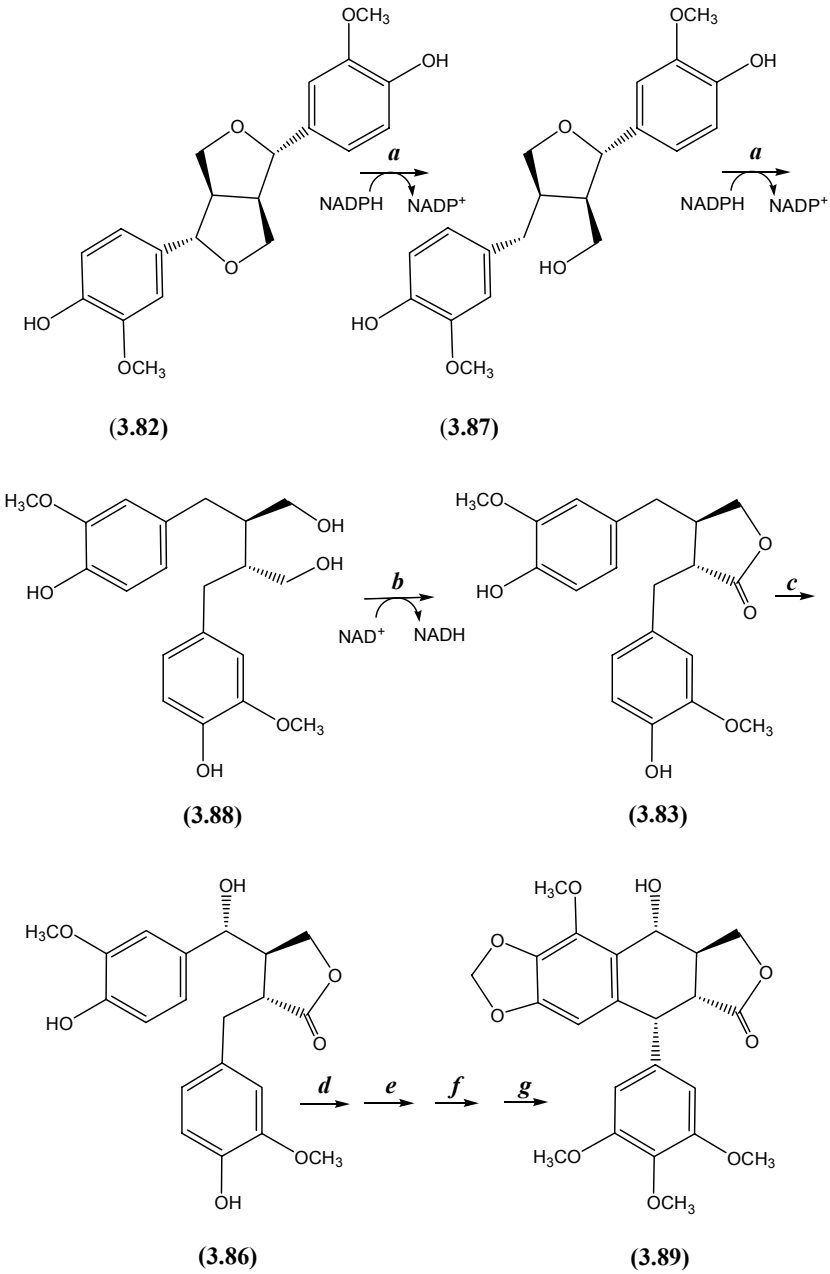
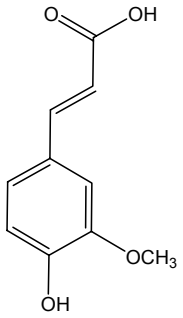
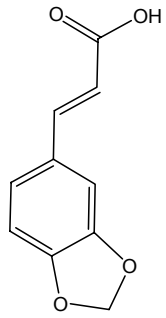


Figure 3-12. Biosynthesis of 5-methoxypodophyllotoxinin in *Linum flavum*, according to Xia et al. (2000). **(*a*)** pinoresinol/lariciresionol reductase, and **(*b*)** secoisolariciresinol dehydrogenase. The enzymes catalyzing steps ***c-g*** have not yet been elucidated, but involve ***(c)*** hydroxylation of C7, ***(d)*** ring closure, ***(e)*** hydroxylation of C5, ***(f)*** methylation of the OH-group on C5, and ***(g)*** formation of a methylene dioxy bridge. The exact order of these reactions is also still unknown.

Synthesis of podophyllotoxin (**3.86**) in cell culture of *Linum album* results in yields comparable to those of the most efficient tissue cultures of *Podophyllum hexandrum*. In order to further improve *L. album* cultures, Seidel et al. (2002) investigated the biosynthesis of podophyllotoxin (**3.86**). They fed a number of labeled compounds that to *L. album* cell cultures to identify which of these compounds could be used as precursors to podophyllotoxin. They determined that the substitution pattern on the benzene ring is critical. The substitution has to be either 3-methoxy, 4-hydroxy, as in ferulic acid (**3.33**), or, alternatively, 3,4-methylenedioxycinnamic acid (**3.90**) can serve as precursor. The precursor of podophyllotoxin in *L. album* appears to be deoxypodophyllotoxin (**3.83**), based on the higher level of isotope incorporation in the latter compound. This means that 7-hydroxymatairesinol, the precursors of 5-methoxypodophyllotoxin in *L. flavum* (Xia et al., 2000), is not a precursor of podophyllotoxin in *L. album*.



(3.33)



(3.90)

12. LIGNIN BIOSYNTHESIS

12.1 Genetic control of lignification

All of the enzymes involved in the biosynthesis of monolignols have been identified and characterized, and the genes encoding these enzymes have been cloned. As discussed in Section 10, the specific role of individual genes that are part of a multigene family needs further investigation. Another aspect that is relatively poorly understood is the regulation of lignin biosynthesis. Given that the biosynthesis of lignin involves a large number of enzymes, the genes encoding these enzymes need to be coordinately expressed. One mechanism to achieve this would be to have one or a few

transcriptional activators that can bind to a common motif present in the promoters of several monolignol biosynthetic genes. Comparison of the promoter sequences of structural genes involved in phenylpropanoid metabolism revealed a motif conserved among several plant species referred to as the PAL-box: CCA (C/A) (A/T) A (A/C) C (C/T) CC; the nucleotides in parentheses are degenerate, meaning that the exact nucleotide that is present varies among promoters. This motif is also referred to as AC-elements and H-box.

Kawaoka et al. (2000) used the PAL-box sequence to isolate a cDNA encoding a transcription factor that could bind to the PAL box. They screened an expression library generated from tobacco (*Nicotiana tabacum*) stem cDNA's with a radio-labeled oligonucleotide comprising of three repeats of a PAL-box. This so-called southwestern screening resulted in a partial cDNA clone corresponding to a LIM transcription factor. This class of transcription factors was first identified in the nematode *Caenorhabditis elegans*. The LIM domain is cysteine rich and contains two DNA-binding zinc fingers separated from each other by two amino acids. Sequence analysis of the full-length cDNA sequence obtained by Kawaoka et al. (2000), referred to as *Ntlim1*, revealed the presence of two LIM domains and one acidic domain. All three domains had to be present to effect transcriptional activation of a target gene based on transient expression studies with tobacco protoplasts. *In vivo* transcriptional activity of the *Nt*LIM1 protein was demonstrated by studies with transgenic plants expressing the *Ntlim1* gene in either the sense of antisense orientation under the control of the strong cauliflower mosaic virus (CaMV) 35S promoter. Introduction of the sense constructs did not enhance expression of the genes encoding PAL, 4CL, and CAD, nor did it result in increased enzyme activity. The antisense constructs, however, resulted in a coordinate down-regulation of the *PAL*, *4CL* and *CAD* genes, along with a reduction in the activities of the enzymes encoded by these genes. Depending on the transgenic line, variations in lignin content and subunit composition was observed.

The *myb* transcription factors had been shown to be involved in the regulation of phenylpropanoid metabolism (*cf.* Grotewold et al., 1994; see Section 9.2). They were first shown to be involved in the regulation of lignin biosynthesis by showing that lignin content and phenolic ester content were reduced in transgenic tobacco in which the snapdragon *AmMYB308* or *AmMYB330* cDNA's were overexpressed (Tamagnone et al., 1998). The reduction rather than increase in phenolic metabolism suggests that these two *myb* transcription factors are normally weak activators, but that their

over-expression out-competes the binding of stronger transcriptional activators that are present. Based on these results Patzlaff et al. (2003) screened a cDNA library made from differentiating loblolly pine (*Pinus taeda*) xylem with a pool of radio-labeled DNA fragments representing conserved sequences of *R2R3 myb* genes. R2R3 *myb* proteins are characterized by two 50- to 52-residue long imperfect repeats. Each of the repeats contains three α-helices, with the second and third helices forming a helix-turn-helix structure when bound to DNA (Sablowski et al., 1994; Dias and Grotewold, 2003). The *PtMYB4* cDNA they isolated was shown to encode an R2R3 *myb* protein able to bind to AC-elements in an electrophoretic mobility shift assay (EMSA) with recombinant *Pt*MYB4 protein and radio-labeled AC-elements based on the bean *PAL2* promoter sequence. Binding of *Pt*MYB4 protein to these targets would reduce their electrophoretic mobility through polyacrylamide gels, as was indeed the case. Analysis of transgenic tobacco over-expressing *PtMYB4* cDNA showed that the expression of the *PAL* gene was reduced, the expression of the *C4H* and *4CL* genes was not changed, and that the expression of the genes encoding C3′H, CCoA-OMT, CCR, COMT and CAD (see Section 10) was increased. Furthermore, the transgenic plants showed ectopic lignification, which means that tissues that are normally not lignified were producing lignin. Given that PAL, C4H and 4CL are part of the general phenylpropanoid pathway (Section 7), whereas the other genes are part of the more specific monolignol biosynthetic pathway (Section 10), these data indicate that the *Pt*MYB4 protein appears to be specifically involved in lignification.

Very similar results were obtained with an R2R3 *myb* cDNA isolated from eucalypt (*Eucalyptus grandis*) xylem (Goicoechea et al., 2005). The *Eg*MYB2 protein was able to bind to radio-labeled promoter fragments from the eucalypt *CAD2* and *CCR* genes. As was the case with the *Pt*MYB4 protein, analysis of transgenic tobacco plants over-expressing the *EgMYB2* gene showed limited effects on the expression of the *PAL*, *C4H* and *4CL* genes, but a 5- to 40-fold increase in the expression levels of the genes encoding HCT, C3′H, CCR, CCoA-OMT, F5H, COMT and CAD.

The Arabidopsis *AtMYB61* gene is homologous to the *PtMYB4* gene and was shown to be mis-expressed in the Arabidopsis *de-etiolated3* (*det3*) mutant, which displays ectopic lignification and is dark photomorphogenic, meaning that the mutant will grow and develop in the dark as if it was growing in the light (Newman et al., 2004). Overexpression of *AtMYB61* in Arabidopsis resulted in a *det3* phenotype, whereas down-regulation of *ATMYB61* in *det3* mutants resulted in a loss of the mutant phenotype. These

data showed that the mis-expression of *AtMYB61* was necessary and sufficient to produce the *det3* mutant phenotype.

Mutants in which lignin deposition is disturbed are valuable sources to obtain information on the mechanisms involved in the spatio-temporal control of lignin deposition. The maize *brown midrib2* (*bm2*) mutant has brown vascular tissue in the leaf blade and leaf sheath. Chemical analysis of dissected vascular tissue obtained from this mutant and the wild-type control indicated an overall reduction in lignin residues derived from both coniferyl alcohol (**3.79**) and sinapyl alcohol (**3.81**), but the most striking difference ws that the developmental-specific gradient in lignin content, whereby older tissues are more heavily lignified than younger tissues, was disrupted as a result of the *bm2* mutation (Vermerris and Boon, 2001). The functional *Bm2* gene has been cloned, but its molecular function remains to be elucidated.

Rogers and Campbell (2004) provided an overview of Arabidopsis mutants in which normal patterns of lignin deposition were altered. This set of mutants includes the *det3*, *ectopic lignification1* (*eli1*) and *pom-pom1* (*pom1*) mutants, all of which synthesize lignin ectopically. These mutants were the subject of study by Rogers et al. (2005) in an attempt to identify whether a common regulatory circuit was disturbed Chemical and histochemical analyses revealed variation in cell wall composition among the three mutants. A comparison of the expression of structural genes involved in monolignol biosynthesis revealed that the expression of some genes (*e.g. PAL1*, *C4H*, *F5H* and *CCoA-OMT*) was increased in all mutants, whereas the expression of the other genes changed only in a subset of the mutants. A microarray analysis was performed to identify genes that were coordinately expressed with the monolignol biosynthetic genes. Microarray experiments are performed with so-called gene chips on which several thousand genes are represented by short gene-specific DNA molecules termed oligonucleotides. The chips are hybridized with fluorescently labeled cDNA pools obtained from a sample of interest (mutant, treatment, developmental stage) and a control. The cDNA pools from the two samples are labeled with different fluorescent dyes, generally Cy-3 and Cy-5, that result in a green and red fluorescence, respectively, upon excitation with a laser (Schulze and Downward, 2001; Wayne and McIntyre, 2002; Gracey and Cossins, 2003). After the hybridization, it is possible to identify whether a given gene was not expressed (no signal), expressed equally in both samples (yellow fluorescent signal from overlap of green and red fluorescence), or higher expression in one or the other sample (red or green signal). Analysis of the promoter sequences of genes coordinately expressed with the subset of lignin biosynthetic genes revealed that the AC-element

(PAL-box) was significantly over-represented. Among these genes were five genes encoding transcriptional activators of the R2R3 *myb*, and Dof classes. Dof – Downstream-of-fibroblast growth factor receptor – transcriptional activators were first identified in the fruit fly (*Drosophila melanogaster*). Given that this set of five transcription factors had expression patterns similar to the monolignol biosynthetic genes, they may be involved in the regulation of lignin deposition.

12.2 Monolignol transport and polymerization

Lignin is a complex polymer generated *via* oxidative coupling of monolignol radicals (see Section 10). The polymerization of lignin occurs in the plant cell wall, so the monolignols need to be transported from the cytosol, where they are synthesized, to the cell wall. There is evidence that in several species, including conifers and Arabidopsis, the monolignols are glycosylated prior to storage and transport, which would reduce their toxicity and make them less reactive. The enzymes responsible for the glycosylation reaction are UDP-glucosyltransferases (UGT's; EC 2.4.1.111) that generate coniferin (coniferyl alcohol 4-*O*-glucoside) and syringin (sinapyl-4-*O*-glucoside) from coniferyl alcohol (**3.79**) and sinapyl alcohol (**3.81**), respectively (Lim et al., 2001). Prior to polymerization these glucosides need to be converted to the corresponding aglycones by specific glucosidases, such as coniferin β-glucosidase (EC 3.2.1.126).

The Arabidopsis genome contains more than 100 UGT genes. In order to determine which of these UGTs were capable of forming either glucosides – involving a linkage between glucose and the hydroxyl group on C4 of the phenylpropanoid – or glucose esters – involving an ester linkage between the carboxyl group of hydroxycinnamic acids and the hydroxyl group of glucose – cDNA's or genomic clones representing 36 Arabidopsis genes were expressed in *E. coli*, and the activity of the resulting recombinant proteins against 11 different intermediates from the monolignol biosynthetic pathway was assayed. This analysis identified three genes encoding UGTs catalyzing the formation of cinnamate glucose esters and two genes – *UGT72E2* and *UGT72E3* – encoding UGTs that catalyze the formation of cinnamate glucosides. The recombinant protein obtained from expression of the *UGT72E* gene showed activity towards both coniferyl alcohol and sinapyl alcohol, whereas the UGT72E3 recombinant protein only showed activity towards sinapyl alcohol.

Polymerization of lignin occurs through an oxidative coupling mechanism, whereby monolignol radicals react with radical sites on the lignin polymer. The two most common monolignols, coniferyl alcohol (**3.79**) and sinapyl alcohol (**3.81**), give rise to guaiacyl (G) and syringyl (S) residues, respectively, in the ligin polymer. *p*-Coumaryl alcohol (**3.70**) gives rise to *p*-hydroxyphenyl (H) residues. The latter is not a very common lignin residue. It is incorporated in compression wood of gymnosperms in response to gravitropic stress (Higuchi, 1985), and to a small extent (~5%) in the lignin of graminaceous species.

There has been some controversy regarding the enzyme catalyzing the formation of the monolignol radicals. Histochemical studies of lignifying tissues with a chromogenic substrate only resulted in oxidation of the substrate in the presence of H_2O_2 (Harkin and Obst, 1973). This observation combined with the broad distribution of peroxidases (EC 1.11.1.7) in the plant kingdom and the more limited distribution of laccases (EC 1.10.3.2) led Higuchi (1985) to propose a dominant role of peroxidases in lignification. Laccases were subsequently shown to be present and active in lignifying tissues of various plant species, including sycamore maple (*Acer pseudoplatanus*; Sterijades et al., 1992) and loblolly pine (*Pinus taeda*; Bao et al., 1993; see also Chapter 2, section 1.8.2). The current view is that both laccases and peroxidases can generate monolignol radicals, but it is unclear whether there is a preference for one or the other as a function of the species, the tissue, the developmental stage, or the environmental conditions. Both peroxidases and laccases are encoded by large multigene families (Raes et al., 2003), which has made it difficult to study the specific role of these enzymes in lignification. Transgenic approaches have been used to down-regulate or over-express *laccase* or *peroxidase* genes, in an attempt to better understand their function. Down-regulation of three *laccase* genes in poplar (*Populus trichocarpa*) using antisense technology did not have an impact on lignin content or composition, but down-regulation of one of these *laccase* genes did result in alterations in xylem cell wall structure and an increase in the level of soluble phenolics (Ranocha et al., 2002). Similarly, a mutation in an Arabidopsis *laccase* gene resulted in irregularly shaped xylem cells (Brown et al., 2005). In contrast, down-regulation of an anionic peroxidase gene from hybrid aspen (*Populus sieboldii x P. gradidentata*) *via* an antisense construct resulted in reduced lignin content and a reduction in G-residues in the lignin (Li et al., 2003).

Given that there are several monolignols and that these monolignols can be linked to each other in a number of different ways (see Chapter 1, section 3.12), the question arises how the plant is able to control lignin subunit

composition and the distribution of inter-unit linkages. The prevailing view is that lignin subunit composition is determined predominantly by the flux of monolignols. In grasses, for example, *p*-coumaryl alcohol (**3.70**) is secreted early on during the lignification process, resulting in the relatively high abundance of H-residues in cell walls that have just begun to lignify. Coniferyl alcohol (**3.79**) and sinapyl alcohol (**3.81**) are secreted into the cell wall later on (Terashima et al. 1993). The distribution of interunit linkages appears to be under chemical control. Based on *in vitro* formation of dehydrogenation polymers (DHPs) the distribution of interunit linkages is affected by the local concentration of monolignol radicals (Syrjanen and Brunow, 2000), the presence of polysaccharides (Terashima et al., 1995), proteins (McDougall et al., 1996), and lignin (Guan et al., 1997). When DHPs were formed in the presence of isolated primary cell walls of maize, the resulting lignin polymer resembled native maize lignin remarkably well (Grabber et al., 1996). The model in which lignin polymerization is considered to be under chemical control is referred to as the 'random coupling' model, or the 'combinatorial coupling' model. According to this model, lignin has little predetermined structure to it. From a biological point of view this is to the plant's advantage when it comes to the role lignin plays in defense against pathogens and insect pests. An invading organism would have to develop a host of hydrolytic enzymes in order to effectively dissolve the barrier created by lignified tissue (Denton, 1998).

An opposing view on the polymerization of lignin was proposed by Lewis and Davin (1998). They reasoned that an abundant and important biopolymer like lignin should be under strict biological control, in a manner analogous to the other biological polymers, including cellulose, chitin, and proteins. According to these authors there are two major pieces of evidence for biological control of lignin polymerization: (**1**) the DHP formed by the *in vitro* polymerization of coniferyl alcohol (**3.79**) in the presence of peroxidase and H_2O_2 displays a very different distribution of interunit linkage than the lignin formed *in planta* (Nimz and Ludemann, 1976) and (**2**) immunohistochemical analyses using antibodies against specific lignin substructures show that the different antibodies recognize different parts of the cell wall, suggesting that lignin subunit composition varies depending on the exact position in the cell wall (Joseleau and Ruel, 1997). Given that both lignans and lignin are synthesized from monolignols, Lewis and Davin (1998) hypothesized that proteins resembling the dirigent protein (see Chapter 1, Section 3.11, and this chapter, Section 11) could be responsible for the polymerization of lignin.

This hypothesis was furthered by Gang et al. (1999), who provided evidence for the existence of epitopes in secondary cell walls of *Forsythia* that were recognized by polyclonal antibodies raised against the dirigent protein. In addition, homology searches of DNA sequence databases revealed the existence of putative genes resembling the gene encoding the *Forsythia* dirigent protein. According to the proposed model, these dirigent-like proteins could stipulate lignin subunit composition, as well as the formation of specific interunit bonds.

The publication by Gang et al. (1999) initiated a heated debate on lignification, because the hypothesis that dirigent-like proteins were involved in lignification was presented as 'a new paradigm'. The existing random coupling model agreed, however, surprisingly well with what was known about lignin polymerization and the establishment of lignin subunit composition, as discussed above. The continuing effort to replace the random coupling model with the new dirigent model (Lewis, 1999), which was touted as the solution to 'the mystery of specificity of radical precursor coupling in lignan and lignin biosynthesis' (Davin and Lewis, 2000), resulted in a series of rebuttals in which the shortcomings of the dirigent model were pointed out (Sederoff et al., 1999; Hatfield and Vermerris, 2001; Ralph et al., 2004). One issue the dirigent model did not adequately address was optical activity. According to the model proposed by Lewis and colleagues, the stereospecificity of the dirigent-like proteins would result in an optically active lignin polymer. Lignin, however, is not optically active (Freudenberg et al., 1965; Ralph et al., 1999; Akiyama et al., 2000). To address this issue, Davin and Lewis (2000) proposed that lignin exists predominantly as a linear polymer, and that the lack of optical activity results from the synthesis of a lignin strand with the opposite chirality, originating from self-replicating polymerization templates. This was schematically represented by Davin and Lewis (2005) based on chemical analyses aimed at resolving the primary structure of a lignin hexamer with electron spray ionization-mass spectrometry (Evtuguin and Amada, 2003). The experiments reported by Evtuguin and Amada (2003) were, however, not based on the assumption that lignin as a whole was a linear polymer. Rather, they were interested in defining how different monolignols were coupled to each other. In addition, mass spectrometry does not allow the distinction between different stereo-isomers, so that stereo-selective coupling cannot be inferred from these data.

The articles dealing with the dirigent model are of interest not only from a scientific but also from a philosophical viewpoint. In the case of the dirigent hypothesis a completely new model was proposed by extrapolating

data obtained from experiments on a different class of compounds (lignans). This came at a time when the existing random coupling model was able to explain the majority of the observations. Based on the way the dirigent hypothesis was presented in review articles (Lewis, 1999; Davin and Lewis, 2000), the newly proposed model was essentially elevated to a theory in a matter of one year.

Theories are formulated after a process during which a new hypothesis is tested. The experiments that allow hypothesis testing are generally set up in such a way that a hypothesis can be disproved. Only after repeatedly being *unable* to disprove a hypothesis can one conclude that all the data support the hypothesis. At that point it becomes possible to predict the outcome of newly designed experiments. If the outcome is indeed consistently predicted correctly, it is appropriate to elevate the hypothesis to a theory.

In the case of the dirigent model, the hypothesis testing phase is far from complete. For example, if dirigent-like proteins are involved in lignin polymerization, then the inactivation of the gene encoding these proteins would be expected to result in a change in lignin distribution or lignin composition. Plants in which specific genes are down-regulated are nowadays relatively easy to obtain through insertional mutagenesis or transgenic approaches. This approach is therefore feasible, and would allow a direct evaluation of the dirigent model. The need for such experiments has been expressed several times (Hatfield and Vermerris, 2001; Boerjan et al., 2003), but no publications in which analyses of such mutants or transgenic plants are discussed have been published to date. Instead, a number of mutants and transgenic plants in which lignin biosynthetic genes were down-regulated (Anterola and Lewis, 2002; Patten et al., 2004) were reanalyzed, and, in many cases, were claimed to be flawed. Until more detailed genetic studies on dirigent-like proteins have been carried out in the context of lignification, the dirigent model needs to be treated solely as a hypothesis.

12.3 Modification of lignin for agro-industrial applications

Lignin biosynthesis has received considerable attention during the past decade, largely because of the economic cost associated with the presence of lignin in agro-industrial feedstocks. Three industrial processes can benefit from the modification of lignin content and lignin subunit composition: **1**) the pulp and paper industry, **2**) the animal feed industry, and **3**) the bio-processing industry. Consequently, transgenic approaches have been

implemented to modify lignin composition in a variety of commercially important species.

12.3.1 Pulp and paper industry

A substantial amount of commercially grown wood is destined for the production of paper. Paper production requires the isolation of the cellulose fibers from the wood. After debarking, the logs are chipped and the chips are subsequently pulped. Pulping results in the physical and/or chemical breakdown of wood so that discrete fibers are liberated. These fibers can then be dispersed in water and reformed in a web (Biermann, 1996). Chemical and/or mechanical methods are used to pulp wood. Mechanical pulping uses grindstones, sometimes in combination with steam, to liberate fibers. Lignin is not removed during this process. The paper yield resulting from mechanical pulping is high, but the paper is weak and exposure to light and air causes the paper to yellow. An example of paper obtained from mechanical pulping is the paper used for newspapers. In contrast, high-quality paper relies on the use of chemicals that can separate cellulose from lignin. Soda pulping uses sodium hydroxide in the cooking liquid to separate cellulose fibers. Sulfite pulping is another chemical pulping method; it relies on mixtures of sulfurous acid and/or its alkali salts to solubilize lignin. Both methods became less popular with the invention of Kraft (or sulfate) pulping, which relies on the action of sodium sulfide and sodium hydroxide at high pH and temperatures between 160 and 180°C. Kraft pulping is useful for any wood species and the resulting paper is strong. A major disadvantage of Kraft pulping is the low yield and the substantial environmental pollution. The black liquor that is left over after the Kraft process contains reduced sulfides which are odiferous compounds. They are released in the air during combustion or evaporation of the liquor. Also the bleaching process, which is used to create white paper, results in pollution, most notably because of the formation of dioxins.

Experimental methods that aim to prevent the environmental polution include organosolv pulping (use of organic solvents to remove lignin) and biological pulping, which uses white rot fungi or lignin degrading enzymes (Biermann, 1996). The removal of lignin is a major bottleneck in the production of high quality paper, regardless of the pulping method that is used. A reduction in lignin content and/or change in lignin chemistry that would allow a more efficient separation of lignin and cellulose would, therefore, be beneficial.

A number of tree species have been transformed with the attempt to either down- or up-regulate lignin biosynthetic genes to improve pulping characteristics. This was considered a relatively uncontroversial application of genetic modification, given that the wood was not intended for human consumption but purely for industrial processing, and that the environmental benefits were apparent. Nonetheless, environmental activists destroyed trials with transgenic poplars in the UK.

Pilate et al. (2002) reported on a long-term study (four years, two sites) to evaluate both the industrial processing and the ecological impacts of transgenic poplar trees (*Populus tremula* x *Populus alba*) in which the *CAD* or the *COMT* gene had been down-regulated via the introduction of antisense constructs. The change in lignin content and composition did not result in increased manifestation of pests and diseases, or in changes in the distribution of species living in the trees. The decomposition of the transgenic trees, however, measured as the release of CO_2 from the soil in which the trees were growing, was more rapid, underscoring the role of lignin in the prevention of decay. Wood harvested from the transgenic trees with down-regulated CAD activity was more easily delignified in the Kraft pulping process, and required 6% less alkali. In contrast, wood obtained from the trees with down-regulated COMT activity required 15% more active alkali to obtain a similar degree of delignification as the the wild-type controls. The poor delignification properties were attributed to a more heavily cross-linked lignin resulting from the reduction of S-residues and the concomitant increase in G-residues. The G-residues, derived from coniferyl alcohol, are able to participate in interunit linkages involving C5, which tend to be more difficult to break. This is consistent with a report by Huntley et al. (2003) in which transgenic poplar trees with enhanced F5H (C5H) activity were evaluated. As a result of the overexpression of the *F5H* gene, the lignin of these trees contained more S-residues. The ratio of S:G residues had increased from 1.90 in the wild-type to 14.17 in the most extreme case. This change drastically increased pulping efficiency. The pulping severity required to reach a certain degree of delignification (indicated by the kappa number) was lower with increasing S:G ratios. The paper obtained from the transgenic trees required considerably less chemicals for bleaching, and when the same pulping and bleaching conditions were used for wild-type and transgenic wood, the paper from the transgenic wood was considerably brighter. The authors concluded that the use of these transgenic poplars could increase pulp throughputs by >60% while at the same time reduce the amount of chemical needed.

12.3.2 Forage and silage quality

The term forage refers to vegetative parts of crop plants that are fed to animals. The animals either graze the forage while it is still in the field, or they eat it after the harvest as dried feed. Examples of the latter include hay and straw obtained from wheat, maize, sorghum, sudan grass, millets and various forage grasses. Silage refers to the material prepared from chopped plants, including both vegetative and reproductive tissues, which were harvested while the plants were still green. The chopped plants are preserved as a result of organic acids secreted by microbes that were either native or that originated from a commercially available product. Maize is the most common source of silage. Both forage and silage provide energy to ruminant animals in the form of cell wall polysaccharides that are hydrolyzed in the rumen. Lignin present in the cell wall has been shown to limit the rate of digestion of forage and silage, largely because it shields the cell wall polysaccharides from hydrolytic enzymes (Ralph and Helm, 1993).

A class of mutants of maize (*Zea mays*), sorghum (*Sorghum bicolor*), and pearl millet (*Pennisetum glaucum*) are known as *brown midrib* mutants. These mutants have reduced lignin content and altered lignin subunit composition, and have therefore been suggested as good candidates to incorporate into breeding programs (Cherney et al., 1990; Cherney et al., 1991; Barrière et al., 1994).

The *brown midrib* mutants of maize, *bm1*, *bm2*, *bm3* and *bm4*, are naturally occurring mutants that have brown vascular tissue in their leaves and stems as a result of changes in lignin content and lignin composition. The *bm1* mutation affects the lignin biosynthetic enzyme cinnamyl alcohol dehydrogenase (CAD), which results in an accumulation of coniferaldehyde end groups in the lignin, and a decrease in S/G ratio (Halpin et al., 1998; Marita et al., 2003). The lignin in the *bm2* mutant (Burnham and Brink, 1932) contains fewer G-residues (Chabbert et al., 1994) and shows a disturbance in the tissue-specific patterns of lignification (Vermerris and Boon, 2001). The *bm3* mutant contains a defective *caffeic acid O-methyltransferase* (*COMT*) gene (Vignols et al., 1995), which results in a reduction in S-residues, and the incorporation of 5-hydroxyguaiacyl residues that participate in the formation of benzodioxane structures (**1.89**; Marita et al., 2003). Based on NMR studies the lignin of stems of the *bm4* mutant is not drastically different from that of the wild-type control, but additional analyses indicated changes in the lignin subunit composition that were similar to the changes in the *bm2* mutant (Barrière et al., 2004).

A collection of 28 *brown midrib* mutants exists in sorghum, and contains both chemically induced (Porter et al., 1978) and spontaneous mutants, some of which have been shown to be allelic (Bittinger et al., 1981; Saballos et al., 2005). The sorghum *brown midrib* mutations are abbreviated as *bmr*, because *bm* in sorghum is the abbreviation for the *bloomless* gene. Some of the *bmr* mutants, notably *bmr6*, *bmr12*, *bmr18*, and *bmr26* have been characterized chemically. The lignin of the *bmr6* mutant contains higher proportions of coniferaldehyde (**3.76**) and sinapaldehyde (**3.78**) and has reduced CAD activity (Pillonel et al., 1991). The *Bmr6* gene, however, has not yet been cloned. The mutants *bmr12*, *bmr18* and *bmr26* are allelic. The lignin composition in these mutants resembles that of the maize *bm3* mutant (Suzuki et al., 1997), and these mutants were shown to carry point mutations resulting in premature stop codons in the *COMT* gene (Bout and Vermerris, 2003).

There are two *brown midrib* mutants in pearl millet, one chemically induced (Cherney et al., 1988) and one spontaneous mutant (Degenhart et al., 1995). Analyses of the cell walls of these mutants *bmr* mutant indicated changes consistent with a reduction in COMT activity (Hartley et al., 1992; Degenhart et al., 1995; Lam et al., 1996), but the corresponding gene has not yet been cloned.

Digestibility studies with *brown midrib* mutants of maize (Barnes et al., 1971; Barrière et al., 1994; 2004; Fontaine et al., 2003), sorghum (Porter et al., 1978; Akin et al., 1986) and pearl millet (Cherney et al., 1990; Akin et al., 1991) have indeed shown that some of the mutants are considerably more easily digestible. The fact that some of the mutants do not show much improvement indicates that not all changes in cell wall composition are automatically of practical use.

12.3.3 Ethanol production from ligno-cellulosic biomass

The global demand for energy has been growing and is expected to continue to grow during the next decades. While oil supplies are currently still abundant, it is anticipated that peak oil production will be reached between 2010 and 2025. After that point, it will become increasingly expensive to pump and distribute oil (Belyaev et al., 2002). This, together with the desire of oil-importing countries to become less dependent on foreign oil, the uncertain political situation in many oil-exporting countries, and the need to curb the emission of greenhouse gases, has stimulated a search for alternative and renewable energy sources, including ethanol and bio-diesel. Ethanol is currently produced from sugar cane in Brazil, and

from corn grain in the US. The corn grain is processed to yield simple sugars that are converted to ethanol through fermentation. In the US this is carried out by a well-developed industry that produced 4 billion gallons (15 million liters) in 2005. Yet the current volume of ethanol represents a mere 2% of the volume of gasoline consumed in the US every year. In order for ethanol to contribute significantly as an alternative fuel, its production needs to go up drastically. This will require the use of alternative sources of fermentable sugars, such as plant-derived lignocellulosic biomass, which is an abundant source of the polysaccharides cellulose and hemicellulose (Ladisch et al., 1978). One of the first lignocellulose substrates to be used in the transition from a starch-based fuel ethanol industry to a lignocellulose-based ethanol industry will likely be corn stover, the agricultural residue resulting from corn grain production. A next step may be the development of specialized biomass crops grown specifically for the production of fermentable sugars.

Chang and Holtzapple (2000) identified lignin removal the dominant factor improving enzyme digestibility. Draude et al. (2001) showed that a removal of 67% of the lignin from softwood pulp resulted in a 174% increase in the ultimate yield of reducing sugars, an 88% increase in the ultimate yield of glucose, and an increase in the initial hydrolysis rate (over the first hour of hydrolysis) by 111%. Charles et al. (2003) confirmed these results in softwood pulps with similar numbers, and Yang and Wyman (2004) have shown similar results with corn stover. There appear to be two primary affects that lignin has on the enzymatic hydrolysis of cellulose within this matrix. The first effect is to prohibit cellulose fiber swelling, which reduces the enzyme accessible surface area. The other significant effect is through the irreversible adsorption of cellulases to lignin, thus preventing their action on the cellulose (Palonen et al., 2004). This 'titration' effect necessitates the use of more enzymes in order to saturate these non-productive adsorption sites on the surface of the biomass. This leads to prohibitively high enzyme costs for processing purposes. Pretreatment strategies aimed at removing lignin from the ligno-cellulosic biomass are currently under development (Mosier et al., 2005).

13. HYDROXYCINNAMIC ACID BIOSYNTHESIS

As discussed in Section 7, the general phenylpropanoid pathway originally included the biosynthesis of the hydroxycinnamic acids caffeic acid (**3.32**), ferulic acid (**3.33**), 5-hydroxyferulic acid (**3.34**), and sinapic acid (**3.35**) from *p*-coumaric acid (**3.30**), as well as the corresponding CoA-esters

through the action of the enzyme 4CL (see Figure 3-4). *In vitro* enzyme activity assays that became possible after the expression of recombinant proteins in heterologous systems demonstrated, however, that the hydroxylation and methoxylation reactions do not occur at the level of the acid, but instead at the more reduced forms such as the CoA esters, aldehydes and alcohols (Humphreys et al., 1999; Osakabe et al. 1999; Humphreys and Chapple, 2002).

The biosynthetic pathway towards ferulic acid and sinapic acid was redrawn after analysis of the *reduced epidermal fluorescence1* (*ref1*) mutant of Arabidopsis. When exposed to UV-radiation, the leaves of this mutant display reduced epidermal fluorescence as a result of lower levels of UV-absorbing sinapate esters. Sequence analysis of the cloned *REF1* gene indicated this gene encodes an $NADP^+$-dependent aldehyde dehydrogenase with *in vitro* enzyme activity against both coniferaldehyde (**3.76**) and sinapaldehyde (**3.78**). Many different species were shown to contain aldehyde dehydrogenase activity against coniferaldehyde and sinapaldehyde (Nair et al., 2004).

So it appears that the hydroxycinnamic acids are – at least in part – synthesized through the oxidation of aldehydes, rather than *via* ring substitutions of the free acids. There is evidence, however, that this is not the exclusive route towards the substitution pattern of the phenyl ring, at least not in all species. The xylem of poplar in which the lignin biosynthetic enzyme caffeoyl-CoA *O*-methyltransferase had been down-regulated through genetic engineering accumulated caffeic acid 3-*O*-glucoside and sinapic acid 4-*O*-glucoside (Meyermans et al., 2000). This suggests that these glucosides are precursors of the Coenzyme A esters in the lignin biosynthetic pathway, and therefore that the ring substitution probably occurred at the level of the free acids.

14. BIOSYNTHESIS OF SINAPOYL ESTERS

Sinapoyl esters are phenolic compounds found in members of the *Brassicaceae*, which includes the model plant *Arabidopsis thaliana*. The two major sinapoyl esters are sinapoyl malate (**3.92**) and sinapoyl choline (**3.93**), which accumulate in leaves and seeds, respectively. Sinapoylmalate plays a role in the protection against UV-radiation (Landry et al., 1995), whereas sinapoyl choline may be used as a storage form of choline in seeds (Shirley and Chapple, 2003). The precursor of these two esters is sinapate (**3.35**).

Sinapate is synthesized via the oxidation of sinapaldehyde (**3.79**) by an aldehyde dehydrogenase, as described in Section 13 of this chapter. Sinapaldehyde, in turn, is derived from the amino acid phenylalanine (**3.27**) *via* the general phenylpropanoid pathway (see Section 7), followed by a number of the hydroxylation and methylation reactions described in Section 10.

The first step in the biosynthesis of sinapoyl esters from sinapate (**3.35**) is glycosylation of the acid by the enzyme UDP-glucose:sinapic acid glucosyltransferase (SGT) to yield 1-*O*-sinapoylglucose (**3.91**). In leaves the enzyme sinapoylglucose:malate sinapoyltransferase (SMT) converts sinapoyl glucose (**3.91**) to sinapoyl malate (**3.92**). In seeds the enzyme sinapoylglucose:choline sinapoyltransferase (SCT) performs a similar reation to yield sinapoylcholine (**3.93**). This latter compound can be converted back to sinapate through the action of the enzyme sinapoylcholinesterase (SCE).

The elucidation of sinapoyl ester metabolism was aided by the availability of mutants. The *sng1* (***si****napoyl* ***g****lucose accumulator 1*) mutant of Arabidopsis had been identified based on a mutant screen for alterations in the composition of fluorescent compounds in the leaves. The screen was performed by thin layer chromatography and revealed that the leaves of the *sng1* mutant contained less sinapoylmalate and instead accumulated the precursor sinapoyl glucose (Lorenzen et al. 1996).

The Arabidopsis *SNG1* gene was cloned using a map-based cloning approach. The gene was shown to encode a serine carboxypeptidase-like (SCPL) protein (Lehfeldt et al., 2000). This class of enzymes has been shown to be involved in protein degradation, whereby the peptide bond between the penultimate and C-terminal amino acid residues of the protein or peptide substrates is cleaved. Expression of the Arabidopsis SNG1 protein in *E. coli* demonstrated SMT activity, and hence that this enzyme catalyzes a transesterification reaction, as opposed to the hydrolytic reaction typical for serine carboxypeptidases.

The SCT gene was also cloned based on the availability of a mutant. The Arabidopsis *sng2* mutant (***si****napoyl****g****lucose accumulator 2*) was identified in a mutant screen whereby seed extracts were analyzed with thin layer chromatography for the accumulation of sinapoylglucose (Shirley et al., 2001). The Arabidopsis *SNG2* gene was cloned via a combination of map-based cloning and a candidate gene approach. The candidate gene in this case was an *SCPL*-gene that had been identified during the sequencing of the

Arabidopsis genome, and that mapped to the same chromosome as the *SNG2* locus. This candidate gene indeed turned out to be the *SNG2* gene, as was demonstrated via complementation of the mutant. Activity towards sinapoylglucose was shown by expressing the *SNG2* cDNA in *E. coli* More detailed kinetic studies of recombinant SCT isolated from yeast are described by Shirley and Chapple (2003).

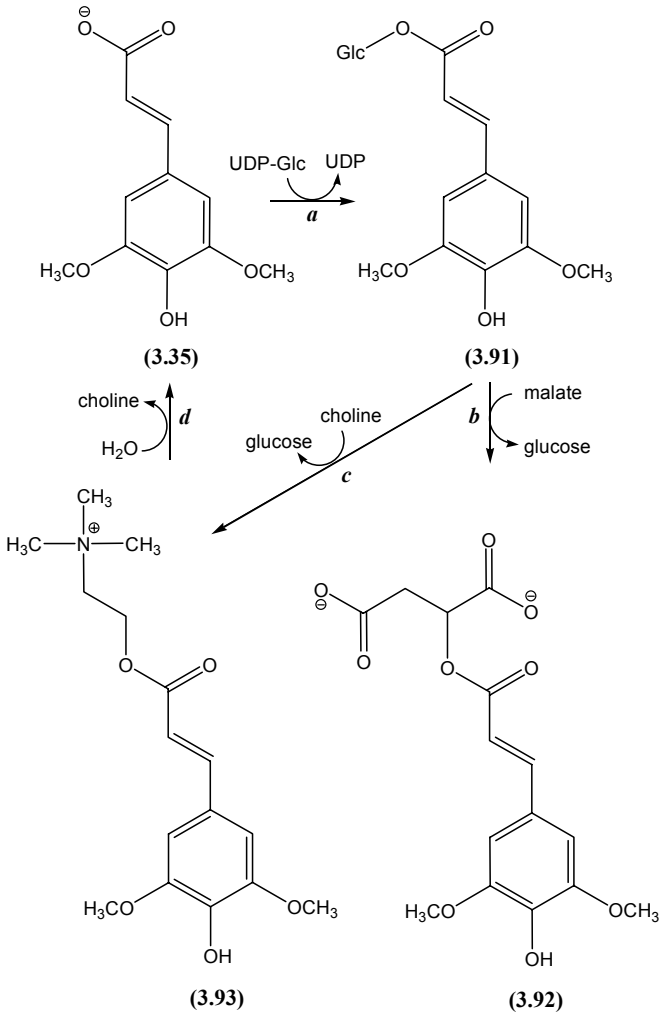


Figure 3-13. Sinapoyl ester metabolism catalyzed by the enzymes (***a***) UDP-glucose:sinapic acid glucosyltransferase (SGT), (***b***) sinapoylglucose:malate sinapoyltransferase (SMT), (***c***) sinapoylglucose:choline sinapoyltransferase (SCT), and (***d***) sinapoylcholinesterase (SCE).

15. COUMARIN BIOSYNTHESIS

In plants coumarins and hydroxycoumarins are believed to be synthesized from *trans*-cinnamic acid (**3.29**) and *trans-p*-coumaric acid (**3.30**), respectively, but the exact mechanism for its synthesis is still unknown. One possible biosynthetic route toward coumarin is *via o*-hydroxylation of **3.29** to give coumaric acid (**3.94**), followed by glycosylation to result in *trans*-coumaric acid-2-*O*-glucoside (**3.95**) (Figure 3-14).

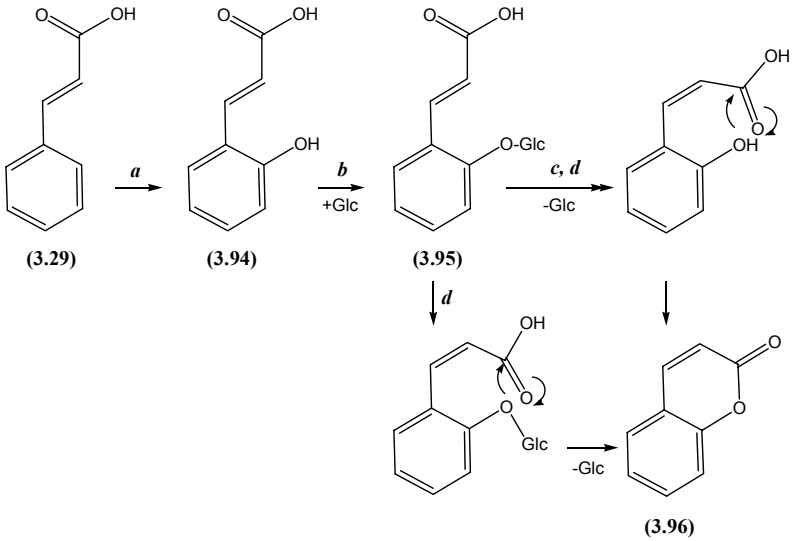


Figure 3-14. Possible biosynthetic route towards coumarin. (***a***) 2-hydroxylase, (***b***) glucosyl transferase, (***c***) β-glucosidase, (***d***) dimethylallyl transferase and/or UV light. The enzymes have not yet been identified.

Either the aglycone resulting from the action of a β-glucosidase or the glucoside, or possibly both, undergo *cis-trans* isomerization under influence of UV-light or possibly mediated by a dimethylallyl transferase. The last step of the biosynthetic pathway is an intramolecular esterification reaction, which can occur spontaneously, to yield coumarin (**3.96**). The enzymes that involved in these reactions have not been purified.

16. STILBENE BIOSYNTHESIS

Stilbene synthase shows similarity to chalcone synthase, which is not surprising given that stilbenes (**3.97**) also originate from the condensation of

p-coumaroyl-CoA (**3.31**) with three malonyl-CoA residues (**3.48**) (Schröder and Schröder, 1990; Tropf et al., 1994) (Figure 3-15). Since stilbene synthesis is relatively straightforward, stilbene synthase has been a target for genetic engineering of disease resistance in plants, as will be discussed in Chapter 6.

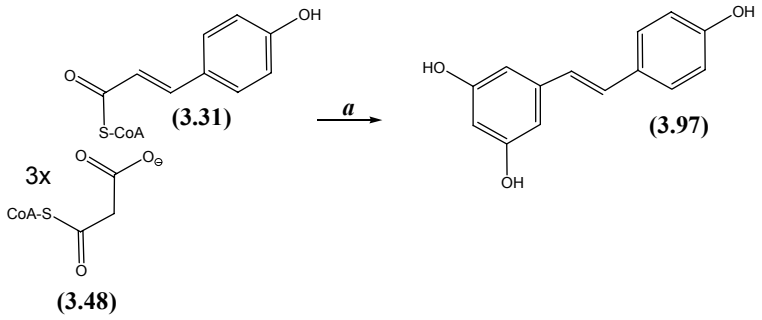


Figure 3-15. Biosynthesis of stilbene from *p*-coumaroyl-CoA with three molecules malonyl-CoA, catalyzed by the enzyme stilbene synthase (***a*;** E.C. 2.3.1.95).

17. BIOSYNTHESIS OF GALLOTANNINS AND ELLAGITANNINS

Gallotannins and ellagitannins make up the hydrolysable tannins and are derived from 1,2,3,4,6-penta-*O*-galloyl-β-D-glucopyranose (**3.102**). In oak leaves, and presumably many other plants that synthesize hydrolysable tannins, this compound is synthesized from β-glucogallin (1-*O*-galloyl-β-D-glucopyranose; **3.98**). The subsequent esterification of gallic acid residues (**3.47**; see Section 8) occurs in a specific sequence: C6-C2-C3-C4 (Figure 3-16). When enzyme preparations obtained from oak leaves were provided with UDP-glucose, β-glucogallin was formed, but also di- and trigalloyl-glucoses (Gross, 1983). This suggested that β-glucogallin (**3.98**) was both an acceptor and a donor of gallic acid residues. The different steps towards the biosynthesis of pentagalloylglucose are catalyzed by different enzymes. With the exception of the enzyme catalyzing the formation of β-glucogallin, these enzymes are very large, with molecular weights between 260 and 450 kDa.

Gallotannins contain an additional 10-12 gallic acid moieties per molecule. This is effectively a continuation of the esterification reactions that resulted in the formation of pentagalloylglucose (**3.99**). The major

difference between the formation of gallotannins and pentagalloylglucose – in terms of the chemistry – is that the additional gallic acid moieties have to react with phenolic hydroxyl groups, as opposed to the aliphatic hydroxyl groups of the glucose molecule. This process results in the formation of characteristic *meta*-depside groups (**1.91**).

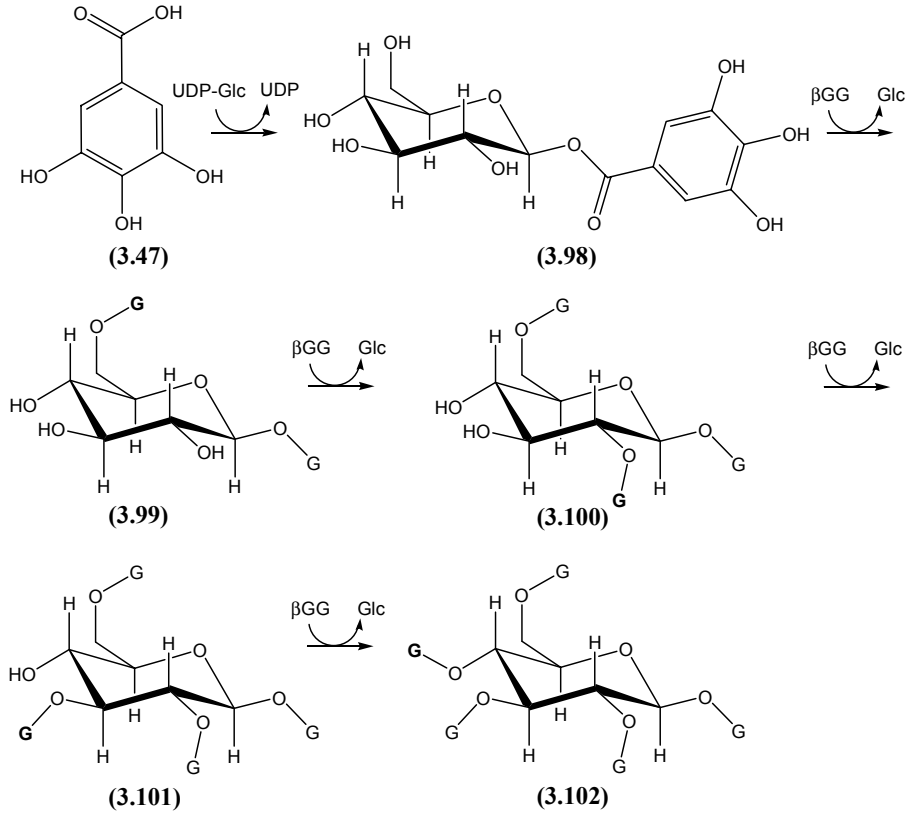


Figure 3-16. Biosynthesis of 1,2,3,4,6-penta-*O*-galloyl-β-D-glucopyranose (**3.102**) from gallic acid (G; **3.47**) and UDP-Glucose. The most recently added gallic acid residue is indicated by a G in bold face. The intermediates are β-glucogallin (βGG; **3.98**), 1,6-di-*O*-galloyl-β-D-glucopyranose (**3.99**), 1,2,6-tri-*O*-galloyl-β-D-glucopyranose (**3.100**), and 1,2,3,6-tetra-*O*-galloyl-β-D-glucopyranose (**3.101**).

Using cell-free extracts from sumac (*Rhus typhina*), Hofmann and Gross (1990) provided evidence that the addition of the gallic acid residues occurred in a manner similar to the acylation of pentagalloylglucose, with β-glucogallin (**3.95**) serving as a donor of gallic acid residues. The biosynthesis of hexa, hepta, and octa galloylated gallotannins appears to be catalyzed by several gallotannin synthesizing β-glucogallin-dependent

galloyltransferases that have a preferred but not unique substrate when it comes to the degree of substitution (penta-, hexa- or hepta- galloylglucose molecules) and the substitution pattern, *i.e.* the location of the *meta*-depside residues. As a consequence, a particular gallotannin molecule can have several biosynthetic origins (Niemetz and Gross, 2005).

Ellagitannins are formed from the oxidative coupling between gallic acid residues in pentagalloylglucose molecules leading to the formation of C-C coupled 3,4,5,3′,4′,5′-hexahydroxydiphenoyl (HHDP) residues (**1.97**; **1.98**). Tellimagrandin II (**3.103**) is a monomeric ellagitannin in which the $^4C_1$ conformation can be observed. The two galloyl residues are coupled *via* a 4,6-linkage.

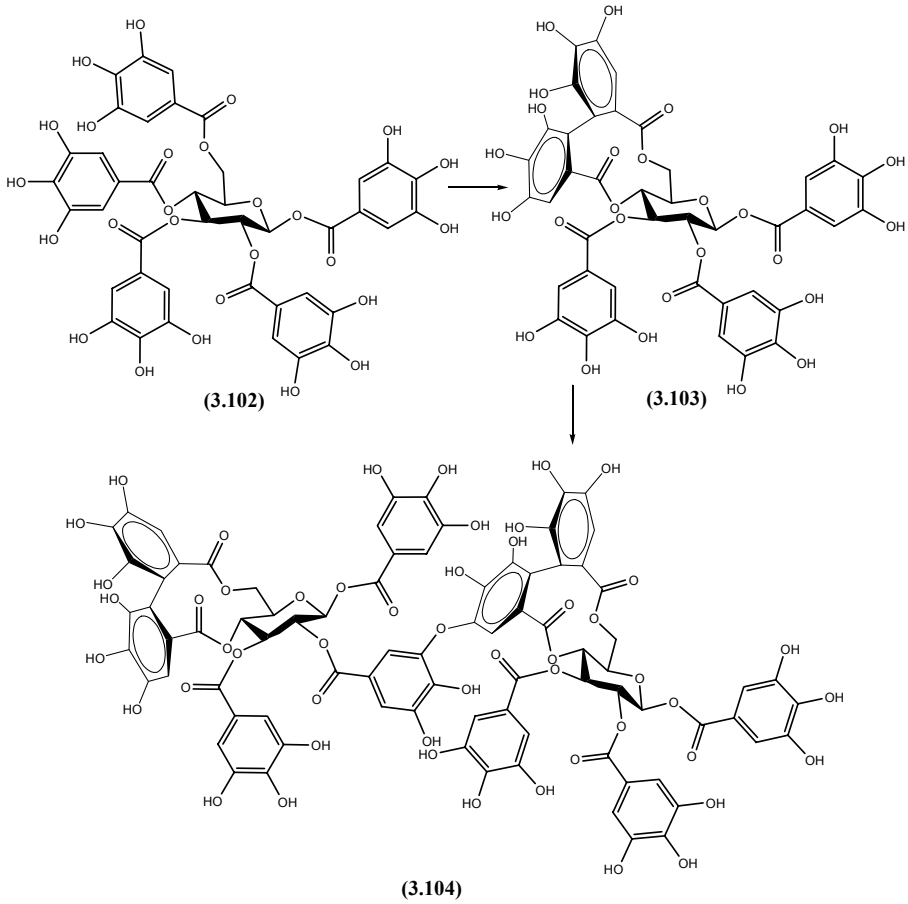


Figure 3-17. Biosynthesis of Tellimagrandin II (**3.103**) and Cornusiin E (**3.104**) from pentagalloylglucose (**3.102**) by polyphenol oxidases of the laccase class.

Other linkages, namely 1,6-, 3,6- and 2,4-*O* HHDP linkages, are possible, but require the less stable $^{1}C_{4}$ conformation of the glucose molecule. Linkages, both C-C and C-O, can also be formed between galloyl residues of different ellagitannin monomers, thereby giving rise to dimers, trimers, and tetramers. A wide range of compounds can thus be synthesized.

The biosynthesis of ellagitannins is not well understood. This is due to the many different compounds that exist in this class, and due to their complex chemical structure, which requires sophisticated analytical tools for identification. Niemetz and Gross (2003) isolated two enzymes that catalyze the formation of Tellimagrandin II (**3.103**) and Cornusiin E (**3.104**) from pentagalloylglucose. Both enzymes, pentagalloylglucose: O_2 oxidoreductase and Tellimagrandin II: O_2 oxidoreductase, require oxygen and were shown to belong to the laccase class polyphenol oxidases.

18. REFERENCES

Akin, D. E., Hanna, W. W., Snook, M. E., Himmelsbach, D. S., Barton, F. E., and Windham, W. R., 1986, Normal-12 and brown midrib-12 sorghum. II. Chemical variation and digestibility, *Agron. J.* **78**: 832-837.

Akin, D. E., and Rigsby, L. L. , 1991, Structure and digestibility of tissues in normal and brown midrib pearl millet (*Pennisetum glaucum*), *J. Sci. Food Agric.* **56**: 523-538.

Akiyama, T., Magara, K., Matsumoto, Y., Meshitsuka, G., Ishizu, A., and Lundquist, K., 2000, Proof of the presence of racemic forms of arylglycerol-β-aryl ether structure in lignin: studies on the stereo structure of lignin by ozonation, *J. Wood Sci.* **46**: 414–415.

Albert, S., Delseny, M., and Devic, M., 1997, BANYULS, a novel negative regulator of flavonoid biosynthesis in the Arabidopsis seed coat, *Plant J.* **11**: 289-299.

Anterola, A. M., and Lewis, N. G., 2002, Trends in lignin modification: a compehensive analysis of the effects of genetic manipulations/ mutations on lignification and vascular integrity, *Phytochem.* **61**: 221-294.

Bao, W., O'Malley, D.M., Whetten, R. and Sederoff, R.R., 1993, A laccase associated with lignification in loblolly pine xylem, *Science* **260**: 672−674.

Barnes, R. F., Muller, L. D., Bauman, L. F., and Colenbrander, V. F. , 1971, *In vitro* dry matter disappearance of brown-midrib mutants of maize, *J. Anim. Sci.* **33**: 881-884.

Barrière, Y., Chabbert, B., Tollier, M. T., and Monties, B., 1994, Breeding silage maize with brown-midrib genes. Feeding value and biochemical characteristics, *Agronomie* **14**: 15-25.

Barrière, Y., Ralph, J., Méchin, V., Guillaumie, S., Grabber, J. H., Argillier, O., Chabbert, B., and Lapierre, C., 2004, Genetic and molecular baisis of grass cell wall biosynthesis and degradability. II. Lessons from brown midrib mutants, *C. R. Biologies* **327**: 847-860.

Biermann, C.J., 1996, *Handbook of Pulping and Papermaking*, 2nd ed. Academic Press, San Diego, pp. 754.

Behringer, F.J., and Medford, J. I., 1992, A plasmid rescue technique for the recovery of plant DNA disrupted by T-DNA insertion, *Plant Mol. Biol. Rep* **10**: 190-198.

Belyaev, L., Marchenko, O. V., Filippov, S. P., Solomin, S. V., Stepanopa, T. B., and Kokorin, A. L., 2002, *World Energy and Transition to Sustainable Development*, Kluwer, Dordrecht, pp. 264.

Bittinger, T. S., Cantrell, R. P., and Axtell, J. D., 1981, Allelism tests of the brown-midrib mutants of sorghum, *J. Heredity* **72**: 147-148.

Boerjan, W., Ralph, J., and Baucher, M., 2003, Lignin biosynthesis, *Annu.*

Rev. Plant Biol. **54**: 519-546.

Bout, S., and Vermerris, W., 2003, A candidate gene-approach to clone the sorghum *Brown midrib* gene encoding caffeic acid *O*-methyltransferase, *Mol. Genet. Genomics* **269**: 205-214.

Broomhead, A. J., Rahman, M. M. A., Dewick, P. M., Jackson, D. E., and Lucas, J.A., 1991, Matairesinol as precursor of *Podophyllum* lignans, *Phytochem*. **30**: 1489–1492.

Brown, D. M., Zeef, L. A. H., Ellis, J., Goodacre, R., and Turnera, S. R., 2005, Identification of novel genes in Arabidopsis involved in secondary cell wall formation using expression profiling and reverse genetics, *Plant Cell* **17**: 2281-2295.

Canel, C., Moraes, R.M., Dayan, F.E., and Ferreira, D., 2000, Molecules of interest: podophyllotoxin, *Phytochem*. **54**: 115-120.

Chabbert, B., Tollier, M. T., Monties, B., Barrière, Y., and Argillier, O., 1994, Biological variability in lignification of maize: expression of the brown midrib bm2 mutation, *J. Sci. of Food Agric.* **64**: 455-460**.**

Chandler, V. L., Radicella, J. P., Robbins, T. P., Chen, J., and Turks, D., 1989, Two regulatory genes of the maize anthocyanin pathway are homologous: isolation of *B* utilizing *R* genomic sequences, *Plant Cell* **1**: 1175-1183.

Chang, C., Bowman, J. L., DeJohn, A. W., Lander, E. S., and Meyerowitz, E. M., 1988, Restriction fragment length polymorphism linkage map for Arabidopsis thaliana, *Proc. Natl. Acad. Sci. USA* **85**: 6856-6860.

Chang, V. S., and Holtzapple, M. T., 2000, Fundamental factors affecting biomass enzymatic reactivity, *Appl. Biochem. Biotechnol*. **84**: 5-37.

Chapple, C., 1998, Molecular-genetic analysis of plant cytochrome P450-dependent monooxygenases, *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **49**: 311-343.

Chapple, C. C. S., Vogt, T., Ellis, B. E., and Sommerville, C. R., 1992, An Arabidopsis mutant defective in the general phenylpropanoid pathway, *Plant Cell* **4**: 1413-1424.

Charles, N., Mansfield, S. D., Mirochnik, O., and Duff, S. J.B., 2003**,** Effect of oxygen delignification operating parameters on downstream enzymatic hydrolysis of softwood substrates, *Biotechnol. Progr.* **19**: 1606-1611.

Chen, C., Meyermans, H., Burggraeve, B., De Rycke, R. M., Inoue, K., De Vleesschauwer, V., Steenackers, M., Van Montagu, M. C., Engler, G. J., and Boerjan, W. A., 2000, Cell-specific and conditional expression of caffeoyl-coenzyme A-3-*O*-methyltransferase in poplar, *Plant Physiol.* **123**: 853-867.

Chen, W., Chao, G., and Singh, K. B., 1996, The promoter of a H_2O_2-inducible, *Arabidopsis* glutathione *S*-transferase gene contains closely linked OBF- and OBP1- binding sites, *Plant J.* **10**: 955-966.

Cheng, W.-H., Taliercio, E. W., and Chourey, P. S., 1996, The *Miniature1* seed locus of maize encodes a cell wall invertase required for normal development of endosperm and maternal cells in the pedicel, *Plant Cell* **8**: 971-983.

Cherney, J. H., Axtell, J. D., Hassen, M. M., and Anliker, K. S., 1988, Forage quality characterization of a chemically induced brown-midrib mutant in pearl millet, *Crop Sci.* **28**: 783-787.

Cherney, D. J. R., Patterson, J. A., and Johnson, K. D., 1990, Digestibility and feeding value of pearl millet as influenced by the brown-midrib, low-lignin trait, *J. Anim. Sci.* **68**: 4345-4351.

Cherney, J. H., Cherney, D. J. R., Akin, D. E., and Axtell, J. D., 1991, Potential of brown-midrib, low-lignin mutants for improving forage quality, *Adv. Agron.* **46**: 157-198.

Cone, K. C., Burr, F. A., and Burr, B., 1986, Molecular analysis of the maize anthocyanin regulatory locus C1, *Proc. Natl. Acad. Sci. USA* **83**: 9631-9635.

Cone, K. C., Cocciolone, S. M., Burr, F. A., and Burr, B., 1993a, Maize anthocyanin regulatory gene *pl* is a duplicate of *c1* that functions in the plant, *Plant Cell* **5**: 1795-1805.

Cone, K. C, Cocciolone, S. M., Moehlenkamp, C. A., Weber, T., Drummond, B. J., Tagliani, L. A., Bowen, B. A., Perrot, G. H., 1993b, Role of the regulatory gene *pl* in the photocontrol of maize anthocyanin pigmentation, *Plant Cell* **5**: 1807-1816.

Davin, B. L., and Lewis, N. G., 2005, Lignin primary structures and dirigent sites, *Curr. Opin. Biotechnol.* **16**: 1-9.

Davin, L. B., and Lewis, N. G., 2000, Dirigent proteins and dirigent sites explain the mystery of specificity of radical precursor coupling in lignan and lignin biosynthesis, *Plant Physiol.* **123**: 453-461.

Davin, L. B., Wang, H.-B., Crowell, A. L., Bedgar, D. L., Martin, D. M., Sarkanen, S., and Lewis, N. G., 1997, Stereoselective bimolecular phenoxy radical coupling by auxilary (dirigent) protein without an active center, *Science* **275**: 362-366.

Debeaujon, I., Peeters, A. J. M., Léon-Kloosterziel, K. M., and Koornneef, M., 2001, The *TRANSPARENT TESTA12* gene of Arabidopsis encodes a multidrug secondary transporter-like protein required for flavonoid sequestration in vacuoles of the seed coat endothelium, *Plant Cell* **13**: 853-872.

Degenhart, N.R., B.K. Werner, and G.W. Burton, 1995, Forage yield and quality of a *brown midrib* mutant in pearl millet, *Crop Sci*. **35**: 986–988.

Dellaporta, S. L., Greenblatt, I., Kermicle, J. L., Hicks, J. B., and Wessler, S. R., 1988, Molecular cloning of the maize *R-nj* allele by transpson tagging with *Ac*, *Stadler Genet. Symp*. 18: 263-282.

Dennis, D. T., and Blakeley, S. D., 2000, Carbohydrate metabolism, in: *Biochemistry and Molecular Biology of Plants*, Buchanan, B. B., Gruissem, W., Jones, R. J., eds., John Wiley & Sons, Somerset, NJ, pp. 630-675.

Denton, F. R., 1998, Beetle juice, *Science* **281**: 1285.

Devic, M., Guilleminot, J., Debeaujon, I., Bechtold, N., Bensaude, E., Koornneef, M., Pelletier, G., Delseny, M., 1999, The *BANYULS* gene encodes a DFR-like protein and is a marker of early seed coat development, *Plant J.* **19**: 387-398.

Dias, A. P., and Grotewold, E., 2003, Manipulating the accumulation of phenolics in maize cultured cells using transcription factors, *Biochem. Eng. J.* **14**: 207-216.

Doerge, R. W., 2002, Mapping and analysis of quantitative trait loci in experimental populations, *Nature Rev. Gen.* **3**: 43-53.

Dooner, H., and Nelson, O. E., 1977, Controlling element-induced alterations in UDPglucose: flavonoid glucosyltransferase, the enzyme specified by the *bronze* locus in maize, *Proc. Natl. Acad. Sci. USA* **74**: 5623-5627.

Draude, K. M., Kurniawan, C. B., and Duff, S. J. B., 2001, Effect of oxygen delignification on the rate and extend of enzymatic hydrolysis of lignocellulosic material, *Biores. Technol*. **79**: 113-120.

Evtuguin, D. V., and Amada, F. M. L., 2003, Application of electrospray ionization mass spectrometry to the elucidation of the primary structure of lignin, *Macromol. Biosci*. **3**: 339-343.

Federoff, N. V., Furtek, D., and Nelson, O. E., 1984, Cloning of the *bronze* locus in maize by a simple and generalizable procedure using the transposable controlling element *Activator* (*Ac*), *Proc. Natl. Acad. Sci. USA* **81**: 3825-3829.

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

Fieldes, M. A., and Dixon, B., 1988, Malate dehydrogenase isozymes in flax genotroph leaves: differences in apparent molecular weight and charge between and within L and S, *Biochem. Genet.* **26**: 249-260.

Fontaine, A. S., Bout, S., Barrière, Y., and Vermerris, W., 2003, Variation in cell wall composition among forage maize (*Zea mays* L.) inbred lines and its impact in digestibility analysis of neutral detergent fiber composition by pyrolysis-gas chromatography-mass spectrometry, *J. Agric. Food Chem.* **51**: 8080-8087.

Franke, R., Hemm, M. R., Denault, J. W., Ruegger, M. O., Humphreys, J. M., and Chapple, C., 2002a, Changes in the secondary metabolism and deposition of an unusual lignin in the *ref8* mutant of Arabidopsis, *Plant J.* **30**: 47-59.

Franke, R., Humphreys, J. M., Hemm, M. R., Denault, J. W., Ruegger, M. O., Cusumano, J. C., and Chapple, C., 2002b, The Arabidopsis *REF8* gene encodes the 3-hydroxylase of phenylpropanoid metabolism, *Plant J.* **30**: 33-45.

Freudenberg, K., 1965, Lignin: Its constitution and formation from *p*-hydroxycinnamyl alcohols, *Science* **148**: 595-600.

Frey, M., Chomet, P., Glawischnig, E., Stettner, C., Grün, S., Winklmair, A., Eisenreich, W., Bacher, A., Meeley, R. B., Briggs, S. P., Simcox, K., and Gierl, A., 1997, Analysis of a chemical plant defense mechanism in grasses, *Science* **277**: 696-699.

Gang, D. R., Costa, M. A., Fujita, M., Dinkova-Kostova, A. T., Wang, H.-B., Burlat, V., Martin, W., Sarkanen, S., Davin, L. B., and Lewis, N. G., 1999, Regiochemical control of monolignol radical coupling: a new paradigm for lignin and lignan biosynthesis, *Chem. & Biol.* **6**: 143-151.

Gelvin, S. B., 2000, *Agrobacterium* and plant genes involved in T-DNA transfer and integration, *Annu. Rev. Plant Physiol. Plant Mol. Biol*. **51**: 223-256.

Gelvin, S. B., 2003, *Agrobacterium*-mediated plant transformation: the biology behind the "gene-jockeying" tool, *Microbiol. Mol. Biol. Rev.* **67**: 16-37.

Goicoechea, M., Lacombe, E., Legay, S., Mihaljevic, S., Rech, P., Jauneau, A., Lapierre, C., Pollet, B., Verhaegen, D., Chaubet-Gigot, N., and Grima-Pettenati, J., 2005, *Eg*MYB2, a new transcriptional activator from *Eucalyptus* xylem, regulates secondary cell wall formation and lignin biosynthesis, *Plant J.* **43**: 553-567.

Goffner, D., Campbell, M.M., Campargue, C., Clastre, M., Borderies, G., Boudet, A. and Boudet, A.M., 1994, Purification and characterization of cinnamoyI-CoA: NADP oxidoreductase in *Eucalyptus gunnii*, *Plant Physiol.* **106**: 625-632.

Goujon, T., Sibout, R., Eudes, A., MacKay, J., and Jouanin, L., 2003, Genes involved in the biosynthesis of lignin precursors in *Arabidopsis thaliana*, *Plant Physiol. Biochem.* **41**: 677-687.

Grabber, J. H., Ralph, J., Hatfield, R. D., Quideau, S., Kuster T., Pell, A. N., 1996, Dehydrogenation polymer-cell wall complexes as a model for lignified grass walls, *J. Agric. Food Chem.* **44**: 1453–1459.

Gracey, A. Y., and Cossins, A. R., 2003, Application of microarray technology in environmental and comparative physiology, *Annu. Rev. Physiol.* **65**: 231-259.

Gross, G.G., 1983, Synthesis of mono-, di- and trigalloyl-β-D-glucose by β-glucogallin-dependent galloyltransferases from oak leaves, *Z. Naturforsch.* **38c**: 519-523.

Grotewold, E., Athma, P, and Peterson, T., 1991, Alternatively spliced products of the maize *P* gene encode proteins with homology to the DNA-Binding domain of *myb*-like transcription factors, *Proc. Natl. Acad. Sci. USA* **88**: 4587-4591.

Grotewold, E., Drummond, B. J., Boen, B., and Peterson, T., 1994, The *myb*-homologous *P* gene controls phlobaphene pigmentation in maize floral organs by directly activating a flavonoid biosynthetic gene subset, *Cell* **76**: 543-553.

Guan, S. Y., Mylnar, J., and Sarkanen, S., 1997, Dehydrogenative polymerization of coniferyl alcohol on macromolecular lignin templates, *Phytochem*. **45**: 911–918.

Halpin, C., Holt, K., Chojecki, J., Oliver, D., Chabbert, B., Monties, B., Edwards, K., Barakate, A., and Foxon, G. A., 1998, *Brown-midrib* maize (*bm1*) – a mutation affecting the cinnamyl alcohol dehydrogenase gene, *Plant J.* **14**: 545-553.

Hartley, R. D., Morrison, W. H., Borneman, W. S., Rigsby, L. L., O'Neill, M., Hanna, W. W., and Akin, D. E., 1992, Phenolic constituents of cell wall types of normal and brown midrib mutants of pearl millet (*Pennisetum glaucum* (L.) R. Br.) in relation to wall biodegradability, *J. Sci. Food Agric.* **59**: 211-216.

Harkin, J. M., and Obst, J. R., 1973, Lignification in trees: indication of exclusive peroxidase participation, *Science* **180**: 296-298.

Hatfield, R., and Vermerris, W., 2001, Lignin formation in plants: the dilemma of linkage specificity, *Plant Physiol.* **126**: 1351-1357.

Higuchi, T., 1985, Biosynthesis of lignin, in: *Biosynthesis and Biodegradation of Wood Components*, T. Higuchi, ed., Academic Press, Orlando, pp. 141-160.

Heller, W., and Kühnl, T., 1985, Elicitor induction of a microsomal 5-*O*-(4-coumaroyl) shikimate 3´-hydroxylase in parsley cell suspension cultures, *Arch. Biochem. Biophys*. **241**: 453-460.

Herrmann, K., and Weaver, L. M., 1999, The shikimate pathway, *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **50**: 473-503.

Hoffmann, L., Besseau, S., Geoffroy, P., Ritzenthaler, Ch., Meyer, D., Lapierre, C., Pollet, B., and Legrand, M., 2004, Silencing of hydroxycinnamoyl-coenzyme A shikimate/quinate hydroxyltransferase affects phenylpropanoid biosynthesis, *Plant Cell* **16**: 1446-1465.

Hoffmann, L., Maury, S., Martz, F., Geoffroy, P., and Legrand, M., 2003, Purification, cloning, and properties of an acyltransferase controlling shikimate and quinate ester intermediates in phenylpropanoid metabolism, *J. Biol. Chem* **278**: 95-103.

Hofmann, A. S., and Gross, G. G., 1990, Biosynthesis of gallotannins: formation of polygalloylglucoses by enzymatic acylation of 1,2,3,4, 6-penta-*O*-galloylglucose, *Arch. Biochem. Biophys*. **283**:530-532.

Humphreys, J. M., and Chapple, C., 2002, Rewriting the lignin road map, *Curr. Opin. Plant Biol.* **5**: 224-229.

Humphreys, J. M., Hemm, M. R., and Chapple, C., 1999, New routes for lignin biosythesis defined by biochemical characterization of recombinant ferulate 5-hydroxylase, a multifunctional cytochrome P450-dependent monooxygenase, *Proc. Natl. Acad. Sci. USA* **96**: 10045-10050.

Huntley, S. K., Ellis, D., Gilbert, M., Chapple, C., and Mansfield, S. D., 2003, Significant increases in pulping efficiency in C4H-F5H-transformed poplars: improved chemical savings and reduced environmental toxins, *J. Agric. Food Chem.* **51**: 6178-6183.

Innes, M. A., Gelfand, D. H., Sinsky, J. J., 1999, *PCR Applications. Protocols for Functional Genomics*, Academic Press, San Diego, pp. 566.

International Rice Genome Sequencing Project, 2005, The map-based sequence of the rice genome, *Nature* **436**: 793-800.

Johnson, C. S., Kolevski, B., and Smyth, D. R., 2002, *TRANSPARENT TESTA GLABRA2*, a trichome and seed coat development gene of Arabidopsis, encodes a WRKY transcription factor, *Plant Cell* **14**: 1359-1375.

Joseleau, J.-P., and Ruel, K., 1997, Study of lignification by noninvasive techniques in growing maize internodes, *Plant Physiol.* **114**: 1123-1133.

Kawaoka, A., Kaothien, P., Yoshida, K., Endo, S., Yamda, K., and Ebinuma, H., 2000, Functional analysis of tobacco LIM protein, Ntlim1, involved in lignin biosynthesis, *Plant J.* **22**: 289-301.

Kruger, N.J., and Von Schaewen, A., 2003, The oxidative pentose phosphate pathway: structure and organization, *Curr. Opin. Plant Biol*. **6**: 236-246.

Kühnl, T., Koch, U., Heller, W., and Wellmann, E., 1987, Chlorogenic acid biosynthesis: characterization of a light induced microsomal 5-*O*-(4-coumaroyl)-D-quinate/shikimate 3'-hydroxylase from carrot (*Daucus carota L.*) cell suspension cultures, *Arch. Biochem. Biophys.* **258**: 226-232.

Kühnl, T., Koch, U., Heller, W., and Wellmann, E., 1989, Elicitor induced *S*-adenosyl-L-methionine: caffeoyl-CoA 3-*O*-methyltransferase from carrot cell suspension, *Plant Sci.* **60**: 21-25.

Ladisch, M. R., Ladisch, C. M., and Tsao, G. T, 1978, Cellulose to sugars: New path gives quantitative yield, *Science* **201**: 743-745.

Lacombe, E., Hawkins, S., Van Doorsselaere, J., Piquemal, J., Goffner, D., Poeydomenge, O., Boudet, A.-M., and Grima-Pettenati, J., 1997, Cinnamoyl CoA reductase, the first committed enzyme of the lignin branch biosynthetic pathway: cloning, expression and phylogenetic relationships, *Plant J.* **11**: 429-441.

Lam, T. B. T., Iayama, K., and Stone, B. A., 1996, Lignin and hydroxycinnamic acids in walls of brown midrib mutants of sorghum, pearl millet and maize stems, *J. Sci. Food Agric.* **71**: 174-178.

Larson, R., and Coe, E. H., 1977, Gene-dependent flavonoid glucosyltransferase in maize, *Biochem. Genet.* **15**: 153-156.

Lechelt, C., Peterson, T., Laird, A., Chen, J., Dellaporta, S., Dennis, E., Starlinger, P., and Peacock, W. J., 1989, Isolation and molecular analysis of the maize *P* locus, *Mol. Gen. Genet*. **219**: 225-234.

Lehfeldt, C., Shirley, A. M., Meyer, K., Ruegger, M. O., Cusumano, J. C., Viitanen, P. V., Strack, D., and Chapple, C., 2000 , Cloning of the *SNG1* gene of Arabidopsis reveals a role for a serine carboxypeptidase-like protein as an acyltransferase in secondary metabolism, *Plant Cell* **12**: 1295-1306.

Leon, J., Shulaev, V., Yalpani, N., Lawton, M. A. and Raskin, I., 1995, Benzoic acid 2-hydroxylase, a soluble oxygenase from tobacco, catalyzes salicylic acid biosynthesis, *Proc. Natl. Acad. Sci. USA* **92**: 10413-10417.

Lewis, N. G., and Davin, L. B., 1998, The biochemical control of monolignol coupling and structure during lignan and lignin biosynthesis. In: *Lignin and Lignan Biosynthesis*, Vol. 697, N. G. Lewis, and S. Sarkanen, eds., ACS Symposium Series, Washington, DC, pp. 334-361.

Lewis, N. G., 1999, A 20th century roller coaster ride: a short account of lignification, *Curr. Opin. Plant Biol.* **2**: 153-162.

Li, L., Cheng, X. F., Leshkevich, J., Umezawa, T., Harding, S. A., and Chiang, V. L., 2001, The last step of syringyl monolignol biosynthesis in angiosperms is regulated by a novel gene encoding sinapyl alcohol dehydrogenase, *Plant Cell* **13**: 1567-1585.

Li, Y. H., Kajita, S., Kawai, S., Katayama, Y., and Morohoshi, N., 2003, Down-regulation of an anionic peroxidase in transgenic aspen and its effect on lignin characteristics, *J. Plant Res*. **116**: 175–182.

Lim, E.-.K, Li, Y., Parr, A., Jackson, R., Ashford, D. A., and Bowles, D. J., 2001, Identification of glucosyltransferase genes involved in sinapate metabolism and lignin synthesis in Arabidopsis, *J. Biol. Chem*. **276**: 4344-4349.

Liu, Y.-G., Mitsukawa, N., Oosumi, T., and Whittier, R. F., 1995, Efficient isolation and mapping of Arabidopsis thaliana T-DNA insert junctions by thermal assymetric interlaced PCR, *Plant J.* **8**: 457-463.

Lorenzen, M., Racicot, V., Strack, D., Chapple, C., 1996, Sinapic acid ester metabolism in wild type and a sinapoylglucose-accumulating mutant of Arabidopsis, *Plant Physiol.* **112**: 1625-1630.

Lourith, N., Katayama, T., and Suzuki, T., 2005, Stereochemistry and biosynthesis of 8-*O*-4′ neolignans in *Eucommia ulmoides*: diastereoselective formation of guaiacylglycerol 8-*O*-4′-(sinapyl alcohol) ether, *J. Wood Sci*. **51**: 370-378.

Marita, J., Vermerris, W., Ralph, J., and Hatfield, R. D., 2003, Variations in the cell wall composition of maize *brown midrib* mutants, *J. Agric. Food Chem.* **51**: 1313-1321.

McClintock, B., 1947, Cytogenetic studies of maize and Neurospora, *Carnegie Inst. Wash. Year Book* **46**: 146-152.

McDougall, G. J., Stewart, D., and Morrison, I. M., 1996, Tyrosine residues enhance cross-linking of synthetic proteins into lignin dehydrogenation products, *Phytochem.* **41**: 43-47.

McLaughlin, M., and Walbot, V., 1987, Cloning of a mutable *bz2* allele of maize by transposon tagging and differential hybridization, *Genetics* **117**: 771-776.

Messen, A., Höhmann, S., Martin, W., Schnable, P.S., Peterson, P.A., Saedler, H., and Gierl, A., 1990, The *En*/*Spm* transposable element of *Zea mays* contains splice sites at the termini generating a novel intron from a *dSpm* element in the *A2* gene. *EMBO J.* **9**: 3051 3057.

Meyer, K., Cusumano, J. C., Sommerville, C., and Chapple, C. C. S., 1996, Ferulate-5-hydroxylase from *Arabidopsis thaliana* defines a new family of cytochrome P-450 dependent monooxygenases, *Proc. Natl. Acad. Sci. USA* **93**: 6869-6874.

Meyermans, H., Morreel, K., Lapierre, C., Pollet, B., De Bruyn, A., Busson, R., Herdewijn, P., Devreese, B., Van Beeumen, J., Marita, J. M., Ralph, J., Chen, C., Burggraeve, B., Van Montagu, M., Messens, E., and Boerjan, W., 2000, Modifications in lignin and accumulation of phenolic glucosides in poplar xylem upon down-regulation of caffeoyl-coenzyme A *O*-methyltransferase, an enzyme involved in lignin biosynthesis, *J. Biol. Chem*. **275**: 36899–36909.

Mosier, N., Wyman, C., Dale, B., Elander, R., Lee, Y. Y., Holtzapple, M., and Ladisch, M., 2005, Features of promising technologies for pretreatment of lignocellulosic biomass, *Biores. Technol*. **96**: 673-686.

Nair, R. B., Bastress, K. L., Ruegger, M. O., Denault, J. W., and Chapple, C., 2004, The *Arabidopsis thaliana REDUCED EPIDERMAL FLUORESCENCE1* gene encodes an aldehyde dehydrogenase involved in ferulic acid and sinapic acid biosynthesis, *Plant Cell* **16**: 544-554.

Nakayama, T., Sato, T., Fukui, Y., Yonekura-Sakakibara, K., Hayashi, H., Tanaka, Y., Kusumi, T., and Nishino, T., 2001, Specifcity analysis and mechanism of aurone synthesis catalyzed by aureusidin synthase, a polyphenol oxidase homolog responsible for flower coloration, *FEBS Lett.* **499**: 107-111.

Nakayama, T., Yonekura-Sakakibara, K., Sato, T., Kikuchi, S., Fukui, Y., Fukuchi-Mizutani, M., Ueda, T., Nakao, M., Tanaka, Y., Kusumi, T., and Nishino, T., 2000, Aureusidin synthase: A polyphenol oxidase homolog responsible for flower coloration, *Science* **290**: 1163-1167.

Nesi, N., Jond, C., Debeaujon, I., Caboche, M., and Lepiniec, L., 2001, The Arabidopsis *TT2* gene encodes an R2R3 MYB domain protein that acts as a key determinant for proanthocyanidin accumulation in developing seed, Plant Cell **13**: 2099-2114.

Newman L. J., Perazza, D. E., Juda, L., Campbell, M. M., 2004, Involvement of the R2R3-MYB, *At*MYB61, in the ectopic lignification and dark-photomorphogenic components of the *det3* mutant phenotype. *Plant J.* **37**: 239-250.

Niemetz, R., and Gross, G. G., 2003, Oxidation of pentagalloylglucose to the ellagitannin, tellimagrandin II, by a phenol oxidase from *Tellima grandiflora* leaves, *Phytochem.* **62**: 301-306.

Niemetz, R., and Gross, G. G., 2005, Enzymology of gallotannin and ellagitannin biosynthesis, *Phytochem*. **66**: 2001-2011.

Niesbach-Klösgen, U., Barzen, E., Bernhardt, J., Rohde, W., Schwarz-Sommer, Z., Reif, H. J., Wienand U., and Saedler, H., 1987, Chalcone synthase genes in plants: A tool to study evolutional relationships, *J. Mol. Evol*. **26**: 213-225.

Nimz, H. H., and Ludemann, H. D., 1976, Carbon 13 nuclear magnetic resonance spectra of lignins: 6. Lignin and DHP acetates. *Holzforsch.* **30**: 33–40.

O'Malley, D. M., Whetten, R., Bao, W., Chen, C.-L., and Sederoff, R. R., 1993, The role of laccases in lignification, *Plant J.* **4**: 751-757.

O'Reilly, C., Shepherd, N. S., Pereira, A., Schwarz-Sommer, Z., Betram, I., Robertson, D. S., Peterson, P. A., and Saedler, H., 1985, Molecular cloning of the a1 locus of *Zea mays* using the transposable elements *En* and *Mu1*, *EMBO J*. **4**: 877-882.

Osakabe, K., Tsao, C. C., Li, L., Popko, J. L., Umezawa, T., Carraway, D. T., Smeltzer, R. H., Joshi, C. P., and Chiang, V. L., 1999, Coniferyl aldehyde 5-hydroxylation and methylation direct syringyl lignin biosynthesis in angiosperms, *Proc. Natl. Acad. Sci. USA* **96**: 8955-8960.

Ossipov, V., Salminen, J., Ossipova, S., Haukioja, E., and Pihlaja, K., 2003, Gallic acid and hydrolyzable tannins are formed in birch leaves from an intermediate compound of the shikimate pathway, *Biochem. Syst. Ecol*. **31**: 3-16.

Palonen, H., Tjerneld, F., Zacchi, G., and Tenkanen, M., 2004, Adsorption of *Trichoderma reesei* CBH I and EG II and their catalytic domains on steam pretreated softwood and isolated lignin, *J. Biotechnol*. **107**: 65-72.

Patten, A. M., Cardenas, C. L., Cochrane, F. C., Laskar, D. D., Bedgar, D. L., Davin, L. B., and Lewis, N. G., 2005, Reassessment of effects on lignification and vascular development in the *irx4 Arabidopsis* mutant, *Phytochem.* **17**: 2092-2107.

Patzlaff, A., Newman, L. J., Dubos, C., Whetten, R. W., Smith, C., McInnis, S., Bevan, M. W., Sederoff, R. R., and Campbell, M. M., 2003, Characterisation of a pine MYB that regulates lignification, *Plant J.* **36**: 743-754.

Paz-Ares, J., Goshal, D., and Saedler, H., 1990, Molecular analysis of the *C1-I* allele from *Zea mays*: A dominant mutant of the regulatory *C1* locus, *EMBO J.* **9**: 315-321.

Paz-Ares, J., Goshal, D., Wienand, U., Peterson, P. A., and Saedler, H., 1987, The regulatory *c1* locus of *Zea mays* encodes a protein with homology to *myb* proto-oncogene products and with structural similarities to transcriptional activators, *EMBO J.* **6**: 3553-3558.

Paz-Ares, J., Wienand, U., Peterson, P. A., and Saedler, H., 1986, Molecular cloning of the *c* locus of *Zea mays*: A locus regulating the anthocyanin pathway, *EMBO J.* **5**: 829-833.

Pelletier, M.K., and Shirley, B.W., 1996, Analysis of flavanone 3-hydroxylase in Arabidopsis seedlings: Coordinate regulation with chalcone synthase and chalcone isomerase, *Plant Physiol*. **111**: 339-345.

Pflieger, S., Lefebvre, V., and Causse, M. 2001, The candidate gene approach in plant genetics: a review, *Mol. Breeding* **7**: 275-291.

Pilate, G., Guiney, E., Holt, K., Petit-Conil, M., Lapierre, C., Leplé, J. C., Pollet, B., Mila, I., Webster, E. A., Marstorp, H. G., Hopkins, D. W., Jouanin, L., Boerjan, W., Schuch, W., Cornu, D., and Halpin, C., 2002, Field and pulping performances of transgenic trees with altered lignification, *Nature Biotechnol.* **20**: 607-612.

Pillonel, C., Mulder, M. M., Boon, J. J., Forster, B., and Binder, A., 1991, Involvement of cinnamyl-alcohol dehydrogenase in the control of lignin formation in *Sorghum bicolor* L. Moench, *Planta* **185**: 538-544.

Plaxton, W. C., 1996, The organization and regulation of plant glycolysis, *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **47**: 185-214.

Plaxton, W., and McManus, M, 2006, *Control of primary metabolism in plants*, Blackwell Publishing, Oxford, pp. 400.

Porter, K. S., Axtell, J. D., Lechtenberg, V. L., and Colenbrander, V. F., 1978, Phenotype, fiber composition, and in vitro dry matter disappearance of chemically induced brown midrib (*bmr*) mutants of sorghum, *Crop Sci.* **18**: 205-209.

Radicella, J. P., Turks, D., and Chandler, V. L., 1991, Cloning and nucleotide sequence of a cDNA encoding B-Peru, a regulatory protein of the anthocyanin pathway in maize, *Plant Mol. Biol*. **17**: 127-130.

Radicella, J. P., Brown, D., Tolar, L. A., and Chandler, V. L., 1992, Allelic diversity of the maize *B* regulatory gene: Different leader and promoter sequences of two *B* alleles determine distinct tissue specificities of anthocyanin production, *Genes Dev*. **6**: 2152-2164.

Raes, J., Rohde, A., Christensen, J. H., Van de Peer, Y., and Boerjan, W., 2003, Genome-wide characterization of the lignification toolbox in Arabidopsis, *Plant Physiol.* **133**: 1051-1071.

Ralph, J., and Helm, R.F., 1993, Lignin/hydroxycinnamic acid/ polysaccharide complexes: synthetic models for regiochemical characterization, in: Jung, H. G., Buxton, D. R., Hatfield, R. D., and Ralph J., eds., *Forage Cell Wall Structure and Digestibility*, ASA- CSSA-SSSA, Madison, WI., pp. 201-246.

Ralph, J., Lundquist, K., Brunow, G., Lu, F., Kim, H., Schatz, P. F., Marita, J. M., Hatfield, R. D., Ralph, S. A., Christensen, J. H., and Boerjan, W., 2004, Lignins: natural polymers from oxidative coupling of 4-hydroxy-phenylpropanoids, *Phytochem.* **3**: 29-60.

Ralph, J., Peng, J., Lu, F., Hatfield, R. D., and Helm, R. F., 1999, Are lignins optically active?, *J. Agric. Food Chem.* **47**: 2991-2996.

Ranocha P., Chabannes, M., Chamayou, S., Danoun, S., Jauneau, A., Boudet, A.-M., and Goffner, D., 2002, Laccase down-regulation causes alterations in phenolic metabolism and cell wall structure in poplar, *Plant Physiol.* **129**: 145-155.

Reddy, G. M., and Coe, E. H., jr., 1962, Intertissue specific complementation: A simple technique for direct analysis of gene sequence, *Science* **138**: 149-150.

Ribnicky, D. M., Shulaev, V., and Raskin, I., 1998, Intermediates of salicylic acid biosynthesis in tobacco, *Plant Physiol.* **118**: 565-572.

Ribot, E. M., 1998, Rapid amplification of transposon ends for the isolation, cloning and sequencing of transposon-disrupted chromosomal genes, *BioTechniques* **24**: 16-20.

Robbins, T. P., Walker, E. L., Kermicle, J. L., Alleman, M., and Dellaporta, S. L., 1991, Meiotic instability of the *R-r* complex arising from displaced intragenic exchange and intrachromosomal rearrangement, *Genetics*, **129**: 271-283.

Roesler, J., Krekel, F., Amrhein, N., and Schmid, J., 1997, Maize phenylalanine ammonia-lyase has tyrosine ammonia-lyase activity, *Plant Physiol.* **113**: 175-179.

Rogers, L. A., and Campbell, M. M., 2004, The genetic control of lignin deposition during plant growth and development, *New Phytol*. 164: 17-30.

Rogers, L. A., Dubos, C., Surman, C., Willment, J., Cullis, I. F., Mansfield, S. D., and Campbell, M. M., 2005, Comparison of lignin deposition in three ectopic lignification mutants, *New Phytol*. **168**: 123.

Saballos, A., Vermerris, W, and Ejeta, G., 2005, Allelic variation among *brown midrib* mutants of sorghum, 50th CSSA Annual Meeting, Salt Lake City, UT.

Sablowski, R. W., Moyano, E., Culianez-Macia, F. A., Schuch, W., Martin, C., and Bevan, M., 1994. A flower-specific Myb protein activates transcription of phenylpropanoid biosynthetic genes, *EMBO J.* **13**: 128-137.

Sagasser, M., Lu, G.-H., Hahlbrock, K., and Weisshaar, B., 2002, *A. thaliana TRANSPARENT TESTA1* is involved in seed coat development and defines the WIP subfamily of plant zinc finger proteins, *Genes Dev.* **16**: 138-149.

Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A., Arnheim, N., 1985, Enzymatic amplification of β-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia, *Science* **230**: 1350-1354.

Saito, K., Kobayashi, M., Gong, Z. Z., Tanaka, Y., and 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**: 181-190.

Schoenbohm, C., Martens, S., Eder, C., Forkmann, G., and Weisshaar, B., 2000, Identification of the Arabidopsis thaliana flavonoid 3´-hydroxylase gene and functional expression of the encoded P450 enzyme, *Biol. Chem*. **381**: 749–753.

Schoch, G., Goepfert, S., Morant, M., Hehn, A., Meyer, D., Ullmann, P., and Werck-Reichhart, D., 2001, CYP98A3 from *Arabidopsis thaliana* is a 3'-hydroxylase of phenolic esters, a missing link in the phenylpropanoid pathway, *J. Biol. Chem.* **276**: 36566-36574.

Schröder, J., and Schröder, G., 1990, Stilbene and chalcone synthases: related enzymes with key functions in plant-specific pathways, *Z. Naturforsch.* **45**: 1-8.

Schulze, A., and Downward, J., 2001, Navigating gene expression using microarrays – a technology review, *Nature Cell Biol*. **3**: E190-E195

Schwarz-Sommer, Z., Shepherd, N., Tacke, E., Gierl, A., Rohde, W., Le Clerq, L., Mattes, M., Berndtgen, R., Peterson, P. A., and Saedler, H., 1987, Influence of transposable elements on the structure and function of the *A1* gene of *Zea mays*, *EMBO J.* **6**: 287-294.

Schwarz-Sommer, Z., Davies, B., and Hudson, A., 2003, An everlasting pioneer: the story of *Antirrhinum* research, *Nature Rev. Genetics* **4**: 656-664.

Sederoff, R. R., Mackay, J. J., Ralph, J., and Hatfield, R. D., 1999, Unexpected variation in lignin, *Curr. Opin. Plant Biol.* **2**: 145-152.

Seidel, V., Windhövel, J., Eaton, G., Alfermann, W. A., Arroo, R. R., Medarde, M., Petersen, M., Woolley, J. G., 2002, Biosynthesis of podophyllotoxin in *Linum album* cell cultures, *Planta* **215**: 1031-1039.

Shirley, A. M., and Chapple, C., 2003, Biochemical characterization of sinapoylglucose:choline sinapoyltransferase, a serine carboxypeptidase-like protein that functions as an acyltransferase in plant secondary metabolism, *J. Biol. Chem.* **278**: 19870-19877.

Shirley, B. W., Hanley, S, and Goodman, H, M., 1992, Effects of ionizing radiation on a plant genome: analysis of two Arabidopsis transparent testa mutations, *Plant Cell* **4**: 333-347.

Sibout, R., Eudes, A., Pollet, B., Goujon, T., Mila, I., Granier, F., Seguin, A., Lapierre, C., and Jouanin, L., 2003, Expression pattern of two paralogs encoding cinnamyl alcohol dehydrogenases in Arabidopsis. Isolation and characterization of the corresponding mutants, *Plant Physiol.* **132**: 848-860.

Sterijades, R., Dean, J. F. D., Gamble, G., Himmelsbach, D. S., and Eriksson, K.-E. L., 1993, Extracellular laccase and peroxidases from sycamore maple (*Acer pseudoplatanus*) cell suspension cultures. Reactions with monolignols and lignin model compounds, *Planta* **190**: 75-87.

Suzuki, S., Lam, T. B. T., and Iiyama, K. 1997, 5-Hydroxyguaiacyl nuclei as aromatic constituents of native lignin, *Phytochem.* **46**: 695-700.

Syrjanen., K, and Brunow, G., 2000, Regioselectivity in lignin biosynthesis: the influence of dimerization and crosscoupling, *J. Chem. Soc. Perkin. Trans. I* **1**: 183-187.

Tamagnone, L., Merida, A., Parr, A., MAckay, S., Culianez-Macia, F. A., Roberts, K., and Martin, C., 1998, The AmMYB308 and AmMYB330 transcription factors from *Antirrhinum* regulate phenylpropanoid and lignin biosynthesis in transgenic tobacco, *Plant Cell* **10**: 135-154.

Tan, B. C., Schwartz, S. H., Zeevaart, J. A. D., and McCarthy, D. R., 1997, Genetic control of abscisic acid biosynthesis in maize, *Proc. Natl. Acad. Sci. USA* **94**: 12235-12240.

Terashima, N., Atalla, R. H., Ralph, S. A., Landucci, L. L., Lapierre, C., and Monties, B., 1996, New preparations of lignin polymer models under conditions that approximate cell well lignification: I. Synthesis of novel lignin polymer models and their structural characterization by 13C NMR., *Holzforsch*. **49**: 521-527.

Terashima, N., Fukushima, K., He, L.-F., and Takabe, K., 1993, Comprehensive model of the lignified plant cell wall, in: *Forage Cell Wall Structure and Digestibility,* Jung, H. G., Buxton, D. R., Hatfield, R. D., and Ralph J., eds., ASA- CSSA-SSSA, Madison, WI., pp. 247-270.

The Arabidopsis Genome Initiative, 2000, Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*, *Nature* **408**: 796-815.

Theres, N., Scheele, T., and Starlinger, P., 1987, Cloning of the *Bz2* locus of *Zea mays* using the transposable element *Ds* as a gene tag, *Mol. Gen. Genet.* **209**: 193-197.

Tobias, C. M, and Chow, E. K., 2005, Structure of the cinnamyl-alcohol dehydrogenase gene family in rice and promoter activity of a member associated with lignification, *Planta* **220**: 678-688.

Tropf, S., Lanz, T., Rensing, S. A., Schröder, J., and Schröder, G., 1994, Evidence that stilbene synthases have developed from chalcone synthases several times in the course of evolution, *J. Mol. Evol*. **38**: 610-618.

Vermerris, W., and Boon, J. J., 2001, Tissue-specific patterns of lignification are disturbed in the *brown midrib2* mutant of maize (*Zea mays* L.), *J. Agric. Food Chem.* **49**: 721-728.

Vignols, F., Rigau, J., Torres, M. A., Capellades, M., and Puigdomènech, P., 1995, The *brown-midrib3* (*bm3*) mutation in maize occurs in the gene encoding caffeic acid *O*-methyl transferase, *Plant Cell* **7**: 407-416.

Weaver, L. M., and Herrmann, K., 1997, Dynamics of the shikimate pathway in plants, *Trends Plant Sci.* **2**: 346-351.

Walbot, V., 1992, Strategies for mutagenesis and gene cloning using transposon tagging and T-DNA insertional mutagenesis, *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **43**: 49-82.

Walker, A. R., Davison, P. A., Bolognesi-Winfield, A. C., James, C. M., Srinivasan, N., Blundell, T. L., Esch, J. J., Marks, M. D., and Gray, J. C., 1999, The *TRANSPARENT TESTA GLABRA1* locus, which regulates trichome differentiation and anthocyanin biosynthesis in Arabidopsis, encodes a WD40 repeat protein, *Plant Cell* **11**: 1337-1350.

Wayne, M. L., and McIntyre, L. M., 2002, Combining mapping and arraying: an approach to candidate gene identification, *Proc. Natl. Acad. Sci. USA* **99**: 14903-14906.

Wegenmayer, H., Ebel, J. and Grisebach, H., 1976, Enzymic synthesis of lignin precursors: Purification and properties of a cinnamoyl-CoA: NADPH reductase from cell suspension cultures of soybean (*Glycine max* L.), *Eur. J. Biochem*. **65**: 529-536.

Werner, I., Bacher, A., and Eisenreich, W., 1997, Retrobiosynthetic NMR studies with ^{13}C-labeled glucose. Formation of gallic acid in plants and fungi, *J. Biol. Chem*. **272**: 25474-25482.

Werner, R. A., Rossmann, A., Schwarz, C., Bacher, A., Schmidt, H.-L., and Eisenreich, W., 2004, Biosynthesis of gallic acid in *Rhus typhina*: discrimination between alternative pathways from natural oxygen isotope abundance, *Phytochem.* **65**: 2809-2813.

Wienand U., Weydeman, U., Niesbach-Klösgen, U., Peterson, P. A., and Saedler, H., 1986, Molecular cloning of the *c2* locus of *Zea mays*, the gene coding for chalcone synthase, *Mol. Gen. Genet*. **203**: 202-207.

Wildermuth M. C., Dewdney, J., Wu, G., and Ausubel, F. M., 2001, Isochorismate synthase is required to synthesize salicylic acid for plant defence, *Nature* **414**: 562-565.

Winkel-Shirley, B., 2001, Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology, *Plant Physiol.* **126**: 485-493.

Xia, Z.-Q., Costa, M. A., Proctor, J., Davin, L. B., and Lewis, N. G., 2000, Dirigent-mediated podophyllotoxin biosynthesis in *Linum flavum* and *Podophyllum peltatum*, *Phytochem*. **55**: 537-549.

Xie, D.-Y., Sharma, S. B., Paiva, N. L., Ferreira, D., and Dixon, R. A., 2003, Role of anthocyanidin reductase, encoded by *BANYULS* in plant flavonoid biosynthesis, *Science* **299**: 396-399.

Yang, B., and Wyman, C. E., 2000, Effect of xylan and lignin removal by batch and flowthrough pretreatment on the enzymatic digestibility of corn stover cellulose, *Biotechnol. Bioeng*. **86**: 88-95.

Ye, Z.-H., Kneusel, R. E., Matern, U., and Varner, J. E., 1994, An alternative methylation pathway in lignin biosynthesis in *Zinnia*, *Plant Cell* **6**: 1427-1439.

For further reading

Davies, K., 2005, *Plant Pigments and Their Manipulation*, Blackwell Publishing, Oxford, pp. 352.

Dennis, D. T., and Blakeley, S. D., 2000, Carbohydrate metabolism, in: *Biochemistry and Molecular Biology of Plants*, Buchanan, B. B., Gruissem, W., Jones, R. J., eds., John Wiley & Sons, Somerset, NJ, pp. 630-675.

Plaxton, W., and McManus, M, 2006, *Control of Primary Metabolism in Plants*, Blackwell Publishing, Oxford, pp. 400.

Verpoort, R., and Alfermann, A.W., 2000, *Metabolic Engineering of Plant Secondary Metabolism*, Kluwer, Dordrecht, pp. 286.

Chapter 4

ISOLATION AND IDENTIFICATION OF PHENOLIC COMPOUNDS

A practical guide

1. INTRODUCTION

In the previous chapters we have discussed the different classes of phenolic compounds, their chemical properties, and their biosynthesis. The characterization of phenolic compounds relied on the ability to isolate them from plant tissues. In this chapter we will discuss methods to isolate and characterize phenolic compounds, and methods to visualize them *in planta*. Chapter 5 focuses on techniques for the identification and characterization of some of these compounds using recently developed mass-spectrometry-based techniques.

2. ISOLATION OF PHENOLIC COMPOUNDS

Soluble phenolic compounds can be isolated easily from plant tissue by extraction into methanol or methanol acidified with 0.1% (v/v) HCl. Since most phenols are present as glycosides one must use caution in their isolation to avoid hydrolysis. Precautionary techniques should include isolation in the dark and under cold conditions. This is particularly true when isolating anthocyanins, as they are easily broken down.

There are several ways to *estimate* the total amount of phenolic compounds present in plant tissue, but it is important to keep in mind that

none of these methods will detect all phenolic compounds. As a consequence, it is often necessary to perform several analyses.

2.1 Total phenolic content: the Folin-Ciocalteu reagent

Some of the commonly used methods to determine the content of phenolics are approximately 100 years old, and were initially developed by Folin and colleagues at Harvard Medical School to study the metabolism of proteins in humans. Folin and Denis (1912b) reported on a colorimetric method to detect the amino acid tyrosine (**3.28**) in protein hydrolysates. This method relied the reduction of a mixture of phosphotungstic (WO_4^{2-})-phosphomolybdic (MoO_4^{2-}) reagent by the phenolic hydroxyl group of tyrosine, resulting in the formation of a blue product. The color intensity could be quantified based on absorbance readings using an early version of the spectrophotometer. The Folin-Denis reagent is prepared by mixing sodium tungstate and (phospho)molybdic acid in phosphoric acid, boiling it for 2 hours, followed by cooling, diluting and filtering it (Folin and Denis, 1912a). This method was subsequently applied to the determination of phenolics in urine (Folin and Denis, 1915). A modification of this method was reported by Folin and Ciocalteu (1927). The modification consisted of the addition of lithium sulfate and bromine to the phosphotungstic-phosphomolybdic reagent at the end of the boiling period, followed by cooling and dilution. The addition of the lithium prevents the formation of a precipitate that would interfere with the quantification of the color intensity. The resulting reagent, referred to as the Folin-Ciocalteu reagent, was used to determine the content of tyrosine and tryptophan in protein hydrolysates, but can be used to determine the content of phenolics from a wide range of sources. Below follows a protocol for this method.

1. Dilute an aliquot of the sample 10:1 with water (9 parts water to 1 part sample). This is not necessary if the phenol content is low.
2. Add 2 ml of freshly prepared 2% (w/v) sodium carbonate (anhydrous) to 0.1 ml of the sample extract (diluted if necessary).
3. Mix vigorously on a Vortex mixer.
4. Let stand for 5 min.
5. While mixing on a Vortex add 0.1 ml of a 1:1 dilution of Folin-Ciocalteu reagent. This reagent can be purchased from chemical supply companies, such as Merck. If the reagent has a green color, it is no longer good and should be replaced.
6. Allow the sample to stand for a minimum of 30 minutes, but not more than one hour.

7. Read the absorbance in a spectrophotometer at 750 nm.

For blanks one can use ethanol, water or methanol, whichever the tissue extract was dissolved in last. Since this is a spectrophotometric assay, it is important to have a standard curve to relate the absorbance value to a concentration. Common compounds used to generate a standard curve are chlorogenic acid (**1.18**) or gallic acid (**1.5**). The concentration of phenolic compounds is then reported as chlorogenic acid or gallic acid equivalents, respectively.

Scalbert et al. (1989) used a slight modification of this method, whereby a 2.5 mL aliquot of the Folin-Ciocalteu reagent (diluted 10 times in water) and 2 mL of a 75g/L solution of sodium carbonate are added to 0.5 mL of the extract (diluted in methanol), followed by a 5 min. incubation in a 50°C waterbath. A potential complication of this method is the deglycosylation of phenolic compounds due to the heating.

The colorimetric assay based on the protocol developed by Folin and Ciocalteu (1927) can be used to determine the concentration of soluble phenolics, such as anthocyanins in the example above, as well as complex phenolics such as hydrolysable and condensed tannins. Swain and Hillis (1959) pointed out that variation in phenolic composition (*e.g.* tannins versus flavonoids) can influence the efficiency of the reduction of the Folin-Ciocalteu reagent, so that comparisons between samples may not always be appropriate, depending on what the origin of the samples is. Apple et al. (2001) investigated the appropriateness of the Folin-Ciocalteu reagent in comparisons of leaf samples obtained from different tree species. They argued that the use of this reagent offers an efficient method to estimate the *reducing capacity* of the sample, but due to variation in phenolic content and composition between samples, comparing phenolic contents across samples may not be meaningful. To demonstrate this, Apple et al. (2001) determined the tannin content of a set of leaves from sixteen tree species, a subset collected at different times of the year, and mixtures of commercially available tannins, using the Folin-Ciocalteu reagent as a measure of total tannin content, the butanol-HCl assay (Bate-Smith, 1977) to determine the content of condensed tannins (see Section 1.1.2.1), and the potassium iodate method (Haslam, 1965; Bate-Smith, 1977; Schultz and Baldwin, 1982; see Section 1.1.3.1) to determine the content of hydrolysable tannins. Since the tannins were extracted from the leaves, they could perform the reactions with the same amount of starting material for each of the samples, and evaluate to what extent the composition of the sample impacted the reducing capacity. The total phenolic content obtained with the Folin-Ciocalteu

reagent was primarily correlated with either the content of hydrolysable tannins, or the content of condensed tannins, or the combination, depending on the actual sample, confirming that the phenolic composition affected the reducing potential of the sample. They concluded that the use of the Folin-Ciocalteu assay resulted in over- or underestimation of the phenolic content values obtained with other methods.

2.2 Determining the content of condensed tannins

Scalbert et al. (1989) described several methods to determine the concentration of hydrolysable and condensed tannins (proanthocyanidins; see Chapter 1, section 3.13.1) in wood extracts. Tannins are complex and heterogeneous: In addition to the distinction between the flavonoid-based condensed tannins and the gallic acid-based hydrolysable tannins, each group can display a large degree of chemical variability that can affect the efficacy of the different assays. Interference of chemically related non-tannin compounds, such as flavonoids, can in certain cases bias the results.

2.2.1 The butanol-HCl assay

A colorimetric assay involves the oxidative cleavage of proanthocyanidins with ferrous sulfate. To 0.5 mL of aqueous plant extract is added a 5-mL portion of an acidic solution of ferrous sulfate (77 mg of $FeSO_4.7H_2O$ dissolved in 500 mL of 2:3 HCl/*n*-butanol). The tubes are loosely covered and placed in a water bath at 95°C for 15 min. The absorbance is read at 530 nm. The concentration of proanthocyanidins is expressed as cyanidin equivalents (used for the standard curve). The molecular extinction coefficient ε_{mol} that can be used to convert the absorbance values to a concentration is equal to 34700 L $mol^{-1}cm^{-1}$.

2.2.2 The vanillin assay

An alternative colorimetric method relies on the reaction with vanillin under acidic conditions. A 2-mL aliquot of a freshly prepared solution of vanillin (1 g/100 mL) in 70% sulfuric acid is added to 1 mL of aqueous plant extract. The mixture is incubated in a 20°C-waterbath and after exactly 15 min. the absorbance at 500 nm read. The concentration of proanthocyanidins is expressed as (+)-catechin equivalents (used for the standard curve). This assay is specific for flavonols. As a consequence, when using this assay to determine the concentration of condensed tannins, widely distributed monomeric flavonols, such as catechin (**1.39**) and epicatechin (**1.90**), can interfere (Hagerman and Butler, 1989).

2.2.3 Precipitation of condensed tannins with formaldehyde

A third method relies on the precipitation of proanthocyanidins with formaldehyde. First, the 'total phenolic content' is measured using the Folin-Ciocalteu reagent as described before. A 0.5 mole equivalent of phloroglucinol (**1.3**) is added for every gallic acid equivalent in the extract. To 2 mL of this plant extract and phloroglucinol is added 1 mL of a 2:5 HCl /H_2O solution and 1 mL of an aqueous solution of formaldehyde (13 mL of 37% formaldehyde diluted to 100 mL in water). After an overnight incubation at room temperature, the unprecipitated phenols are estimated in the supernatent by the Folin-Ciocalteu method. The precipitate contains the proanthocyanidins and the known amount of phloroglucinol, which is always quantitatively precipitated.

2.3 Determining the content of gallotannins

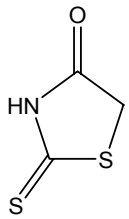
Gallotannins are hydrolysable tannins and contain a gallic acid residue esterified to a polyol (see Chapter 1, section 3.13.2).

2.3.1 The potassium iodate assay

Gallotannins, can be detected quantitatively with the potassium iodate assay. This assay was first described by described Haslam (1965), and is based on the reaction of potassium iodate (KIO_3) with galloyl esters, which will form a red intermediate and ultimately a yellow compound. The concentration of the red intermediate can be measured spectrophotometrically at 550 nm. The reaction was initially performed by adding 1.5 mL of a saturated potassium iodate solution to 3.5 mL of tannin, followed by a 40-min. incubation at 0°C for (Haslam, 1965). Since the red intermediate turns yellow over time, it is important to be consistent in terms of the time and temperature of the reaction. Bate-Smith (1977) recommended performing the reaction at 15°C until a maximum absorbance was reached (regardless of the time). The formation of the precipitate – resulting from the absence of free hydroxyl groups on the pentagalloyl parent residue – can be avoided by changing the solvent of the reaction. Aqueous acetone (20% (v/v) or methanol above 10% (v/v) generally resulted in less precipitation than just water. Since most standard curves for this assay are made with tannic acid, the concentration of hydrolysable tannins is expressed as tannic acid equivalents (TAE).

2.3.2 The rhodanine assay

Despite the changes Bate-Smith (1977) made to the original protocol, the time- and temperature sensitivity of the iodate assay, as well as the cross-reactivity with ellagitannins negatively impact the reproducibility of this method. Inoue and Hagerman (1988) developed the rhodanine assay. Rhodanine (2-thio-4-ketothiazolidine; **4.1**) reacts with the vicinal hydroxyl groups of gallic acid to produce a red complex that can be detected spectrophotometrically at 520 nm. This reaction is specific for gallic acid, and can thus be used for the detection of gallotannins. This requires acid hydrolysis of the gallotannins before the reaction with rhodanine. Tannins are extracted from the plant tissue in 1 mL 70% (w/v) acetone in water per 100 mg dry sample in a sonicator at 4°C. The extract is then filtered through a sintered glass filter. The filtrate containing the tannins is collected into a glass ampule. The residue in the filter is then washed with 5 mL 2N sulfuric acid, which is added to the ampule. The ampule is frozen and vacuum-sealed, and then heated at 100°C for 26 h. to hydrolyze the gallotannins.



(4.1)

The contents of the ampule are diluted in water to a final volume of 50 mL. A 1-mL sample is then taken for the assay. To this sample 1.5 mL 0.667% (w/v) rhodanine in methanol is added. After exactly 5 min. 1 mL 0.5N KOH is added. After 2.5 min. water is added to a final volume of 25 mL. The absorbance is read at 520 nm after a 5-10 min. incubation. A standard curve is made by reaction of gallic acid in 0.2N sulfuric acid with the rhodanine solution. Hagerman and Butler (1989) argued that this assay is more suitable than the potassium iodate assay for the determination of hydrolysable tannins, although it has to be kept in mind that the rhodanine assay is sensitive to any gallic acid ester, including those in non-tannin compounds.

A helpful review of the different methods for the analysis of tannins and the rationale for choosing one over the other was presented by Hagerman and Butler (1989). They also pointed out that it was critical to choose a suitable standard to compare the experimental data with. Given the

difficulty of isolating chemically pure tannins, they recommended using either non-tannin standards (such as gallic acid or cathechin), in which case the data are expressed in 'equivalents of the standard', or using commercially available tannic acid preparations. In the latter case it is critical to indicate the source of the standard, since the commercial preparations vary from each other considerably.

2.4 Determining the content of ellagitannins

Ellagitannins are characterized by the presence of hexahydroxydiphenoyl esters (HHDP; **1.94**) with a polyol such as glucose. Acid hydrolysis of ellagitannins will result in the release of HHDP units, which will spontaneously form ellagic acid (**1.96**), as described in Chapter 1, section 3.13.3. Hence, the quantification of ellagic acid reflects the content of ellagitannins present in the sample.

2.4.1 Nitrous acid oxidation

Ellagic acid concentration can be determined through an oxidation reaction with nitrous acid, and two methods have been described by Bate-Smith (1977) and by Wilson and Hagerman (1990), respectively.

Ellagitannins can be isolated from plant tissue by extraction of freeze-dried and ground tissue in 2N sulfuric acid (5-10 mg of sample/mL of acid), followed by freezing in a dry ice/2-propanol bath, vacuum sealing, and heating at 100°C for 10 h. in glass ampules. After hydrolysis, the ampules are cooled to room temperature, opened, cooled in an ice bath for 10 min. and filtered through a membrane filter. The residue on the filter (containing the ellagic acid) is then washed with several volumes of ice-cold wash solvent (acetone/H_2O/concentrated HC1 (70:30:1 v/v/v) and air-dried. When the residue is dry, both sample and filter are transferred to a test tube and dissolved in 10 mL pyridine. Undissolved material is removed by a second filtration. Samples are filtered through a glass-fiber filter supported on a fine sintered glass filter to remove insolubles (undissolved membrane filter and plant material). The final filtrate is assayed for ellagic acid (Wilson and Hagerman, 1990).

Scalbert et al. (1989) summarized the method of Bate-Smith (1977): In a tube sealed with a Teflon-lined screw cap, 0.2 mL of the aqueous plant extract (1 mL if the extract is not very concentrated) is added to 1.8 mL of 1:1 methanol/water (1 mL of 9:1 methanol/water if the sample is not concentrated) and 0.16 mL of aqueous 6% (v/v) acetic acid. Nitrogen is

bubbled for 5-10 min and 0.16 mL of an aqueous solution of 6% (w/v) sodium nitrite ($NaNO_2$) is added. Nitrogen is bubbled through the solution for a few more seconds, after which the tube is sealed and incubated in a 25°C water bath for 100 min. A blue product is formed during the reaction, which is quantified by reading the absorbance with a spectrophotometer at 590 nm. The concentration of HHDP esters is expressed as 4, 6-hexahydroxydiphenoyl-glucose equivalents (ε_{mol}= 2169 L $mol^{-1}cm^{-1}$).

2.4.2 The $NaNO_2$/HCl assay

Wilson and Hagerman (1990) described an alternative spectrophoto metric procedure to quantify ellagic acid. This method has three advantages over the procedure developed by Bate-Smith (1977): **1**) it is not sensitive to gallic acid, which can result in an overestimate of the ellagitannin content, **2**) it is more sensitive, and **3**) it is more convenient because oxygen does not interfere with the reaction.

To maximize the formation of a red-colored product in the assay developed by Wilson and Hagerman (1990), (the sample containing) ellagic acid is dissolved in pyridine (which needs to be kept in a fume hood) to a volume of 2.1 mL. A volume of 0.1 mL concentrated HCl is added, and the mixture is brought to 30°C. After a volume of 0.1 mL 1% (w/v) $NaNO_2$ is added, the mixture is quickly mixed and the absorbance at 538 nm is read immediately afterwards ($A_{538,t0}$). A red product that decays over time forms while the sample is incubated. After 36 min. at 30°C the red product reaches its maximum concentration, and the absorbance ($A_{538,t36}$) is recorded again. The ellagic acid concentration in the sample is a function of the difference between $A_{538,t0}$ and $A_{538,t36}$. Glass or quartz cuvettes need to be used, because pyridine will dissolve plastic cuvettes. The assay is also sensitive to residual detergent in the test tubes, so that new glass test tubes need to be used for each assay.

This procedure is based on the formation of the electophile NO^+, which can react with an ellagic acid residue (**4.2**) at two sites, either *via* a substitution reaction which results in **4.3**, or an addition reaction that results in the nitrosyl dienone **4.4**. These compounds can decay to form the quinine oxime **4.5**, which under alkaline conditions forms the red product **4.6**. When related compounds, such as gallic acid, phloroglucinol, hydroxycinnamic acids, and phenol are subjected to this assay, a yellow-brown product is formed, which does not interfere with the spectrophotometric detection of ellagic acid.

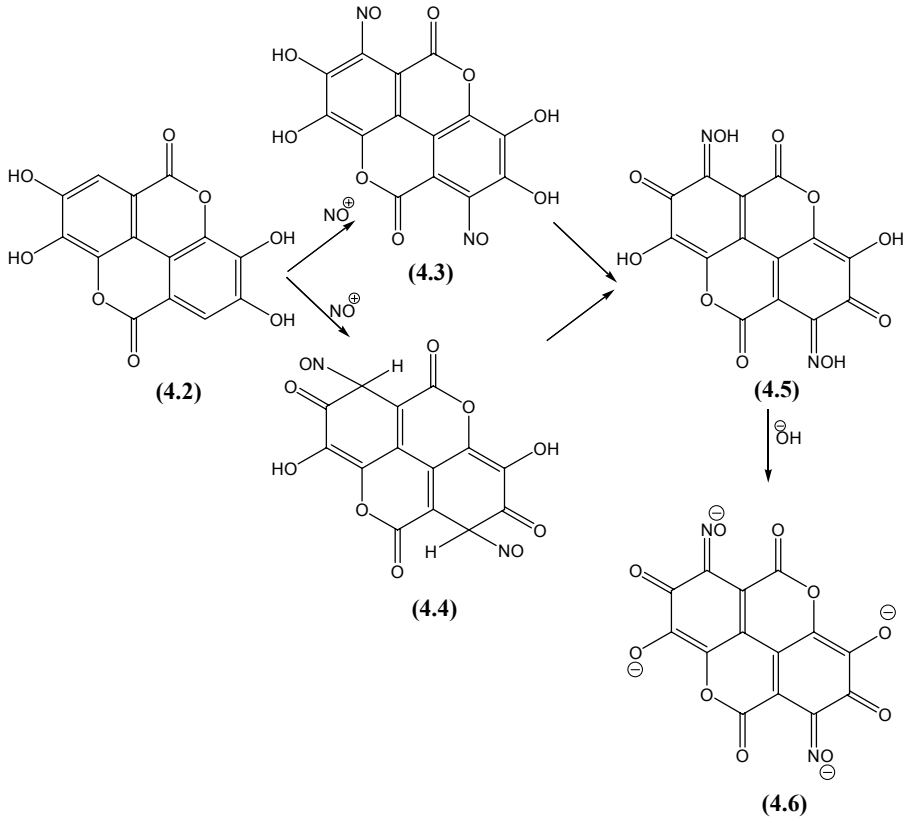


Figure 4-1. Formation of a red chromophore (**4.6**) from the reaction of ellagic acid with the electrophile NO^+, based on Wilson and Hagerman (1990).

2.5 Determining lignin content

Lignin is a complex phenolic cell wall polymer that is chemically cross-linked with hemicellulose and cell wall proteins. Most of the methods to determine lignin content are based on the removal of all other cell wall constituents, typically through acid hydrolysis, which will readily remove hemicellulose under mild conditions, and non-crystalline cellulose under more severe conditions. Several different methods will be discussed below. The different methods have also been extensively reviewed and compared by Hatfield et al. (1994), Brinkmann et al. (2002), Fukushima and Hatfield (2004), and Hatfield and Fukushima (2005).

2.5.1 Klason lignin

Klason lignin is the fraction of lignin that remains as a solid residue after the polysaccharides in a cell wall extract have undergone hydrolysis in sulfuric acid. The cell wall extract can be obtained by several different methods, which have in common that soluble sugars, proteins, and unbound phenolics are removed from dried and ground plant tissue. Ground woody tissue is commonly extracted in a 2:1 (v/v) mixture of benzene (or the less noxious toluene) and ethanol, followed by extraction in ethanol, and finally water. Theander and Westerlund (1986) recommended an extraction protocol for herbaceous species. This involves extraction of 1-3 gram dried and ground samples in an ultrasonic waterbath, first in 80% (w/v) ethanol (3 x 75 mL; 15 min. each) and then in petroleum ether (boiling point of 60-70°C; 2 x 50 mL, 10 min. each). After removal of the solvent, the samples are dried. The starch is removed from the samples by resuspending 1.5-2 g of the dried residue in 75 mL 0.1 M sodium acetate buffer pH 5.0, containing α-amylase (100 L, Termamyl 120L). The suspension is kept in a glass tube with a screwcap and placed in a boiling water bath, with occasional shaking. After cooling to 60°C, the content in the tube is incubated with 500 µL amyloglucosidase suspension and kept capped at 60°C overnight in a shaking water bath. After centrifugation for 15 min. the supernatant is removed and filtered into an evaporation flask. The insoluble residue remaining in the tube is washed successively by suspension (ultrasonication, 5 min.) and centrifugation in water (2 x 50 mL), ethanol (2 x 50 mL), and acetone (50 mL). The remaining insoluble residue (containing the water-insoluble fiber fraction) is dried with warm air and weighed.

Klason lignin is a measure of the total amount of lignin; it is not informative with respect to lignin composition. Despite attempts to extract proteins prior to the lignin determination, Klason lignin tends to be contaminated with proteins (Monties, 1989). In addition, covalently linked hydroxycinnamic acids can be included as part of the lignin. In order to measure Klason lignin in herbaceous plants, the proteins and hydroxycinnamic acids should ideally be removed by an acidic detergent before the sulfuric acid treatment is applied. This then leads to acid detergent lignin (see 1.1.6) and is also referred to as ‘core lignin’, and acid-soluble lignin, or ‘non-core lignin’. The distinction between core and non-core lignin should not be used, because the non-core lignin fraction contains mostly hydrolyzed hydroxycinnamic acids (Ralph and Helm, 1993; Jung and Deetz, 1993), which are distinct from the actual lignin polymer.

The Klason lignin procedure was first developed by Klason in the early 1900s and is based on the acid hydrolysis of all cell wall polysaccharides, with the remaining residue being the lignin. A protocol optimized for small samples based on the Klason method was described by Hatfield et al. (1994), which in turn was based on a protocol described by Theander and Westerlund (1986). The procedure involves the hydrolysis of all polysaccharides in a 100-mg sample of extracted cell walls, by treating it with 1.5 mL 12 M sulfuric acid that has been chilled on ice. The mixture is kept on ice for 30 min. and then shaken at 30°C for 2 h. Deionized water (9.75 mL) is added to the mixture, after which a secondary hydrolysis is performed for 1 h. in an autoclave at 121°C. The black liquid that results is filtered through a Whatman GF/A glass filter placed in a Hirsch funnel. The solid residue on the filter is washed 3 times with 5 mL of hot deionized water, and then dried for 48 h. in a 55°C oven. The dried residue contains the isolated lignin and minerals. The filter and residue are weighed. The lignin is then removed by a 5-hour incubation in a 500°C furnace. The samples are weighed again, and the loss in mass resulting from the incubation in the furnace is attributed to lignin. The Klason lignin assay is therefore a gravimetric method, and lignin content is reported as mg/g.

2.5.2 Acid detergent lignin

Acid detergent lignin (ADL) is part of a sequential analysis developed to determine characteristics related to digestibility and feed quality of forage samples, such as alfalfa and various grasses (Van Soest, 1967). This analysis includes the gravimetric determination of neutral detergent fiber (NDF), acid detergent fiber (ADF) and ADL (Van Soest, 1963; Van Soest and Wine, 1967). NDF is a measure of the total cell wall content (cellulose, hemicellulose and lignin). Subtracting ADF from NDF results in an estimate of the hemicellulose content, whereas subtracting ADL from ADF results in an estimate of the cellulose content (Jung, 1997).

NDF content is determined by extracting a 0.5-g sample in boiling neutral detergent solution (3% (w/v) sodium dodecyl sulfate (SDS; sodium lauryl sulfate)), 18 mM sodium tetraborate decahydrate ($Na_2B_4O_7 \cdot 10\ H_2O$), 32 mM sodium phosphate (dibasic; Na_2HPO_4), 50 mM Na_2-EDTA, and 1% (v/v) ethylene glycol (2-ethoxyethanol) for 1 h. The suspension is filtered, and the cell wall residue on the filter is washed in hot water and acetone. After drying overnight at 100°C the sample is weighed.

ADF is determined on the dried NDF sample, by boiling for 1 h. in 200 mL acid detergent solution (0.5 M sulfuric acid, 2% (w/v) CTAB

(hexadecyltrimethylammonium bromide)), followed by filtration, washing in hot water, and rinsing in acetone. The resulting ADF residue is weighed after overnight incubation at 100°C. The ADF residue is then hydrolyzed in 12 M sulfuric acid for 3 h. The insoluble residue that remains is collected by filtration, washed with hot water, rinsed with acetone, and dried overnight at 100°C. ADL is represented by the difference in weight between this dried, residue containing lignin and minerals, and the residue remaining after heating for 5 h. at 450°C, containing only minerals.

Hatfield et al. (1994) compared the Klason and ADL methods on samples derived from a variety of species representing different maturities (and hence degrees of lignification). While both methods showed in increase in lignin content with increasing maturity, the Klason lignin content was consistently higher than the ADL content for all samples. This difference could be the result of protein contamination of the Klason lignin. As part of the ADL determination, proteins are removed by the neutral detergent solution. In order to investigate this in more detail, Hatfield et al. (1994) determined N-content of the different samples and reported a negative correlation between N-content and lignin content, suggesting that proteins were not responsible for the higher Klason lignin values. This was further substantiated by performing the procedure on isolated cell wall polysaccharides to which nitrogen (either bovine serum albumine, lysine or ammonia sulfate) was added. The addition of nitrogen did not affect the amount of residue that was recovered. The N-content of both Klason lignin and ADL was attributed to cell wall proteins cross-linked with the other cell wall constituents. The difference between ADL and Klason lignin is therefore most likely the result of some of the lignin being solubilized in the acid detergent solution (Hatfield and Fukushima, 2005). Based on pyrolysis-gas chromatography-mass spectrometry (see 1.3.4), lignin subunit composition did not significantly differ between the two types of residues. The ADL contained some contaminants from the extraction procedure. Hatfield et al. (1994) concluded that while both methods are of value for the characterization of forage samples, the Klason lignin determination resulted in a better representation of the actual lignin content.

2.5.3 Thioglycolic acid lignin

The reaction of thioglycolic acid (**4.7**) with the benzyl alcohol groups in lignin (**4.8**) under acidic conditions and at elevated temperatures results in the formation of thioethers that render the derivatized lignin (**4.9**) soluble in an alkaline solution (Figure 4.2).

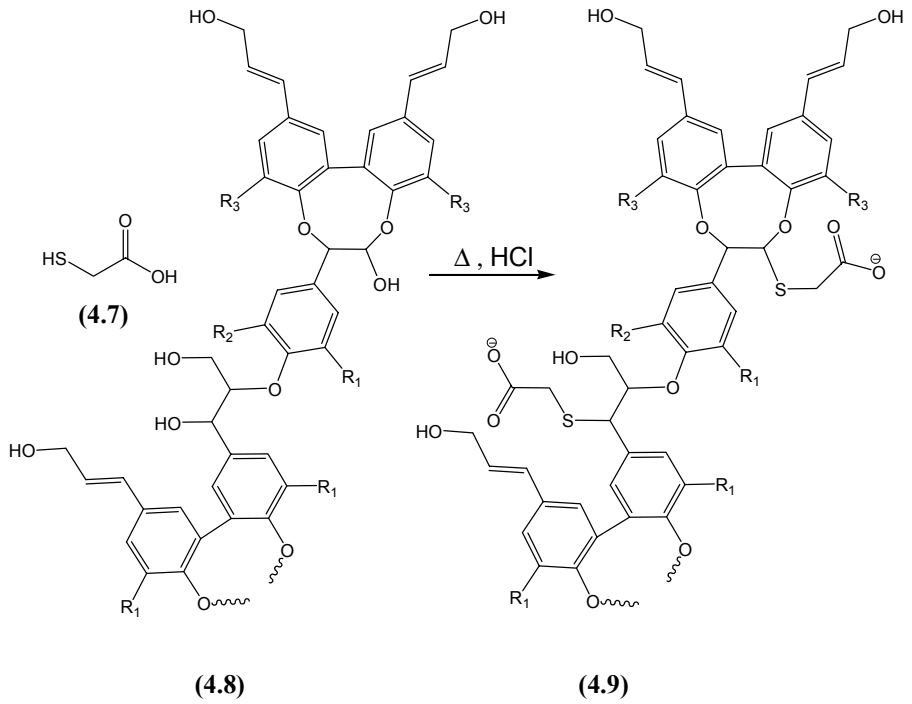


Figure 4-2. Reaction of thioglycolic acid with lignin. Adapted from Hatfield and Fukushima (2005).

The method was first described by Browning (1967) for 40-gram samples and then adapted for smaller samples (15 mg) by Bruce and West (1989). In this modified protocol, which is the most practical, the sample is incubated for 4 h. at 95°C in 1 mL 2 M HCl and 0.2 mL thioglycolic acid. The insoluble residue containing the lignin is recovered by centrifugation, washed three times in water, and the pellet is dissolved in 0.5 M NaOH. Non-lignin materials will not dissolve and are removed by centrifugation. The supernatant containing the thioglycolate lignin is then recovered by precipitation after addition of HCl and a 4-h. incubation at 4°C. The thioglycolate lignin is again recovered by centrifugation, dried, and dissolved in 0.5 M NaOH. An aliquot is diluted 40 times in 0.5 M NaOH and the absorbance at 280 nm is read. An *in-vitro* generated dehydrogenation lignin polymer (DHP) is used as a standard. The DHP is made by polymerizing coniferyl alcohol in the presence of hydrogen peroxide and horseradish peroxidase.

It is unclear how accurate this method is for the determination of lignin content in grasses, since a fraction of the lignin from grasses appears to be readily soluble in acidic solutions, such as used in this procedure. This would result in underestimation of the actual lignin content (Hatfield and Fukushima, 2005).

2.5.4 Acetylbromide lignin

The acetyl bromide procedure was developed by Johnson et al. (1961) to determine lignin content in small samples from woody species. This method uses acetyl bromide (**4.10**) to acetylate unbound hydroxyl groups in the lignin (**4.11**), while the hydroxyl group on the α-carbon is substituted with a bromine group. The substituted lignin derivative (**4.12**) is soluble under acidic conditions, and its concentration can be measured with a spectrophotometer at 280 nm.

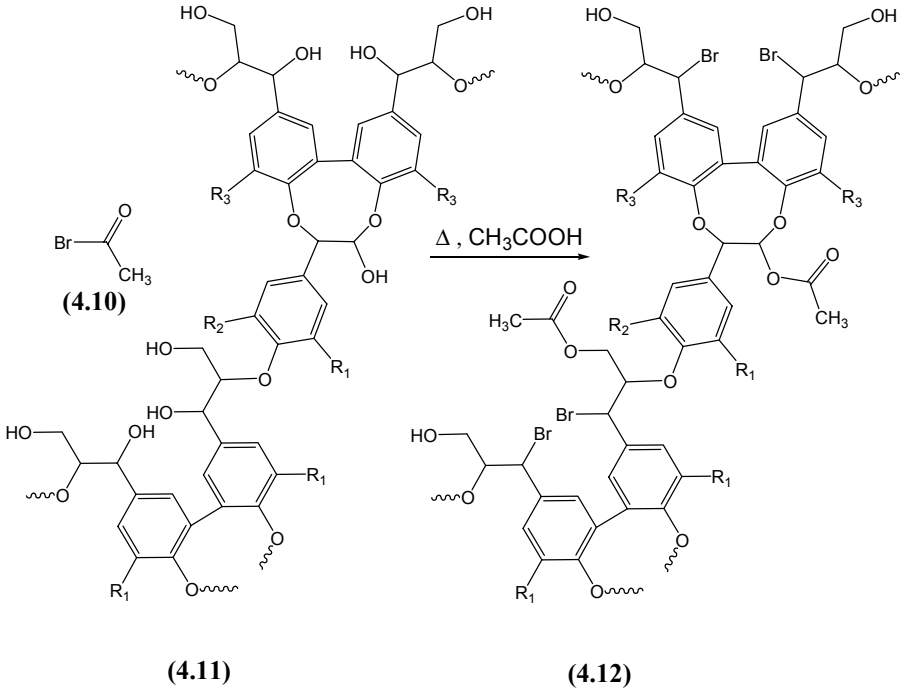


Figure 4-3. Reaction of acetyl bromide with lignin. Adapted from Hatfield and Fukushima (2005).

Acetyl bromide lignin is obtained from isolated cell walls. These can be obtained by extraction of ground woody samples in ethanol and chloroform.

The isolated cell walls or the dioxane lignin are dissolved by incubation in 2.5 mL of 25% (w/v) acetyl bromide in glacial acetic acid at 70°C for 30 min. The lignin is extracted by adding the dissolved cell walls to a flask containing 10 mL 2 M sodium hydroxide and 12 mL glacial acetic acid. A 1.75-mL volume of 0.5 M hydroxylamine is then added, and the total volume is brought up to 50 mL with acetic acid. The lignin content is determined spectrophotometrically by measuring the absorbance at 280 nm.

Modifications have been made to the original protocol in order to apply this method to herbaceous species. Iiyama and Wallis (1990) added 100 µL perchloric acid ($HClO_4$) to improve the dissolution of wall material. This decreased the overall time required for the procedure and made the use of hydroxylamine unnecessary.

Hatfield et al. (1999) reexamined the different protocols used for the determination of lignin content in herbaceous species, using alfalfa stems and corn rind as representative samples. They performed the initial incubation at 50°C for up to 4 h., at 60°C for up to 2 h., or at 70°C for up to 1 h., and with or without 100 µL perchloric acid. They concluded that the addition of perchloric acid indeed raises the overall absorbance values, but that this increase is the result of oligomeric xylan-degradation products that also absorb light in the 250-300-nm range. In order to avoid excessive xylan degradation, they recommended performing the extraction in acetyl bromide at 50°C for a period of 2-4 h, in the absence of perchloric acid.

As for any spectrophotometric method, a standard curve needs to be made. Since lignin is very heterogeneous in nature (variation in subunit composition and interunit linkages within and between species), finding a good standard is not trivial. Fukushima and Hatfield (2001) recommended using lignin extracted with dioxane-HCl, rather than isolated cell walls, because it contains little protein and polysaccharide residues. Isolating dioxane lignin requires adding 100 mL of acidified dioxane (90 mL dioxane and 10 mL 2 N HCl) to 5 g dry cell wall material. The flask is then connected to a reflux condenser, nitrogen gas is blown onto the liquid surface, and the solution is refluxed for 30 min. under nitrogen gas. The solution is then cooled and filtered through a Whatman GF/C glass fiber filter, and collected in a flask. The residue on the filter is washed in 20 mL 96% dioxane, and then added to the filtrate. In order to neutralize the solution, 4 g sodium bicarbonate ($NaHCO_3$) is added to the flask, followed by a 3-min. incubation on a rotary shaker until the pH was neutral. The solution is filtered through a 0.45-µm nylon membrane, and subsequently concentrated to 10-15 mL under reduced pressure on a rotary evaporator

using a waterbath set at 40°C. The solution is then added drop-wise to a 250-mL centrifuge bottle containing 200 mL of rapidly stirring distilled water. Any insoluble residue remaining in the flask is washed with 2.0 mL 96% dioxane solution and added drop-wise to the water. The drop-wise addition to the water should result in the precipitation of the lignin, which can be stimulated by the addition of 2 g anhydrous Na_2SO_4. The lignin precipitate can be collected by centrifugation (9000g for 20 min.). The pellet is then dissolved in 100% dioxane, filtered through a 0.45-µm nylon membrane, and added drop-wise to 200 mL stirring anhydrous ether in a centrifuge bottle. The lignin precipitates and is collected by centrifuging at 9000g for 15 min. at 0°C. In order to remove all hydrophobic non-lignin contaminants, the lignin pellet is again dissolved in dioxane, filtered, and added to ether. The ether is then removed and 60 mL petroleum ether is added to wash the lignin. After removal of the petroleum ether, the lignin residue is freeze dried.

Lignin content determined with the acetyl bromide method using dioxin lignin as a standard, showed the best correlation with *in vitro* dry matter digestibility for a diverse set of forage samples (Fukushima and Hatfield, 2004). Even though negative correlations between lignin content and digestibility were identified for the other methods, including Klason lignin and acid detergent lignin, the correlation coefficients were not as high. Fukushima and Hatfield (2004) pointed out, however, that it is important to calculate the lignin content of a given sample based on the standard derived from that same type of sample.

3. IDENTIFICATION AND CHARACTERIZATION OF PHENOLIC COMPOUNDS

In the previous section methods to isolate (certain classes of) phenolic compounds were described. In general, however, these methods do not provide information on specific chemical composition. In order to characterize mixtures of phenolic compounds, a variety of separation and identification methods exist. They will be described below.

3.1 Thin layer chromatography

A common, simple, inexpensive and relatively fast method for the separation of phenolic compounds from a mixture is thin layer chromatography (TLC). A small amount of the extract (40-100 µl) is applied approximately 2 cm from the bottom of a thin layer chromatography

plate, which is a matrix (typically cellulose or silica gel) attached to an inert carrier material, such as glass or plastic. The solvent is allowed to dry, either at room temperature, under a gentle stream of nitrogen, or with the use of warm air (a hair dryer can be convenient for this purpose). Application of multiple small amounts with drying in between applications will result in tighter spots and better resolution later on. A sharp pencil can be used to scratch off the matrix in order to clearly delineate individual lanes running along the length of the TLC plate. This will avoid samples from mixing later on.

The TLC plate is then placed in a glass container with a solvent filled to approximately 1 cm from the bottom. The solvent will move to the top of the TLC plate as a result of capillary action. Since each compound in the mixture will have a unique way of interacting with the matrix and the solvent, some compounds will move faster towards the top of the TLC plate than others. The R_f-value is the ratio of the distance of the compound has migrated divided by the distance the solvent has migrated, and has by definition a maximum value of 1. The R_f-value tends to be constant for a given combination of compound, solvent, and matrix so that comparisons can be made between separations performed at different times. If a given compound is colored, it is easy to determine the R_f-value. For non-colored compounds staining methods are available (see section 1.3).

The identification of phenolic compounds separated by TLC is somewhat challenging. The most common strategy is to include a set of reference compounds on the TLC plate. These compounds are applied individually, and if the mixture contains any of the reference compounds, they can be identified based on the R_f-value. Note that this approach always leaves some room for uncertainty, because two different compounds can have the same R_f-value. Further characterization is necessary to establish compound identity with more confidence. This can be achieved by scraping off the area on the TLC plate where the compound of interest has migrated to, followed by solvent extraction of the matrix, and more detailed chemical analyses, such as, for example gas chromatography-mass spectrometry or mass spectrometry (see Section 1.5 and Chapter 5).

Below is a step-by-step protocol for TLC. In this example the goal is to separate anthocyanins isolated from flower petals.

1. Pick the petals and place them in 1 ml methanol acidified with 0.1 or 1% (v/v) HCl.

2. Let the petals remain in the methanol overnight in the refrigerator. Alternatively, the tissue can be gently crushed with a glass rod or pipette.

3. Remove the supernatant with a Pasteur pipette and put it into a clean glass vial or a microfuge tube. Keep the material in the dark and in the cold as the compounds may break down easily.

4. Cellulose TLC plates are recommended for the separation of anthocyanidins, but silica gel will also work. If you use silica gel plates, it is helpful to wash the plates first with a mixture of chloroform and methanol (1:1 v/v). Simply run the solvent up the plates and let them air dry. This helps to remove some of the debris that is often associated with the plates. The plates will work even if you do not take the time to do this, but using pre-washed plates often provides a cleaner separation of compounds.

5. Spot a thin, fine band of pigment on the plate. To see as many spots as possible you need to put quite a lot of material onto the plate. It is best to apply the pigments in a band, while removing the solvent with a gentle flow of nitrogen.

6. There are many possible carriers, most of them mixtures of organic solvents. Table 4.1 lists some options that can be used for the separation of anthocyanins from petals.

Table 4-1. Solvent mixtures for thin layer chromatograpy of phenolic compounds

Solvent mixture	**Ratio**	**Layer**
n-butanol-acetic acid-water	4:1:5	upper layer
acetic acid-HCl-water (Forestal solvent)	30:3:10	miscible
ethyl acetate-formic acid-water	85:6:10	upper layer
ethyl acetate-water-formic acid-HCl	85:8:6:1	upper layer
n-butanol-2N hydrochloric acid	1:1	upper layer
n-butanol-acetic acid water	4:1:1	upper layer

The oxidation reaction of phenolic compounds can be used for the purpose of detection. While the phenolic compound is oxidized, a reagent is reduced, and the reduction can be monitored by a change in color. Two common reagents are ammoniacal silver nitrate and the Folin-Denis reagent.

Reactions of phenolic compounds with ammoniacal silver nitrate result in the formation of metallic silver. A simple procedure is to mix equal volumes

of 0.1 N NH_4OH and 0.1 N $AgNO_3$ and to apply this as a spray to a thin layer chromatogram at room temperature. The oxidized phenols appear as brown spots because of the silver. The reaction is, however, not specific to phenolic compounds. The Folin-Denis reagent, discussed in Section 1.1.1, and produces a blue color upon reaction with phenolic compounds.

An alternative way to visualize certain phenolic compounds is with the use of a hand-held UV lamp in the dark. The presence of the phenolic compounds can be observed by the fluorescence.

3.2 Liquid chromatography: HPLC and LC-MS

The principle of liquid chromatography was described in Chapter 3, Section 2. Liquid chromatography can be used for the separation of proteins, as well as for the separation of individual phenolic compounds from complex mixtures, based on the compounds' variation in affinity for a resin packed in a column. Changing the pH and/or ionic strength of the solution will allow all compounds of interest to elute, ideally in a sequential manner. The identification of the compounds is based on a combination of the retention time and either a UV/vis-spectrum or a mass spectrum. The absorbance spectrum in the UV and visible range of the light spectrum for each of the eluted compounds can be obtained by so-called diode array detectors, which will record the absorbance of a compound across a wide range of the spectrum. The mass spectrum is obtained with a mass spectrometer that is connected to the liquid chromatography equipment. Mass spectrometry is described in detail in Chapter 5. Large searchable databases exist to aid with the identification of compounds based on mass spectral data. The ultimate way to confirm a compound's identity is to show co-elution of the compound identified in the sample and the candidate reference compound.

Scalbert et al. (1989) described the chromatographic separation of cyanidin and delphinidin from methanolic extracts of oak heartwood using a C-18 Novapak column. The elution solvent was a mix of two solvents A and B that changed in composition from 0-100% B in a linear fashion over a period of 20 min. Solvent A was a mixture of 94:5:1 H_2O/methanol/H_3PO_4, and solvent B was 99:1 methanol/H_3PO_4, at a flow rate of 1.7mL/min. In this case a dual-channel spectrophotometer was used with two wavelengths selected for detection: 280 and 530 nm. Under these conditions, and based on reference compounds, delphinidin eluted after 13.9 min and cyanidin after 14.8 min.

Tamagnone et al. (1998) described the impact of over-expression of Myb transcription factors in tobacco on the phenolic composition of leaf extracts, and used HPLC coupled to a diode array detector to evaluate change in chemical composition. Leaf material frozen in liquid nitrogen and ground to a fine powder was extracted in cold methanol (chilled on dry ice). Cell debris (mostly cell walls) was removed by centrifugation (5 min. at 50g). The supernatant was further clarified by additional centrifugation (2,000g for 30 min.) and then analyzed by HPLC using a C-18 column with a flow rate of 1 mL/min. The solvents consisted of 10% (v/v) methanol in water with 1 mM trifluoric acid (TFA) (solvent A), and 80% (v/v) methanol in water with 1 mM TFA (solvent B). The solvent gradient started out with 90% solvent A and 10% solvent B, and ended 60 min. later with 100% solvent B. The diode array detector provided a full absorbance spectrum on each of the eluted compounds.

Morreel et al. (2004) extracted wood from transgenic poplars in which the monolignol biosynthetic gene caffeic acid *O*-methyltranferase was down-regulated. They used methanol to extract frozen (liquid N_2) and ground xylem tissue from debarked stems. After removing cell debris by centrifugation, the supernatant was freeze-dried and extracted with 1:1 cyclohexane:water acidified with 0.1% (v/v) TFA. The samples were then separated on a C-18 column with 1% aqueous triethylammonium acetate (TEAA) as solvent A, and a 25:75 mixture of methanol and acetonitrile containing 1% TEAA as solvent B, at a flow rate of 0.3 ml/min, going from 100% solvent A to 100% solvent B in 40 min. The eluted compounds were detected with a diode array spectrophotometer which obtained the absorbance spectrum of each compound in the range between 200 and 450 nm. In addition, a mass spectrum was obtained for each of the compounds after the eluate was vaporized.

These three examples illustrate technology developments over time (dual-channel detector, diode array detector, mass spectrometer). Note that while the overall methodology is very similar (methanolic extracts, methanol-based, acidified solvents used for HPLC, detection of eluted compounds), the exact conditions for successful separation need to be defined for each system.

3.3 Gas chromatography

The principle of gas chromatography (GC) is similar to that of liquid chromatography or TLC, in that compounds in mixtures are separated from each other based on their affinity for a resin. GC is performed on volatile

samples, using a long (20-30 meter) and very thin (<0.5 mm internal diameter) column coated on the inside with a silica-based solid phase applied as a thin (~0.2 μm) film. The column is placed in an oven in which the temperature can be changed quickly, and with great accuracy. The samples are generally dissolved in a volatile organic solvent that has little affinity for the solid phase. A small volume (1-5 μL) is loaded onto the column through a septum using a syringe with a thin needle. The sample is then moved through the column with an inert carrier gas such as helium. The interaction of the compound with the solid phase is temperature-dependent, and elution of the compounds of interest is achieved by increasing the temperature. Detection of the compounds at the end of the column is nowadays mostly based on mass spectrometry (see Chapter 5), although flame ionization detectors (FID) are also used for routine analyses of samples in which there are only a limited number of compounds that can be easily identified based on retention time. In order to increase the volatility of phenolic compounds, chemical modifications such as methylation, acetylation or silylation can be employed.

An example of the use of GC for the identification of phenolic compounds can be found in Dias and Grotewold (2003). In order to identify the role of the ZmMyb-IF35 transcription factor, the phenolic composition of transgenic maize cell lines in which the gene encoding this transcription factor was over-expressed was examined. They extracted phenolic compounds by homogenizing the tissue in methanol, followed by 15 min. centrifugation at 13,000g to remove insoluble debris. The methanol was then evaporated and the extract was hydrolyzed by boiling in 2 M HCl for 20 min. The acid was evaporated and the pellet was redissolved in methanol, adjusted for weight, so that the final concentration was 0.1 mg/μl. One μl of this extract was injected in a GC, with a mass spectrometer as a detector. The initial temperature of the GC column was 40°C for 1 min., and the temperature was increased to 310°C at a rate of 11°C/min. This resulted in a gas chromatogram with more than 40 peaks representing the compounds in the methanol extract. Some of these peaks were uniquely associated with the presence of the transgene, thus providing clues regarding the function of this transcription factor in phenolic metabolism.

3.4 Methods for the identification of lignin subunit composition

Chemical analyses to determine lignin content and composition are difficult because of the complex nature of lignin. In addition, they often result in loss of information regarding the original structure. A large number

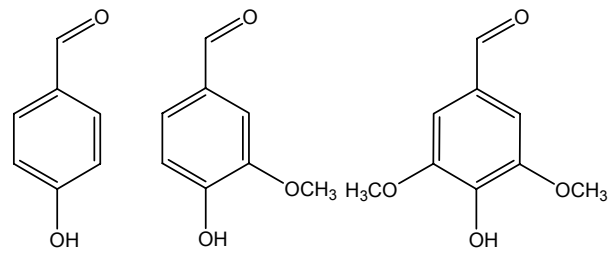
of analytical techniques have been developed to investigate the content and composition of lignin. Reviews and protocols of various techniques can be found in Monties (1989), Lewis and Yamamoto (1990), Chen (1991), Lin and Dence (1992), and Sjöström and Alen (1999). This section presents a brief description of the different methods available.

3.4.1 The nitrobenzene oxidation

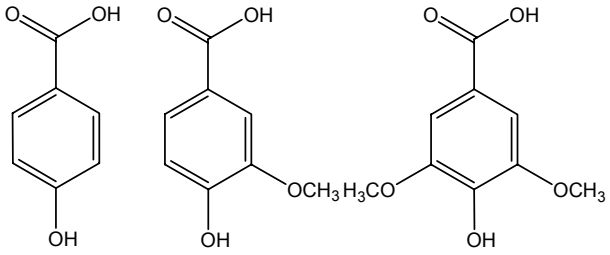
Nitrobenzene oxidation involves the treatment of cell wall material with sodium hydroxide containing nitrobenzene. The method was developed by Freudenberg (1939). Nitrobenzene oxidation degrades the phenylpropane structure from a C_6-C_3 unit to a C_6-C_1 unit. After incubation the soluble fraction is acidified and extracted with ether. The degradation products – *p*-hydroxybenzaldehyde (**4.13**), vanillin (**4.14**), and syringaldehyde (**4.15**), and the corresponding acids *p*-hydroxybenzoic acid (**4.16**), vanillic acid (**4.17**), and syringic acid (**4.18**) – can be separated on an HPLC column to determine the lignin composition. Alternatively, gas chromatography (GC) or gas chromatography-mass spectrometry (GC-MS) can be used.

The lignin composition can be expressed as the S/V ratio, whereby the 'V' (vanillin (**4.14**), vanillic acid (**4.17**)) reflects guaiacyl residues and the 'S' (syringaldehyde (**4.15**), syringic acid (**4.18**)) reflects syringyl residues. (Monties, 1989). Nitrobenzene oxidation does not affect the aromatic C-C bonds, but only the bonds from the so-called uncondensed units, so that the estimates of lignin composition can be somewhat biased (Lewis and Yamamoto, 1990). In grasses these methods cannot distinguish between aromatic compounds derived from lignin and from *p*-coumaric and ferulic acid esterified to the cell wall (Lapierre, 1993), so that this method is seldom used for the analysis of lignin from grasses.

An example of the use of the nitrobenzene oxidation to elucidate differences in lignin subunit composition between wild-type and mutant Arabidopsis plants can be found in Chapple et al. (1992). They dried and ground stem tissue of Arabidopsis and extracted 50 mg with methanol (three times 1 mL) and deionized water (twice, 1 mL) at 60°C. Esterified phenolics



(4.13) **(4.14)** **(4.15)**



(4.16) **(4.17)** **(4.18)**

were saponified in 1 M NaOH (37°C, 60 min.), after which the tissue was rinsed twice and resuspended in water. Sodium hydroxide solution was added to an aliquot representing 20 mg of dry tissue so that the final concentration was 2 M NaOH in a volume of 980 µL. Nitrobenzene (20 µL) was added to this suspension, followed by a 2-hour incubation at 160°C. After cooling the sample, a 250-µL aliquot was added to 750 µL 1 M acetic acid to neutralize the sample. A 100-µL aliquot was then reduced (addition of 1 mL 2% (w/v) sodium borohydride in DMSO; 90 min. at 40°C), acetylated to increase the volatility (4 mL of acetic anhydride plus 250 µL of 1-methylimidazole as catalyst; 10 min. at room temperature), and extracted in methylene chloride (CH_2Cl_2). The break-down products of the lignin were subsequently analyzed on a GC with an FID, using *p*-hydroxybenzaldehyde (**4.13**), vanillin (**4.14**), 5-hydroxyvanillin, and syringaldehyde (**4.15**) as reference compounds.

3.4.2 Thioacidolysis

Thioacidolyis is an acidolytic degradation method used to analyze lignin subunit composition. The method relies on the use of boron trifluoride etherate ($(C_2H_5)_2$-O-BF_3) and ethanethiol (C_2H_5-SH) to cleave β-*O*-4 bonds (**1.82**) in the lignin (**4.19**). The reaction mechanism is shown in Figure 4.4.

First the α-carbon is substituted by BF_3, resulting in compound **4.20**. The activated α-carbon makes this compound reactive with ethanethiol, resulting in the formation of **4.21**. The thioethyl group now attacks the activated β-carbon of **4.21**, resulting in the intermediate **4.22**. Attack of ethanethiol results in the formation of **4.23**, and through a similar mechanism the γ-carbon is substituted, resulting in the formation of 1,2,3-trithioethane phenylpropanoid monomer **4.25** (Lapierre et al., 1986; Chen, 1991; Rolando et al., 1992; Lapierre, 1993).

The thioacidolysis reagent is prepared fresh for each analysis by mixing 2.5 mL BF_3-etherate and 10 mL of ethanethiol in a flask. The volume is adjusted to 100 mL with dioxane. The cell wall extract (20 mg) is added to 10 mL of the thioacidolyis reagent, along with an internal control in order to be able to quantify the products later on. Docosane dissolved in methylene chloride is a suitable internal standard. The reaction is performed in glass tubes that can be closed with Teflon-lined screw caps. This is important, because the reagent is very odiferous. The actual thioacidolysis is performed for 4 h. at 100°C in an oil bath. Deionized water is added to the cooled reaction to get a volume of 30 mL. Sodium bicarbonate (0.4 M $NaHCO_3$) is added to increase the pH to a value between 3 and 4. The reaction mixture is then extracted in 3 x 30 mL methylene chloride. Water is removed from the combined organic extracts by addition of Na_2SO_4 and the methylene chloride is removed *via* evaporation at 40°C. The resulting residue is redisolved in 0.5 mL methylene chloride, silylated, and injected in a GC-MS for quantitation. The internal standard is used to correct for loss of compounds during the extraction.

Thioacidolysis allows the distinction between products derived from lignin and products derived from *p*-coumaric and ferulic acids, and the distinction between products derived from cinnamaldehydes and cinnamyl alcohols. Recent improvements have made it possible to estimate the fraction of free phenolic groups in uncondensed lignin (see Section 1.3.1), and to depolymerize the dimers, so that they can be included in the analysis of the lignin composition.

Thioacidolysis allows the distinction between products derived from lignin and products derived from *p*-coumaric and ferulic acids, and the distinction between products derived from cinnamaldehydes and cinnamyl alcohols. Recent improvements have made it possible to estimate the fraction of free phenolic groups in uncondensed lignin (see Section 1.3.1), and to depolymerize the dimers, so that they can be included in the analysis of the lignin composition.

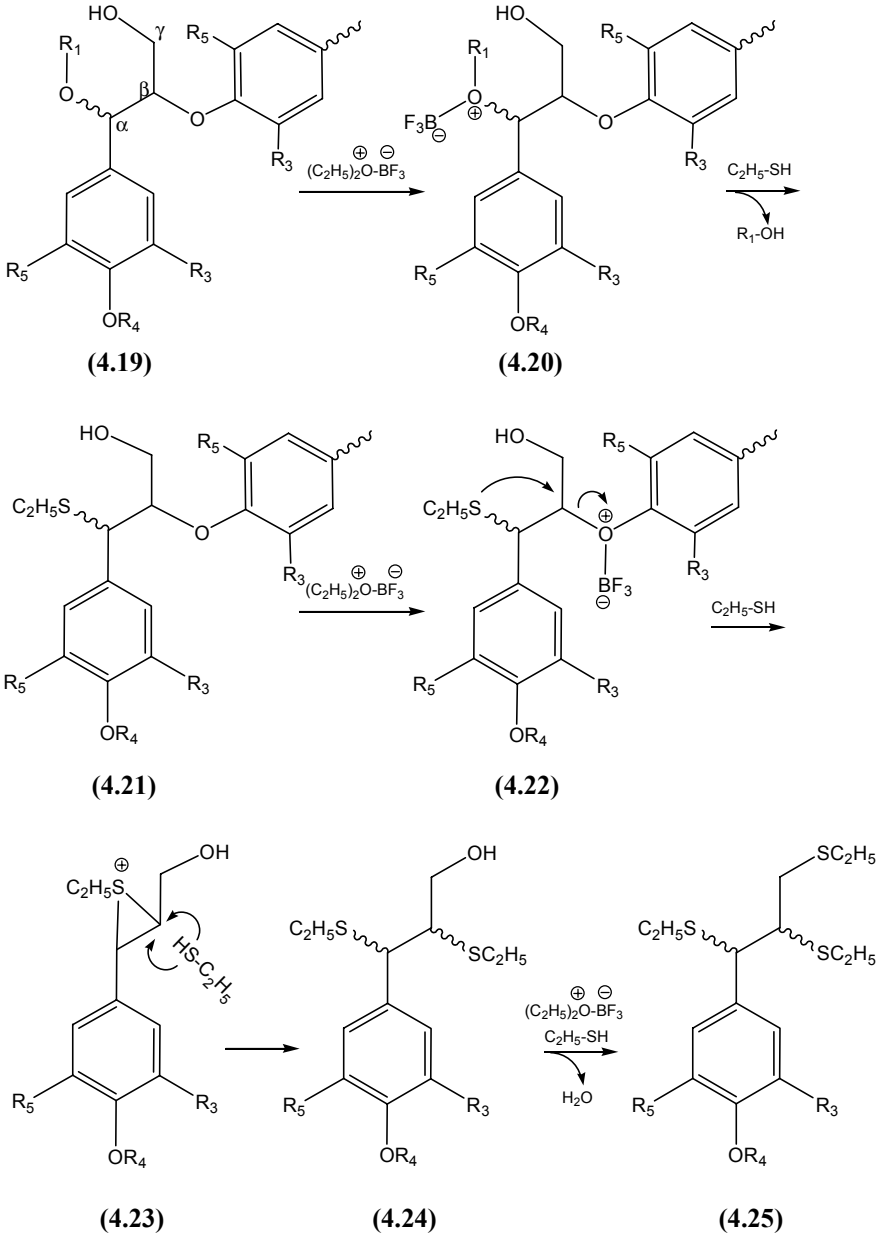


Figure 4-4. Reaction mechanism for the formation of 1,2,3- trithioethane phenylpropanoid monomers (**4.25**) from lignin (**4.19**). R_1 is either an aryl group or a hydrogen atom. In H-residues both R3 and R5 are hydrogen atoms, in G-residues R_3 is a methoxyl group and R_5 is a hydrogen atom, and in S-residues both R_3 and R_5 are methoxyl groups. R_4 is either a hydrogen atom or an alkyl group. The wavy bonds indicate that both the *S*- and *R*-stereoisomers are present.

Examples of the use of thioacidolysis can be found in Hoffmann et al. (2004), who describe the effects on lignin subunit composition resulting from silencing the gene encoding hydroxycinnamoyl-CoA shikimate/quinate hydroxyltransferase (HCT) in tobacco, and in O'Connell et al. (2002) who investigated the effects of down-regulating *cinnamoyl-CoA reductase* in tobacco on lignin subunit composition.

3.4.3 Derivatiation Followed by Reductive Cleavage

A recently developed method based on acetylbromide cleavage is 'Derivatization Followed by Reductive Cleavage' (DFRC; Lu and Ralph, 1997). This method uses acetylbromide to acetylate alcohols and phenols in lignin (**4.26**). In addition, the α-carbon is brominated (**4.27**). Cleavage of the β-*O*-4 linkage is catalyzed by zinc. The monomeric residues (**4.28**) that are formed are acetylated to yield diacetylated monolignols (**4.29**) and separated using gas chromatography. The DFRC reaction is presented in Figure 4-5.

DFRC is achieved by heating the sample in 2.5 mL acetyl bromide with HCl, followed by removal of the reagent by evaporation. The residue that remains is dissolved in 2.5 mL of dioxane/acetic acid/H_2O (5:4:1). To this solution is added 50 mg zinc dust, to catalyze the cleavage of β-bromo-ethers. After stirring for 30 min. the zinc is removed by filtration through glass wool. Dichloromethane (10 mL) is then added to the filtrate along with 10 mL saturated NH_4Cl solution, and the internal standard tetracosane is added. After mixing, phase separation, and removal of the organic phase, the aqueous phase is extracted twice more in dichloromethane. The combined dichloromethane extracts are then combined and dried. The resulting residue is dissolved in 1.5 mL dichloromethane, acetylated with 0.2 mL acetic anhydride and 0.2 mL pyridine. Ethyl acetate (2.5 mL) is added after 90 min. and the mixture is evaporated. Ethanol is added during the evaporation to get rid of the pyridine. Pyridine-free samples are dissolved in 200 μL dichloromethane and a 1-2 μL is injected in a gas chromatograph for analysis using the reference compounds *p*-acetoxycinnamyl acetate (**4.29a**), coniferyl diacetate (**4.29b**), and sinapyl diacetate (**4.29c**) for peak identification.

Advantages of the DFRC method over thioacidolysis include the better yield, simpler and milder reaction conditions, and the fact that esters are not cleaved. Especially in the analysis of lignin in grasses, with large amounts of esterified *p*-coumaric acid (**1.13**), this is an important consideration.

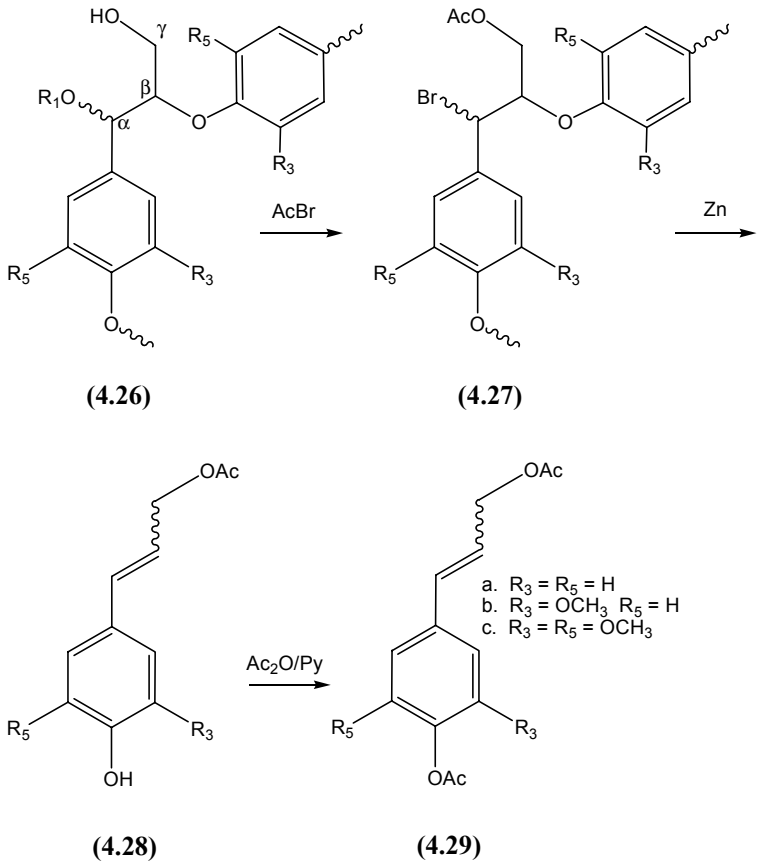


Figure 4-5. Reaction mechanism for the derivatization followed by reductive cleavage based on Lu and Ralph (1997). Reaction of the lignin with acetylbromide (AcBr) results in the acetylation of the γ-carbon, while the α-carbon is brominated. Zinc (Zn) catalyzes the cleavage of the ether bond between the β-carbon of one residue and the *O*-4 position of the adjacent residue. The resulting monomer is acetylated with acetic anhydride (Ac_2O) and pyridine (Py). R_1 can be a proton or an aryl group. In H-residues R_3 and R_5 are protons, in G-residues R_3 is a methoxyl group and R_5 is a proton, whereas in S-residues both R_3 and R_5 are methoxyl groups. The wavy bonds indicate that both the *S*- and *R*- (**4.26**, **4.27**) or *E*- and *Z*-stereo-isomers (**4.28**, **4.29**) are present.

An example of the use of the DFRC method can be found in Zhang et al. (2006), who analyzed the lignin subunit composition in the *gold hull and internode2* mutant of rice. This mutant has a defective *OsCAD2* gene and accumulates coniferaldehyde residues in its lignin.

3.4.4 Analytical pyrolysis

Pyrolysis (Py) is the rapid thermal degradation of a compound in the absence of oxygen. Pyrolysis of plant cell wall material results in the release of a pyrolysate of low molecular weight compounds that are break-down products of both polysaccharides and lignin. Lignin breaks down to substituted monomeric phenols; the propanoid chain is often reduced to one or two carbons or is lost altogether (Meier and Faix, 1992). The pyrolysate is generated in an oven or through the use of a heated filament to which the sample is applied. The resulting pyrolysate is led directly into a mass spectrometer (Py-MS) or a gas chromatography coupled to a mass spectrometer (Py-GC-MS) (Boon, 1989; Meier and Faix, 1992; Lapierre, 1993).

Analytical pyrolysis requires only small samples (10-1,000 mg) and is fast compared to most other analytical methods (Mulder *et al*., 1991; Lapierre, 1993). This makes it an attractive method for lignin analysis. Another major advantage is that it is not necessary to isolate the lignin prior to pyrolysis, because there is little similarity between carbohydrate and lignin pyrolytic fragments and their mass spectra are distinctly different (Mulder *et al*., 1991; Meier and Faix, 1992; Boon, 1992). Pyrolytic disintegration of lignin is mostly due to the cleavage of ether-bonds. This may bias the quantitation of S and G units since S-units are more often involved ether-bonds than G units (Lapierre, 1993). In general, the reproducibility of results obtained with pyrolysis is good when the same experimental unit is used. Results from different experimental settings tend to be less similar.

Examples of the use of Py-MS or Py-GC-MS for the characterization of mutants in which lignin biosynthesis was disrupted can be found in Vermerris and Boon (2001) and Bout and Vermerris (2003), whereas Fontaine et al. (2003) used Py-GC-MS to investigate the chemical basis of forage quality.

3.4.5 Nuclear magnetic resonance

Nuclear magnetic resonance (NMR) is considered the most informative method to analyze lignin structure, because it does not rely on chemical degradation of the lignin prior to the analysis. As a consequence, NMR analysis of lignin provides information on the different types of interunit linkages, including the C-C bonds that are typically not broken by the chemical degradation methods or pyrolysis.

NMR is a spectroscopic technique that is based on the ability to influence the *spin* of certain nuclei. Spin is an electromagnetic property of electrons, protons and neutrons that can be described in quantum chemical terms. Its value is +½ or -½. A nucleus with a non-zero net spin can be considered as a small magnet. If a population of such nuclei are placed in a magnetic field, the nuclei will align with the magnetic field, and an equilibrium will be established whereby the majority of the nuclei will be in the energetically lower state (alignment of the North-pole of the nuclei with the South-pole of the magnet), hence creating a net magnetic field in the direction of the external magnetic field.

Transition from the low- to the high-energy state can occur by the absorption of a photon with energy in the radio frequency range. The exact amount of energy required depends on the nucleus itself, but also on its immediate environment, since the nucleus experiences a shielding effect from neighboring nuclei that also act like small magnets. The principle of NMR is based on the detection of nuclei with a net non-zero spin, by placing a sample in a strong magnetic field and subjecting it to energy in the radio frequency range that will reverse the net magnetization of the nuclei. The relaxation time (T_1) is the time required to let the nuclei return to their equilibrium state after the radio frequency energy has been absorbed, and provides information on the chemical environment of the nucleus of interest. The NMR spectrum that is obtained after several mathematical manipulations of the output signal reflects the different energy levels associated with nuclei in their specific environments. The variable on the x-axis of the spectrum is the *chemical shift*, indicated by the symbol δ, and measured in part per million (ppm). The chemical shift of a nucleus is defined as the difference between the resonance frequency of the nucleus and a standard, relative to the standard. Common standards are tetramethylsilane (TMS; $Si(CH_3)_4$) and deuterated d_6-acetone ($(CD_3)_2CO$). The use of chemical shift allows standardization of spectra across different experimental platforms (manufacturers, models, temperatures, etc.).

NMR can detect nuclei with a net non-zero spin, as long as the abundance of that nucleus is high enough. For example, the ^{13}C nucleus has a net spin of ½, and is the only carbon nucleus with a non-zero spin. Its natural abundance is, however, only 1.1%, which limits the sensitivity. As a consequence, large samples need to be analyzed in order to get detailed compositional data, or the sample needs to be analyzed for prolonged periods, or, if the sample is derived from plants, the plants can be grown in an environment enriched with $^{13}CO_2$. For biological materials in general, ^{1}H NMR, ^{13}C NMR or ^{31}P NMR are most commonly used. For the analysis of

lignin ^{13}C NMR, or a combination of ^{1}H and ^{13}C NMR is used, the latter being referred to as 2D-NMR.

The low abundance of ^{13}C in biological samples has resulted in the development of new techniques that have enhanced the sensitivity of NMR. This includes high frequency NMR spectrometers in combination with recent advances such as Cross Polarization/Magic Angle Spinning (CP/MAS). The cross-polarization technique transfers magnetic impulses from abundant nuclei such as protons onto rare nuclei (such as ^{13}C), which enhances the signal from the rare nucleus. Magic angle spinning, whereby the sample is at a 54.7° angle with respect to the magnetic field, sharpens the bandwidths. Structural information can also be obtained through heteronuclear single quantum coherence (HSQC), HSQC-total correlation spectroscopy (HSQC-TOCSY), and heteronuclear multiple bond correlation (HMBC) experiments. A detailed description of these techniques goes beyond the scope of this book, but several excellent references exist. For a more detailed general description, see Friebolin (2004) or Keeler (2005), whereas Ralph et al. (1999) provide a description of NMR specifically for the characterization of lignin.

^{13}C-NMR is the most informative technique for the elucidation of chemical structures, but generally requires large sample sizes as a result of the low abundance of ^{13}C. As a consequence, micro-sampling to obtain tissue specific information is practically impossible. In addition, the amount of time required for the acquisition of NMR spectra makes this technique unsuitable for the analysis of large samples.

NMR analysis of cell walls traditionally relied on the isolation of lignin from the cell wall, with the associated risk that the NMR spectrum was not completely representative of native lignin. A method to dissolve cell walls in their entirety was recently developed (Lu and Ralph, 2003), and this has made analysis of cell walls using NMR appreciably easier. Ground cell wall samples are extracted with organic solvents to remove non-cell wall materials, followed by ball milling. Milled wood lignin (MWL) is obtained by extracting the ball-milled residue in a mixture of dioxane and water (96:4), followed by freeze-drying, and washing with water. Approximatley 600 mg of the MWL is suspended in 10 mL dimethylsulfoxide (DMSO), and 5 mL of *N*-methylimidazole is added. A clear solution is formed in up to 3 h. Acetylation of the lignin is achieved by adding 3 mL acetic anhydride followed by 1.5 h. of stirring. The brown solution is added to 2 L of water and allowed to sit overnight, after which the precipitate is recovered by filtration through a 0.2 μm filter. After washing the residue with 250 mL of

water, it is dried, dissolved in deuterochloroform ($CDCl_3$) to a concentration of 150 mg/mL, and subjected to NMR spectroscopy.

Advanced NMR techniques have resulted in the elucidation of novel lignin structures, involving, for example, the benzodioxan structure involving β-*O*-4 and α-*O*-5 linkages between two monolignol residues (**1.89**; Ralph et al., 2001; Marita et al., 2003). In addition, NMR has been used to determine the impact of mutations and introduction of transgenic constructs on lignin subunit composition (Ralph et al., 1997; Marita et al., 1999; Marita et al., 2003).

3.4.6 Fourier-transform infrared spectroscopy and near infrared reflectance spectroscopy

Fourier-transform infrared spectroscopy (FT-IR) and near infrared reflectance (NIR) spectroscopy are methods that rely on the detection of vibration of molecular bonds. These vibrations can either involve the distance between the atoms (stretch) or the angle between the atoms (bend). Changes between two vibration energy states that also impact the dipole moment can be induced by absorption of infrared light, and this absorption can be monitored and represented in a spectrum. Various functional groups have characteristic absorptions and their presence can be used for diagnostic purposes. FT-IR spectra of ligno-cellulosic materials are typically complex, because of overlap in signal between various functional groups.

FTIR spectra can be acquired in transmission mode with an FTIR microscope after drying a cell wall suspension on a BaF_2 microscope slide (Sené et al., 1994), or by pressing a thin potassium bromide pellet containing finely ground plant tissue and measuring the absorbance of infrared light going through the pellet with the use of an FTIR spectrometer (Vermerris et al., 2002). Alternatively, a suspension of isolated cell walls can be placed on a gold-plated reflective surface under a microscope equipped with an FTIR spectrometer (Yong et al., 2005).

NIR spectra are acquired using dried tissue, either whole or ground. The reflectance is measured relative to a so-called white reference, which is a highly reflective surface, such as Gore-Tex® or white ceramic. Near-infrared reflectance spectra, or absorbance spectra, defined as log(1/R), with R being the reflectance, look rather flat. For analysis purposes the second derivative of the spectra is often used, because it enhances certain spectral features.

NIR spectroscopy is popular for the analysis of large numbers of samples because of the speed at which the spectra can be obtained. The lack of distinct spectral features, however, necessitates the need for calibration using a set of standards that have been analyzed *via* wet-chemical methods. Brinkmann et al. (2002) described the use of NIR spectroscopy to predict the lignin content leaves and stems of beech (*Fagus sylvatica*). They used lignin content determined with either acid detergent lignin method (see Section 1.1.5.2) or the thioglycolic acid method (see Section 1.1.5.3), and in both cases were able to define a calibration model that allowed accurate prediction of the lignin content using NIR spectroscopy. Interestingly, different variables (absorbances at given wavelengths) were selected for the model, depending on the way lignin content had been established, indicating that the two lignin determinations rely on the detection of different structural features. Indeed, the acid detergent lignin was shown to contain considerable amounts of protein.

When the spectra of several (groups of) samples are compared, it is often helpful to analyze the difference spectrum, which is obtained by digitally subtracting one spectrum from the other. The analysis of signal that thus remains can help to establish what is different between the two samples (Faix, 1992; Sené et al., 1994). Recent developments include the application of multivariate statistics to the analysis of FTIR and NIR spectra (Johnson and Wichern, 1992; Kemsley, 1998; Chen et al., 1998; Mouille et al., 2003). This makes sense given the high-dimensionality of the spectral data. There are typically over 1000 measurements per sample, with a high degree of correlation between certain data points. The spectral data are typically first compressed by principal components analysis (PCA) or partial least squares (PLS). These manipulations result in the generation of a small (<10) set of new variables in which the original variables (absorbance values across the entire spectrum) are combined, taking into account either the variance-covariance structure of the data, or correlations between variables. These new variables can then be used in further analyses, such as discriminant analysis, in which a sample of unknown identity is classified into pre-defined groups based on its spectral features and a classification model derived from a large group of samples with known identity. This approach is now applied to the identification of maize and Arabidopsis mutants with altered cell wall composition (Yong et al., 2005). A useful reference book on the application of NIR spectroscopy for the analysis of biological samples is written by Siesler et al. (2002).

4. VISUALIZATION OF PHENOLIC COMPOUNDS *IN PLANTA* USING HISTOCHEMICAL STAINS

4.1 An overview of histochemical staining protocols

A number of histochemical stains are available to visualize phenolic compounds in the plant, either in thin sections, or applied to whole tissues. De Neergaard (1997) published a series of detailed protocols that provides an excellent source of information. Below follows a summary of the available reagents and the specific compounds they detect.

Aniline sulfate is dissolved in 0.1N sulphuric acid in aqueous solution or 60% (v/v) ethanol to a final concentration of 1-6 % (w/v). This stain reacts with lignin, which will turn yellow.

Chlorine sulfite can be used to detect syringyl lignin. The specimen is placed above a tissue soaked in a sodium hypochlorite ('bleach') solution for 30 minutes. A 3% (w/v) solution of sodium sulfite (Na_2SO_3) is applied to the specimen for 5-10 minutes. Lignin will stain orange to red.

4′,6 diamino 2-phenylindole (**DAPI**) is a fluorescent stain that reacts with DNA and phenolic compounds. It can be used on live tissue at a concentration of 0.002% (w/v) or in fixed tissue at a concentration of 0.01-0.05% (w/v). Several buffers, such as TRIS or potassium phosphate can be used, depending on the exact application; the pH is generally around 7.5. The staining solution needs to be stored at 4°C. The specimens are incubated in the dark for a minimum of 30 minutes, up to an overnight incubation (4°C) and then exposed to UV light with an excitation wavelength of 365 nm.

Ethidium bromide is a fluorescent dye that can be used for the visualization of DNA as well as lignin and other phenolic substances. An aqueous solution of 0.1 % (w/v) is prepared and applied to tissue sections for 5-10 minutes. The specimen is exposed to UV light, which results in an orange fluorescence as a result of the ethidium bromide.

Ferric chloride ($FeCl_3$) is prepared as a 2% (w/v) aqueous solution, although for embedded tissues a 10% solution is recommended. The specimen is incubated in the solution for 5-10 min. Phenolic compounds, including tannins will stain yellow or orange with this reagent.

The **Mäule reagent** is used to distinguish between guaiacyl and syringyl residues in the lignin. A freshly prepared aqueous solution of 1% (w/v) potassium permanganate ($KMnO_4$) is applied to the specimen for at least 30 min. After washing in distilled water for 2 min. and removing all liquid, a 15-20% aqueous solution of HCl is applied. The HCl is removed, and a 10% solution of NH_3 is applied. Syringyl residues result in the development of a red color. Absence of syringyl residues results in the development of a yellow color. Chapple et al. (1992) used this reagent to demonstrate the lack of syringyl residues in the lignin of the Arabidopsis *fah-1* mutant, in which the activity of the monolignol biosynthetic enzyme ferulate 5-hydroxylase (coniferaldehyde/coniferyl alcohol 5-hydroxylase) is reduced. Overexpression of a wild-type copy of the defective *F5H* gene in the *fah1* mutant resulted in a predominance of syringyl residues, as evident from the red color of the lignified tissues in the stem (Meyer et al., 1998).

Methylene Blue is used to detect pectin and phenolic compounds. It is applied to the specimen as a 0.01-0.15% (w/v) aqueous solution, incubated for 30 min. at room temperature, or for 5 min. at 60°C, and then removed by washing in distilled water. Methylene Blue is a blue dye, but it produces a *red* coloration as a result of the reaction with the phenolic or pectic substances. This phenomenon is referred to as metachromasia, and the cell or tissue components that exhibit it are called metachromatic.

Toluidine Blue O is dissolved in a buffer with a pH between 4 and 8, or in 70% ethanol or in a 0.05% (w/v) borax solution and kept at 4°C. The specimen is stained for 1-10 min. washed for 1-2 min. in distilled water, or in the buffer used to dissolve the Toluidine Blue O, or in ethanol until all excess stain is removed. When the specimen is viewed under the microscope, different cell components produce colors different colors (metachromasia): DNA is bluish-green, RNA is violet, the middle lamella is red, non-lignified cell walls are red-violet or blue-violet, and polymerized phenolics such as lignin are green or blusih-green. This stain is therefore of general use.

Phloroglucinol-HCl is a 2% (w/v) solution of phloroglucinol (**4.30**) dissolved in a 2:1 mixture of ethanol and concentrated HCl. This reagent reacts with cinnamaldehyde end groups (**4.31**) in the lignin, resulting in the cationic chromophore (**4.32**), and which appears as a burgundy-red compound (Adler et al., 1948; Geiger and Fuggerer, 1979; Pomar et al., 2002; see Figure 4-6). This procedure is known in the literature as the **Wiesner reaction** and is also used as a general stain for lignin. This may not be appropriate if samples differing in cinnamaldehyde content are being

compared. This is probably the most common stain used to detect lignin, because the reagent is easy to make and the reaction can be readily monitored. Pictures of the use of the Wiesner reaction can be found in Halpin et al. (1998), Vermerris et al. (2002), and Pomar et al. (2002).

Vanillin-HCl is prepared by making a 0.05% (w/v) solution in 50% ethanol. This is then mixed in a 2:5 ratio with concentrated HCl. The mix is applied to the specimen for 1 min. Alternatively, the sample is immersed in

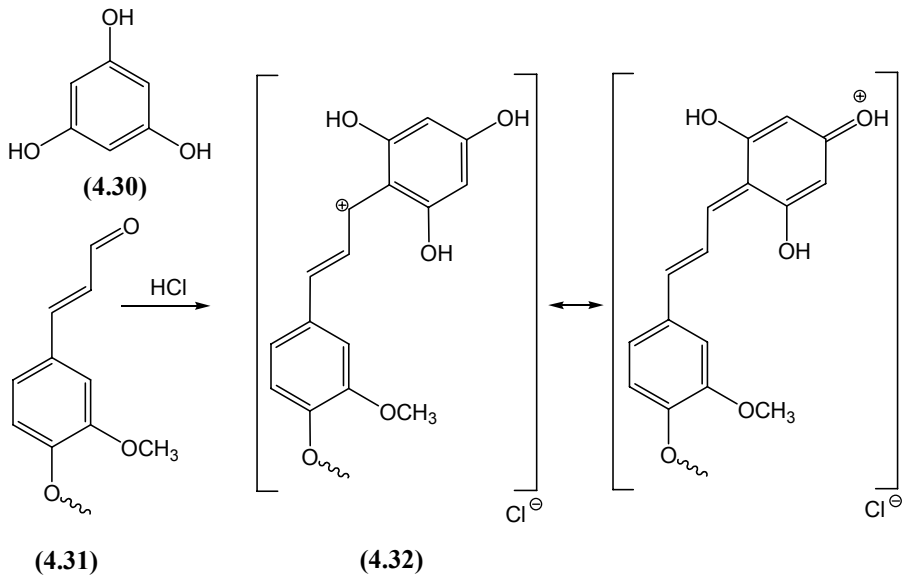


Figure 4-6. Reaction of phloroglucinol with coniferaldehyde end-groups in the lignin under acidic conditions results in the formation of a red chromophore.

a saturated solution of vanillin in 95% ethanol for 15-30 min. The section is then placed on a microscope slide, one drop of 9N HCl is added and the section is immediately viewed under the microscope. This stain will produce a red color in the presence of tannins. When pictures need to be taken, it is important to realized that this stain is not permanent.

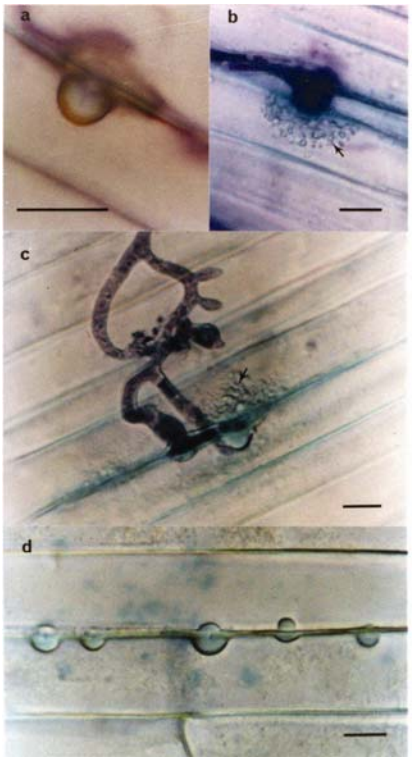
4.2 Visualizing plant-pathogen interactions involving phenolics with histochemical stains

As most phenols are found as esters or glycosides, knowledge of their location in a cell or tissue is essential. It is typical that such phenols are sequestered or stored in the cell vacuole. This is important since all phenols are weak acids (see Chapter 2) and as such they are relatively toxic even to

the cells in which they are synthesized. The toxicity of the compounds is significantly reduced when they are present as either esters or glycosides. The enzymes that cleave them are sometimes specific esterases and glycosidases but nonspecific enzymes may also cleave them. The enzymes are also located within the cell vacuole and in some cases have been shown to be present within the cell wall. Under environmental stresses, such as heat, cold and high light intensity, appropriate enzymes can be produced resulting in compound breakdown. The most typical situation, however, is that of disease where an invading pathogen such as a fungus perturbs cells or tissues causing them to respond by the synthesis of enzymes. It is also not uncommon for a single cell of the host to respond to the presence of an invading pathogen, a condition often referred to as a *hypersensitive response*. Although the hypersensitive response is a restricted phenomenon with the involvement of a very limited number of host cells, there is substantial evidence that host cells that surround the site of attempted infection are themselves stimulated or activated to a condition that allows for a much more rapid response if they themselves are perturbed by a pathogen that is attempting to infect the cell or tissue. The best-documented examples of this phenomenon are found in the barley powdery mildew disease interaction. The powdery mildew pathogen is *Erysiphe graminis* (*Blumeria graminis*), which is an obligate pathogen. It was shown that when a conidium of a compatible pathogen attempts to penetrate into a susceptible host, the penetration attempt is successful. However, if a conidium of an incompatible pathogen attempts to penetrate, the host cell responds in an incompatible manner and expresses resistance (Shiraishi et al., 1995).

Papillae are structures made by the plant in an attempt to contain and eliminate a pathogenic fungus. Papilla formation in maize leaves inoculated with *Colletotrichum graminicola* was reported by Politis and Wheeler (1973). These authors reported that even 'massive' papillae were unable to prevent penetration by the fungus. Their data were, however, not based on a time-sequence study of the events leading up to and including penetration and they did not present data on the number of successful penetrations. Papilla formation was also reported in isolated maize root cap cells after inoculation with *C. graminicola* (Sherwood, 1985). Differences in the frequency of papilla formation and penetration by the fungus were found among different maize lines.

Inoculation of maize roots with the non-pathogen *Phytophthora cinnamomi* also resulted in production of papillae (Hinch and Clarke, 1980). Callose was reported as the major component of the papillae, and carbohydrates but not proteins were identified. Lignin did not appear to be



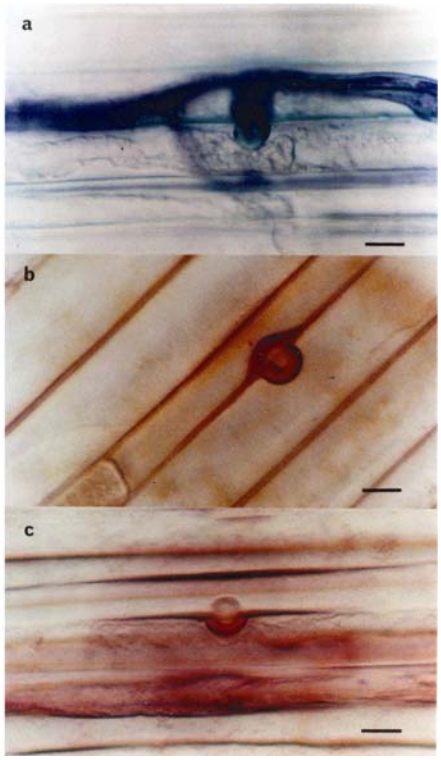


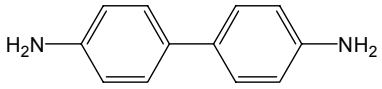
Figure 4-7 (left). Characteristics of papilla formation by the epidermis of maize mesocotyls in response to attempted penetration by *Colletotrichum graminicola* and *Helminthosporium maydis*. a) Papilla formed by inbred $B73_{Ht}$ in response to *C. graminicola,* b) granulation (arrow) in $B73_{Ht}$ in response to *H. maydis,* c) granulation (arrow) and papilla formation in a single epidermal cell of the inbred $B73_{Htrhm}$ in response to *H. maydis.* All specimens stained with Toluidine Blue-O. Bars represent 10 µM. Reprinted from Physiol. Mol. Plant Pathol. Vol. 31, Cadena-Gomez, G., and Nicholson, R.L., Papilla formation and associated peroxidase activity: A non-specific response to attempted fungal penetration of maize, pages 51-67, 1987, with permission from Elsevier.

Figure 4-8 (right). Histochemical reactions of maize ($B73_{Htrhm}$) papillae. a) Specimen stained with Toluidine Blue-O. Blue-green color of the papilla and associated host cell wall indicate the presence of lignin. b) Specimen stained with phloroglucinol-HCl. Red to red-brown color indicates the presence of cinnamaldehyde end groups in lignin. c) Specimen stained for the presence of peroxidase by the syringaldazine procedure. Red color indicates the presence of peroxidase. Specimens were taken between 12 and 24 h after inoculation. Reprinted from Physiol. Mol. Plant Pathol. Vol. 31, Cadena-Gomez, G., and Nicholson, R.L., Papilla formation and associated peroxidase activity: A non-specific response to attempted fungal penetration of maize, pages 51-67, 1987, with permission from Elsevier.

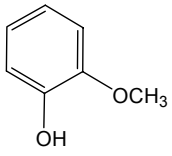
present. The authors reported that papillae formed after penetration of roots may represent a wound response that allows repair of damaged cells or provides a barrier to toxic products produced by the fungus.

In maize a variety of stain procedures were used with light microscopy to demonstrate that papillae formed in response to attempted infection by *Colletotrichum graminicola* and *Helminthosporium maydis* are composed of lignin (Cadena-Gomez and Nicholson, 1987; Figures 4-7 and 4-8).

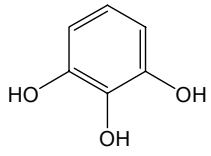
Different histochemical tests have been used for peroxidase identification. Benzidine (**4.33**) has been used as a staining reagent, as well as guaiacol (**4.34**) and pyrogallol (**4.35**). However, until the 1970s, no reliable methods were known that allowed a sharp discrimination between oxidase and peroxidase activities (Maehly and Chance, 1954). Harkin and Obst (1973) reported the syringaldazine (**4.36**) histochemical test for peroxidase. This test permitted the proof of exclusive peroxidase participation in the lignification process.



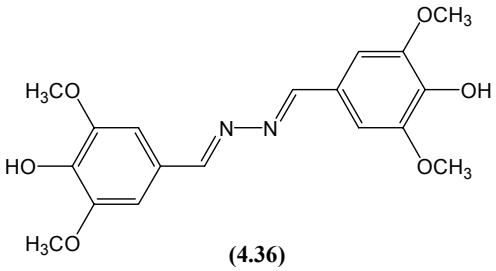
(4.33)



(4.34)



(4.35)



(4.36)

Cytochemical identification of peroxidase was conducted in infected wound margins in the cell walls and degenerating cytoplasm of wheat leaves with 3,3′-diaminobenzidine (Thorpe and Hall, 1984). The same substrate

was used to demonstrate peroxidase in the epidermis of wheat roots (Smith and O'Brian, 1979), and in roots and hypocotyls of cotton seedlings (Mueller and Beckman, 1978), and in bean leaves infected by the rust fungus *Uromyces apendiculatus* (Deising et al., 1992).

Peroxidase activities were also shown to occur during the bean rust infection process (Mendgen, 1975). In reed canary grass (*Phalaris arundinacea* L.), peroxidase was identified with the pyrogallol test (Vance and Sherwood, 1976).

5. REFERENCES

Adler, E., Björkqvist, K. J., and Häggroth, S., 1948, Über die Ursache der Farbreaktionen des Holzes, *Acta Chem. Scand*. **2**: 93-94.

Bate-Smith, E. C., 1977, Astringent tannins of *Acer* species, *Phytochem*. **16**: 1421-1426.

Boon, J. J., 1992, Analytical pyrolysis mass spectrometry: new vistas opened by temperature-resolved in-source PYMS. *Int. J. Mass Spectrom. Ion Process.* **118/119**: 755-787.

Boon, J. J., 1989, An introduction to pyrolysis mass spectroscopy of lignocellulosic material: case studis of barley straw, corn stem and Agropyron, in: *Physico-chemical Characterisation of Plant Residues for Industrial and Feed Use*, A. Chesson, and E. R. Ørskov, eds., Elsevier Applied Science London, pp. 25-49.

Bout, S., and Vermerris, W., 2003, A candidate-gene approach to clone the sorghum *Brown midrib* gene encoding caffeic acid *O*-methyltransferase. *Mol. Genet. Genomics* **269**: 205-214.

Brinkmann, K., Blaschke, L., Polle, A., 2002, Comparison of different methods for lignin determination as a basis for calibration of near-infrared reflectance spectroscopy and implications of lignoproteins, *J. Chem. Ecol.* **28**: 2483-2501.

Browning, B. L., 1967, *Methods of Wood Chemistry*, Vol. II, Wiley-Interscience, New York, pp. 785-793.

Bruce, R., and West, C., 1989, Elicitation of lignin biosynthesis and isoperoxidase activity by pectic fragments in suspension cultures of castor bean, *Plant Physiol.* **91**: 889–897.

Cadena-Gomez, G., and Nicholson, R. L., 1987, Papilla formation and associated peroxidase activity: A non-specific response to attempted fungal penetration of maize, *Physiol. Mol. Plant Pathol.* **31**: 51-67.

Chapple, C. C. S., Vogt, T., Ellis, B. E., and Sommerville, C. R., 1992, An Arabidopsis mutant defective in the general phenylpropanoid pathway, Plant Cell **4**: 1413-1424.

Chen, L., Carpita, N. C., Reiter, W.-D., Wislon, R. H., Jeffries, C., and McCann, M. C. 1998, A rapid method to screen for cell-wall mutants using discriminant analysis of Fourier transform infrared spectra, *Plant J.* **16**: 385-392.

De Neergaard, E., 1997, *Methods in Botanical Histopathology*, Danish Government Institute of Seed Pathology for Developing Countries, Copenhagen, pp. 216.

Deising, H., Nicholson, R. L., Haug, M., Howard, R. J., and Mendgen, K., 1992, Adhesion pad formation and the involvement of surface localized cutinase and esterase in the attachment of Uredospores to the host cuticle, *Plant Cell* **4**: 1101-1111.

Dias, A. P., and Grotewold, E. Manipulating the accumulation of phenolics in maize cultured cells using transcription factors, *Biochem. Eng. J.* **14**: 207-216.

Faix, O., 1992, Fourier-transform infrared spectroscopy, in: *Methods in Lignin Chemistry*, S. Y. Lin, and C. W. Dence, eds. Springer-Verlag, Berlin, pp. 81-109.

Folin, O., and Ciocalteu, V., 1927, On tyrosine and tryptophan determinations in proteins, *J. Biol. Chem*. **73**: 627-650.

Folin, O., and Denis, W., 1912a, On phosphotungstic-phospomolybdic compounds as color reagents, *J. Biol. Chem*. **12**: 239-243.

Folin, O., and Denis, W., 1912b, Tyrosine in proteins as determined by a new colorimetric method, *J. Biol. Chem*. **12**: 245-251.

Folin, O., and Denis, W., 1915, A colorimetric method for the determination of phenols (and phenol derivatives) in urine, *J. Biol. Chem*. **22**: 305-308.

Fontaine, A.-S., Bout, S., Barrière, Y., and Vermerris, W., 2003. Variation in cell wall composition among forage maize inbred lines and its impact on digestibility: Analysis of neutral detergent fiber composition by pyrolysis-gas chromatography-mass spectrometry, *J. Agric. Food Chem*. **51**: 8080-8087.

Freudenberg, K., 1939, Über lignin. *Angew. Chem*. **52**: 362-363.

Friebolin, H., 2004, *Basic One- and Two-Dimensional NMR Spectroscopy*, John Wiley & Sons, pp. 430.

Fukushima, R. S., and Hatfield, R. D., 2001, Extraction and isolation of lignin for utilization as a standard to determine lignin concentration using the acetyl bromide spectrophotometric method, *J. Agric. Food Chem*. **49**: 3133-3139.

Fukushima, R. S., and Hatfield, R. D., 2004, Comparison of the acetylbromide spectrophotometric method with other analytical lignin methods for the determination of lignin concentration in forage samples, *J. Agric. Food Chem*. **52**: 3713-3720.

Geiger, H., and Fuggerer, H., 1979, Über den Chemismus der Wiesner-Reaktion auf Lignin, *Z. Naturforsch*. **34b**: 1471-1472.

Hagerman, A., And Butler, L. G., 1989, Choosing appropriate methods and standards for assaying tannin, *J. Chem. Ecol*. **15**: 1795-1810.

Halpin, C., Holt, K., Chojecki, J., Oliver, D., Chabbert, B., Monties, B., Edwards, K., Barakate, A., and Foxon, G. A., 1998, *Brown-midrib* maize (*bm1*) - a mutation affecting the cinnamyl alcohol dehydrogenase gene, *Plant J.* **14**: 545-553.

Harkin J. M., and Obst, J. R., 1973, Lignification in trees: indication of exclusive peroxidase participation, *Science* **180**: 296-298.

Haslam, E., 1965, Galloyl esrers in the *Aceraceae*, *Phytochem*. **4**: 495-498.

Hatfield, R. D., Grabber, J., Ralph, J., and Brei, K., 1999, Using the acetyl bromide assay to determine lignin concentrations in herbaceous plants: some cautionary notes, *J. Agric. Food Chem*. **47**: 628-632.

Hatfield, R. D., Jung, H.-J. G., Ralph, J., Buxton, D. R., Weimer, P. J., 1994, A comparison of the insoluble residues produced by the Klason and acid detergent lignin procedures, *J. Sci. Food Agric*. **65**: 51-58.

Hatfield, R. D., and Fukushima, R. S., 2005, Can lignin be accurately measured, *Crop Sci*. **45**: 832-839.

Hinch, J. M., and Clarke, A. E., 1980, Callose formation in *Zea mays* as a response to infection with *Phytophthora cinnamomi, Physiol. Plant Pathol.* **21**: 113-124.

Hoffmann, L., Besseau, S., Geoffroy, P., Ritzenthaler, C., Meyer, D., Lapierre, C., Pollet, B., and Legrand, M., 2004, Silencing of hydroxycinnamoyl-coenzyme A shikimate/quinate hydroxyltransferase affects phenylpropanoid biosynthesis, *Plant Cell* **16**: 1446-1465.

Iiyama, K., and Wallis, A. F. A., 1990, Determination of lignin in herbaceous plants by an improved acetyl bromide procedure, *J. Sci. Food Agric.* **51**: 145-162.

Inoue, K. H., and Hagerman, A., 1988, Determination of gallotannin with rhodanine, *Anal. Biochem*. **169**: 363-369.

Johnson, D. B., Moore, W. E., and Zank, L. C., 1961, The spectro-photometric determination of lignin in small wood samples, *Tappi* **44**: 793-789.

Johnson, R. A., and Wichern, D. W., 1992, *Applied Multivariate Statistical Analysis*, Prentice Hall, Englewood Cliffs, NJ, pp. 767.

Jung, H. G., 1997, Analysis of forage fiber and cell walls in ruminant nutrition, *J. Nutr*. **127**: 810S-813S.

Jung, H. G., and Deetz, D. A., 1993, Cell wall lignification and degradability, in: *Forage Cell Wall Structure and Digestibility*, H. G. Jung, D. R. Buxton, R. D. Hatfield, and J. Ralph, eds., ASA, CSCA, SSSA, Madison, WI, pp. 315-346.

Keeler, J., 2005, *Understanding NMR Spectroscopy*, John Wiley & Sons, pp. 476.

Kemsley, E.K., 1998, *Discriminant Analysis and Class Modelling of Spectroscopic Data*, John Wiley & Sons Ltd, Chichester, UK, pp. 179.

Lapierre, C., 1993, Application of new methods for the investigation of lignin structure, in: *Forage Cell Wall Structure and Digestibility*, H. G. Jung, D. R. Buxton, R. D. Hatfield, and J. Ralph, eds., ASA, CSCA, SSSA, Madison, WI, pp. 133-166.

Lewis, N. G., and Yamamoto, E., 1990, Lignin: occurrence, biogenesis and biodegradation, *Ann. Rev. Plant Phys. Plant Mol. Biol.* **41**: 455-496.

Lin, S. Y., and Dence, C. W., 1992, *Methods in Lignin Chemistry*, Springer-Verlag, Berlin, pp. 578.

Lu, F., and Ralph, J., 1997, Derivatization followed by reductive cleavage (DFRC method), a new method for lignin analysis: protocol for analysis of DFRC monomers, *J. Agr. Food Chem*. **45**: 2590-2592.

Lu, F., and Ralph, J., 2003, Non-degradative dissolution and acetylation of ball-milled cell walls: high-resolution solution-state NMR, *Plant J.* **35**: 535-544.

Maehly, A. C., and Chance, B., 1954, The assay of catalases and peroxidase, in: *Methods of Biochemical Analysis*, Vol. 1, D. Glick, ed., Interscience Publishers, New York, pp. 357-408.

Marita, J. M., Ralph, J., Hatfield, R. D., and Chapple, C., 1999, NMR characterization of lignins in Arabidopsis altered in the activity of ferulate 5-hydroxylase, *Proc. Natl. Sci. Acad. USA* **96**: 12328–12332

Marita, J., Vermerris, W., Ralph, J., and Hatfield, R. D., 2003, Variations in the cell wall composition of maize *brown midrib* mutants, *J. Agric. Food Chem.* **51**: 1313-1321.

Meier, D., and Faix, O., 1992, Pyrolysis-gas chromatography-mass spectroscopy, in: *Methods in Lignin Chemistry*, S. Y. Lin and C. W. Dence, eds., Springer-Verlag, Berlin, pp.177-199.

Mendgen, K., 1975, Ultrastructural demonstration of different peroxidase activities during the bean rust infection process, *Physiol. Plant Pathol*. **6**: 2275-2282.

Meyer, K., Shirley, A. M., Cusumano, J. C., Bell-Lelong, D. A., and Chapple, C., 1998, Lignin monomer composition is determined by the expression of a cytochrome P450-dependent monooxygenase in Arabidopsis, *Proc. Natl. Acad. Sci. USA* **95**: 6619-6623.

Monties, B., 1989, Lignins, in: *Methods in Plant Biochemistry*, *Vol. 1 Plant Phenolics*, J. B. Harborne, ed., Academic Press, London, pp. 113-157.

Morreel, K., Ralph, J., Lu, F., Goeminne, G., Busson, R., Herdewijn, P., Goeman, J. L., Van der Eycken, J., Boerjan, W., and Messens, E. 2004, Phenolic profiling of caffeic acid *O*-methyltransferase-deficient poplar reveals novel benzodioxane oligolignols., *Plant Physiol.* **136**: 4023-4036.

Mouille, G., Robin S., Lecomte, M., Pagant S., and Höfte, H., 2003, Classification and identification of Arabidopsis cell wall mutants using Fourier-Transform InfraRed (FT-IR) microspectroscopy, *Plant J.* **35**: 393-404.

Mueller, W. C., and Beckman, C. H., 1978, Ultrastructural localization of polyphenoloxidase and peroxidase in roots and hypocotyls of cotton seedlings, *Can. J. Bot*. **56**: 1579-1587.

Mulder, M. M., Van der Hage, E. R. E., and Boon, J. J., 1992, Analytical in source pyrolytic methylation electron impact mass spectroscopy of phenolic acids in biological matrices, *Phytochem. Anal*. **3**: 165-172.

O'Connell, A., Holt, K., Piquemal, J., Grima-Pettenati, J., Boudet, A., Pollet, B., Lapierre, C., Petit-Conil, M., Schuch, W., and Halpin, C. 2002, Improved paper pulp from plants with suppressed cinnamoyl-CoA reductase or cinnamyl alcohol dehydrogenase, *Transgen. Res.* **11**: 495-503.

Politis, D. J., and Wheeler, H., 1973, Ultrastructural study of penetration of maize leaves by *Colletotrichum graminicola, Physiol. Plant Pathol*. **3**: 465-471.

Pomar, F., Merino, F., and Ros Barceló, A., 2002, *O*-linked coniferyl and sinapyl aldehydes in lignifying cell walls are the main targets of the Wiesner (phloroglucinol-HCl) reaction, *Protoplasma* **220**: 17-28.

Ralph, J., and Helm, R. F., 1993, Lignin/hydroxycinnamic acid/ polysaccharide complexes: synthetic models for regiochemical characterization, in: *Forage Cell Wall Structure and Digestibility*, H. G. Jung, D. R. Buxton, R. D. Hatfield, and J. Ralph, eds., ASA, CSCA, SSSA, Madison, WI, pp. 201-246.

Ralph, J., Lapierre, C., Marita, J. M., Kim, H., Lu, F., Hatfield, R. D., Ralph, S., Chapple, C., Franke, R., Hemm, M. R., Van Doorsselaere, J., Sederoff, R. R., O'Malley, D. M., Scott, J. T., Mackay, J. J., Yahiaoui, N., Boudet, A. M., Pean, M., Pilate, G., Jouanin, L., and Boerjan, W., 2001, Elucidation of new structures in lignins of CAD- and COMT-deficient plants by NMR, *Phytochem.* **57**: 993-1003.

Ralph, J., MacKay, J. J., Hatfield, R. D., O'Malley, D. M., Whetten, R. W., and Sederoff, R. R., 1997, Abnormal lignin in a loblolly pine mutant, *Science* **277**: 235-238.

Ralph, J., Marita, J. M., Ralph, S. A., Hatfield, R. D., Lu, F., Ede, R. M., Peng, J., Quideau, S., Helm, R. F., Grabber, J. H., Kim, H., Jimenez-Monteon, G., Zhang, Y., Jung, H.-J. G., Landucci, L. L., MacKay, J. J., Sederoff, R. R., Chapple, C., and Boudet, A.-M., 1999, Solution-state NMR of lignins, in: *Advances in Lignocellulosics Characterization*, Argyropoulos, D. S. and Rials, T., eds., Technical Association of the Pulp and Paper Industry, Atlanta, GA, pp. 55-108.

Scalbert, A., Monties, B., and Janin, G., 1989, Tannins in wood: comparison of different estimation methods*, J. Agric. Food Chem.* **37**: 1324-1329.

Schultz, J. C., and Baldwin, I. T., 1982, Oak leaf quality declines in response to defoliation by gypsy moth larvae, *Science* **217**: 149–151.

Sené, C. F. B., McCann, M. C., Wilson, R. H., Grinter, R., 1994, Fourier-transform Raman and Fourier-transform infrared spectroscopy, *Plant Physiol.* **106**: 1623-1631.

Shiraishi, T., Yamada, T., Nicholson, R. L., and Kunoh, H., 1995, Phenylalanine ammonia-lyase in barley: Activity enhancement in response to *Erysiphe graminis* f.sp. *hordei* (race 1) a pathogen, and *pisi,* a non-pathogen, *Physiol. Mol. Plant Pathol*. **46**: 153-162.

Sherwood, R. T., 1985, Papilla formation in isolated root-cap cells of maize inoculated with *Colletotrichum graminicola, Phytopathol*. **75**: 1313-1314.

Siesler, H. W., Ozaki, Y., Kawata, S., and Heise, H. M., 2002, *Near Infrared Spectroscopy. Principles, Instruments, Applications*, Wiley-VCH, Weinheim, Germany, pp. 348.

Sjöström, E., and Alen, R., 1999, *Analytical Methods in Wood Chemistry, Pulping and Papermaking*, Springer-Verlag, Berlin, pp. 316.

Smith, M. M., and O'Brian, T. P., 1979, Distribution of auto fluorescence and esterase and peroxidase activities in the epidermis of wheat roots, *Austr. J. Plant Physiol*. **6**: 201-219.

Swain, T., and Hillis, W. E., 1959, The phenolic constituents of *Prunus domestica.* I. The quantitative analysis of phenolic constituents, *J. Sci. Food Agric.* **10**: 63-68.

Tamagnone, L., Merida, A., Parr, A., MAckay, S., Culianez-Macia, F. A., Roberts, K., and Martin, C. 1998, The AmMYB308 and AmMYB330 transcription factor from Antirrhinum regulate phenylpropanoid and lignin biosynthesis in transgenic tobacco, *Plant Cell* **10**: 135-154.

Theander, O., and Westerlund, A., 1986, Studies on dietary fiber. 3. Improved analysis of dietary fiber, *J. Agric. Food Chem*. **34**: 330-336.

Thorpe, J. R., and Hall, J. L., 1984, Chronology and elicitation of changes in peroxidase and phenylalanine ammonia-lyase activities in wounded wheat leaves in response to inoculation with *Botrytis cinerea*, *Physiol. Plant Pathol.* **25**: 363-379.

Van Soest P.J., 1963, Use of detergents in analysis of fibrous feeds. 2. A rapid method for the determination of fiber and lignin, *J. Assoc.Off. Agr. Chem.* **46**: 829-835.

Van Soest, P. J., 1967, Development of a comprehensive system of feed analyses and its application to forages, *J. Anim. Sci.* **26**: 119-128.

Van Soest P.J., Wine R.H., 1967, Use of detergents in analysis of fibrous feeds. Determination of plant cell-wall constituents, *J. Assoc. Off. Agr. Chem.* **50**: 50-55.

Vance, C. P., and Sherwood, R. T., 1976, Cyclohexamide treatments implicate papilla formation in reed canarygrass to fungi, *Phytopathol.* **66**: 498-502.

Vermerris, W., and Boon, J. J., 2001, Tissue-specific patterns of lignification are disturbed in the *brown midrib2* mutant of maize (*Zea mays* L.), *J. Agric. Food Chem.* **49**: 721-728.

Vermerris, W., Thompson, K. J., and McIntyre, L. M., 2002, The maize *Brown midrib1* locus affects cell wall composition and plant development in a dose-dependent manner, *Heredity* **88**: 450-457.

Viriot, C., Scalbert, A., Lapierre, C., and Moutounet, M. 1993, Ellagitannins and lignins in aging of spirits in oak barrels, *J. Agric. Food Chem.* **41**: 1872-1879.

Wilson, T. C., and Hagerman, A. E., 1990, Quantitative determination of ellagic acid, *J. Agric. Food Chem.* **38**: 1678-1683.

Yong, W., Link, B., O'Malley, R., Tewari, J., Hunter, C. T., III, Lu, C.-A., Li, X., Bleecker, A. B., Koch, K. E., McCann, M. C., McCarty, D. R., Patterson, S. E., Reiter, W.-D., Staiger, C., Thomas, S. R., Vermerris, W., and Carpita, N. C., 2005, Genomics of plant cell wall biogenesis, *Planta* **221**: 745-751.

Zhang, K., Qian, Q., Huang, Z., Wang, Y., Li, M., Hong, L., Zeng, D., Gu, M., Chu, C., and Cheng, Z., 2006, *GOLD HULL AND INTERNODE2* encodes a primarily multifunctional cinnamyl-alcohol dehydrogenase in rice, *Plant Physiol*. **140**: 972-983.

Recommendations for further reading

De Neergaard, E., 1997, *Methods in Botanical Histopathology*, Danish Government Institute of Seed Pathology for Developing Countries, Copenhagen, pp. 216.

Friebolin, H., 2004, *Basic One- and Two-Dimensional NMR Spectroscopy*, John Wiley & Sons, pp. 430.

Johnson, R. A., and Wichern, D. W., 1992, *Applied Multivariate Statistical Analysis*, Prentice Hall, Englewood Cliffs, NJ, pp. 767.

Keeler, J., 2005, *Understanding NMR Spectroscopy*, John Wiley & Sons, pp. 476.

Kemsley, E.K., 1998, *Discriminant Analysis and Class Modelling of Spectroscopic Data*, John Wiley & Sons Ltd, Chichester, UK, pp. 179.

Siesler, H. W., Ozaki, Y., Kawata, S., and Heise, H. M., 2002, *Near Infrared Spectroscopy. Principles, Instruments, Applications*, Wiley-VCH, Weinheim, Germany, pp. 348.

Sjöström, E., and Alen, R., 1999, *Analytical Methods in Wood Chemistry, Pulping and Papermaking*, Springer-Verlag, Berlin, pp. 316.

Chapter 5

ANALYSIS OF PHENOLIC COMPOUNDS WITH MASS SPECTROMETRY

Karl V. Wood
Department of Chemistry, Purdue University

1. THE PRINCIPLES OF MASS SPECTROMETRY

Mass spectrometry is a very valuable analytical tool based on the simple premise of determining the molecular weight of the compound of interest. In fact, mass spectrometry involves the measurement of the mass (m) of a compound as a function of charge (z), m/z. In most mass spectrometer experiments the charge on an ion is one, such that the molecular weight of the ion is equal in value to m/z. Mass spectrometry can be divided into three steps: ionization, mass analysis and detection. Sample introduction is really a distinct component for obtaining quality mass spectra and will be discussed separately.

1.1 Ionization

Once the sample has been introduced into the mass spectrometer, it must be ionized. There are a number of different ionization techniques which span a wide-range of different compound classes. The original ionization technique was termed electron impact (EI). EI involves the interaction of the compound of interest (M) with 70 eV electrons that have been emitted from a filament wire (1 eV = $1.6*10^{-19}$ Joule). This results in a relatively energetic reaction in which the compound of interest (M) loses an electron, thereby forming a radical cation.

$$M + e^- \rightarrow M^{+\cdot} + 2e^- \quad (1)$$

Using appropriate electric fields the radical cation $M^{+\cdot}$ can now be sent to the mass analyzer for subsequent mass analysis. This radical cation can also undergo fragmentation.

$$M^{+\cdot} \rightarrow F^+ + e^- \quad (2)$$

The fragment ions (F^+) formed can be used to provide additional structural information about the original compound.

Another conventional ionization technique termed chemical ionization (CI), utilizes a reagent gas (such as isobutane, methane, ammonia) to form reagent ions (RH^+) which can undergo ion-molecule reactions with the compound of interest to form protonated molecules.

$$M + RH^+ \rightarrow (M+H)^+ + R \quad (3)$$

The protonated molecules are much less energetic than the related radical cations formed during electron impact and are less prone to fragmentation. Both EI and CI require volatile samples. This puts a tremendous limitation on the types of molecules that can be ionized, in particular, compounds found in biological systems.

1.2 Mass Analysis

After the compound of interest has been ionized the resulting ion must be mass analyzed. There are a number of commercially available mass analyzers, including but not limited to quadrupole, time-of-flight, ion trap and double-focusing sector analyzers. A quadrupole mass analyzer utilizes four concentric quadrupole rods with the opposite rods having opposite polarity. With appropriate direct current (DC) and radio frequency (*rf*) voltages, ions can be made to pass through the rods as a function of their mass-to-charge ratio.

An ion trap mass analyzer has a variety of differing physical arrangements of its electrodes, but the primary objective remains the same to allow the ions to enter and then to trap them in space between the electrodes. Unlike the fly-through mass analysis scheme of a quadrupole, the ion trap mass analyzer stores the ions. They are then ejected to the detector as a function of the mass-to-charge ratio, typically by scanning the *rf* voltage.

The time-of-flight mass analyzer utilizes the physical characteristics of the ions themselves. Smaller ions travel faster than larger ions. The ions formed are accelerated such that ions with the same m/z have similar kinetic energies. This enables the arrival times of the ions at the detector to be related to their mass-to-charge ratio. Two improvements for reducing kinetic energy spread – and ultimately resolution – have been introduced in conventional linear time-of-flight mass analyzers: delayed extraction and the reflectron. In delayed extraction there is a measured delay in the time between initial ion formation and acceleration down the flight tube, resulting in reduced kinetic energy spread of ions with the same m/z. A reflectron is an electrostatic ion focusing element. Ions of the same m/z but with different kinetic energies require more or less time in order to be reflected, with, as the net result, simultaneous arrival at the detector.

The double-focusing sector mass spectrometer is used primarily to obtain high-resolution mass measurements, that is, determining the mass to better than 5 ppm accuracy. This mass analyzer is comprised of both a magnetic sector and an electric sector, utilized in combination. The magnetic sector separates the ions based on momentum and the electric sector separates the ions based on kinetic energy. The result is ions that have very little spread in energy, thus enabling their mass-to-charge ratios to be measured very accurately.

1.3 Detectors

Once the ions of the compound of interest have been mass analyzed, the ions are directed to the detector of the mass spectrometer. Most detectors are based on the premise that the incoming ion will impact the surface of the detector, thereby forming many secondary electrons. Depending on the geometry of the detector these secondary electrons then impact another part of the detector surface. This process continues resulting in a 10^5 and higher amplification of the original ion beam. Parameters such as mass and velocity of the impacting ion as well as chemical characteristics of the impacted surface affect the ultimate current amplification in any detector.

The information obtained from the detector is used to generate the mass spectrum. The mass spectrum is a plot of the intensity of the individual mass-analyzed ions plotted as a function of m/z. Usually the most intense ion, termed the base peak, is given a relative abundance of 100% and the rest of the ions in the mass spectrum are normalized to this intensity. Figure 5.1 shows the EI mass spectrum of 2-methoxy-4-vinyl phenol (molecular weight 150 Daltons). The base peak is the molecular ion at m/z 150, a radical

cation. The three most intense fragment ions are m/z 135 (loss of $CH_3^{\cdot}$), m/z 107 (loss of $CH_3^{\cdot}$ and CO), m/z 77 ($C_6H_5^+$; loss of $^{\cdot}OCH_3^{\cdot}$, C_2H_2, and $^{\cdot}OH$), and m/z 51 ($C_4H_3^+$; loss of C_2H_2 from m/z 77).

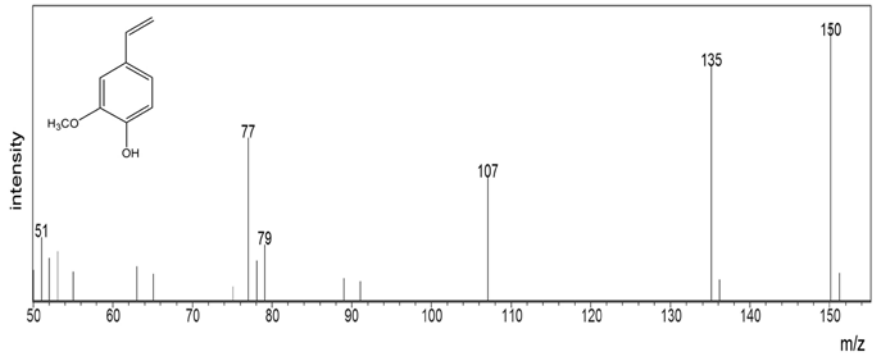


Figure 5-1. Mass spectrum of 2-methoxy 4-vinylphenol.

1.4 Sample introduction

Samples submitted for analysis can be introduced into the mass spectrometer in many forms. Probably the most common means are (1) direct introduction, (2) gas chromatography (GC), and (3) liquid chromatography (LC).

In direct introduction the sample can be introduced via a sample probe or plate through a vacuum lock, and can subsequently be ionized via EI, CI or matrix-assisted laser desorption ionization (MALDI; see Section 2.4). Alternatively, the sample can be introduced as a liquid stream into an ion source at atmospheric pressure, after which it is subjected to electrospray ionization (ESI; see Section 2.3). Direct injection does not offer any form of sample separation.

In contrast, both gas and liquid chromatography enable the samples of interest to be separated into individual components prior to introduction into the mass spectrometer ion source. Gas chromatography involves sample introduction with the requisite that the sample components must be volatilized prior to separation, and results in a gas sample being introduced to the mass spectrometer (*i.e.* EI, CI). Figure 5-2 shows the chromatogram obtained after a mixture of three simple phenolic compounds – phenol,

o-cresol and 2,5-dimethylphenol – was injected in the gas chromatograph. This analysis utilized a 15-m long DB5 capillary column with an initial column temperature of 50°C heated to 320°C at 15°C per minute. The retention time is the time that elapses between injection of the mixture and the elution (detection) of an individual compound and depends on the volatility of the compound combined with the affinity of the compound for the stationary phase inside the capillary column. Liquid chromatography involves a liquid sample being separated into individual components and these components being introduced into a mass spectrometer ion source at atmospheric pressure (*i.e.* ESI).

It is important to remember the few restrictions imposed by electrospray when considering an LC-MS analysis. Common solvents like methanol, water, acetonitrile and volatile salts (below 25 mM) like ammonium acetate and ammonium bicarbonate are acceptable in the mobile phase, whereas phosphate salts/buffers, mineral acids or other nonvolatile components cannot be used. Unfortunately, this conflicts with many of the routine mobile phases used for the analysis of phenolic compounds and anthocyanins, necessitating changes in methods when going from LC to LC-MS analyses.

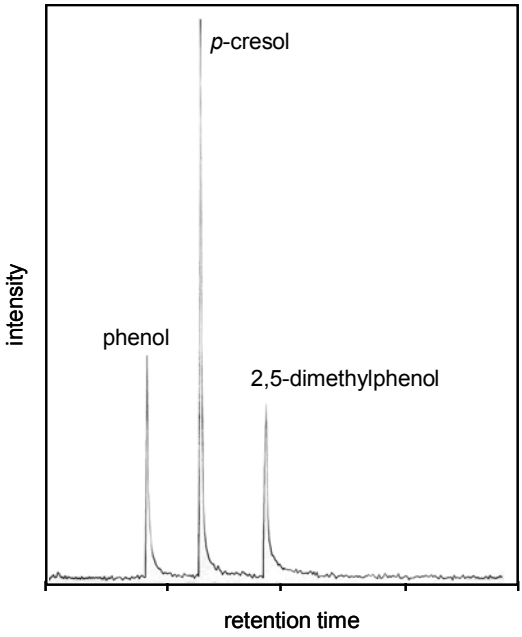


Figure 5-2. Gas chromatogram obtained after injecting a mixture of three phenolic compounds into a gas chromatograph equipped with a 15-m long DB5 column. The temperature was increased from 50 to 320°C at 15°C per minute.

Many other analytical techniques can be coupled to mass spectrometers. These so-called hyphenated techniques, like GC-MS and LC-MS, include but are not limited to ICP-MS (inductively coupled argon plasma), SCF-MS (supercritical fluid), NMR-MS (nuclear magnetic resonance) and IR-MS (infrared).

2. NEW DEVELOPMENTS IN MASS SPECTROMETRY

Direct mass spectrometric analysis of thermally labile and nonvolatile molecules, like many of the anthocyanins and other phenolic compounds, has not always been possible. In the days of EI and CI it was necessary to derivatize these molecules before obtaining mass spectrometric data. However, with the advent of soft ionization techniques - initially fast atom bombardment (FAB; Fenselau, 1983) and plasma desorption (PD; Sundqvist and Macfarlane, 1985) - followed by the development of electrospray ionization (ESI; Cole, 2000) and matrix-assisted laser desorption ionization (MALDI; Zenobi and Knochenmuss, 1998), the mass spectrometry of these compound classes became possible, even routine. These techniques provide molecular weight information through formation of $(M+H)^+$ or $(M+Na)^+$ ions, or in the case of pre-charged ions like anthocyanins, the cations themselves. Mass spectrometry not only provides information on the molecular weight of a compound of interest, but with the advent of tandem mass spectrometric techniques structurally significant information can also be generated. In tandem mass spectrometry (ms/ms) the ion of interest is selected and given excess energy. This can be done in a variety of ways, including collisions with neutral atoms (Busch et al., 1988). The excited molecule can then undergo fragmentation, with the resulting ions providing structural information about the original precursor ion.

2.1 Fast atom bombardment

FAB mass spectrometric analyses require a high-energy atom beam, usually 6-10 keV. The atom beam, typically xenon, is directed at the sample which is dissolved in a matrix. Typical matrices include glycerol, thioglycerol, *m*-nitrobenzyl alcohol and a mixture of dithiothreitol and dithioerythritol. The continual bombardment of the sample/matrix mixture results in desorption of both species. Ions are formed, either as pre-formed ions from the matrix or in the gas phase immediately above the sample surface.

2.2 Plasma desorption ionization

Plasma desorption mass spectrometry utilizes a ^{252}Cf (californium) ionizing source which produces MeV fission fragments. The interaction of the fission fragments with the sample produces ions which are mass analyzed. The samples are applied to a nitrocellulose-coated mylar target, either by droplet or electrospraying, allowed to adsorb, and then washed with a 0.1% trifluoroacetic acid solution. Typically the samples are allowed to air dry prior to being introduced into the mass spectrometer. The sample ions that are formed, are accelerated into a time-of-flight mass spectrometer for mass analysis.

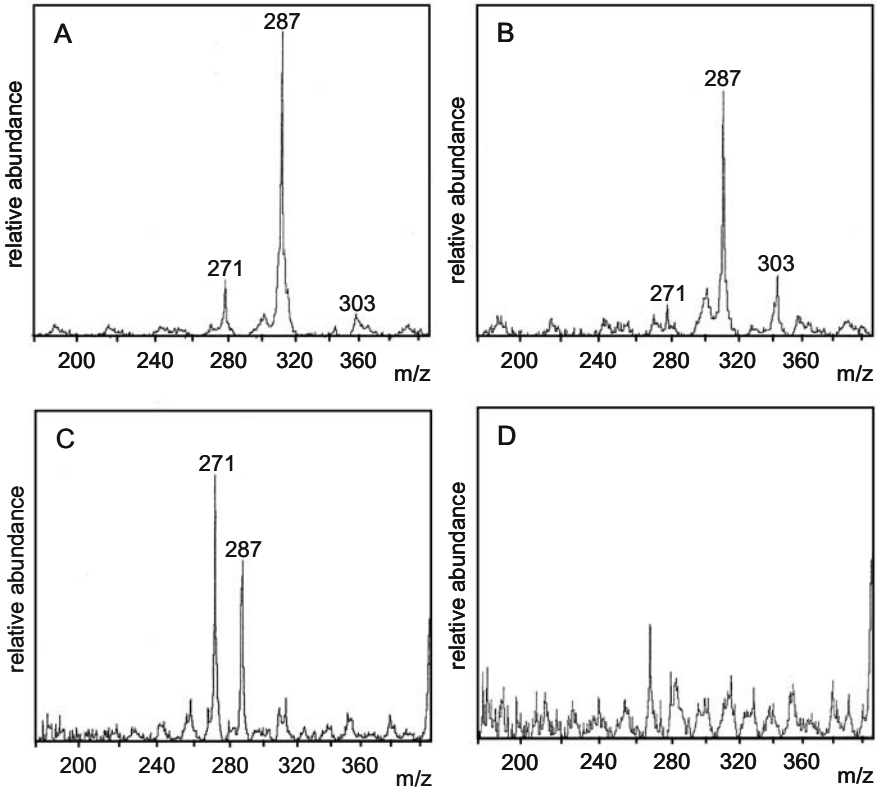


Figure 5-3. Plasma desorption mass spectra of anthocyanidins extracted from **A**. chrysanthemum, **B**. begonia, **C**. carnation, and **D**. phlox. The data in this figure was published in the article 'Plasma desorption mass spectrometry of anthocyanidins', Rap. Comm. Mass Spectrom. 7:400-403, by Wood, K. V., Bonham, C. C., Ng, J., Hipskind, J. and Nicholson, R. L. 1993. Copyright John Wiley and Sons. Reproduced with permission.

Accelerating potentials between 15-20 keV are typically used, with data being collected from anywhere between fifteen minutes to an hour depending on the sample. The useful mass range for PDMS is up to m/z 5000. Typically the observed ion is the protonated molecule. PDMS, however, has proven to be particularly well-suited for the analysis of pre-charged species like anthocyanidins. This can be seen in Figure 5-3, which shows the plasma desorption mass spectrum of anthocyanidins extracted from chrysanthemum, begonia, carnation and phlox (Wood et al., 1993). Note the presence of cyanidin (m/z 287; **1.40**), as a dominant ion in the first three mass spectra. Pelargonidin (m/z 271; **1.39**) and delphinidin (m/z 303; **1.42**) are also observed. The phlox sample analyzed was white phlox which explains the absence of any anthocyanidins.

2.3 Electrospray ionization

FAB and PD have been replaced by electrospray ionization (ESI) and matrix-assisted laser desorption ionization (MALDI) in the analytical mass spectrometry laboratory, because both of these newer techniques have a wider mass range of analysis and have lower detection limits. ESI and MALDI have become invaluable ionization techniques for nonvolatile components. This is particularly true for a wide range of biological molecules including proteins, peptides, nucleic acids, etc. Samples can be analyzed by ESI using either direct injection or introduction through liquid chromatography.

Typically ESI forms protonated molecules with little or no fragmentation. High molecular weight molecules can also be ionized using ESI. In this case the molecule becomes multiply-charged and the molecule of interest is observed at its respective m/z. For example, a protein with a molecular weight of 20,000 Da that has 20 protons attached to it, would be detected at m/z 20,020/20, which is equal to 1,001. Figure 5-4 shows the ESI mass spectrum of horse heart myoglobin, with a molecular weight 16,951 Daltons. The $M+16H^+$ and $M+15H^+$ ions result in the peaks at m/z 1060.5 ((16,951 + 16)/16) and 1131.1((16,951 + 15)/15), respectively.

In the case of a protein of unknown molecular weight, the molecular weight can be determined by using equations (4) and (5) and knowing the m/z values of two adjacent multiply-charged ions.

$$n = (X_1 - 1)/(X_2 - X_1) \qquad (4)$$

X_1 is the mass-to-charge ratio of the ion with the lower m/z, X_2 is the mass-to-charge ratio of the ion with the higher m/z, and *n* is the charge on the ion with the higher m/z. Knowing the charge on a given m/z ion, the molecular weight of the protein can be calculated by using the following equation.

$$MW = nX_2 - n \quad (5)$$

In the example of the horse heart myoglobin, we select X_1 = 1060.5 and X_2 = 1131.1, *n* = 15, and MW = 15*1131.1 – 15 = 16,951.5. This is consistent with the known molecular weight of the protein.

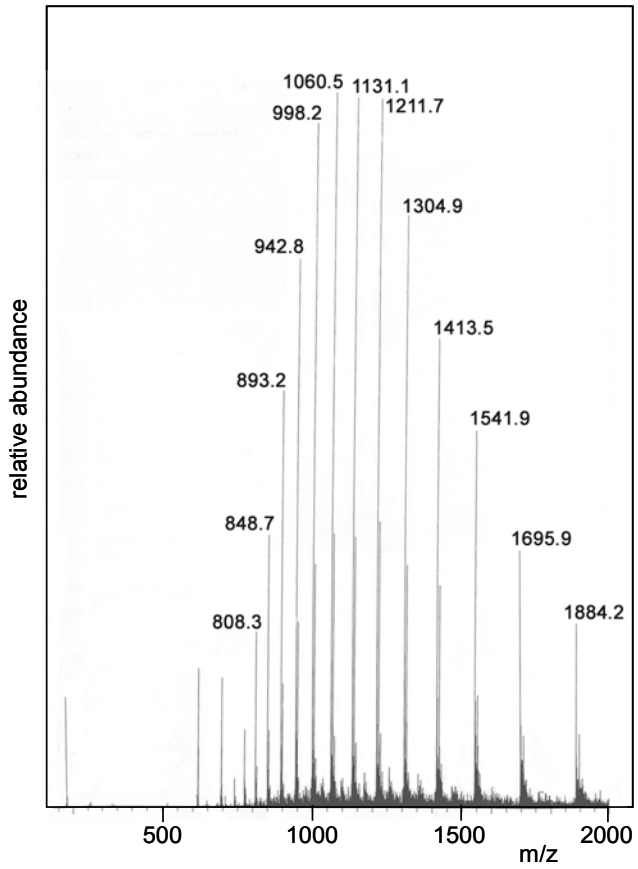


Figure 5-4. Electrospray ionization mass spectrum of horse heart myoglobin.

The choice of solvents and related components, which must be volatile, is very important for obtaining high-quality electrospray spectra. Even though sample introduction is through a liquid medium, the solvent eventually needs to be volatilized. If a nonvolatile solvent or buffer is used, the result is frequent plugging of the capillary tubing and/or some of the beam-defining components. In ESI the solvated sample is passed through a needle held at a high potential (3-10 kV). As the molecule exits from the needle, the resulting spray undergoes electrostatic nebulization, which places (a) charge(s) on the droplet. The charged droplet passes through a variety of focusing elements, which are differentially pumped. One result is desolvation of the droplet. Depending on the size and chemical makeup of the analyte, the resulting stable ion can have a single charge, or depending on the repulsive forces, may be multiply charged. Mass analysis of this ion can be carried out with any type of mass analyzer, including magnetic sector, quadrupole, ion trap or time-of-flight. Sample concentrations that can routinely be analyzed are sub-picomole/microliter. The detection limit for many components, however, is considerably lower. Table 5-1 gives the results of the ms/ms analysis of a set of betaine analogs. Notice the structurally diagnostic fragment ions available for each of these individual betaines (Wood et al., 2002).

2.4 Matrix-assisted laser desorption ionization

MALDI typically utilizes a nitrogen laser at 337 nm as the ionization source. The sample is mixed with a matrix, and allowed to dry prior to insertion into the mass spectrometer. Crystallization of the sample within the matrix is an important component of successful MALDI analysis. A variety of matrices, present in great excess relative to the sample amount, are used to span the range of compound classes amenable to MALDI mass analysis. Formation of sample ions, upon laser irradiation, involves a proton transfer reaction involving the matrix (which absorbs the UV photon) and the analyte. The ions are then accelerated into a time-of-flight mass analyzer for mass analysis. Typical matrices include α-cyano-4-hydroxycinnamic acid, sinapinic acid and 2,5-dihydroxybenzoic acid. MALDI can be done routinely to m/z 100,000 and there are many examples of analyses going well above this mass range. Figure 5-5A shows a MALDI mass spectrum of grape anthocyanins over the mass range m/z 450 to 550 (Sugui et al., 1999).

Two series of ions are evident, representing monoglucosides (1-5) and acetylglucosides (6-10) of the following aglycones: cyanidin (1, 6; **1.40**), peonidin (2, 7; **1.41**), delphinidin (3, 8; **1.42**), petunidin (4, 9; **1.43**) and malvidin (5, 10; **1.44**). These results were obtained with a first-generation (continuous mode) MALDI mass spectrometer that did not have delayed

extraction capabilities. Figure 5-5B shows a similar mass range for the analysis of anthocyanins from a different grape variety, with improved mass resolution using a MALDI mass spectrometer having delayed extraction capabilities. Note the presence of the monoglucosides of cyanidin (m/z 449), peonidin (m/z 463), petunidin (m/z 479) and malvidin (m/z 493). Both MALDI mass spectra were obtained in the reflector mode.

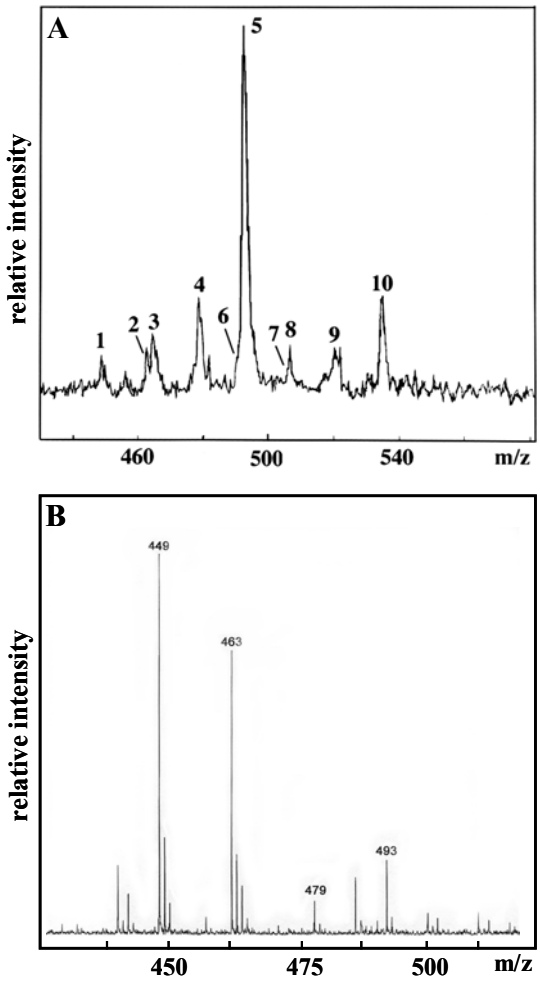


Figure 5-5. **A.** MALDI mass spectrum of the anthocyanin pigments in the grape variety Marechal Foch. This figure is from the article 'Matrix-assisted laser desorption ionization mass spectrometry analysis of grape anthocyanins', Am. J. Enol. Vitic. 50:199-203 by Sugui, J. A., Wood, K. V., Yang, Z., Bonham, C. C. and Nicholson, R. L. 1999. Reprinted by permission of the American Society for Enology and Viticulture. **B.** MALDI mass spectrum of grape anthocyanins acquired on a MALDI mass spectrometer with delayed extraction capabilities. Peak identities are discussed in the text.

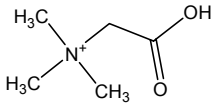
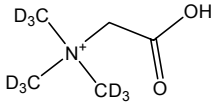
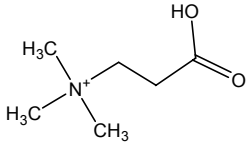
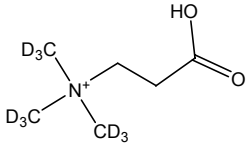
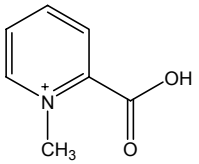
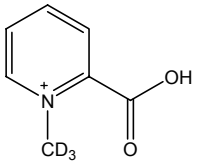
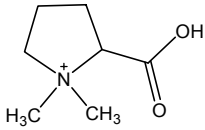
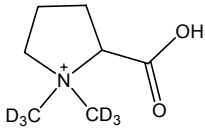
MALDI is considered the most sensitive of these techniques, with detection limits in the femtomoles/microliter range being relatively behind these other two, and this probably explains the general lack of continued interest in them in recent years. Of course statements concerning the relative sensitivities of the four ionization techniques have to take into account the chemical differences between compound classes, which plays a major part in the ionization of a given compound by a given ionization technique.

3. QUANTITATION

In addition to providing molecular weight confirmation and structural information, mass spectrometry can also provide quantitative results (Hoffmann and Stoobant, 2001; Watson, 1997). This is most often done using either a calibration curve (external calibration) or an internal standard, like an isotopically labeled analog or a compound with a closely related structure (internal calibration). Standard addition, the addition of known amounts of the compound of interest to the unknown sample is another frequently used quantitative method. Mass spectrometry is a very universal and sensitive analytical technique. As the methodology has matured over the years it has been used to solve a myriad of far-ranging analytical problems including those involving complex plant phenolics such as anthocyanins.

The data presented in Table 5-1 (next page) were reprinted from *Phytochemistry* **59**: Wood K. V., Bonham, C. C., Miles, D., Rothwell, A. P., Peel, G., Wood, B. C., and Rhodes, D., Characterization of betaines using electrospray MS/MS, p. 759-765, copyright 2002, with permission from Elsevier.

Table 5-1. Electrospray MS/MS spectra of a series of betaine analogs. The product ions in bold face are the base peaks of the MS/MS spectra.

Compound	Structure	Parention [M] (m/z)	Productions (m/z)
glycine-betaine		118	**59** [M-CH_2CO_2H] 58 [M-CH_3CO_2H]
d_9-glycine-betaine		127	**68** [M-CH_2CO_2H] 66 [M-CH_2DCO_2H]
β-alanine-betaine		132	**60** [M-H_2CHCO_2H] 73 [M-$N(CH_3)_3$]
d_9-β-alanine-betaine		141	**69** [M-H_2CHCO_2H] 73 [M-$N(CD_3)_3$]
Trigonelline		138	**94** [M-CO_2] 110 [M-CO]
d_3-trigonelline		141	**97** [M-CO_2] 113 [M-CO]
proline-betaine		144	**84** [M-C_3H_6-H_2O] 102 [M-C_3H_6] 98 [M–HCO_2H] 58 $[CH_2N(CH_3)_2]^+$
d_6-proline-betaine		150	**90** [M-C_3H_6-H_2O] 108 [M-C_3H_6] 104 [M–HCO_2H] 64 $[CH_2N(CD_3)_2]^+$

4. REFERENCES

Busch, K. L., Glish, G. L. and McLuckey, S. A., 1988, *Mass Spectrometry/Mass Spectrometry: Techniques and Applications of Tandem Mass Spectrometry*, VCH Publishers, New York.

Cole, R. B., 2000, Some tenets pertaining to electrospray ionization mass spectrometry, *J. Mass Spectrom*. **35**: 763-772.

Fenselau, C., 1983, Fast atom bombardment (review), in *Ion Formation from Organic Solids: Proceedings of the Second International Conference, Springer Series in Chem. Phys Vol.25.*, A. Benninghoven, ed., Springer-Verlag New York, pp 90-100.

Sugui, J. A., Wood, K. V., Yang, Z., Bonham, C. C. and Nicholson, R. L., 1999, Matrix-assisted laser desorption ionization mass spectrometry analysis of grape anthocyanins, *Am. J. Enol. Vitic*. **50**: 199-203.

Sundqvist, B. and Macfarlane, R. D., 1985, ^{252}Cf-plasma desorption mass spectrometry, *Mass Spectrom. Rev*. **4**: 421-460.

Wood, K. V., Bonham, C. C., Miles, D., Rothwell, A. P., Peel, G., Wood, B. C. and Rhodes, D., 2002, Characterization of betaines using electrospray MS/MS, *Phytochem*. **59**: 759-765.

Wood, K. V., Bonham, C. C., Ng, J., Hipskind, J. and Nicholson, R. L., 1993, Plasma desorption mass spectrometry of anthocyanidins, *Rap. Comm. Mass Spectrom*. **7**: 400-403.

Zenobi, R. and Knochenmuss, R., 1998, Ion formation in MALDI mass spectrometry, *Mass Spectrom. Rev.* **17**: 337-366.

For further reading we suggest several excellent books on the specifics of mass spectrometry:

Herbert, C. G. and Johnstone, R. A. W., 2003, *Mass Spectrometry Basics*, CRC Press, Boca Raton, FL.

Hoffmann, E. de and Stroobant, V., 2001, *Mass Spectrometry: Principles and Applications*, John Wiley, New York.

Watson, J. T., 1997, *Introduction to Mass Spectrometry,* 3rd Edition, Lippincott-Raven, Philadelphia.

Chapter 6

THE ROLE OF PHENOLS IN PLANT DEFENSE

1. PREFORMED ANTIMICROBIAL AND INSECTICIDAL METABOLITES

When considering substances produced by plants that act as agents that protect the plant from pathogens and insect pests, we must first consider whether the compounds are present prior to the time of infection or whether they are synthesized in response to infection. When compounds are present prior to attempted infection they are known as *preformed antimicrobial metabolites*. Such preformed compounds are part of a *passive* resistance mechanism. In general, such preformed metabolites are toxic to a broad spectrum of fungi and bacteria, but the compounds have a relatively low level of toxicity. Thus, preformed compounds are present in all plant species and help plants to ward off pathogens that are not considered as highly aggressive organisms. They are also referred to as *phytoanticipins* (Van Etten et al., 1994).

When one considers resistance expression in plants, it is necessary to consider whether resistance expression is part of a passive or active response system. There are several situations that could arise:

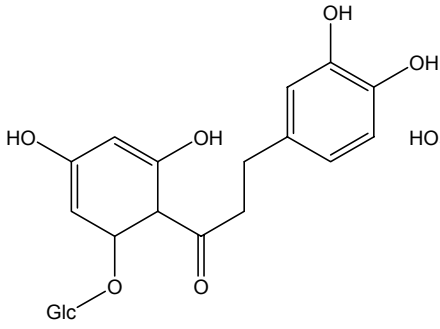
1. Compound "A" is present in the plant and is toxic to the potential pathogen. The compound is present in cells or tissues that the pathogen must come into contact with, at some time during the attempted infection. The compound is not further metabolized, but may or may not be changed by the pathogen, and the compound as such then accounts

for toxicity to the pathogen. This example represents a form of *passive* resistance.

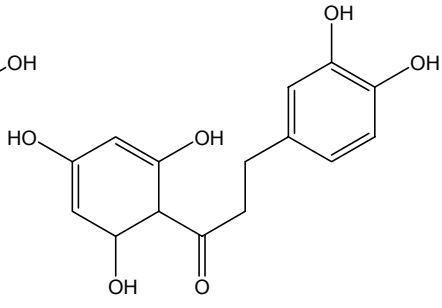
2. A preformed substance is degraded or metabolized to a different compound by the host in direct response to the pathogen and it is this compound that accounts in part for toxicity to the pathogen. Because the host changes the compound, this would be considered a mechanism of *active* resistance.

There are several criteria that must be satisfied before it can be decided whether a particular compound plays a significant role in the resistance to a pathogen. These are as follows:

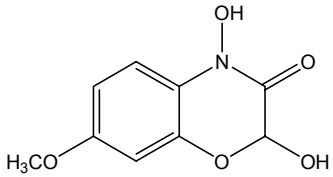
1. The compound must be present in those parts of the plant where the pathogen will come in contact with it. For example, apple leaves contain phloridzin (**6.1**) and its aglycone phloretin (**6.2**), but these compounds are not present in the fruit.
2. The compound must be present at concentrations high enough to affect the pathogen. For example, maize contains the preformed cyclic hydroxamic acid DIMBOA (2,4-dihydroxy-7-methoxy-1,4-benzoxazin-3-one; **6.3**). The concentration of DIMBOA, however, decreases over time and is initially not high enough for it to serve as a fungi-toxic agent.
3. Directly related to the criterion above is the requirement that the compound be "available" in the host in a form or place where it can express its toxicity. This may ultimately be a question of whether the compound is changed when it is extracted. Is the compound compartmentalized in a cell and possibly inaccessible to the pathogen? Is it in only one tissue?
4. Another criterion that must be met deals with the time of appearance of the compound. Thus, even if preformed, the compound must be at a sufficient *in vivo* concentration by the time resistance is being expressed.



(6.1)



(6.2)



(6.3)

Phenolic substances are often thought of when referring to preformed rsistance compounds. The following classes of metabolites, however, should also be included: alkaloids, carbohydrates that can bind bacteria, proteins that act as lectins, amino acids, terpenoids, and polyacetylenes. Some of these compounds will be included in the discussion below.

1.1 Chlorogenic acid

Chlorogenic acid (**6.4**) is an example of a preformed compound that has a relatively low level of toxicity to most microorganisms. It is effective against microorganisms considered as weak pathogens of potato. In potato tubers, chlorogenic acid is present in the periderm (Kojima et al., 1985) and is toxic to the organism that causes potato scab, *Streptomyces scabies*. In general, there is more of the compound in tubers of cultivars that are resistant to the pathogen than in tubers of cultivars that are susceptible. Furthermore, more of the compound is present during the time of tuber expansion. Chlorogenic acid also affects the growth of the vascular pathogen *Verticillium albo-atrum* and is present in the vascular tissue of the potato (Dao and Friedman, 1994). The mechanism by which *ortho*-dihydroxy phenolic compounds such as chlorogenic acid provide defense against insect pests was studied by Felton et al. (1989). They investigated the fate of chlorogenic acid in tomato (*Lycopersicon esculentum*) on the feeding behavior of the tomato fruit worm (*Heliothis zea*) and beet army worm (*Spodoptera exigua*). Upon feeding by the insects, polyphenol oxidases that are compartmentally separated from chlorogenic acid in the plant, come in contact with their substrate and convert chlorogenic acid to the toxic chlorogenoquinone (**6.5;** Figure 6-1). This quinone is a highly reactive electrophile and will react with nucleophilic –SH and $–NH_2$ moieties in proteins (as indicated by structure **6.6;** Matheis and Whitaker, 1984). This results in the cross-linking of proteins with chlorogenic acid,, which reduces the availability of free amino acids and proteins to the insect.

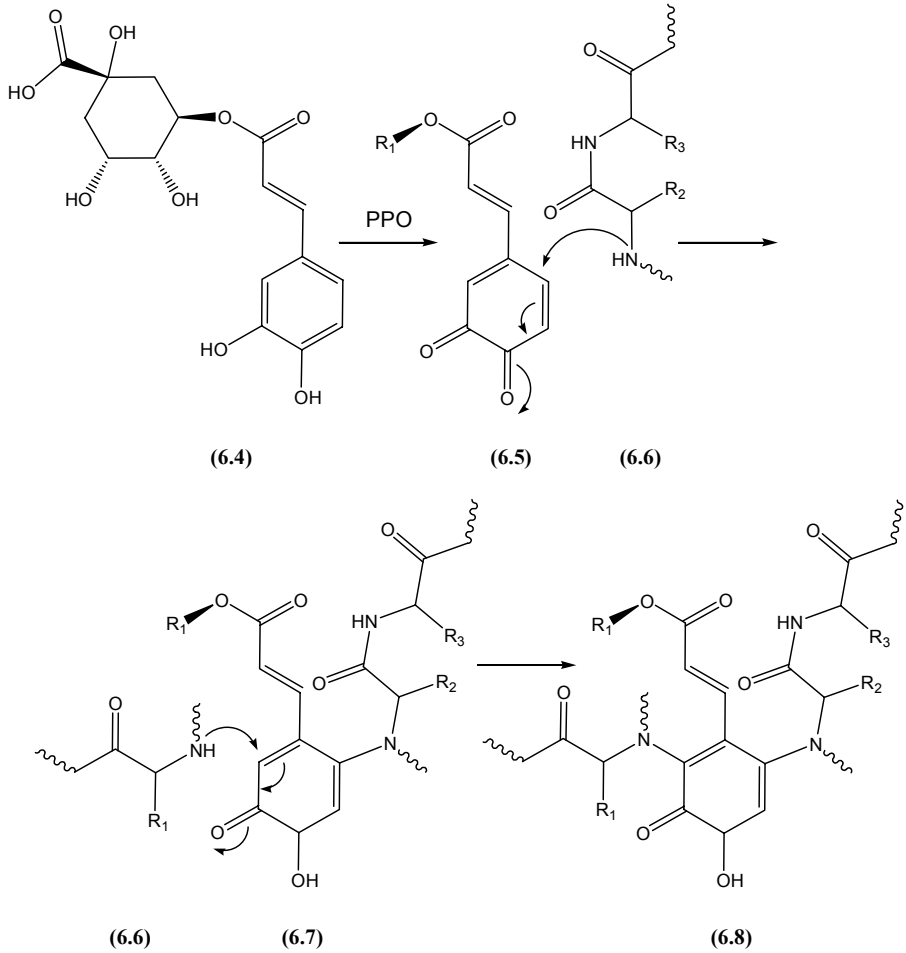


Figure 6-1. Oxidation of chlorogenic acid (**6.4**) by polyphenoloxidase (PPO), resulting in chlorogenoquinone (**6.5**), which can react with nucleophilic groups in proteins (**6.6**) to give the cross-linked compound **6.7**, which can react with another protein molecule to yield **6.8**. The quinate residue in structures **6.5**, **6.7** and **6.8** is represented by R, whereas R_1 and R_2 indicate different amino acid residues.

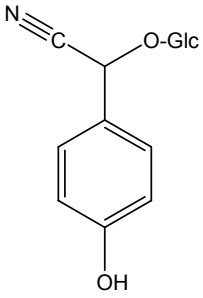
1.2 Phloridzin and phloretin

In apple, the glycoside phloridzin (**1.27**) and its aglycone phloretin are thought to inhibit the apple scab fungus *Venturia inaequalis.* Phloridzin is also an *ortho*-dihydroxyphenolic compound, and, like chlorogenic acid, can also be easily converted to a reactive quinone upon attack by a pathogen. Raa (1968) demonstrated that oxidation products of phloridzin inhibit fungal

pectinases. Fungal pectinases hydrolyze pectin, a cell wall compound that is abundant in the middle lamella and plays a role in cell adhesion. Thus, by inhbiting pectinases, the ability of the fungus to hydrolyze and invade the plant cell wall would be compromised. Although phloridzin and phloretin are toxic at high concentrations, based on the poor correlation between resistance to scab and the concentration of phenolic compounds such as phloridzin, they are probably not the factors that account for the actual resistance of apple cultivars to this fungus (Nicholson and Rahe, 2004). This was further substantiated by Leser and Treutter (2005), who investigated the effect of nitrogen supply on the scab susceptibility of the susceptible apple cultivar 'Golden Delicious' and the resistant cultivar 'Rewena'. Increased nitrogen supply was hypothesized to stimulate growth and decrease the levels of phenolic compounds. This was shown to indeed be the case. Consistent with this hypothesis, the susceptible cultivar became more susceptible under high nitrogen supplements. The resistant cultivar, however, did not become susceptible, even though the levels of phenolic compounds, including phloridzin, decreased.

1.3 Cyanogenic glycosides

Many plants contain cyanogenic glycosides. Toxicity of the cyanogenic glycoside results when the compound is enzymatically cleaved to release hydrogen cyanide (HCN) that is toxic to the pathogen. Sorghum contains the cyanogenic glycoside dhurrin (**6.9**). This compound is of interest to both pathologists and entomologists as an example of a preformed resistance compound and acts as an insect feeding deterrent and as a fungicidal agent (Starr et al., 1984; Adewusi, 1990).



(6.9)

1.4 Tuliposides

Tulips contain preformed compounds known as tuliposides, (tuliposide A, **6.10**, and tuliposide B, **6.11**). Hydrolysis of the tuliposides results in formation of aglycones (**6.12**) and (**6.13**) which will form butyrolactones (**6.14**) and (**6.15**). These lactone forms are inhibitory to fungi.

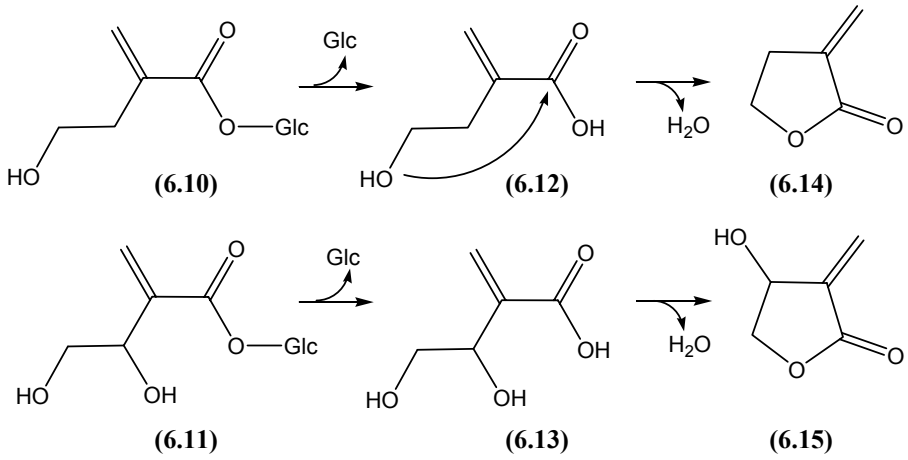
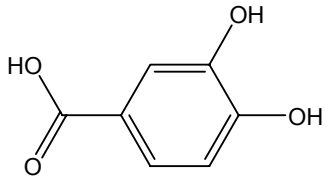


Figure 6-2. Formation of antifungal butyrolactones from tuliposides *via* an internal esterification reaction.

Some fungi are able to metabolize the lactone form to a butyric acid form that is not inhibitory to a variety of organisms, including the tulip pathogen *Fusarium oxysporum*. Two important diseases of tulip are bulb rot and grey mold. *Botrytis tulipae* infects all parts of the tulip, including the pistils, macerating the tissue. In contrast, *Botrytis cinerea* does not develop on tulips in the field, but eventually infects various parts of the plant when kept in a high humidity chamber. *B. cinerea* never infects the flower pistils because they contain exceptionally high concentrations of tuliposides. Tuliposides are stored in cell vacuoles. Importantly, growth of the pathogen is at first intercellular. Under these conditions the pathogen does not encounter the tuliposides. It is only when the tuliposides are released from the vacuoles that the effects of their toxicity can be expressed. *B. cinerea* converts tuliposides into inhibitory lactones. *B. tulipae* converts tuliposides into hydroxycarboxylic acids which are non-toxic (Schönbeck and Schroeder, 1972).

1.5 Protocatechuic acid

Probably the most commonly referred to compound which accounts for a form of passive resistance of a chemical nature is protocatechuic acid (**6.16**) which is found in yellow and red skinned onions and prevents the germination of spores of the onion smudge fungus, *Colletotrichum circinans*. Thus, protocatechuic acid serves as a barrier to infection prior to penetration. Once penetration occurs the compound is ineffective.



(6.16)

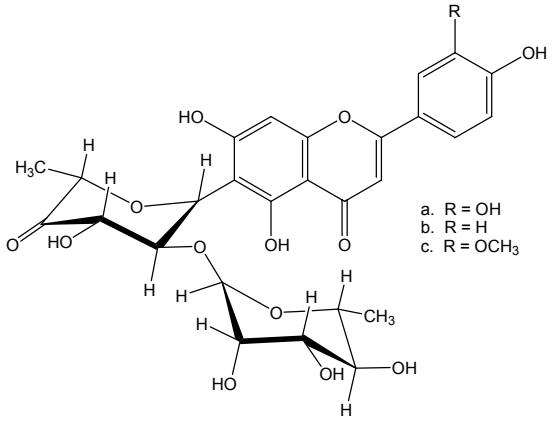
1.6 Lignin

The cell wall polymer lignin (see Chapter 1, Section 3.12, and Chapter 3, Section 12) has been implicated in defense against pests and pathogens, as a preformed, passive defense compound. It is a suitable defense compound in that respect, because it hardens the cell wall and thus creates a physical barrier against invasion. In addition, the chemical structure of lignin is very complex because of the different lignin subunits, and the many different types of chemical bonds that exist between subunits. Hydrolysis would, therefore, require a battery of enzymes, which is something most pests and pathogens do not have access to (Vance et al., 1980; Denton, 1998). Exceptions are white rot fungi, which can oxidize lignin in order to degrade it (Chen and Chang, 1985). This is how fallen tree logs eventually disintegrate. There is evidence that lignin can also be synthesized *de novo*. This lignin is synthesized locally, and specifically in response to pathogenic attack. This mechanism resembles papilla formation (Cadenagomez and Nicholson, 1987) and will be discussed in the next section.

1.7 C-glycosyl flavones

Maysin (2′′-*O*-α-l-rhamnosyl-6-C-(6-deoxy*xylo*-hexos-4-ulosyl)-luteolin; **6.17a**), apimaysin (**6.17b**) and methoxymaysin (**6.17c**) are C-glycosyl flavones that confer resistance against the corn earworm (*Helicoverpa zea* (Boddie)), a major silk- and kernel-feeding insect pest in the United States.

These three compounds differ from each other in the substitution pattern of the B-ring: apimaysin has one hydroxyl group on the 4′ position of the B-ring, maysin has a 3′,4′-dihydroxy substitution pattern, and methoxymaysin has a 4′-hydroxy, 3′-methoxy substitution pattern. These compounds accumulate in the silks of maize (*i.e.* the styles attached to the ovules) and are thought to act in a manner similar to chlorogenic acid (see Figure 6-1) when insects damage the silks.



(6.17)

Maysin is generally the most abundant of these three compounds, and is typically present at concentrations of 0.3% fresh silk weight, which is very high for a single compound. As a consequence, the C-glycosyl flavones can be considered preformed defense compounds.

The concentration of C-glycosyl flavones varies considerably among different maize lines. Since the concentration of the C-glycosyl flavones does not have a discrete value, but rather varies along a continuum, it can be considered a quantitative trait (see Chapter 3, Section 3.4). In order to identify loci controlling maysin concentration in silks, Byrne et al. (1996) investigated the role of a number of structural and regulatory genes known to play a role in flavonoid biosynthesis. They generated an F2 population derived from the maize inbred lines GT119 and GT114, which had low (0.031%) and high (0.56%) maysin levels, respectively. They determined the genotype at a number of loci at or near flavonoid biosynthetic genes, and concluded that the *P1* ('P-one'; see Chapter 3, Section 9.2) and a locus referred to as *recessive enhancer of maysin* (*rem1*) near the *Brown pericarp1* (*Bp1*) locus accounted for 58 and 11% of the variance, respectively. In addition, a QTL near the centromere of chromosome 1 was uncovered, but there were no obvious candidate genes at this locus.

The effects of the *P1* gene and the tightly linked homolog *P2* on maysin biosynthesis were further investigated by Zhang et al. (2003). These researchers used introduced the *P1* and *P2* cDNA's under control of the ubiquitin promoter in cultured Black Mexican Sweet maize cells using microprojectile bombardment. This is a method in which small gold or tungsten particles coated with an expression construct are introduced into target cells using a burst of pressure. Based on gene expression studies, P1 and P2 activate the same genes, including *phenylalanine ammonia lyase* (Chapter 3, Section 7), *chalcone synthase*, *chalcone isomerase*, and *dihydroflavonol 4-reductase* (Chapter 3, Section 9), but not genes involved in the biosynthesis of flavonols and anthocyanins. Increased levels of flavones were detected in extracts obtained from the transformed cells. Further evidence for a role of *both* genes in maysin biosynthesis came from the observation that maize plants in which both *P1* and *P2* were deleted did not produce maysin, whereas plants in which *P1* was deleted, but *P2* was still present, still produced maysin, albeit at reduced levels.

Maysin is two times more effective in its ability to inhibit growth of the corn earworm, which is attributed to the fact that two neighboring hydroxyl groups (such as on the on the 3′ and the 4′ positions in maysin) will result in the efficient formation of a toxic quinone, whereas the quinone formation from apimaysin and methoxymaysin is less efficient (Elliger et al., 1980; Snook et al., 1994). The genetic basis of the substitution reactions of the B-ring have been the subject of several studies. Using an F2 population derived from the maize inbred lines GT114 (moderately high levels of maysin, negligible levels of apimaysin) and NC7A (moderately high levels of apimaysin, maysin, and chlorogenic acid (**6.4**), Lee et al. (1998) determined that the *rem1* locus identified by Byrne et al. (1996) explained 55% of the variance for maysin, whereas a QTL that mapped near the *Pr1* gene, which is thought to encode flavonoid 3′ hydroxylase (F3′H), explained 65% of the variance for apimaysin. Furthermore, the levels of maysin and apimaysin were independent of each other, suggesting these two compounds are synthesized *via* different pathways. Surprisingly, a functional *Pr1* gene was not required for maysin production. Lee et al. (1998) speculated that the actual gene responsible for maysin biosynthesis may be near *Pr1*, but does not have to be *Pr1*, that a second *F3′H* gene is responsible for maysin production, or that the hydroxylation at C3′ occurs at the level of the hydroxycinnamoyl CoA ester rather than at the level of the flavone.

The genetic control of the substitution of the C-glycosyl flavones was investigated in further detail by Cortés-Cruz et al. (2003). Two F2 populations were generated from maize inbred lines that differed from each

other in the relative concentrations of maysin, apimaysin, and methoxymaysin. In both F2 populations the main QTL associated with levels of chlorogenic acid, maysin and methoxymaysin was located on the short arm of chromosome 4, whereas the main QTL associated with levels of apimaysin was located on the long arm of chromosome 5. Presence of a specific allele in the QTL on chromosome 4 resulted in higher levels of methoxymaysin and lower levels of maysin and chlorogenic acid. The fact that a single QTL affects the concentrations of three compounds (methoxymaysin, maysin and chlorogenic acid) suggests that there may be a regulatory gene underlying the QTL, or that there is a branched rather than a linear biosynthetic pathway leading to these different compounds. The QTL for apimaysin on chromosome 5 coincided with the *Pr1* locus, consistent with the data reported by Lee et al. (1998).

The biosynthetic pathway leading to maysin starts with flavanone (**6.18**), which is hydroxylated by flavone 3´ hydroxylase to yield di-hydroxyl flavanone (**6.19**) and is the reduced by flavone synthase to the flavone luteolin (**6.20**). The next steps were recently investigated in more detail by McMullen et al. (2004) using two *salmon silk* mutants, *sm1* (Anderson, 1921) and a newly discovered mutant *sm2*. These mutants have salmon colored silks instead of green silks as a result of pigment accumulation throughout the shaft of the silks, as opposed to only in the silk hairs, but do require a functional *P1* gene in order for the mutant phenotype to be apparent (see also Chapter 3, Section 9.2).

Detailed chemical analyses of flavone composition in the silks in wild-type, *sm1*, *sm2* and *sm1-sm2* plants revealed that isoorientin (**6.21**) is the only flavone accumulating in *sm1-sm2* double mutants, indicating the synthesis of this compound precedes the action of the gene products of the functional *Sm1* and *Sm2* genes. Isoorientin (**6.21**) is present at high levels in *sm2* but not *sm1* mutants, so that a functional *Sm2* gene is required for the formation of rhamnosyl-isoorientin (**6.22**) from isoorientin. The *Sm2* gene may encode a rhamnosyl transferase, or control the expression of a *rhamnosyl tranferase* gene. Rhamnosyl-isoorientin (**6.22**) accumulates in *sm1* mutants, suggesting that *Sm1* encodes a protein that catalyzes the formation of maysin (**6.17a**), or otherwise controls the expression of gene(s) encoding the necessary enzymes.

Taking all of the abovementioned data into consideration, the most likely biosynthetic pathway leading to maysin is as shown in *Figure 6-2*, although further research is needed to fully elucidate the pathway and the regulatory genes.

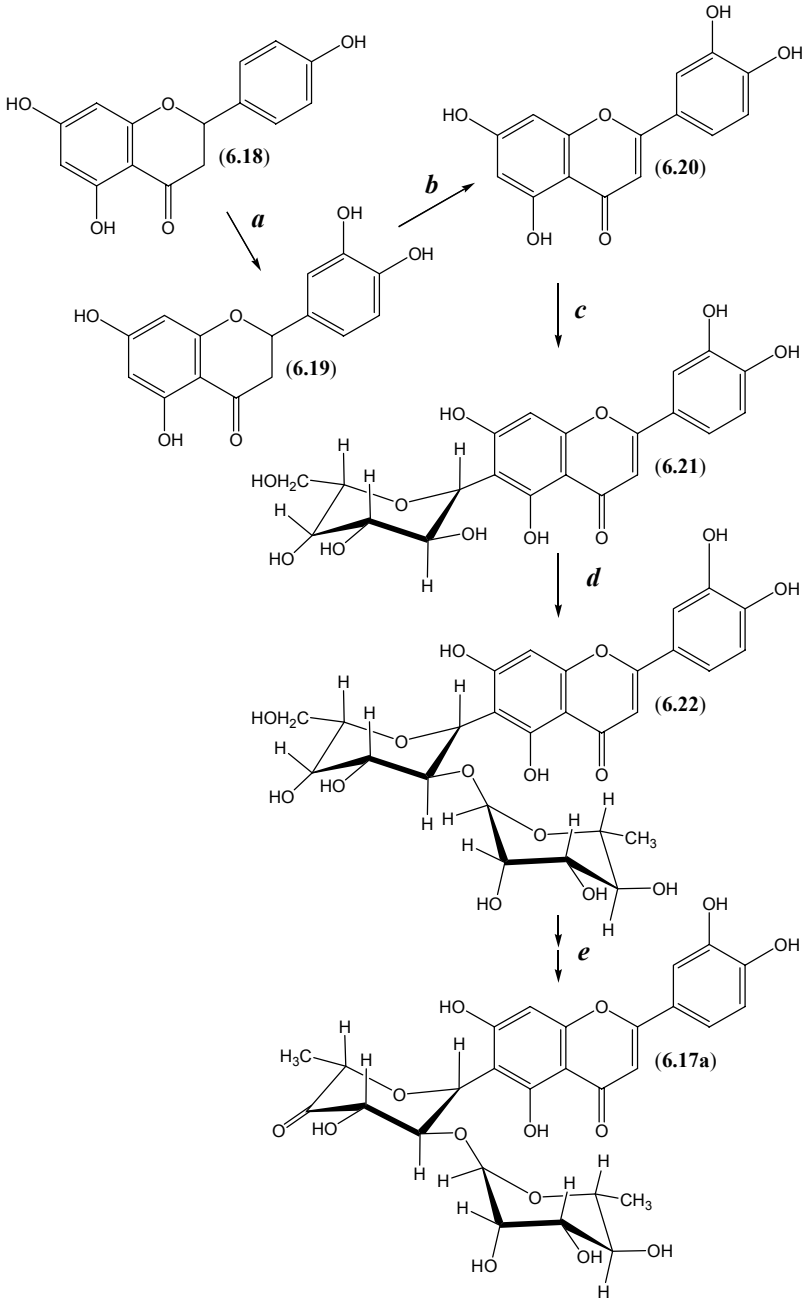


Figure 6-2. Biosynthesis of maysin proposed by McMullen et al. (2004) based on the analysis of flavones in the silks of maize *salmon silk* mutants. **a.** flavone 3′ hydroxylase (encoded by the maize *Pr1* gene), **b.** flavone synthase, **c.** *C*-glucosyltransferase, **d.** putative rhamnosyl transferase (encoded by the *Salmon silk2* gene), **e.** the step(s) controlled by the *Salmon silk1* gene.

2. COMPOUNDS FORMED IN RESPONSE TO PATHOGEN ATTACK

Compounds formed in response to stress may occur in at least two ways. In one response, the plant may form compounds throughout the tissue at a considerable distance from the infection site (Hammerschmidt, 1999). In another response, the plant may form compounds specifically at the infection site. This may include only a few cells and in rare cases, as few as one or two cells. (Snyder and Nicholson, 1990; Nicholson and Wood, 2001). In general, such compounds are referred to as either stress metabolites or more often as *phytoalexins*. By definition phytoalexins are formed in response to infection (Aguero et al., 2002; Lo et al., 2002; Hammerschmidt and Nicholson, 2001; Lo and Nicholson, 1998). Phytoalexins often exhibit toxicity to specific pathogens. In this case there is a genetic relationship between the expression of phytoalexin synthesis and the organism that induces that synthesis (Essenberg et al., 1985).

2.1 3-Deoxyanthocyanidins

The 3-deoxyanthocyanidins are a class of phytoalexins found in sorghum. These compounds are so fungi-toxic that they are effective at femtogram levels (Snyder and Nicholson, 1990; Nicholson and Wood, 2001). The synthesis of these compounds is initiated on the endoplasmic reticulum. Compounds are then trafficked in subcellular inclusions. The inclusions appear similar to vesicles, but there is no evidence that membranes surround the inclusions (Snyder and Nicholson, 1990). Nielsen et al. (2004) recently summarized this defense response. The cytological response commences when clear, colorless inclusions (less than 0.1 μm in diameter) accumulate in leaf cells under fungal attack. The inclusions eventually are seen as red bodies at the infection site. When the 3-deoxyanthocyanidins enter the apoplast, the host cell collapses. The phytoalexins then accumulate in the pathogen and cause its death. Excess phytoalexins are trapped in host cell walls at infection sites (Lo et al., 1998; 1999).

In the publication by Nielsen et al. (2004) images of the pigmented inclusions that contain the phytoalexins were prepared by confocal microscopy. This provided a three-dimensional perspective of inclusion body formation and visualization of the phytoalexins. A representation of deoxyanthocyanidin accumulation is shown in *Figure 6-3* where inclusions begin to form by 5 to 8 hours after a fungal appressorium was formed by a hypha.

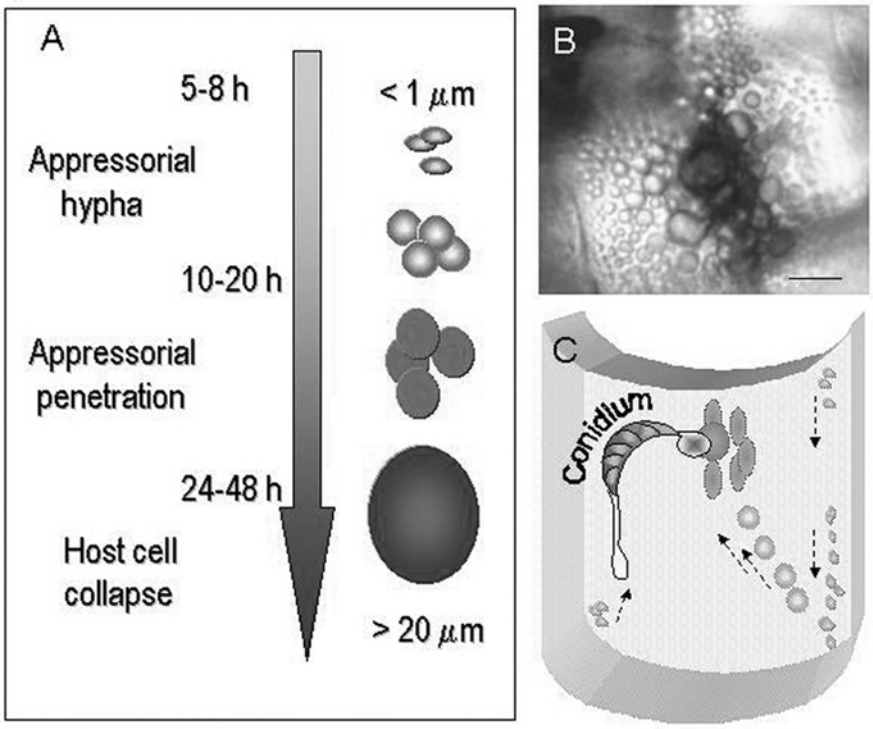


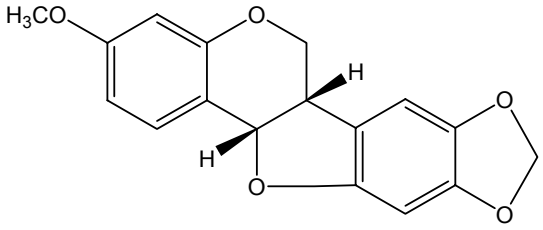
Figure 6-3. Cell-specific accumulation of 3-deoxyanthocyanidins in *Sorghum bicolor* in response to attempted fungal attack. (**A**) Illustration of changes in inclusion morphology in cells under fungal attack in response to formation of infectious structures 0–48 h after inoculation. (**B**) Site-specific accumulation of 3-deoxyanthocyanidins at site of incipient penetration, before host cell collapse 24 h after inoculation. (**C**) Illustration of site-specific trafficking (arrows) of inclusions in relation to position of fungal infectious structures. Reprinted from Phys. Mol. Plant Pathol., 65, Nielsen, K. A., Gottfredsen, C. H., Buch-Pedersen, M. J., Ammitzbøll, H., Mattsson, O., Duus, J. Ø., and Nicholson, R. L., Anti-microbial flavonoid 3-deoxyanthocyanidins in *Sorghum bicolor* self-organize into spherical structures, 187-196, Copyright 2004, with permission from Elsevier.

Note that the inclusions at this early time are not pigmented; rather they are colorless bodies that move through the cytoplasm toward the site of appressorial attack. Over time, the inclusions take on a yellow color and eventually become deep red in pigmentation. Inclusion size changes from less than 1 μm to 20 μm or even larger. Inclusions move to the penetration site and cluster in the area where the penetration peg has made physiological contact with the host cell. When the appressorium begins the process of penetration the inclusions burst, releasing their contents into the cytoplasm. The deoxyanthocyanidins kill the host cell itself *and* are taken up by the

pathogen. This is possible because these deoxyanthocyanidins are soluble in both water and organic solvents. In this manner the pathogen is also killed and prevented from causing extensive damage and cell death of the host.

2.2 Pisatin

Pisatin (**6.23**) is an isoflavonoid phytoalexin that is synthesized by pea (*Pisum sativum* L.) as a response to infection (Preisig et al., 1989). Subsequently, it was shown that pathogens capable of demethylating pisatin were tolerant of this phytoalexin. The enzyme responsible for demethylation is a specific cytochrome P450 mono-oxygenase released by the fungus *Nectria haematococca* (Delserone et al., 1999).

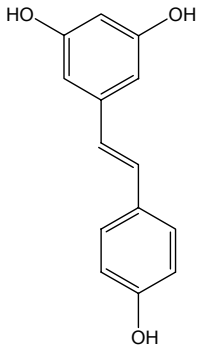


(6.23)

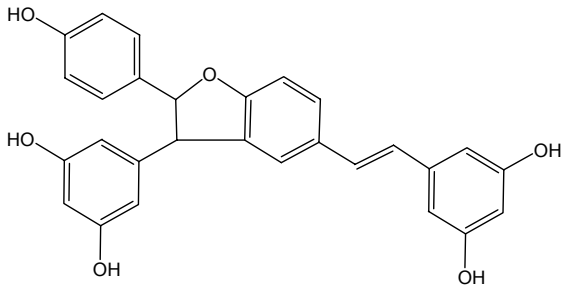
2.3 Stilbenes

Aside from being UV-protectants, in a number of species certain stilbenes act as phytoalexins. Resveratrol (**6.24**; *trans*-3,5,4′-trihydroxy-stilbene), its *cis*-isomer, as well as their glucosides and dehydrodimer *trans*-ε-viniferin (**6.25**) are present in grape leaves and berries and play a role in the defense against gray rot caused by the fungal pathogen *Botrytis cinerea*.

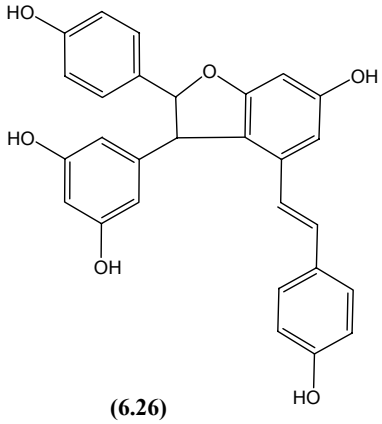
Viniferin is synthesized by a grape peroxidase (Morales et al., 1997). The fungus in return is able to inactivate resveratrol through the action of a laccase-like stilbene oxidase (Breuil et al., 1998). This results in the formation of resveratrol *trans*-dehydrodimer (**6.26**), as well as the corresponding *cis*-dehydrodimer (**6.27**), both of which structurally resemble viniferin.



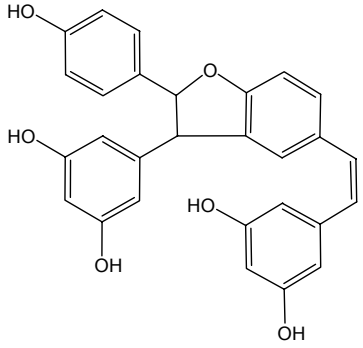
(6.24)



(6.25)



(6.26)



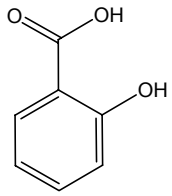
(6.27)

2.4 Salicylic acid

When a specific part of a plant is attacked by a fungal pathogen, distant parts of the plant may display an enhanced state of resistance involving the accumulation of pathogenesis-related (PR) proteins, a class of plant proteins that are normally not present but that are induced upon pathogen attack (reviewed by Van Loon and Van Strien, 1999). In addition, salicylic acid (SA; **6.28**) and hydrogen peroxide accumulate at the wound site and in other parts of the plant. This response is referred to as *systemic acquired resistance* (SAR) and is thought to be mediated by one or more signaling molecules in the phloem.

SA was hypothesized to be one of those signaling molecules, because **1**) SA accumulation was shown to be correlated with SAR and resistance (Uknes et al., 1993), **2**) exogenous SA applied to an uninfected plant induced SAR and resistance in a manner similar to that of an infected plant (Ward

et al., 1991), and **3**) transgenic plants expression the *nahG* gene from *Pseudomonas putida*, which encodes a salicilate hydroxylase, were unable to display SAR (Gaffney et al., 1993).



(6.28)

SA was hypothesized to be one of those signaling molecules, because **1**) SA accumulation was shown to be correlated with SAR and resistance (Uknes et al., 1993), **2**) exogenous SA applied to an uninfected plant induced SAR and resistance in a manner similar to that of an infected plant (Ward et al., 1991), and **3**) transgenic plants expression the *nahG* gene from *Pseudomonas putida*, which encodes a salicilate hydroxylase, were unable to display SAR (Gaffney et al., 1993). While this latter study demonstrated the role of SA in initiating SAR, it did not address whether SA was the signaling molecule that transmitted the SAR signal through the phloem to other parts of the plant. Vernooij et al. (1994) performed a series of elegant experiments to investigate the role of SA in signaling. They grafted a scion from a transgenic tobacco plant expressing the *NahG* gene onto the root stock of an untransformed tobacco plant. In addition, an untransformed scion was grafted onto a transgenic rootstock. Ungrafted plants and plants where the scion was grafted back on the rootstock from which it came were used as controls. The root stocks were inoculated with the viral pathogen tobacco mosaic virus (TMV) to induce SAR. The degree of SAR was evaluated by challenging the scions of the inoculated plants 7 days later with TMV or the fungal pathogen *Cercospora nicotianae*. In the untransformed graft control, the lesions induced by TMV were 41% smaller compared to a mock-inoculated control. This indicated that the SAR signal was not hampered by the graft. When the scions of the transgenic grafted plants were inoculated, the lesion size was the same as in the corresponding mock-inoculated control. This indicated that the expression of the *NahG* gene prevented SAR, as had been shown by Gaffney et al. (1993). The TMV-inoculated transgenic scions on untransformed root stocks behaved similarly as the mock-inoculated controls, indicating that SA was required to induce SAR in the scions. When untransformed scions on transgenic root stocks were inoculated with TMV, however, they displayed SAR. This reveals that the *NahG*-expressing tissues were able to transmit the signal required for

SAR. Similar results were obtained when the SAR response was induced by *C. nicotianae*, suggesting that SAR is a response to a broad range of pathogens. These experiments thus demonstrated that SA is required for the induction of SAR, but that it is not the actual signaling molecule.

After an additional 10 years of research it is still not entirely clear what the signaling molecule is. Van Bel and Gaupels (2004) recently reviewed the possible signaling molecules that could induce SAR. The list includes jasmonic acid, lipid-derived molecules, reactive oxygen species (see Chapter 2, Section 1.9), oligosaccharides, mRNA molecules, calcium, and various peptides.

2.5 Lignin

There is evidence that lignin can be synthesized *de novo*. This lignin is synthesized locally, and specifically in response to pathogen attack. There is new evidence that this lignin requires different biosynthetic enzymes, which results in a different subunit composition than the lignin of the vascular tissue.

Wheat cultivars resistant to *Puccinia recondita* f. sp. *tritici*, a fungal pathogen causing leaf rust, were shown to accumulate more lignin than susceptible cultivars, based on histochemical stains and a quantitative assay to detect total phenolics (Southerton and Deverall, 1990). Similar results were reported for the response of wheat to infection with *Fusarium graminearum*, which causes Fusarium head blight. In this case immunogold labeling against lignin was used to evaluate the accumulation of lignin in inoculated and non-inoculated spikes of a resistant and susceptible cultivar. Labeling densities were significantly higher in inoculated spikes of the resistant cultivar, compared to either non-inoculated spikes and inoculated spikes of the susceptible cultivar (Kang and Buchenauer, 2000).

Several studies have focused on the role lignin biosynthetic enzymes play in response to pathogenic attack. Moerschbacher et al. (1990) used specific inhibitors of the enzymes phenylalanine ammonia lyase (PAL) and cinnamyl alcohol dehydrogenase (CAD). These inhibitors were applied to wheat cultivars highly resistant to stem rust (*Puccinia graminis* Pers f.sp. *tritici* Erics. & E. Henn.). Regardless of the inhibitor that was used, *de novo* lignification was decreased and fungal development increased. Thus, a strict correlation between resistance and lignification was demonstrated.

Lignification in response to infection has been found to be associated by an increase activity of several enzymes of the lignin biosynthetic pathway. When a wheat lines carrying the stem rust resistance gene *Sr5* was compared with a near-isogenicline without this resistance gene, different patterns of enzymatic activities were observed (Moerschbacher et al., 1988). Both lines had an early activation of the enzymes PAL, 4-coumarate:CoA ligase (4CL), and CAD, but only the resistant genotype exhibited a *second* significant increase at the time of the hypersensitive response. Similarly, Mitchell et al. (1994) measured *p*-coumaryl, coniferyl and sinapyl alcohol dehydrogenase activity in lignifying leaves of wheat. Lignification induced by wounding or elicitors was found to be specifically associated with an increase in sinapyl alcohol dehydrogenase activity, which is expected to result in a lignin rich in syringyl units. In a subsequent study Mitchell et al. (1999) investigated the role of CAD in the defense response of wheat, mimicked by wounding or the application of the elicitors chitosan and chitin. Three major forms of CAD were identified, but only one, CAD-C, was found to be induced during lignification at the wound margin. This particular form had a substrate preference for sinapyl alcohol. Thus, in agreement with the previous study, *de novo* lignin could contain a high level of syringyl units.

Deborah et al. (2001) studied PAL and peroxidase (PO) activity as well as accumulation of lignin in response to inoculation of rice leaf sheaths with a pathogen and a non-pathogen. Infection with the non-pathogen *Pestalotia palmarum* resulted in a stronger increase in PO and PAL activity and a higher accumulation of lignin than after inoculation with the pathogen *Rhizoctonia solani*.

Altogether, these results indicate that if lignin plays a role in the resistance against a pathogen, a coordinated activation of lignin biosynthetic enzymes is required.

Two recent studies, one in maize and one in Arabidopsis, provided evidence for the existence of multiple copies of the same gene and a divergence in their function. Pichon et al. (1998) cloned two different maize cDNA's encoding cinnamoyl-CoA reductase (CCR) and found that the corresponding genes are differentially expressed in different parts of the plant. A similar situation was observed in Arabidopsis (Lauvergeat et al., 2001). The expression of the two Arabidopisis genes was studied during plant development, and in response to infection with the pathogenic bacteria *Xanthomonas campestris pv. campestris*. The *AtCCR1* gene was mainly expressed in lignifying tissues during development. In contrast, the *AtCCR2* gene had a low expression level during development, but was induced when

the plant was challenged with the bacteria. With the availability of whole genome sequences, it is apparent that many of the lignin biosynthetic genes are respresented by multiple copies, some of which are likely to be involved in defense responses (Raes et al., 2003).

3. REFERENCES

Adewusi, S. R. A., 1990, Turnover of dhurrin in green sorghum seedlings, *Plant Physiol*. **94**: 1219-1224.

Aguero, M. E., Gevens, A., and Nicholson, R.L., 2002, Interaction of *Cochliobolus heterostrophus* with phytoalexin inclusions in *Sorghum bicolor*. *Physiol. Mol. Plant Pathol*. **61**: 267-271.

Anderson, E. G., 1921, The inheritance of salmon silk color in maize, *Cornell Univ. Agric. Exp. Stn. Memoir* **48**: 535–554.

Breuil A.-C., Adrian M., Pirio N., Meunier P., Bessis R., and Jeandet P., 1998, Metabolism of stilbene phytoalexins by *Botrytis cinerea*: 1. Characterization of a resveratrol dehydrodimer, *Tetrahed. Lett.* **39**: 537-540.

Byrne, P. F., McMullen, M. D., Snook, M. E., Musket, T. A., J. Theuri, M., Widstrom, N. W., Wiseman, B. R., and Coe, E. H., 1996, Quantitative trait loci and metabolic pathways: Genetic control of the concentration of maysin, a corn earworm resistance factor, in maize silks, *Proc. Natl. Acad. Sci. USA* **17**: 8820-8825.

Cadenagomez, G, and Nicholson, R. L., 1987, Papilla formation and associated peroxidase-activity – a nonspecific response to attempted fungal penetration of maize, *Phys. Mol. Plant Pathol.* **31**: 51-67.

Chen, C.-L., and Chang, H.-M., 1985, Chemistry of lignin biodegradation, in: *Biosynthesis and Biodegradation of Wood Components*, T. Higuchi, ed., Academic Press, Orlando, FL, pp. 535-556.

Cortés-Cruz, M., Snook, M., and McMullen, M. D., 2003, The genetic basis of C-glycosyl flavone B-ring modification in maize (*Zea mays* L.) silks, *Genome* **46**: 182-194.

Dao, L., and Friedman, M., 1994, Chlorophyll, chlorogenic acid, glycoalkaloid, and protease inhibitor content of fresh and green potatoes, *J. Agric. Food Chem.* **42**: 633-639.

Deborah, S. D., Palaniswami, A., Vidhyasekaran, P., and Velazhahan, R., 2001, Time-course study of the induction of defense enzymes, phenolics and lignin in rice in response to infection by pathogen and non-pathogen, *J. Plant Dis. Protect.* **108**: 204-216.

Delserone, L. M., McCluskey, K., Matthews, D. E., and VanEtten, H. D., 1999, Pisatin demethylation by fungal pathogens and nonpathogens of pea: association with pisatin tolerance and virulence, *Physiol. Mol. Plant Pathol*. **55**: 317-326.

Denton, F.R., 1998, Beetle juice, *Science* **281**: 1285.

Elliger, C. A., Chan, B. G., Waiss, A. C. Jr., 1980, Flavonoids as larval growth inhibitors, *Phytochem*. **19**: 293-297.

Essenberg, M., Pierce, M., Shevell, J. L., Sun, T. J. and Richardson, P. E., 1985, Sesquiterpenoid phytoalexins and resistance of cotton to *Xanthomonas campestris* pv. *Malvacearum*, *Curr. Comm. Mol. Biol.* **152**: 145-149.

Felton, G. W., Donata, K., Del Vecchio, R. J., Duffey, S. S., 1989, Activation of plant foliar oxidases by insect feeding reduces nutritive quality of foliage for noctuid herbivores, *J. Chem. Ecol.* **15**: 2667-2694.

Gaffney, T., Friedrich, L., Vernooij, B., Negmtto, D., Nye, G., Uknes, S., Ward, E., Kessmann, H., and Ryals, J., 1993, Requirement of salicylic acid for the induction of systemic acquired resistance, *Science* **261**: 754-756.

Hammerschmidt, R., 1999, Phytoalexins: What have we learned after 60 years?, *Annu. Rev. Phytopathol*. **37**: 285-306.

Hammerschmidt, R., and Nicholson, R. L., 2001, A survey of plant defense responses to pathogens, in: *Inducible plant defenses against pathogens and herbivores: Biochemistry, ecology, and agriculture.* Am. Phytopathol. Soc., St. Paul, MN, pp. 57-71.

Kang, Z., and Buchenauer, H., 2000, Ultrastructural and immunochemical investigation of pathogen development and host responses in resistant and susceptible wheat spikes infected by *Fusarium culmorum*, *Physiol. Mol. Plant Pathol*. **57**: 255-268.

Kojima, M., and Kondo, T., 1985, An enzyme in sweet potato root which catalyzes the conversion of chlorogenic acid, 3-caffeoylquinic acid, to isochlorogenic acid, 3,5-dicaffeoylquinic acid, *Agric. Biol. Chem.* **49**: 2467-2469.

Lee, E. A., Byrne, P. F., McMullen, M. D., Snook, M. E., Wiseman, B. R., Widstrom, N. W., and Coe, E. H., 1998, Genetic mechanisms underlying apimaysin and maysin synthesis and corn earworm antibiosis in maize (*Zea mays* L.), *Genetics* **149**: 1997-2006.

Leser, C., and Treutter, D., 2005, Effects of nitrogen supply on growth, contents of phenolic compounds, and pathogen (scab) resistance of apple trees, *Physiol. Plant*. **123**: 49-56.

Lauvergeat, V., Lacomme, C., Lacombe, E., Lasserre, E., Roby, D., and Grima-Pettenati, J., 2001, Two cinnamoyl-CoA reductase (CCR) genes from *Arabidopsis thaliana* are differentially expressed during development and in response to infection with pathogenic bacteria, *Phytochem.* **57**: 1187-1195.

Lo, C., Coolbaugh, R. C., and Nicholson, R. L., 2002, Molecular characterization and *in silico* expression analysis of a chalcone synthase gene family in *Sorghum bicolor*, *Physiol. Mol. Plant Pathol*. **61**: 179-188.

Lo, S.-C. C., Hipskind, J. D., and Nicholson, R. L., 1999, cDNA cloning of a sorghum pathogenesis-related rotein (PR-10) and differential expression

of defense-related genes following inoculation with *Cochliobolus heterostrophus* or *Colletotrichum sublineolum, Mol. Plant Microb. Inter.* **12**: 479-489.

Lo, S. C. C., and Nicholson, R. L., 1998, Reduction of light-induced anthocyanin accumulation in inoculated sorghum mesocotyls: Implications for a compensatory role in the defense response. *Plant Physiol*. **116**: 979-989.

Matheis, G., and Whitaker, J. R., 1984, Modification of proteins by polyphenol oxidase and peroxidase and their products, *J. Food Biochem.* **8**: 137-162.

McMullen, M. D., Kross, H., Snook, M. E., Cortés-Cruz, M., Houchins, K. E., Musket, T. A., and Coe, E. H., Jr, 2004, *Salmon silk* genes contribute to the elucidation of the flavone pathway in maize (*Zea mays* L.), *J. Hered*. **95**: 225-233.

Mitchell, H.J., Hall, J.L., and Barber, M.S., 1994, Elicitor-induced cinnamyl alcohol dehydrogenase activity in lignifying wheat (*Triticum aestivum* L.) leaves, *Plant Physiol*. **104**: 551-556.

Mitchell, H.J., Hall, S.A., Stratford, R., Hall, J.L., and Barber, M.S. 1999. Differential induction of cinnamyl alcohol dehydrogenase during defensive lignification in wheat (*Triticum aestivum* L.): characterization of the major inducible form, *Planta* **208**: 31-37.

Moerschbacher, B.M., Noll, U.M., Flott, B.E., and Reisener, H.J., 1988, Lignin biosynthetic enzymes in stem rust infected resistance and susceptible near-isogenic wheat lines, *Physiol. Mol. Plant Pathol*. **33**: 33-46.

Moerschbacher, B.M., Noll, U.M., Gorrichon, L., and Reisener, H.J., 1990, Specific inhibition of lignification breaks hypersensitive resistance of wheat to stem rust, *Plant Physiol*. **93**: 465-470.

Morales, M, Alcántara, J., and Barceló, A. R. (1997) Oxidation of *trans*-resveratrol by a hypodermal peroxidase isoenzyme from Gama rouge grape (*Vitis vinifera*) berries, *Am. J. Enol. Vitic*. **48**: 33-38.

Nicholson, R. L., and Rahe, J. E., 2004, Apple scab and its management, in: *Fruit and vegetable diseases* Mukerji., K. G., ed., Kluwer Academic Publishers, Dordrecht, pp. 41-58.

Nicholson, R. L., and Wood, K. V., 2001, Phytoalexins and secondary products, where are they and how can we measure them?, *Physiol. Mol. Plant Pathol*. **59**: 63-69.

Nielsen, K. A., Gottfredsen, C. H., Buch-Pedersen, M. J., Ammitzbøll, H., Mattsson, O., Duus, J. Ø., and Nicholson, R. L., 2004, Anti-microbial flavonoid 3-deoxyanthocyanidins in *Sorghum bicolor* self-organize into spherical structures, *Phys. Mol. Plant Pathol*. **65**: 187-196.

Pichon, L. Courbou, I, Beckert, M, Boudet, A-M., and Grima-Pettenati, J., 1998, Cloning and characterization of two maize cDNAs encoding cinnamoyl-CoA reductase (CCR) and differential expression of the corresponding genes, *Plant Mol. Biol*. **38**: 671-676.

Preisig, C. L., Matthews, D. E., VanEtten, H. D.,1989, Purification and characterization of S-adenosyl-L-methionine: 6a-hydroxymaackiain 3-*O*-methyltransferase from *Pisum sativum, Plant Physiol*. **91**: 559-566.

Raa, J., 1968, Polyphenols and natural resistance of apple leaves against *Venturia inaequalis, Eur. J. Plant Path*.**74**: 37-45.

Raes, J., Rohde, A., Christensen, J. H., Van de Peer, Y., and Boerjan, W., 2003, Genome-wide characterization of the lignification toolbox in Arabidopsis, *Plant Physiol.* **133**: 1051-1071.

Schönbeck, F., and Schroeder, C., 1972, Role of antimicrobial substances (tuliposiodes) in tulips attacked by *Botrytis spp*., *Physiol. Plant Pathol.* **2**: 91-99.

Snook, M. E., Widstrom, N. W., Wiseman, B. R., Gueldner, R. C., Wilson, R. L., Himmelsbach, D. S., Harwood, J. S., Costello, C. E., 1994, New flavone C-glycosides from corn (*Zea mays* L.) for the control of the corn earworm (*Helicoverpa zea*), in: *Bio-Regulators For Crop Protection And Pest Control*, P. A. Hedin, ed., Symposium Series 557 of the American Chemical Society, Washington, DC, pp. 122-135.

Snyder, B. A. and Nicholson, R. L., 1990, Synthesis of phytoalexins in sorghum as a site-specific response to fungal ingress, *Science* **248**: 1637-1639.

Southerton, S. G., and Deverall, B. J., 1990, Histochemical and chemical evidence for lignin accumulation during expression of resistance to leaf rust fungi in wheat. *Physiol. Mol. Plant Pathol*. **36**: 483-494.

Starr, J.L., Newton, R. J., and Miller, F.R., 1984, Presence of dhurrin in sorghum root tissue and the effect of pathogenesis on hydrogen cyanide potential, *Crop Sci*. **24**: 739-742.

Southerton, S. G., and Deverall, B. J., 1990, Histochemical and chemical evidence for lignin accumulation during expression of resistance to leaf rust fungi in wheat, *Physiol. Mol. Plant Pathol*. **36**: 483-494.

Uknes, S., Winter, A., Delaney, T., Vernooij, B., Morse, A., Friedrich, L., Nye, G., Potter, S., Ward, E., and Ryals, J., 1993, Biological induction of systemic acquired resistance in Arabidopsis, *Mol. Plant-Microbe Interact.* **6**: 692-698.

Vance C.P, Kirk T.K. and Sherwood R.T., 1980, Lignification as a mechanism of disease resistance, *Annu. Rev. Phytopath*. **18**: 259-288.

Van Bel, A. J. E., and Gaupels, F., 2004, Pathogen-induced resistance and alarm signals in the phloem, *Mol. Plant Pathol*. **5**: 495-504.

VanEtten, H.D., Mansfield, J.W., Bailey, J.A., and Farmer, E., 1994, Two classes of plant antibiotics: Phytoalexins versus "phytoanticipins", *Plant Cell* **6**: 1191-1192.

Van Loon, L.C., and Van Strien, E.A., 1999, The families of pathogenesis-related proteins, their activities, and comparative analysis of PR-1 type proteins, *Physiol. Mol. Plant Pathol*. **55**: 85-97.

Vernooij, B., Friedrich, L., Morse, A., Reist, R., Kolditz-Jawhar, R., Ward, E., Uknes, S., Kessmann, H., and Ryals, J. (1994) Salicylic acid is not the translocated signal responsible for inducing systemic acquired resistance but is required in signal transduction, *Plant Cell* **6**: 959-965.

Ward, E.R., Uknes, S.J., Williams, S.C., Dincher, S.S., Wiederhold, D.L., Alexander, D.C., Ahl-Goy, P., Metraux, J.-P., and Ryals, J.A., 1991, Coordinate gene activity in response to agents that induce systemic acquired resistance, *Plant Cell* **3**: 1085-1094.

Zhang, P., Wang, Y., Zhang, J., Maddock, S., Snook, M., and Peterson, T., 2003, A maize QTL for silk maysin levels contains duplicated Myb-homologous genes which jointly regulate flavone biosynthesis, *Plant Mol. Biol*. **52**: 1-15.

Chapter 7

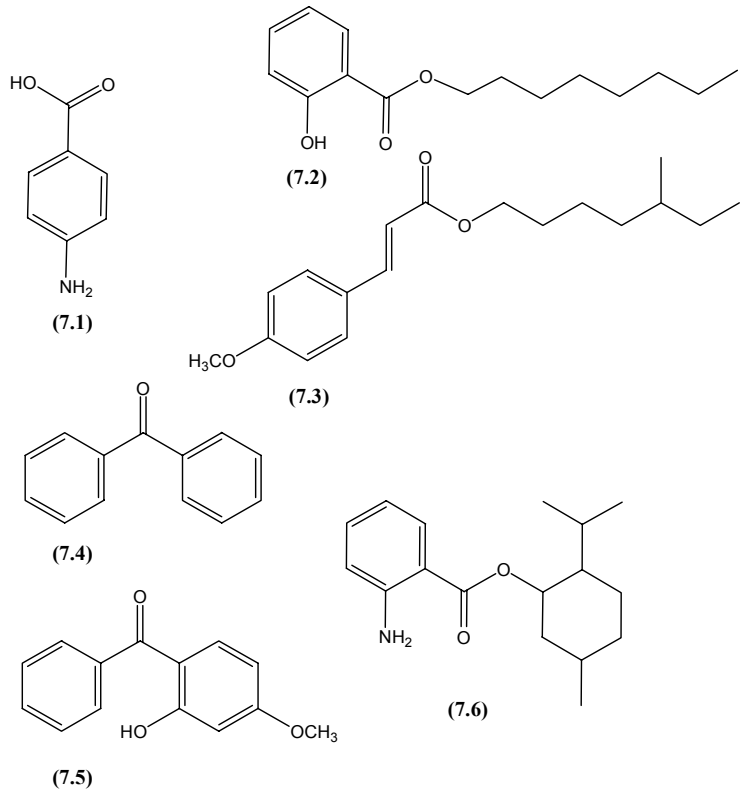
PHENOLIC COMPOUNDS AND THEIR EFFECTS ON HUMAN HEALTH

1. INTRODUCTION

There is a huge body of evidence that phenolic compounds have effects on human health, and that is the topic of this chapter. Perhaps the oldest medical application of phenolic compounds is the use of phenol (**1.1**) as an antiseptic. Because of its negative side effects on living tissues, including blister formation, especially at higher concentrations, it is no longer used in this capacity. The modern antiseptic agents effective against the bacterium *Staphylococcus aureus* are, however, still compared to a 5% (w/v) solution of phenol. Phenol is still used as an oral anesthetic in throat lozenges, at a typical concentration of 1.4%.

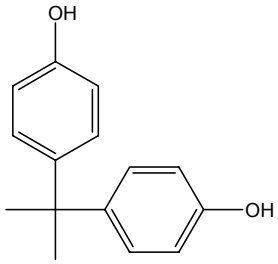
Another very common use of phenolic compounds is in sunscreens. The presence of the aromatic ring results in the effective absorbance of the UV-B radiation (between 280 and 315 nm) from the sun and thus prevents sunburns. The most common active ingredient in many sunscreens is *p*-aminobenzoic acid (PABA; **7.1**), which is actually not a phenolic compound. This compound has been widely used since the 1970's but is less popular nowadays due to the formation of skin rashes and acne. As a result many sunscreens are now PABA-free. Alternative active ingredients include salicylates such as octylsalicylate (**7.2**), cinnamates such as octyl methylcinnamate (**7.3**), benzophenone (**7.4**) and the related compound oxybenzone (**7.5**), and anthranilates, such as menthylanthranilate (**7.6**). Octyl methylcinnamate (**7.3**) is insoluble in water and is therefore commonly

used in water-proof sunscreens. The most recently developed sunscreens use titanium dioxide and/or zinc oxide which reflect the light, rather than absorb it, and are considered more effective. They work best in relatively thick layers, which is less desirable from a cosmetic perspective. The more traditional sunscreens with phenolic compounds that absorb UV radiation are therefore still very common.



A concern of the widespread use of phenolic compounds is the estrogenic activity these compounds may display, which impacts the hormone balance and may result in breast cancer in women. In order to investigate this, Miller et al. (2001) used recombinant yeast in an estrogen assay to assess the activity of 73 phenolic additives in sunscreens, preservatives, perfumes, disinfectants, antioxidants and flavorings. Thirty-two compounds were shown to have activity in this assay. Twenty-two exhibited potencies relative to 17b-estradiol that ranged from 1/3,000 to 1/3,000,000. Forty-one compounds were inactive. The major criteria for estrogenic activity were the presence of an unimpeded phenolic OH group in a *para*-position and a molecular weight of 140-250 amu.

Kawamura et al. (2003) performed a similar study, also using a yeast-based assay to detect estrogenic activity, with UV-absorbing compounds in food plastics, as well as with benzophenone (**7.4**) derivatives used in sunscreens. They reported estrogenic activity higher than the known endocrine disrupting compound bisphenol A (**7.7**) for several benzophenone derivatives, including 2,4-dihydroxyphenone and 4-hydroxybenzophenone. Based on the specific estrogenic activity, they concluded that a hydroxyl group on the phenol ring of benzophenone has the biggest impact when it is present at the *para*-position, followed by the *meta*- and *ortho*-positions, which is consistent with the data reported by Miller et al. (2001).



(7.7)

Aside from medical applications, polyphenols, including the flavonoids and tannins, are an integral part of human and animal diets, because they represent one of the most numerous and ubiquitous groups of plant metabolites (Bravo, 1998). Although traditionally regarded as anti-nutrients, because of their bad taste, unappealing color, or cause of browning of tissues, polyphenols and other food phenolics are the subject of increasing interest because of their possible beneficial effects on health.

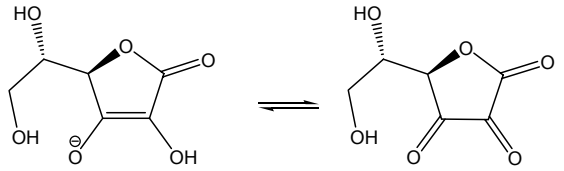
This chapter will highlight the positive effects of phenolic compounds on human health. This is by no means meant to be an exhaustive presentation of all the literature available on this topic. An entire book could be dedicated to it, and such books indeed exist. Several references for further reading are provided at the end of this chapter.

2. ANTIOXIDANT PROPERTIES

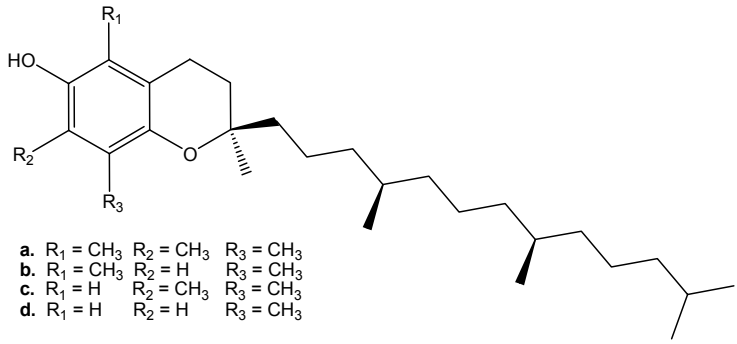
As part of normal metabolism, radicals are generated. Radicals, as discussed in Chapter 2, are compounds with free (*i.e.* unpaired) electrons. Radicals are very reactive, and, when left unchecked, can cause oxidative

damage to the molecules in a cell, and hence have negative impacts on the cellular metabolism. An excess of radicals can cause oxidative stress.

Compounds that can scavenge radicals are also referred to as anti-oxidants. The best known anti-oxidants are vitamin C and vitamin E. Vitamin C is L-ascorbate (**7.8**), a good reducing agent that prevents oxidation of other molecules. The oxidized form of L-ascorbate is L-dehydroascorbic acid (**7.9**). Vitamin E is a mixture of α-, β-, γ-, and δ-tocopherol (**7.10a-d**). Of these four compounds, α-tocopherol is the most effective. Vitamin E is lipid-soluble and has the ability to disrupt the chain reaction during lipid peroxidation (see Chapter 2, Section 1.9).



(7.8) **(7.9)**



a. $R_1 = CH_3$ $R_2 = CH_3$ $R_3 = CH_3$
b. $R_1 = CH_3$ $R_2 = H$ $R_3 = CH_3$
c. $R_1 = H$ $R_2 = CH_3$ $R_3 = CH_3$
d. $R_1 = H$ $R_2 = H$ $R_3 = CH_3$

(7.10)

A lack of vitamin C in the diet results in the disease scurvy. Scurvy's symptoms include purple lesions on the skin, rotten gums, and, as a consequence, loss of teeth. This disease was common among 16th and 17th century sailors who relied on preserved foods and an overall unbalanced diet on their long journeys. The intake of fresh fruits and vegetables, which are rich in vitamin C, can effectively prevent scurvy. The biochemical basis of scurvy is the reduced activity of the enzyme prolyl hydroxylase (E.C. 1.14.11.2), probably because the iron atom that is part of the enzyme cannot be maintained in its active, ferrous state due to the lack of vitamin C. This reduced enzyme activity then results in insufficient hydroxylation of collagen, a structural protein that gives elasticity to the skin and blood

vessels. The lack of elasticity causes the lesions in the skin and the rupture of blood vessels (Stryer, 1988).

In the current era scurvy is a rare disease, but many other diseases can arise from low levels of these vitamins. For example, low plasma levels of α-tocopherol and L-ascorbate correlate with an increased incidence of myocardial infarction and of some forms of cancer (Gey et al., 1987). In fact, many diseases are thought to be associated with higher levels of radicals in the cell (Halliwell, 1991). An example includes rheumatoid arthritis (RA), whereby joint tissues have an excess of activated neutrophiles that secrete radicals, such as O_2^-. Under normal circumstances the radicals are used to kill pathogenic microorganisms, but in the case of RA, the excess of activated neutrophiles contribute to the inflammation and swelling, and hence aggravate the disease symptoms. Other diseases in which reactive oxygen species are implicated include atherosclerosis, adult respiratory distress syndrome (ARDS), myocardial infarction and some forms of cancer. Thus, the ability to scavenge radicals can prevent the onset of a disease, slow down the progress of the disease, or alleviate its symptoms.

Aside from vitamins C and E many other compounds present in fruits and vegetables have been shown to have anti-oxidant properties. Among these compounds there are several classes of phenolic compounds. Aside from preventing scurvy these compounds have a positive influence on cardiovascular health.

An interesting case is the prevention of cardio-vascular diseases as a result of the consumption of wine. Like most fruits grapes are rich in polyphenols, and the process of wine making results in the concentration of polyphenols. Wine polyphenols are considered to have beneficial effects on coronary heart disease and atherosclerosis. The presence of polyphenols in wine are thought to be the reason for the 'French paradox': France was shown to have a coronary mortality rate close to that of China and Japan in spite of the high amount of saturated fat and cholesterol levels in the French diet. The consumption of red wine in France, however, is considerably higher than in either China or Japan (Staggs, 1996).

Wollin and Jones (2001) investigated the effects and mechanisms of action of consumption of red wine compared to other alcoholic beverages on the risk of cardiovascular disease. Of particular interest was the form and quantity of alcohol consumed. The relationship between alcohol consumption and mortality is supported by epidemiologic studies suggesting that different forms of alcohol alter the relative risk for mortality. Evidence

from various epidemiologic and experimental studies suggests a protective effect against the development of cardiovascular disease by moderate consumption of red wine. They point out that components of red wine that are thought to be responsible for the protective effects include various phenolic compounds *as well* as alcohol content.

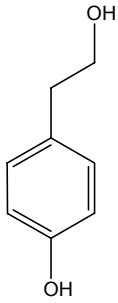
The effects of wine and its polyphenol constituents on early indicators of coronary heart disease such as elevated levels of plasma lipids, platelets and serum antioxidant activity were discussed in a review by Cooper et al. (2004). This review also addressed whether the polyphenols or alcohol are responsible for the beneficial effects of wine on cardio-vascular health. The authors conclude that red wine polyphenols have little effect on plasma lipid concentrations, but that wine consumption reduces the susceptibility of low-density lipoprotein (LDL) cholesterol to oxidation and increase serum antioxidant capacity. These effects, however, do depend on the amount of wine that is consumed and the period of supplementation. It was suggested that specific polyphenols appear to have endothelium-dependent vaso-relaxing abilities. Red wine phenolics also have an inhibitory effect on platelet aggregation. Evidence suggests that alcohol has a positive synergistic effect with wine polyphenols on some atherosclerosis risk factors. Thus, evidence that wine drinking is beneficial for cardiac health appears positive.

Flavonoids may benefit health in cardiovascular disease by adjusting adhesion of monocytes (large mononuclear leukocytes in the blood) in the inflammatory process of atherosclerosis. Most *in vitro* studies have used types of flavonoids present in food rather than those that appear in plasma after food ingestion. Koga and Meydani (2001) tested the effects of plasma metabolites of the flavonoids (+)-catechin (**1.39**) and quercetin (**1.43**) on the alteration of monocyte adhesion to human aortic endothelial cells and on the production of reactive oxygen species. Plasma extracts of flavonoid metabolites were prepared after administration of pure compounds to rats. The plasma preparations contained sulfate or glucuronide conjugates or both, as well as methylated forms. Adhesion of U937 monocyte cells to human aortic endothelial cells was measured, and the production of reactive oxygen in the endothelial cells was monitored, when the cells were pretreated with either pure compounds or plasma extracts from control or treated rats. Adhesion assays were performed with endothelial cells stimulated with interleukin or cells activated with phorbol myristylacetate. Reactive oxygen species were measured after challenging the human aortic endothelial cells with interleukin-1b (IL-1b) or hydrogen peroxide.

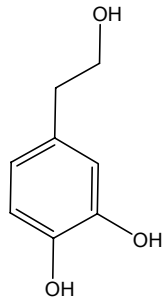
Pretreatment of endothelial cells with (+)-catechin (**1.39**) *metabolites* inhibited U937 cell adhesion to interleukin IL-1b-stimulated endothelial cells, whereas pretreatment with intact (+)-catechin had no effect. Generation of reactive oxygen species in hydrogen peroxide-stimulated cells was inhibited by (+)-catechin, its metabolites, and control plasma extract, whereas reactive oxygen species generation in IL-1b-stimulated cells was inhibited by (+)-catechin metabolites only. In contrast, quercetin inhibited U937 cell adhesion to interleukin IL-1b-stimulated cells, whereas its metabolites were not effective. The authors concluded that metabolic conversion of flavonoids such as (+)-catechin and quercetin modifies the biological activity of the flavonoids. It was suggested that metabolites of flavonoids, rather than their intact forms, contribute to the effects of flavonoids on reducing the risk of cardiovascular disease.

The seeds of fenugreek (*Trigonella* spp.) are rich in phenolic compounds. Kaviarasan et al. (2004) evaluated fenugreek seeds for their potential to protect erythrocytes from oxidation induced by hydrogen peroxide (H_2O_2). Human erythrocytes from diabetic and non-diabetic subjects were incubated with increasing amounts of fenugreek seed extract and challenged with H_2O_2. They were then analyzed for hemolysis (release of hemoglobin) and lipid peroxidation. Erythrocytes from diabetic subjects were more susceptible to hemolysis and lipid peroxidation than those from non-diabetic subjects. It was significant that incubation of the cells with the polyphenol-rich seed extract significantly reduced the oxidative modifications in both cell groups. The inhibition of lipid peroxidation was concentration-dependent. The extract contained 0.75 mM gallic acid-equivalents (**1.41**) of phenolic compounds. The findings demonstrated the potent antioxidant properties of the phenol-rich fenugreek seeds.

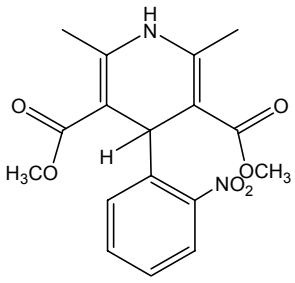
Palmerini et al. (2005) reported that phenols in the Mediterranean diet are free radical scavengers and have antioxidant properties. Yet the mechanisms of their effects are not fully understood. Palmerini and co-workers hypothesized an effect on the concentration of Ca^{2+}, which plays an important role in intracellular signaling and regulates various processes. To test this hypothesis they incubated human lymphomonocytes with two phenolic compounds isolated from olive oil: 3,4-dihydroxyphenyl-ethanol (**7.11**) and *p*-hydroxyphenyl-ethanol (**7.12**). They showed that both of these compounds increased the intracellular concentration of Ca^{2+} in a dose-dependent manner, both in the presence and in the absence of calcium in the extracellular medium. This effect was antagonized by the drug nifedipine (**7.13**), a calcium channel blocker administered as a muscle relaxing agent to patients suffering from chest pain.



(7.11)

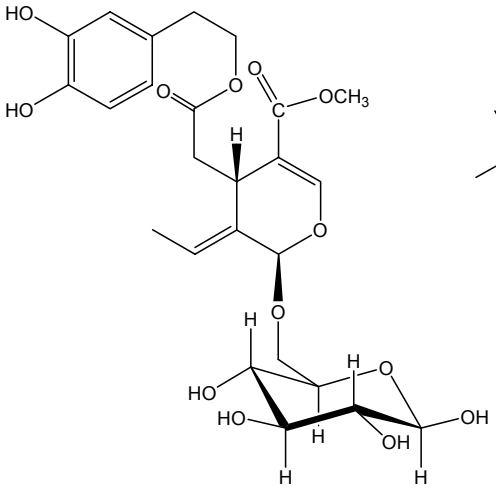


(7.12)

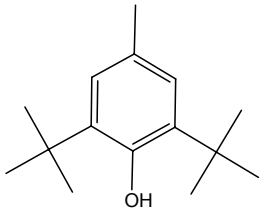


(7.13)

The olives themselves contain many phenolic compounds with antioxidant properties. Bouaziz et al. (2005) investigated the olive cultivar "Chemlali" from Tunisia. Oleuropein (**7.14**), a bitter glycoside esterified with a phenolic acid, was the major compound present. Phenolic monomers and twelve flavonoids were also identified. The antioxidant activity of the extract was evaluated. Acid hydrolysis of the extract enhanced its antioxidant activity. *p*-Hydroxyphenyl-ethanol (**7.12**) and quercetin (**1.43**) showed antioxidant activities similar to that of 2,6-di-*tert*-butyl-4-methyl phenol (**7.15**), a reference compound with known antioxidant properties. It was suggested that a hydroxyl group at the *ortho*-position on the flavonoid B ring could contribute to the antioxidant activity of the flavonoids.



(7.14)



(7.15)

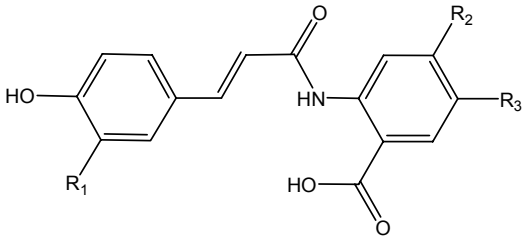
In addition to the contribution of olives and olive oil to antioxidant activity in the Mediterranean diet, spices and dressings have also been shown to have health-promoting activity. In a study by Ninfali et al. (2005), twenty-seven vegetables, fifteen aromatic herbs and some spices consumed in Central Italy were studied to determine total phenolic flavonoid content as well as their antioxidant capacity measured by the oxygen radical absorbance capacity method. A comparison of antioxidant capacity was made between different salads, as well as between salads to which aromatic herbs had been added. Lemon balm and marjoram at a concentration of 1.5% (w/w) increased the antioxidant capacity of a salad by 150% and 200%, respectively. A 200-gram portion of a salad enriched with marjoram corresponded to an intake of 200 mg phenolics and 4000 oxygen radical absorbance capacity units. Olive oils, wine and apple vinegars were salad dressings that provided the highest increase in antioxidant capacity. Of the spices tested, cumin and ginger made the most significant contribution to the antioxidant capacity.

Both intact compounds and their metabolites – formed either in human tissues or meabolized in the colon by microflora – might explain the effects on health of dietary polyphenols. In order to assess the importance and biological activities of microbial metabolites *in vivo*, Gonthier et al. (2003) measured the microbial metabolites formed in rats fed a diet supplemented with three levels of catechin (**1.43**) or red wine powder containing proanthocyanidins, phenolic acids, flavanols, anthocyanins and flavonols. This was compared to rats fed an unsupplemented diet. Aromatic acid metabolites in urine were assayed by an HPLC-electrospray ionization-mass spectrometry method. The primary metabolites formed from the catechin diet were 3-hydroxyphenylpropionic acid, 3-hydroxybenzoic acid and 3-hydroxyhippuric acid. Their total urinary excretion accounted for 4.7% (w/w) of the catechin ingested, and that of intact catechins for 45.3% (w/w). When the diet was supplemented with red wine powder, the same metabolites observed with the catechin diet were identified in urine, along with *p*-hydroxybenzoic (**1.4**), 3-hydroxyphenylacetic (see **1.11** for a comparable structure), *p*-coumaric (**1.13**), vanillic (**1.8**), and hippuric (N-benzoylglycine) acids. These aromatic acids accounted for 9.2% (w/w) of the total wine polyphenols ingested, whereas intact catechins accounted for only 1.2% (w/w). It was suggested that the higher excretion of aromatic acids by rats fed wine polyphenols was due to their poor absorption in the proximal part of the gut. Some of the microbial metabolites still contained a reducing phenolic group and should also prevent oxidative stress in inner tissues. The authors suggested that attention be given in the future to these

microbial metabolites and their biological properties to help explain the effects of polyphenols that are not easily absorbed through the gut.

Whole grains provide another source of phenolic antioxidants, but whole grains contain many other compounds that have a positive effect on human health. They have high concentrations of dietary fiber, starch, and oligosaccharides, and contain phytate, phyto-oestrogens such as lignans, plant stanols and sterols, vitamins and minerals. Epidemiological studies have shown that whole-grain intake is protective against cancer, cardiovascular disease, diabetes, and obesity (Slavin et al., 2004). Despite recommendations to consume three servings of whole grains daily, usual intake in Western countries is only about one serving. Feeding studies show that consumption of whole grains improves biomarkers such as weight loss, blood-lipid levels, and the concentration of antioxidants. Although it is difficult to separate the protective properties of whole grains from dietary fiber and other components, the disease protection seen from whole grains in prospective epidemiological studies far exceeds the protection from isolated nutrients and phytochemicals in whole grains

An interesting question is whether the health-promoting properties of phenolic compounds is consistent, or whether there are effects of the culture practice during crop production, the location of the field where the crop is grown, and the specific cultivar that was selected. Emmons and Peterson (2001) investigated whether cultivar and location had an effect on phenolic contents and antioxidant activities of alcohol-soluble extracts from groats (*i.e.* the edible part of the grain) of oat (*Avena sativa* L.). Antioxidant activities (AOA) and concentrations of eight phenolic compounds having AOA were measured in three cultivars grown at seven locations in Wisconsin during 1998. The phenolic compounds included *p*-coumaric acid (**1.13**), ferulic acid (**1.15**) and avenanthramides (**7.16a-d**). Avenanthramides are phytoalexins found in oats. There are several different compounds – avenanthramide A, B, D and G – that differ in the substitution pattern of the two aromatic rings, as is shown below. Avenanthramide L (**7.17**) contains an additional carbon in the chain linking the two aromatic rings (Okazaki et al., 2004).



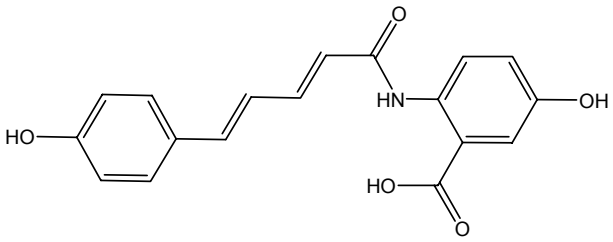
a. Avenanthramide A: R_1 = H R_2 = H R_3 = OH

b. Avenanthramide B: R_1 = OMe R_2 = H R_3 = OH

c. Avenanthramide D: R_1 = H R_2 = H R_3 = H

d. Avenanthramide G: R_1 = H R_2 = OH

(7.16)



(7.17)

There were significant differences among cultivars for AOA concentrations of all of the phenolic compounds measured, except *p*-coumaric and ferulic acids, and for total free phenolic contents (FPC). Location significantly affected the concentrations of five of the phenolics and total FPC, but did not affect AOA. There were significant cultivar x location interactions for the concentrations of avenanthramides and for total FPC. The presence of this interaction means that the cultivar with the highest FPC level in one location, does not produce the highest FPC levels in another location. The unexpectedly high concentrations of avenanthramides from the Sturgeon Bay location were confirmed by analysis of groats from 1999 and 2000. Based on these observations it should be possible to improve the AOA and phenolic concentrations of oat as quantitative traits in a cultivar development program (see Chapter 3, Section 3.4), but significant location effects may slow down progress.

Similarly, Tarozzi et al. (2004) assessed the impact of cultivation practices, commercial processing, and storage of fruits and vegetables on phenolic antioxidants. They investigated the influence of commercial cold-storage periods on antioxidant properties of apples grown by organic or

integrated systems. Regardless of the production method, total phenolics and total antioxidant activity decreased in apples with the peel intact only in the first three months of storage. It was suggested that cold storage rapidly depletes antioxidant properties in apple skin but not in the pulp. Antioxidant activity was assessed *in vitro* in terms of intracellular antioxidant, cytoprotective, and anti-proliferative activities in human colon carcinoma (Caco-2) cells. Time-related decreases in antioxidant activity after six months cold storage were found regardless of the cultivation method. These data suggest that cold storage should be taken into account when evaluating the cancer-preventive benefits of fruits and vegetables. Furthermore, the authors concluded that organic production methods of apples do not provide health benefits. This latter conclusion is in contrast with a study by Halweil (2003), who concluded that organic produce was richer in health-promoting phenolic compounds.

3. DISEASE PREVENTION

From the previous section it is apparent that the antioxidant activity of polyphenols offers many health benefits. There are also cases where the impact of phenolic compounds in the diet may not be related to their antioxidant activity. For example, phenols might exert effects within the gastrointestinal tract. Such effects could include binding of iron, scavenging reactive nitrogen, chlorine, and oxygen species, and inhibition of cyclooxygenases and lipoxygenases (Halliwell et al., 2005). This section therefore focuses on the role of phenols in disease prevention, where the precise mechanism is either unknown, or related to activity other than antioxidant activity.

The stilbene *trans*-resveratrol (*trans*-3,5,4′-trihydroxystilbene; **1.60**) is present in grapes and wines. Resveratrol has been shown to have cancer chemopreventive activity in the three major stages of carcinogenesis: promotion, initiation, and progression (Jang et al., 1997). In *in vitro* assays performed on tumor cell lines, resveratrol was shown to effectively inhibit cyclo-oxygenase (COX), an enzyme that catalyzes the formation of pro-inflammatory compounds implicated in tumor cell growth and the suppression of immune surveillance mechanisms. Furthermore, rats in which inflammation had been induced through injection of carrageenan showed less severe symptoms after they were treated with resveratrol. The anti-promotion activity of resveratrol was demonstrated by showing a dose-dependent inhibitory effect on the formation of free radicals in cultured HL-60 promyelocytic leukemia cells treated with the inflammation-inducing compound 12-*O*-tetradecanoylphorbol-13-acetate (TPA). In addition,

resveratrol inhibited in a dose-dependent manner the mutagenic effects of 7,12-dimethylbenz(*a*)anthracene (DMBA) in assays performed on *Salmonella typhimurium* strain TM677. Finally, the anti-progression activity of resveratrol was demonstrated by showing that cultured HL-60 cells treated with resveratrol were no longer able to proliferate indefinitely, but rather developed symptoms of terminal differentiation to a non-proliferative phenotype.

The anti-carcinogenic effects of resveratrol were also tested *in vivo*, by examination of mammary glands of mice treated with DMBA and TPA. A dose-dependent reduction in the formation of tumors was observed when the mice were treated with resveratrol. Resveratrol is present in high concentrations in the skins of grapes (50-100 μg/g) and, consequently, in red wine (1.5-3 mg/l). Based on these experiments, the consumption of grapes and grape products appears to have beneficial effects.

Resveratrol has also been reported to offer protection against cardiovascular disease, such as coronary heart disease. The effects of resveratrol on factors implicated in the development of coronary heart disease – human platelet aggregation and the synthesis of eicosanoids (lipids) from arachidonate by platelets – were investigated and compared with the actions of other wine phenolics – catechin (**1.39**), epicatechin (**7.18a**), and quercetin (**1.43**) – and the antioxidants α-tocopherol (**7.10a**), hydroquinone and butylated hydroxytoluene. Resveratrol and quercetin demonstrated a dose-dependent inhibition of platelet aggregation, whereas the other compounds tested were inactive. Resveratrol also inhibited the synthesis of the eicosanoids in a dose-dependent manner, whereas the other phenolics were less effective of not effective at all. Removal of the alcohol from the wine did not diminish the effect on platelet aggregation (Pace-Asciak et al., 1995; Goldberg et al., 1995).

While consumption of products rich in resveratrol appears to be beneficial for the reasons discussed above, little is known about resveratrol bioavailability in humans and animals. Emilia et al. (1999) developed an analytical method to measure stilbene present in blood. Resveratrol administered orally to rats was detected in plasma. Excellent HPLC-based separation of *trans*-resveratrol from other compounds in the blood was achieved, allowing a rapid analysis of the sample for absorption, distribution, and metabolism studies.

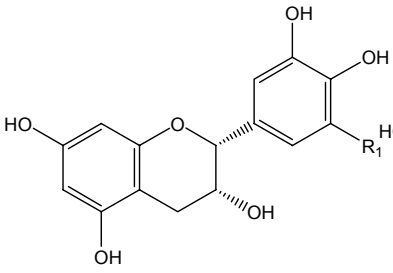
A high dietary intake of fruits and vegetables has consistently been associated with a reduced risk of various human cancers, including those of the lung, breast, prostate, and colon. It is unknown which bioactive

compound or compounds in plant foods provide these protective effects, but flavonoids have been of special interest. There are numerous animal model studies that suggested that flavonoids influence important cellular mechanisms related to carcinogenesis, including cell cycle control and apoptosis. There are limited data from studies with humans. In a review artice by Neuhouser (2004) four studies are reviewed, in which associations of flavonoid intake with cancer risk were examined. There is substantial evidence that flavonoids, especially quercetin (**1.43**), reduce the risk of lung cancer.

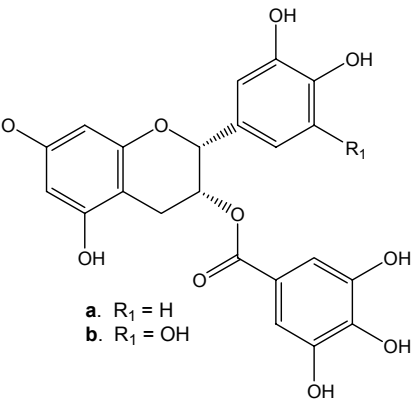
Green and black tea, obtained from dried leaves of the Chinese tea plant (*Camellia sinensis*), are sources of different polyphenolic compounds. The main polyphenols of green tea are (-)-epicatechin (**7.18a**), (-)-epicatechin 3-gallate (**7.18b**), (-)-epigallocatechin (**7.19a**), and (-)-epigallocatechin 3-gallate (**7.19b**), whereas in black tea theaflavin (**7.20**) and thearubigin (**7.21**) are the most abundant. Green tea is prepared by steaming or pan-frying freshly picked tea leaves. This treatment inactivates oxidases, so that the catechins in the leaves remain stable and thus contribute to the characteristic color and smell of green tea. For the production of black tea, the leaves are left to dry down to approximately 55% of the fresh weight, rolled and then crushed. Theaflavin and thearubigen arise from oxidation of catechins during this process. Oolong tea is produced in a similar manner as black tea, except that the oxidation process is stopped by firing the dried and rolled leaves (Mukhtar and Ahmad, 2000).

Haqqi et al. (1999) investigated the impact of polyphenols present in green tea on rheumatoid arthritis (RA). The study used mice as test subjects and showed that although 92% of ordinary mice developed RA when injected with a compound that induced RA, less than half of the mice that consumed the green tea polyphenols developed RA after a similar injection.

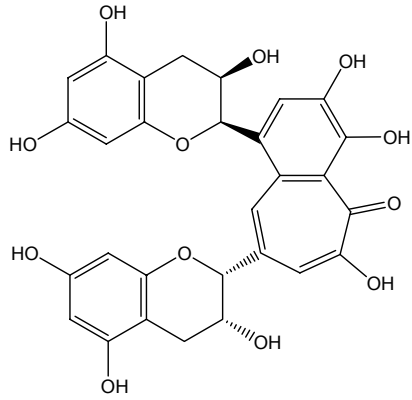
The health benefits of both green and black tea were summarized by Mukhtar and Ahmad (2000). Consumption of tea and its polyphenolic constituents offers protection against skin cancer induced by either chemical carcinogens or ultraviolet radiation in mice. Tea consumption also provides protection against cancers induced by chemical carcinogens that involve the lung, forestomach, esophagus, duodenum, pancreas, liver, breast, colon, and skin in mice, rats, and hamsters. As was shown for resveratrol from grapes, polyphenolics from green tea do not only have anti-carcinogenic properties, but also protect against coronary heart disease and atherosclerosis, as well as inflammation. The latter effect was demonstrated by showing



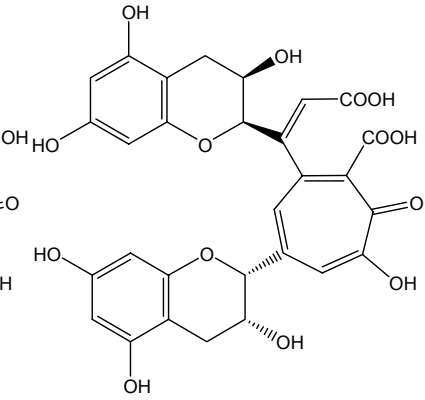
(7.18)



(7.19)



(7.20)



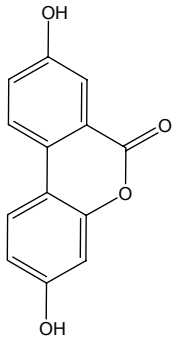
(7.21)

anti-inflammatory effects of green tea in response to UV-B radiation and the inflammation-inducing compound TPA.

The mechanisms underlying the protective role of tea polyphenols have been investigated in a number of studies. Green tea polyphenols were shown to transcriptionally activate a signaling cascade that resulted in the activation of detoxifying enzymes involved in the elimination of chemical carcinogens (Yu et al., 1997). Apoptosis (programmed cell death) is a way to clean up cells that are no longer needed or cells that have aged to the point where they start to malfunction. Certain types of cancer have been shown to arise from a disturbance in the process of apoptosis, whereby damaged cells do not get eliminated, or whereby carcinogens result in the elimination of functional cells that are critical. A report by Ahmad et al. (1997) showed that green tea

polyphenolics induced apoptosis and cell cycle arrest in human epidermoid carcinoma cells. This apoptotic response was specific to cancer cells, because the induction of apoptosis was also observed in several other types of cancer cell lines, but not, for example, in normal human epidermal cell lines. Dong et al. (1997) used a mouse epidermal cell line to examine the anti-tumor promotion effects of polyphenolics from green and black teas. They showed that these compounds inhibited the formation of tumor cells resulting from the application of epidermal growth factor or TPA, and this occurred in a dose-dependent manner. They also showed that polyphenolics from tea inhibited the transcriptional activation of genes regulated by the transcription factor AP-1. This transcription factor has been shown to be important for the formation of tumor cells.

Ellagitannins are dietary polyphenols containing ellagic acid (**1.96**) subunits that are thought to act as cancer chemo-preventive agents. Thus, they may have properties that contribute to health benefits in humans. Little is known, however, of their metabolic fate. Cerdá et al. (2005) investigated the metabolism of different dietary ellagitannin derivatives in humans. Healthy volunteers in four groups consumed, in a single dose, a different ellagitannin-containing food. The reported consumption was strawberries (250 g), red raspberries (225 g), walnuts (35 g), and oak-aged red wine (300 ml). Urine fractions were collected at 8, 16, 32, 40, and 56 hours after consumption. Neither ellagitannins nor ellagic acid were detected in the urine samples by liquid chromatography-MS/MS analysis. However, a microbial metabolite 3,8-dihydroxy-6H-dibenzo-[*b,d*]-pyran-6-one (urolithin B; **7.22**) conjugated with glucuronic acid was detected in all study participants independent of the consumed food.



(7.22)

These authors reported that considerable differences among and between individuals were found, which identified individuals who excreted high and

low metabolite levels. The variation in excretion of these metabolite levels was likely reflective of the micro-flora in the colon during ellagitannin metabolism. These results indicate that urolithin B is a marker of exposure to dietary ellagitannins and may be useful in intervention studies with ellagitannin-containing products.

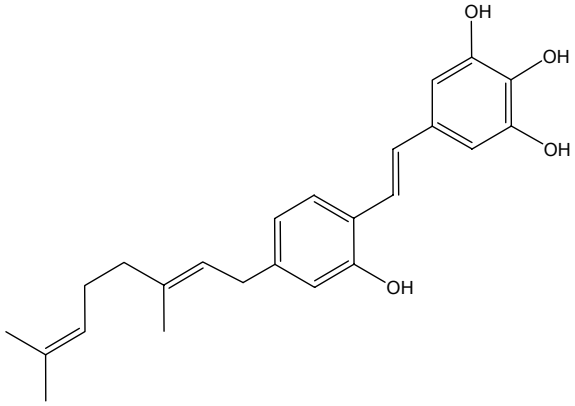
Proanthocyanidins and tannin-like compounds are complex flavonoid polymers naturally present in cereals, legume seeds and are frequently found in some fruits and fruit juices. They share some common structural features with phenolic polymers found in black tea and red wine. The polymeric nature of proanthocyanidins makes their analysis in food difficult. Thus, little is known about their consumption, although they probably contribute a large part of the daily polyphenol consumption. They also share common physico-chemical properties: they form stable complexes with metal ions and proteins and are good reducing agents. Many of their biological attributes of nutritional interest derive from these properties. As metal ion chelators, they influence the availability of several minerals. The nutritional significance of the property of forming non-specific complexes with proteins is not clear. As reducing agents, they may participate in prevention of cancers of the digestive tract and inner organs. They may also protect low-density lipids (LDLs) against oxidation and inhibit platelet aggregation. This property would act as a preventative of cardiovascular diseases (Santos-Buelga and Scalbert, 2000).

Since epidemiological studies have suggested a correlation between high flavonoid consumption and decreased risk of cancer, cardiovascular disease, and other age-related diseases, enhancing flavonoid biosynthesis in crops may result in foods with benefits to human health. Verhoeyen et al. (2002) attempted to generate transgenic tomatoes with increased levels of flavonoids. They achieved a 78-fold increase in fruit flavonols achieved through ectopic expression of the gene encoding chalcone isomerase. Furthermore, they observed that chalcone synthase and flavonol synthase transgenes were found to act synergistically to up-regulate flavonol biosynthesis in tomato tissues.

4. ACTIVITY AGAINST TOXINS

Plant phenolic compounds have also been suggested to provide a means for preventing the adverse affects that fungal toxins (mycotoxins) have on human health as well as serving in their detoxification (Beekrum et al., 2003). These authors investigated the impact of the plant phenolic

compounds chlorophorin (4-homogeranyl-2,3',4',5'-tetrahydroxystilbene; **7.23**), iroko, maakianin, benzoic acid, caffeic acid (**1.14**), ferulic acid (**1.15**), and vanillic acid (**1.8**) on the growth of *Fusarium verticillioides* and its production of the mycotoxin fumonisin B-1. The stilbene chlorophorin was the most effective compound in inhibiting fungal growth and in reducing toxin production. As a group these phenols reduced fumonisin levels by more than 90%. The authors suggested that the widespread occurrence of fumonisins and the lack of adequate prevention measures require 'biologically safe' alternatives to prevent the transfer of fungi and their hazardous toxins into foods.



(7.23)

5. REFERENCES

Ahmad, N., Feyes, D. K., Nieminen, A.-L., Agarwal, R., and Mukhtar, H., 1997, Green tea constituent epigallocatechin-3-gallate and induction of apoptosis and cell cycle arrest in human carcinoma cells, *J. Natl. Cancer Inst.* **89**:1881-1886.

Beekrum S., Govinden R., Padayachee T., and Odhav B., 2003, Naturally occurring phenols: a detoxification strategy for fumonisin B-1, *Food Add. Contam.* **20**: 490-493.

Bouaziz M., Grayer R. J., Simmonds, M. S. J., Damak, M., and Sayadi, S., 2005, Identification and antioxidant potential of flavonoids and low molecular weight phenols in olive cultivar Chemlali growing in Tunisia *J. Agric. Food Chem.* **53**: 236-241.

Bravo, L., 1998. Polyphenols: chemistry, dietary sources, metabolism, and nutritional significance, *Nutr. Rev*. **56**: 317-333.

Cerdá, B., Tomás-Barberán, F. A., and Espín, J. C. (2005) Metabolism of antioxidant and chemo-preventive ellagitannins from strawberries, raspberries, walnuts, and oak-aged wine in humans: Identification of biomarkers and individual variability, *J. Agric. Food Chem.* **53**: 227-235.

Cooper, K. A., Chopra, M., and Thurnham D. I., 2004, Wine polyphenols and promotion of cardiac health, *Nutr. Res. Rev.* **17**, 111-129.

Dong, Z., Ma, W.-Y., Huang, C.,Yang, C. S., 1997, Inhibition of tumor promoter induced activator protein 1 activation and cell transformation by tea polyphenols, (-)-epigallocatechin gallate and theaflavins, *Cancer Res*. **57**: 4414-4419.

Emilia, J. M., Lamuela-Raventos, R. M., and de la Torre-Boronat, C. M., 1999, Determination of *trans*-resveratrol in plasma by HPLC, *Anal. Chem.* **71**:747-750.

Emmons, C. L., and Peterson, D. M., 2001, Antioxidant activity and phenolic content of oat as affected by cultivar and location, *Crop Sci.* **41**: 1676-1681.

Gey, K. F., Brubacher, G. B., Stahelin, H. B., 1987, Plasma levels of antioxidant vitamins in relation to ischemic heart disease cancer. *Am. J. Clin. Nutr*. **45**: 1368-1377.

Goldberg, D. M., Hahn, S. E., and Parkes, J. G., 1995, Beyond alcohol: Beverage consumption and cardiovascular mortality, *Clin. Chim. Acta* **237**: 155.

Gonthier, M.-P., Cheynier, V., and Donovan, J. L., 2003, Microbial aromatic acid metabolites formed in the gut account for a major fraction of the polyphenols excreted in urine of rats fed red wine polyphenols, *J. Nutr.* **133**:461-467.

Jang, M., Cai, L., Udeani, G. O., Slowing, K. V., Thomas, C. F., Beecher, C. W. W., Fong, H. H. S., Farnsworth, N. R., Kinghorn, A. D., Mehta, R. G., Moon, R. C., and Pezzuto, J. M., 1997, Cancer chemopreventive activity of resveratrol, a natural product derived from grapes, *Science* 275: 218-220.

Halliwell, B., 1991, Reactive oxygen species in living systems, *Am. J. Med.* **91**: 14S

Halliwell, B., Rafter, J., and Jenner A., 2005, Health promotion by flavonoids, tocopherols, tocotrienols, and other phenols: direct or indirect effects? Antioxidant or not? *Am. J. Clin. Nutr.* **81**: 268S-276S (*Suppl.*)

Halweil, B., 2003, Organic produce found to be higher in health-promoting compounds, *World Watch* **16**: 9.

Haqqi, T. M., Anthony, D. D., Gupta, S., Ahmad, N., Lee, M. S., Kumar, G. K., and Mukhtar, H, 1999, Prevention of collagen-induced arthritis in mice by a polyphenolic fraction from green tea, *Proc. Natl. Acad. Sci. USA* **96**: 4524-4529.

Kaviarasan, S., Vijayalakshmi, K., and Anuradha, C.V., 2004, Polyphenol-rich extract of fenugreek seeds protect erythrocytes from oxidative damage, *Plant Foods Hum. Nutrition* **59**: 143-147.

Kawamura, Y., Ogawa, Y., Nishimura, T., Kikuchi, Y., Nishikawa, J., Nishihara, T., and Tanamoto, K., 2003, Estrogenic activities of UV stabilizers used in food contact plastics and benzophenone derivatives tested by the yeast two-hybrid assay, *J. Health Sci*. **49**: 205-212.

Koga, T., and Meydani, M., 2001, Effect of plasma metabolites of (+)-catechin and quercetin on monocyte adhesion to human aortic endothelial cells, *Am. J. Clin. Nutrition* **73**: 941-948.

Miller, D, Wheals, B. B., and Beresford, N., 2001, Estrogenic activity of phenolic additives determined by an in vitro yeast assay, *Env. Health Persp.* **109**: 133-138.

Mukhtar, H., and Ahmad, N., 2000, Tea polyphenols: prevention of cancer and optimizing health, *Am. J. Clin. Nutrition* **71** (*suppl*.): 1698S-1702S.

Neuhouser, M. L., 2004, Dietary flavonoids and cancer risk: evidence from human population studies, *Nutr. Cancer* **50**: 1-7.

Ninfali P, Mea G, Giorgini S, Rocchi M, and Bacchiocca M. 2005. Antioxidant capacity of vegetables, spices and dressings relevant to nutrition, *Br. J. Nutr.* **93**: 257-266.

Okazaki , Y., Isobe T., Iwata, Y., Matsukawa, T., Matsuda, F., Miyagawa, H., Ishihara, A., Nishioka, T., and Iwamura, H., 2004, Metabolism of avenanthramide phytoalexins in oats, *Plant J.* **39**: 560-572.

Pace-Asciak, C. R., Hahn, S., Diamandis, E. P., Soleas, G, and Goldberg, D. M., 1995, The red wine phenolics *trans*-resveratrol and quercetin block

human platelet aggregation and eicosanoid synthesis: Implications for protection against coronary heart disease, *Clin. Chim. Acta* **237**: 207-219.

Palmerini C. A., Carlini, E., Saccardi, C., Servili, M., Montedoro, G., and Arienti, G., 2005, Activity of olive oil phenols on lymphomonocyte cytosolic calcium, *J. Nutr. Biochem.* **16**: 109-113.

Santos-Buelga C, and Scalbert, A., 2000, Proanthocyanidins and tannin-like compounds - nature, occurrence, dietary intake and effects on nutrition and health, *J. Sci. Food Agric.* **80**: 1094-1117.

Slavin, J., 2004. Whole grains and human health, *Nutr. Res. Rev.* **17**: 99-110.

Staggs, B., 1996, Why wine is really better, *Health* **10**: 30.

Stryer, L.,1988, Biochemistry, W. H. Freeman & Co, New York. pp. 1089.

Tarozzi, A, Marchesi, A, and Cantelli-Forti, G., 2004. Cold-storage affects antioxidant properties of apples in Caco-2 Cells, *J. Nutr.* **134**: 1105-1109.

Verhoeyen M. E., Bovy, A., Collins, G., Muir, S., Robinson, S., de Vos, C. H., and Colliver, S., 2002, Increasing antioxidant levels in tomatoes through modification of the flavonoid biosynthetic pathway, *J. Exp. Bot.* **53**: 2099-2106.

Wollin, S,D,, and Jones, P. J. H., 2001, Alcohol, red wine and cardiovascular disease, *J. Nutr.* **131**: 1401-1404.

Yu, R. Jiao, J.-J., Duh, J.-L., Gudehithlu, K., Tan, T.-H., and Kong, A.-N. T. 1997, Activation of mitogen-activated protein kinases by green tea polyphenols: potential signaling pathways in the regulation of antioxidant-responsive element-mediated Phase II enzyme gene expression, *Carcinogen.* **18**: 451-456.

For further reading

Bagchi, D., 2000, *Resveratrol and Human Health*, Keats Publishing, Lincolnwood, IL, pp. 48.

Goldberg, G., 2003, *Plants: Diet and Health, Blackwell Science*, Oxford, pp. 347.

Packer, L., Hiramatsu, M., Yoshikawa, T., 1999, *Antioxidant Food Supplements in Human Health*, Academic Press, San Diego, pp. 511.

Watson, R. R., 2000, *Vegetables, Fruits, and Herbs in Health Promotion*, CRC Press, Boca Raton, FL., pp. 360.

Appendix

Alphabetical list of the compounds represented by chemical structures

The compounds are listed in alphabetical order, whereby numbers, prefixes such as *para* or *O*- have been ignored. For example, '*p*-coumaric acid' is listed under 'C', and '1,2-dihydroxybenzene' is listed under 'D'. When applicable, the compound may be listed multiple times. In the example above, '1,2-dihydroxybenzene' is also listed under 'B', as 'benzene, 1,2-dihydroxy'.

A

B

C

D

R

S

T

Index

Wilfred Vermerris
Ralph Nicholson
Phenolic Compound Biochemistry

How does the consumption of red wine reduce heart disease? How does a plant battle an invading fungus using physical and chemical defense mechanisms? How are tannins used in leather production synthesized?

These are just a few examples that illustrate the chemical diversity and use of phenolic compounds, the topic of *Phenolic Compound Biochemistry*.

This book is written for researchers, instructors, advanced undergraduate students and beginning graduate students in the life sciences who wish to become more familiar with these and many other intriguing aspects of phenolic compounds. Topics covered include nomenclature, chemical properties, biosynthesis, including an up-to-date overview of the genetics controlling phenolic metabolism, isolation and characterization of phenolic compounds, phenolics used in plant defense, and the impact of phenolics on human health. The book is written in an accessible style, and assumes only basic knowledge of organic chemistry, biochemistry and cell physiology. More than 300 chemical structures and reaction schemes illustrate the text.

Wilfred Vermerris is Associate Professor of Agronomy at the University of Florida Genetics Institute in Gainesville, FL. His research focuses on the genetic control of phenolic compounds that impact agro-industrial processing of crop plants.

Ralph Nicholson is Professor of Botany and Plant Pathology at Purdue University in West Lafayette, IN. He is an expert on phenolic compounds involved in the plant's defense against pathogenic fungi and bacteria.

ISBN 978-4020-5163-0

springer.com

Vermerris · Nicholson

Phenolic Compound Biochemistry