BIOSYNTHESIS OF CUTICULAR ALKYLRESORCINOLS IN SELECTED GRASS SPECIES *BRACHYPODIUM DISTACHYON* AND *SECALE CEREALE*

by

Ruonan Yao

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

The Faculty of Graduate Studies

(Botany)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

October 2011

© Ruonan Yao, 2011

Abstract

Alkylresorcinols are phenolic lipids which occur in diverse plant species as well as microorganisms. In plants, alkylresorcinols are usually deposited at or near the surfaces where they are thought to serve as a first line of defense. Earlier work in our lab had shown the surface accumulation of alkylresorcinols in *Secale cereale* leaves was mainly restricted to the cuticle. However, direct evidence showing the protective role of these bioactive compounds at the surface is still insufficient.

The current work was to investigate the biosynthesis of cuticular alkylresorcinols in order to get a better understanding of their biological function. This research focused on *S. cereale*, since it had previously been shown to contain relatively large amounts of alkylresorcinols, and on *Brachypodium distachyon*, a closely related genetic model system with completely sequenced genome. First, chemical analyses revealed that the cuticular wax covering leaves of *B. distachyon* included 5% of alkylresorcinols with alkyl chains varying from C_{17} to C_{25}. Therefore, it was hypothesized that both species have genes encoding alkylresorcinol synthases (ARSs). A central goal of this work was to clone and characterize potential ARSs.

One ARS (BdARS) was cloned from *B. distachyon* by mining the Brachypodium expressed sequence tag libraries and one ARS (ScARS) was cloned from *S. cereale* using a homology-based cloning strategy. *In vivo* biochemical characterization in yeast *Saccharomyces cerevisiae* demonstrated that both enzymes were capable of using C_{10} to C_{22} fatty acyl-CoAs with malonyl-CoA to generate a broad range of alkylresorcinols. Organ-specific expression in leaves but not in roots was observed for both *BdARS* and *ScARS*. Additionally, the expression pattern of *ScARS* matched the time-course of cuticular alkylresorcinol accumulation along the leaf of *S. cereale*. An investigation into their subcellular localization revealed that both ARSs were likely localized to the endoplasmic reticulum membrane. All these results taken together support the idea that BdARS and ScARS are the enzymes responsible for the biosynthesis of cuticular alkylresorcinols, and that the cuticular alkylresorcinols are indeed biosynthesized for a protective function associated with the wax lining the surface of grass leaves.

Table of Contents

Abstract ii

Table of Contents iii

List of Tables vi

List of Figures vii

List of Abbreviations ix

Acknowledgements x

Dedication xi

Chapter 1 Introduction to alkylresorcinols and plant cuticles 1

1.1 Overview of alkylresorcinol occurrence, localization and biological activities ... 1

1.2 Structure, composition and function of plant cuticle and alkylresorcinols in the cuticle 4

1.3 Biosynthesis of alkylresorcinols via polyketide pathway 9

1.4 Biosynthesis of VLC aliphatics 14

1.5 Research objectives and questions 16

Chapter 2 Materials and methods 19

2.1 Plant materials and growth conditions 19

2.1.1 Plant growth conditions 19

2.1.2 Plant materials 19

2.2 Wax analyses 21

2.2.1 Wax extraction and derivatization 21

2.2.2 Chemical analyses using GC-FID and GC-MS 21

2.3 Genomic DNA extraction, RNA isolation and reverse transcription 22

2.3.1 Genomic DNA preparation and RNA isolation 22

2.3.2 Reverse transcription 22

2.4 Cloning of putative alkylresorcinol synthase (ARS) genes 22

2.4.1 Cloning of *BdARS* from *B. distachyon* 22

2.4.2 Cloning of *ScARS* from *S. cereale* 23

2.5 Phylogenetic analyses 25

2.6 Heterologous expression of ARSs in yeast *Saccharomyces cerevisiae* 27

2.6.1 Yeast strains, transformation and expression 27

2.6.2 Yeast lipid extraction and analyses 28

2.6.3 Thin layer chromatography (TLC) analyses 28

2.7 Semi-quantitative RT-PCR and quantitative RT-PCR analyses 29
2.7.1 Semi-quantitative RT-PCR 29
2.7.2 Quantitative RT-PCR..... 30
2.8 Constructs and plant transformation 31
2.8.1 Plasmid constructs 31
2.8.2 Agrobacterium-mediated transformation 31
2.9 Light microscopy and laser scanning confocal microscopy 32
Chapter 3 Accumulation of cuticular alkylresorcinols in *Brachypodium distachyon* leaves 33
3.1 Introduction 33
3.2 Results..... 34
3.2.1 Total leaf wax of *B. distachyon* 34
3.2.2 Identification of alkylresorcinols in the total leaf wax of *B. distachyon*... 37
3.2.3 GC-MS analysis of alkylresorcinols in leaves and roots of *S. cereale* 41
3.2.4 Two novel compounds identified as methylated alkylresorcinols in the total leaf wax from *B. distachyon* 42
3.3 Discussion..... 45
Chapter 4 Cloning of putative alkylresorcinol synthases (ARSs) from *Brachypodium distachyon* and *Secale cereale* 51
4.1 Introduction 51
4.2 Results..... 52
4.3 Discussion..... 55
4.3.1 Phylogenetic relationships of ARSs and other type III PKSs..... 56
4.3.2 Biochemical function of ARSs 60
Chapter 5 Biochemical and biological characterization of alkylresorcinol synthases (ARSs) from *Brachypodium distachyon* and *Secale cereale* 64
5.1 Introduction 64
5.2 Results..... 65
5.2.1 Functional expression of BdARS and ScARS in wild type yeast..... 66
5.2.2 Functional expression of BdARS and ScARS in yeast mutant..... 71
5.2.3 Gene expression studies of *BdARS* and *ScARS* 72
5.2.4 Subcellular localization of BdARS and ScARS 75
5.3 Discussion..... 78
5.3.1 Biochemical characterization in yeast 78
5.3.2 Gene expression analyses 81

5.3.3 Subcellular localization of BdARS and ScARS 83

Chapter 6 Conclusion and future directions 85

6.1 Identification of alkylresorcinols and other cuticular wax compounds in *B. distachyon* leaves 86

6.2 Cloning and characterization of ARSs from *B. distachyon* and *S. cereale* 88

References 94

List of Tables

Table 1.1 Sources of alkylresorcinols and their derivatives. 3

Table 3.1 Qualitative isomer composition of alkyl esters identified in *B. distachyon* total leaf wax..... 37

Table 4.1 Amino acid sequence identities between selected ARSs and CHSs. 55

List of Figures

Figure 1.1 Structure of 5-*n*-alkylresorcinols. 1

Figure 1.2 Schematic cross-section of a plant cuticle devoid of wax crystals. 5

Figure 1.3 Accumulation of cuticular alkylresorcinols during the development of *S. cereale* leaves. 9

Figure 1.4 Illustration of the tetraketide cyclization mechanisms catalyzed by plant type III PKSs. 11

Figure 2.1 First fully expanded leaves of *B. distachyon* (**A**) and *S. cereale* (**B**) used for wax compositional analyses. 20

Figure 2.2 Homology-based cloning strategy of gene cloning from *S. cereale*. 24

Figure 3.1 Wax compound classes from both sides of *B. distachyon* leaves. 35

Figure 3.2 Chain length distributions within compound classes in the total wax mixture from both sides of *B. distachyon* leaves. 36

Figure 3.3 GC-MS analysis of the TMSi-derivatized standard 5-*n*-nonadecylresorcinol.38

Figure 3.4 Extracted ion chromatogram (*m/z* 268) of TMSi-derivatized alkylresorcinols from the total leaf wax of *B. distachyon*. 40

Figure 3.5 Mass spectra of TMSi derivatives of individual alkylresorcinol homologs in *B. distachyon* total leaf wax. 41

Figure 3.6 Chain length distributions of alkylresorcinols in the total leaf waxes of *B. distachyon* and *S. cereale*. 41

Figure 3.7 Extracted chromatograms of fragments *m/z* 268 and *m/z* 282 in the total leaf wax of *B. distachyon*. 43

Figure 3.8 Mass spectra of TMSi-derivatized methylated alkylresorcinol homologs. .. 44

Figure 3.9 Relative abundance of methylated alkylresorcinols 19:0 and 21:0 in the total leaf wax of *B. distachyon*. 45

Figure 3.10 Possible positions of the methyl group in alkylresorcinols. 48

Figure 4.1 Phylogenetic analysis of *B. distachyon* candidate proteins related to selected type III PKSs. 53

Figure 4.2 Gene structures of *BdARS* from *B. distachyon* (**A**) and *ScARS* from *Secale cereale* (**B**). 54

Figure 4.3 Conserved nucleotide sequences of *ARS*s used to design the forward and reverse primers for the homology-based cloning of a putative *ARS* from *S. cereale*. .. 54

Figure 4.4 Phylogenetic relationships between ARSs and other related type III PKSs.. 59

Figure 5.1 TLC analysis of yeast total lipids. 66

Figure 5.2 GC-MS analysis of alkylresorcinols produced by recombinant wild type yeast expressing BdARS and ScARS. 68

Figure 5.3 Mass spectra of TMSi derivatives of individual alkylresorcinol homologs... 69

Figure 5.4 Profiles of alkylresorcinols from recombinant yeast lines. 71

Figure 5.5 RT-PCR analyses of gene expression patterns in *B. distachyon* (**A**) and *S. cereale* (**B**). 73

Figure 5.6 Quantitative RT-PCR analysis of relative gene expression levels of *ScARS* and *ScCHS* in *S. cereale* as a function of position along the first true leaf. 75

Figure 5.7 Prediction of transmembrane domains for BdARS from *B. distachyon* (**A**) and ScARS from *S. cereale* (**B**). 76

Figure 5.8 Light microscopy of transient expression of 35S:BdARS-sGFP (**A** and **C**) and 35S:ScARS-sGFP (**B** and **D**) in *N. benthamiana* leaves..... 77

Figure 5.9 Subcellular localization of BdARS (**A**) and ScARS (**B**). 78

List of Abbreviations

ARS	alkylresorcinol synthase
bp	base pair
BSTFA	bis-N,O-(trimethylsilyl)trifluoroacetamide
CER	*eceriferum*
CHS	chalcone synthase
CoA	coenzyme A
ECR	enoyl-CoA reductase
ER	endoplasmic reticulum
EST	expressed sequence tag
FAE	fatty acid elongase
FID	flame ionization detector
GC	gas chromatography
GFP	green fluorescent protein
GUS	β-glucuronidase
HCD	β-hydroxylacyl-CoA dehydratase
KCR	β-ketoacyl-CoA reductase
KCS	β-ketoacyl-CoA synthase
MS	mass spectrometry
PKS	polyketide synthase
RACE	rapid amplification of cDNA ends
STS	stilbene synthase
TLC	thin layer chromatography
TMSi	trimethylsilyl
UV	ultraviolet
VLC	very-long-chain

Acknowledgements

Apart from the efforts that I have taken, this research project would not have been possible without the support of many people. I would like to express my sincerest gratitude to my supervisor, Dr. Reinhard Jetter, who offered invaluable assistance and guidance, generous flexibility and encouragement. He also offered substantive information and contributions which helped me shape this project and materialize this thesis into its current form. Many warm and special thanks go out to the members of the supervisory committee, Dr. Carl Douglas and Dr. Ljerka Kunst, whose admirable input and excellent comments made this project go well.

My deep thanks to the past and present lab members including Dr. Zhonghua Wang, Dr. Fengling Li, Dr. Christopher Buschhaus, Yan Cao, Xiufeng Ji, Chen Peng and Mariya Skvortsova, for sharing their great advice and inspiration as well as providing a good laboratory environment. In addition, my appreciation is due to members of Dr. Ljerka Kunst's and Dr. Carl Douglas' labs, especially Lin Shi, Dr. SungSoo Kim, Dr. Etienne Grienenberger and Dr. Eryang Li, for their technical help and constructive thoughts. Also, thanks to Dr. Chris Ambrose in Dr. Geoffrey Wasteneys' lab and Dr. Etienne Grienenberger in Dr. Carl Douglas' lab for microscopy imaging. Images were obtained in the UBC Bioimaging facility and Dr. Geoffrey Wasteneys' lab.

I acknowledge the financial support given by the China Scholarship Council (CSC) and the Natural Sciences and Engineering Research Council of Canada (NSERC). Thanks to Department of Botany, the University of British Columbia, and China Agricultural University.

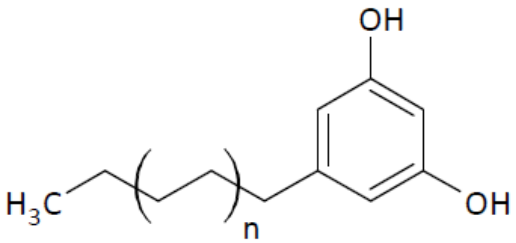
Finally, and most importantly, I would also avail myself of this opportunity to express my heartfelt thanks to my cherished friends and my beloved family. It is their constant support and endless love that pulls me through my doubt and frustration, and encourages me to keep going to the completion of this project and throughout my life.

To my dearest mother who knows me the best and has always been there for me.

Chapter 1 Introduction to alkylresorcinols and plant cuticles

1.1 Overview of alkylresorcinol occurrence, localization and biological activities

Alkylresorcinols (a.k.a. 1,3-dihydroxy-5-alkylbenzenes or 5-*n*-alkylresorcinols) are a class of polyketide-derived phenolic lipids, which have been identified in diverse plant species as well as fungi and bacteria, but only rarely in animals (Figure 1.1). They are amphiphilic compounds due to the combination of a hydrophilic resorcinol ring derived from polyketide biosynthetic pathway, and a hydrophobic alkyl chain biosynthesized during fatty acid production. Alkylresorcinols usually occur as homologous series with side chains varying in length from C_5 to C_{29} (Kozubek and Tyman, 1999). To specify the number of carbons in the side chain and the degree of unsaturation analogous to the fatty acid nomenclature, the individual homologs are typically designated as alkylresorcinol 5:0, alkylresorcinol 17:1, *etc*.



Chapter 2 Figure 1.1 **Structure of 5-*n*-alkylresorcinols (n=1, 2…13).**

Alkylresorcinols have been detected widely in plants since they were initially identified in the 1930s (Anderson *et al.*, 1931; Wasserman and Dawson, 1948). Alkylresorcinols and their derivatives so far have been identified in a wide range of species spread over twelve higher plant families (Table 1.1) (Kozubek and Tyman, 1999). For instance, in *Ginkgo biloba* (ginkgo) alkylresorcinols were identified as homologs with the alkyl side chains ranging from C_{15} and C_{17}, accumulating in leaves (27 to 87 µg/g) and the outermost covering of seeds (34 to 454 µg/g) (Żarnowska *et al.*, 2000). In *Anacardium occidentale* (cashew) alkylresorcinols were present as a mixture of homologs with the alkyl side chain of C_{15} in the shell of nuts (Wasserman and Dawson, 1948). Within the same family, a homologous series of alkylresorcinols

with the alkyl side chains ranging from C_{15} to C_{19} were ascertained in *Mangifera indica* (mango), where they were found to be restricted to fruit peels (Knödler *et al.*, 2007) and the latex oozing out from unripe fruit (Bandyopadhyay *et al.*, 1985).

Alkylresorcinols were found to be commonly present in cereal species from the grass family Poaceae, with a variety of homologs that have side chains ranging from C_{13} to C_{29}. There are numerous reports on alkylresorcinol occurrence especially in grains, where they thus appear to accumulate at relatively high concentrations (Verdeal and Lorenz, 1977; Kozubek and Tyman, 1995; Kozubek and Tyman, 1999; Ross *et al.*, 2003; Kulawinek and Kozubek, 2008). Although variations of cultivars analyzed and techniques used for extraction might have occurred, a comparison between the different studies on cereal grains showed that the highest levels of alkylresorcinols were found in *Secale cereale* (rye; 360-3200 µg/g), intermediate levels in *Triticum aestivum* (wheat; 317-1430 µg/g) and *Hordeum vulgare* (barley; 41-210 µg/g), and low levels in other species including *Sorghum bicolor* (sorghum), *Oryza sativa* (rice) and *Zea mays* (maize) (Żarnowski *et al.*, 2002; Ross *et al.*, 2004). Dissection analyses further demonstrated that the bran fraction had much higher amounts of alkylresorcinols than the flour fraction (Kozubek and Tyman, 1995; Chen *et al.*, 2004). Apart from grains, alkylresorcinols were also detected in other organs of selected cereal species, such as roots of *S. bicolor* (Cook *et al.*, 2010) and leaves of *Secale cereale* (Ji and Jetter, 2008), as well as in seedlings of *O. sativa* (Suzuki *et al.*, 1996; Suzuki *et al.*, 2003), *S. cereale* (Deszcz and Kozubek, 2000; Magnucka *et al.*, 2001), *T. aestivum* and *Z. mays* (Suzuki and Yamaguchi, 1998).

In addition to the original alkylresorcinols, some further derivatives also occur in diverse plant species, some of which have substantial biological and pharmaceutical importance. For example, in *Cannabis sativa* alkylresorcinol 5:0 serves as a first intermediate on the cannabinoid biosynthetic pathway leading to tetrahydro-cannabinol in expanding leaves and flowers (Taura *et al.*, 2007; Taura *et al.*, 2009). In *Sorghum* spp., alkylresorcinol 15:3 is the intermediate for formation of sorgoleone, a secondary metabolite that is exclusively present in exudates of root hairs (Baerson *et al.*, 2008). Glucosides of alkylresorcinols 3:0 and 5:0 were isolated from leaves of

Grevillea robusta (Yamashita *et al.*, 2008; Yamashita *et al.*, 2010). As another example, bis-5-alkylresorcinols, *i.e.* alkylresorcinols having dihydroxybenzene rings at both ends of an alkyl chain varying from C_{14} to C_{22}, were found in several plant species such as *Hakea trifurcate* (Lytollis *et al.*, 1995), *Oncostemon bojerianum* (Chaturvedula *et al.*, 2002), *Panopsis rubescens* (Deng *et al.*, 1999) and *Secale cereale* (Suzuki *et al.*, 1999).

Chapter 3 Table 1.1 **Sources of alkylresorcinols and their derivatives. Table adapted from Kozubek and Tyman (1999) and updated to present.**

Source	Family	Genus
Higher plants	Anacardiaceae	*Anacardium*, *Mangifera*, *Melanorrhoea*
	Araceae	*Monstera*, *Philodendron*
	Cannabaceae	*Cannabis*
	Compositae	*Artemisia*, *Baccharis*, *Conyza*, *Senecio*
	Cyperaceae	*Eriophorum*, *Rhynchospora*, *Trichophorum*
	Fabaceae	*Genista*, *Lathyrus*, *Ononis*, *Pisum,*
	Ginkgoaceae	*Ginkgo*
	Iridaceae	*Iris*
	Myristicaceae	*Knema*, *Myristica*, *Virola*
	Myrsinaceae	*Ardisia*, *Lysimachia*, *Rapanea*
	Poaceae	*Agropyron, Alopecurus*, *Arrhenatherum*, *Bromus*, *Dactylis*, *Elymus*, *Festuca*, *Hordeum*, *Oryza*, *Secale*, *Sorghum*, *Triticale*, *Triticum*
	Proteaceae	*Cardwellia*, *Grevillea*, *Hakea*, *Opistholepis*, *Persoonia*, *Protea*
Algae		*Apatococcus*, *Botryococcus*, *Caulocystis*, *Cystophora*
Mosses		*Lobaria*, *Physcomitrella*, *Sphaerophorus*
Fungi		*Aspergillus*, *Corticium*, *Merulius*, *Neurospora*, *Phlebia*, *Phoma*, *Pulcherricium*, *Stemphylium*, *Streptomyces*, *Verticicladiella*
Bacteria		*Azotobacter*, *Mycobacterium*, *Pseudomonas*
Animals		*Haliclona*

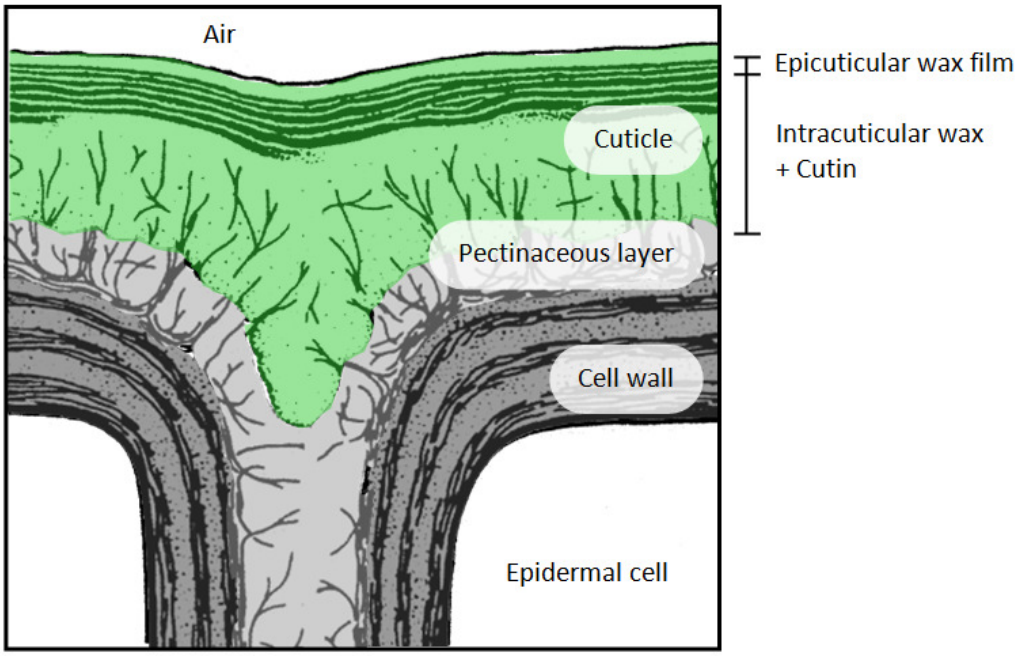
Besides plant occurrences, alkylresorcinols have been identified in microorganisms as well (Table 1.1). For instance, a mixture of alkylresorcinol homologs with the alkyl side chains ranging from C_{15} to C_{17} was reported for the fungus *Merulius incarnatus* (Jin and Zjawiony, 2006). A few bacterial species from the genera of *Azotobacter* and *Pseudomonas* were found to produce alkylresorcinol homologs in the outer layer of cysts during encystment, which is a process of transforming vegetative cells into dormant cysts for protection against adverse conditions such as desiccation and heat stress. In contrast, the bacterial alkylresorcinols reported so far have exclusively saturated alkyl chains ranging from C_{13} to C_{27} and they can accumulate up to 56,300

µg/g of dry weight (Kozubek *et al.*, 1996; Segura *et al.*, 2003; Funa *et al.*, 2006).

Potent biological and pharmaceutical activities of alkylresorcinols have been revealed in *in vitro* assays (Alonso *et al.*, 1997; Kozubek and Tyman, 1999). For instance, isolated alkylresorcinols from *Mangifera indica* (Droby *et al.*, 1986), *Hordeum vulgare* (García *et al.*, 1997) and *S. cereale* (Reiss, 1989; Suzuki and Yamaguchi, 1998) showed antifungal activity, inhibiting the growth of a range of pathogens. Extracted alkylresorcinols from *Ginkgo biloba* (Itokawa *et al.*, 1989) and *Lysimachia japonica* (Arisawa *et al.*, 1989) possessed antitumor activity. The beneficial antifungal and antibacterial activities have led to the general assumption that alkylresorcinols play a defensive role during plant growth, even though they occur as minor components. This idea has been fostered by the notion that plant alkylresorcinols seemingly always accumulate at or near the tissue surfaces, such as the shell of *Anacardium occidentale* nuts, the peel of *M. indica* fruit and the bran fraction of whole grains in cereal species (as mentioned earlier). Therefore, further information on plant surface structures will be provided in the next section.

1.2 Structure, composition and function of plant cuticle and alkylresorcinols in the cuticle

The surfaces of all primary aerial organs of land plants are covered by a hydrophobic layer called cuticle. It is produced by epidermal cells, and serves as the first barrier against multiple biotic and abiotic stresses at the plant surface (Figure 1.2). Epicuticular wax is the outermost layer of the cuticle, consisting of either a wax film or of wax crystals in different shapes protruding from a fine wax film into the surrounding environment. The layer underneath the epicuticular wax film is composed of intracuticular wax embedded within the cutin matrix. A pectinaceous layer may be present between the intracuticular wax layer and the cell wall, but the occurrence and distribution of this layer is unclear.



Chapter 4 Figure 1.2 **Schematic cross-section of a plant cuticle devoid of wax crystals. Figure adapted from Jetter *et al.* (2000).**

Cutin is the major component of the plant cuticle, accounting for 40-80% of weight of the cuticle (Heredia, 2003). It is an insoluble polymer containing C_{16} and C_{18} fatty acid monomers cross-linked via in-chain hydroxyl groups originating from hydroxylation and epoxidation, as well as a small portion of glycerol. Different from cutin, cuticular wax is soluble and can be removed from the surface by submerging the tissue into organic solvents such as chloroform, and thus used for wax analysis. The composition and percentage of wax constituents is variable depending on plant species, organs of the same species, or even different developmental stages of the same organ. Overall, the cuticular wax mixture possesses mainly aliphatic compounds as well as cyclic compounds. The aliphatics include wax compound classes of very-long-chain fatty acids (VLCFAs) and their derivatives such as primary alcohols, alkyl esters, aldehydes, alkanes, secondary alcohols and ketones. These VLC aliphatics are typically of 20 to 36 carbons, except for esters of 38 to 70 carbons (Kunst and Samuels, 2003; Jetter *et al.*, 2006; Samuels *et al.*, 2008). Besides VLC aliphatics, cuticular wax also consists of cyclic compounds such as triterpenoids and phenolic lipids (Kunst and Samuels, 2003; Jetter *et al.*, 2006; Samuels *et al.*, 2008). Triterpenoids are C_{30} hydrocarbons made up by six

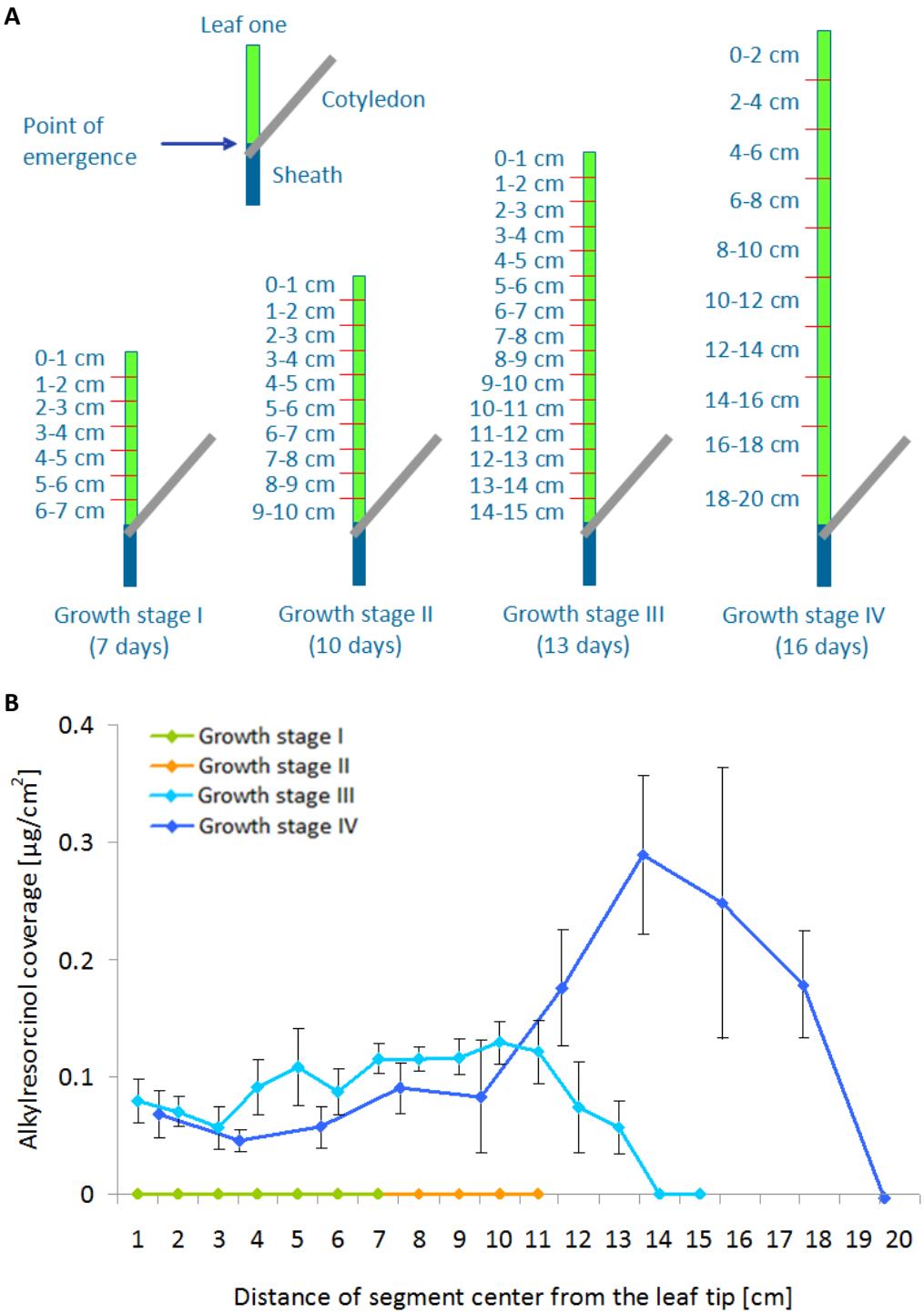
of C_5 isoprene units. In some instances, they may accumulate as dominant compounds in wax mixtures of plants (Guhling *et al.*, 2006). Phenolic lipids contain one or more phenolic groups. It should be noted that alkylresorcinols studied in the current work belong to this wax compound class.

Being the first barrier at the surface, the hydrophobic cuticle plays pivotal roles interfering between the plant and its environment. Most importantly, the plant cuticle is an effective barrier in limiting uncontrolled water loss. Additionally, it also protects against ultraviolet (UV) radiation and mechanical damages, interacts with pathogens and herbivores as well as reduces the adhesion of dust and other particles (Riederer, 2006).

The relationship between structure, composition and function of plant cuticles is still largely unknown. It needs to be well explained why certain plant species have certain wax components, how the wax components partition between epicuticular and intracuticular wax layers, and how an individual component in the designated layer contributes to the function of the cuticle. Further efforts should be made to understand the physiological and ecological roles of wax components within the plant cuticle.

As surveyed in Chapter 1.1, alkylresorcinols seemed to be localized at or near the surfaces of different plant species according to the various phytochemical investigations. More detailed studies had shown that the alkylresorcinols may well accumulate within the cuticle, as part of the cuticular wax mixture. In *H*. *vulgare* grains, alkylresorcinols were found in the wax and their antifungal activity was revealed by *in vitro* studies. However, the chain length distribution of these compounds was not determined (García *et al.*, 1997). A recent study of *S*. *cereale* leaves in our lab showed surface accumulation of alkylresorcinols in detail (Ji and Jetter, 2008). Alkylresorcinols with a chain length distribution from C_{19} to C_{27}, *i.e.* exclusively VLC homologs, were found to be restricted mainly to the cuticle of *S. cereale* leaves. The series of cuticular alkylresorcinols accounted for 3% of the total wax coverage, and 21:0 (33%), 23:0 (36%) and 25:0 (22%) were the prevalent

homologs thereof. Moreover, the deposition of these compounds was monitored over time along the leaf of *S. cereale* (Ji and Jetter, unpublished data). This time-course analysis demonstrated that cuticular alkylresorcinols were present at relatively abundant levels at growth stage IV, where they accumulated to particularly high levels in the leaf region 10-18 cm away from the tip (*i.e.* the basal region of the leaf blade 2-8 cm away from the point of emergence) (Figure 1.3). In contrast, relatively small amounts of alkylresorcinols were detected in leaves at growth stage III, and they were absent in younger leaves at growth stage I and II (Figure 1.3B). The spatial and temporal distributions showed that the series of cuticular alkylresorcinols was formed in a restricted time period relatively late during leaf development in comparison to other wax compound classes (data not shown; Ji and Jetter, unpublished data). Thus, a detailed picture of surface accumulation of alkylresorcinols and their wax context was available for further studies into the biosynthesis and function of these surface compounds in *S. cereale*.


A
Leaf one
Cotyledon
Point of emergence
Sheath
0-1 cm
1-2 cm
2-3 cm
3-4 cm
4-5 cm
5-6 cm
6-7 cm
7-8 cm
8-9 cm
9-10 cm
10-11 cm
11-12 cm
12-13 cm
13-14 cm
14-15 cm
0-2 cm
2-4 cm
4-6 cm
6-8 cm
8-10 cm
10-12 cm
12-14 cm
14-16 cm
16-18 cm
18-20 cm
Growth stage I (7 days)
Growth stage II (10 days)
Growth stage III (13 days)
Growth stage IV (16 days)
B
Growth stage I
Growth stage II
Growth stage III
Growth stage IV
Alkylresorcinol coverage [µg/cm²]
0.4
0.3
0.2
0.1
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Distance of segment center from the leaf tip [cm]

Chapter 5 Figure 1.3 **Accumulation of cuticular alkylresorcinols during the development of *S. cereale* leaves.** A**, Schematic overview of the sampling design employed for wax analyses.** B**, Distribution of alkylresorcinols along the first true leaf of *S. cereale* at four growth stages. The alkylresorcinol coverages are given as mean values (*n* = 6) ± SD. Original data adapted from Ji and Jetter (unpublished).**

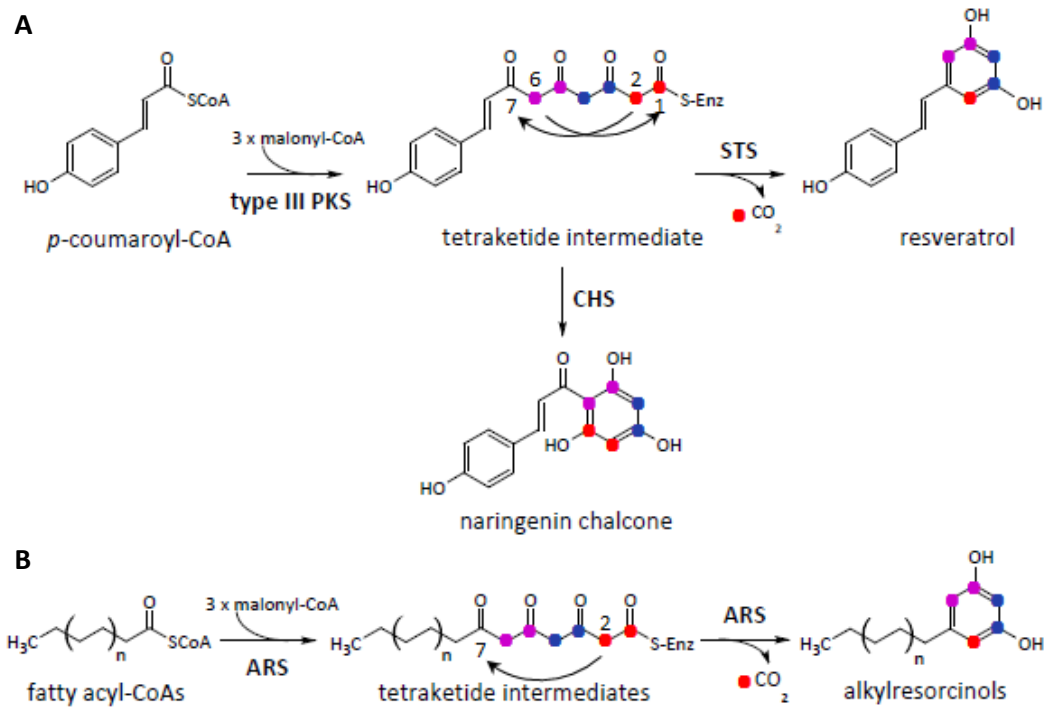
1.3 Biosynthesis of alkylresorcinols via polyketide pathway

Alkylresorcinols, as well as other types of plant phenolic lipids, are synthesized by type III polyketide synthases (PKSs). Type III PKSs are homodimeric enzymes that catalyze various cycles of decarboxylative condensation reactions with malonyl-CoA extenders to a variety of acyl-CoA starter substrates, yielding a broad spectrum of natural products (Austin and Noel, 2003).

Chalcone synthase (CHS) is the first discovered and most well-known enzyme in the family of plant type III PKSs. It catalyzes the first committed step in flavonoid biosynthesis. Accepting one molecule of *p*-coumaroyl-CoA derived from the phenylpropanoid pathway as a starter unit, CHS carries out three iterative decarboxylative condensation reactions with three molecules of malonyl-CoA as extenders, to build up a tetraketide intermediate backbone, and then performs the intramolecular cyclization via C6→C1 Claisen condensation to yield naringenin chalcone (Austin and Noel, 2003). This resulting naringenin chalcone is then modified by downstream enzymes in branching biosynthetic pathways generating flavonoids that are crucial as anthocyanins for flower pigmentation (Winkel-Shirley, 2001), as antimicrobial agents for plant defense (Cushnie and Lamb, 2005), and as UV absorptive compounds for photoprotection (Winkel-Shirley, 2002). Apart from CHS, stilbene synthase (STS) is another well-studied enzyme belonging to the type III PKSs. Contrary to CHS, which is ubiquitously present *in planta*, STS occurs only in a certain number of plants where it is involved in the biosynthesis of stilbenoids. STS acts in the same manner as CHS to produce the tetraketide intermediate, and only differs in the mechanism for the final ring folding. Instead of Claisen condensation, STS catalyzes intramolecular cyclization via C2→C7 aldol condensation, and yields resveratrol, which is modified by downstream enzymes for the biosynthesis of phytoalexin

stilbenes (Gorham, 1995; Austin *et al.*, 2004) (Figure 1.4A).

In addition to CHS and STS, other enzymes in the family of type III PKSs have been characterized over the years. Thus, it has become clear that the enzymes synthesizing alkylresorcinols belong to this family of enzymes. Alkylresorcinols are biosynthesized by alkylresorcinol synthases (ARSs) by three sequential decarboxylative condensation reactions of fatty acyl-CoA starter substrates with three molecules of malonyl-CoA extenders, and consecutive C2→C7 aldol condensation (Figure 1.4B). Historically, the formation of alkylresorcinols via polyketide biosynthesis was hypothesized a hundred years ago. In a study by Suzuki *et al*. (2003), the biosynthesis of alkylresorcinols was then investigated experimentally in seedlings of *Oryza sativa*. Feeding and labeling experiments were carried out based on predictions for 6-methylsalicylic acid biogenesis. To prove the incorporation of supplied fatty acid substrates into the corresponding alkylresorcinols, odd-numbered fatty acids with 11 to 19 carbons were used that could be easily distinguished from the endogenous fatty acids having even carbon numbers. Thus, it was shown that fatty acids (or their endogenously formed derivatives) served as direct substrates for alkylresorcinol biosynthesis, and that they formed the alkyl chains of the alkylresorcinols together with the ring carbon to which they are attached. Consistently high levels of homologous alkylresorcinols 13:0, 15:0 and 17:1 in the product mixture, regardless of the amounts of exogenous fatty acids with varying chain lengths that had been supplied, demonstrated that the ARS enzymes exhibited substrate specificity.



Chapter 6

Chapter 7 Figure 1.4 **Illustration of the tetraketide cyclization mechanisms catalyzed by plant type III PKSs.** A**, CHS catalyzes C6→C1 Claisen condensation leading to naringenin chalcone using *p*-coumaroyl-CoA as the substrate, and STS catalyzes C2→C7 aldol condensation leading to resveratrol using *p*-coumaroyl-CoA as the substrate.** B**, ARS catalyzes C2→C7 aldol condensation leading to alkylresorcinols using fatty acyl-CoAs as substrates (n=1, 2…13).**

The study of alkylresorcinol biosynthesis cannot rival that of CHS and STS, which has been undertaken for decades. In the limited studies, the proposed biosynthetic route leading to alkylresorcinols has been confirmed from a few microorganisms and several plants. The first ARS enzyme, designated as ArsB, was identified from the bacterium *Azotobacter vinelandii* in 2006 (Funa *et al.*, 2006). Using the CHS sequence from *Medicago sativa* (alfalfa) as a query, two type III PKSs were identified and one of them was characterized as an ARS enzyme using *in vitro* assays. It was shown that it transformed a series of C_{10} to C_{22} fatty acyl-CoA starter substrates into the corresponding alkylresorcinol products with C_9 to C_{21} alkyl side chains. This activity partially matched the profile of alkylresorcinols in *A. vinelandii*, where alkylresorcinols

with the alkyl side chains ranging from C_{21} and C_{23} had been found in the metabolically dormant cysts as a replacement of phospholipids. Similarly, other ARSs, including SrsA from *Streptomyces griseus* (Funabashi *et al.*, 2008) and ORAS (2'-oxoalkylresorcylic acid synthase)/PKSIII*Nc* in *Neurospora crassa* (Funa *et al.*, 2007; Goyal *et al.*, 2008), were cloned and characterized. Interestingly, *in vivo* characterization showed that SrsA used both methlymalonyl-CoA and malonyl-CoA extenders to yield methylated alkylresorcinols.

In planta, PpCHS11 isolated from *Physcomitrella patens* (moss) was found to synthesize alkylresorcinols using long-chain fatty acyl-CoAs *in vitro*, although the occurrence of alkylresorcinols in this species had not been reported (Jiang *et al.*, 2008). As mentioned in Chapter 1.1, alkylresorcinol 5:0 (a.k.a. olivetol) is the first intermediate during biosynthesis of cannabinoids, such as the well-known psychoactive compound tetrahydrocannabinol. An ARS, referred to as olivetol synthase (OLS), was cloned from *C. sativa* and assayed *in vitro*, showing that the enzyme produced alkylresorcinols 3:0 and 5:0 from short-chain C_4 and C_6 fatty acyl-CoAs (Taura *et al.*, 2009). In another case, alkylresorcinol 15:3 derived from an unusual starter substrate $16{:}3\Delta^{9,12,15}$ fatty-acyl CoA is the intermediate formed by *Sorghum* spp. root hairs for the biosynthesis of sorgoleone, an allelochemical considered as a potent inhibitor against the growth of other crop species (Baerson *et al.*, 2008; Cook *et al.*, 2010). Very recently, two ARSs were cloned from *S. bicolor*, and shown to have *in vitro* activities accepting a variety of fatty acyl-CoA starter substrates with different chain lengths from C_6 to C_{20} and different degrees of unsaturation to produce alkylresorcinols (Cook *et al.*, 2010). Similar biochemical results were also observed in the characterization of three ARSs from 2-week-old seedlings of *O. sativa* in the same study.

The bacterial ArsB from *A. vinelandii*, fungal ORAS from *N. crassa* and one plant ARS from *O. sativa* were all further characterized in a follow-up study performed by Miyanaga and Horinouchi (2009). Interestingly, the three selected ARS enzymes showed the ability of producing bis-5-alkylresorcinols that had not been discovered in previous characterizations. The bis-5-alkylresorcinols were formed via a two-step

conversion when given the appropriate substrates, which, in this case, were alkanedioic acid *N*-acetylcysteamine (NAC) dithioesters with chain lengths from C_{10} to C_{16}. NAC derivatives can mimic acyl CoAs. Nevertheless, it is unclear whether bis-5-alkylresorcinols are also *in vivo* products of ArsB in *A. vinelandii*, ORAS in *N. crassa* and one of the ARSs in *O. sativa*, as these unusual natural products have not yet been isolated from any of these species. Conversely, bis-5-alkylresorcinols were identified from a number of other plant species (see Chapter 1.1), but no ARS enzymes have been reported from them.

One thing to be noted is that, while forming alkylresorcinols from fatty acyl-CoA substrates, ARSs may also produce alkylpyrones as by-products via intramolecular lactonization. The enzymes have been shown to exhibit differential preferences for various fatty acyl-CoAs to generate alkylresorcinols versus triketide/tetraketide pyrones (Funa *et al.*, 2006; Funa *et al.*, 2007; Funabashi *et al.*, 2008; Goyal *et al.*, 2008; Cook *et al.*, 2010). In assays of ARSs from *S. bicolor*, 9% of triketide pyrones derived from C_8 to C_{14} fatty acyl-CoAs were identified as derailment products within the total products, dominated by triketide pyrones with C_7, C_9 and C_{11} side chains. However, alkylresorcinols with C_5 to C_{19} side chains were found to be the major products, and their profile was dominated by the C_{11}, C_{13} and C_{15} homologs (Cook *et al.*, 2010). Moreover, even for the same enzyme, different *in vitro* studies can lead to different results. Funa *et al.* (2007) found that ORAS from *N. crassa* produced alkylresorcylic acids as the prevalent products, accompanied by minor amounts of alkylresorcinols, triketide and tetraketide pyrones with varying chain lengths. In contrast, Goyal *et al.* (2008) showed that the same enzyme (named PKSIII*Nc* in their study) produced mainly alkylresorcinols in addition to triketide and tetraketide pyrones. This discrepancy might have resulted from different conditions and methods applied during biochemical analyses. It also indicated that type III PKSs can produce unnatural novel products if using non-physiological conditions and/or starter substrates during biochemical characterization.

Even though breakthroughs have been achieved regarding the biosynthetic route leading to alkylresorcinols in a few species, many details in the biosynthesis of plant

alkylresorcinols are still unclear. First, although several microbial ARS enzymes have been functionally characterized, these sequences share only low identity with their plant counterparts. Second, it is still pending whether the same pathway, as shown for a few species, is indeed occurring throughout the plant kingdom. Third, it is not clear whether differences in chain length profiles of alkylresorcinols occurring in various plant species and organs are due to differences in substrate availability or in enzyme specificity. So far, the characterized plant ARS enzymes have shown distinct specificities for short- to long-chain substrates, and they are, therefore, thought to be responsible for synthesizing distinct alkylresorcinols in their respective contexts. For instance, the OLS expressed in expanding leaves and flowers of *C. sativa* is the enzyme likely involved in the formation of the short-chain alkylresorcinol 5:0 en route to tetrahydrocannabinol (Taura *et al.*, 2009). The two ARSs preferentially expressed in root hairs of *S. bicolor* are the enzymes likely producing the intermediate alkylresorcinol 15:3 during sorgoleone biosynthesis (Cook *et al.*, 2010). However, the alkylresorcinols found in the cuticle of *Secale cereale* leaves have exclusively very long side chains, which by analogy should be biosynthesized from the VLC fatty acyl-CoA starter substrates. It is not clear whether such VLC alkylresorcinols are formed by specific ARSs with narrow VLC substrate preference, by low-specificity ARSs that accept VLC substrates in addition to starter substrates with shorter chains, or even by STSs that are promiscuous enough to accept aliphatic CoAs with widely varying chain lengths in addition to the aromatic CoA substrates. In this context, the availability of VLC fatty acyl CoAs as starter substrates for potential ARSs is of central importance. Thus, the following section will describe the formation and localization of available pools of VLC fatty acyl-CoAs.

1.4 Biosynthesis of VLC aliphatics

VLC aliphatic compounds are originally derived from VLC fatty acyl-CoAs during fatty acid biosynthesis that occurs ubiquitously in plants. The elongation of VLCFAs happens at the endoplasmic reticulum (ER) where VLC C_{20} to C_{36} fatty acyl-CoAs are formed by elongation of preexisting C_{16} and C_{18} fatty acyl-CoAs after export from the plastids.

The elongation of VLCFAs is catalyzed by a multi-enzyme complex (fatty acid elongase, FAE) at the ER. It is a four-step reaction proceeding in iterative cycles by introducing additional C_2 units derived from malonyl-CoAs to the current saturated fatty acyl chain by four enzymes. First, a fatty acyl-CoA is condensed with a malonyl-CoA to form a β-ketoacyl-CoA by β-ketoacyl-CoA synthase (KCS). Second, the resulting β-ketoacyl-CoA is reduced to a β-hydroxyacyl-CoA by β-ketoacyl-CoA reductase (KCR). Third, the β-hydroxyacyl-CoA is dehydrated to an enoyl-CoA catalyzed by β-hydroxyacyl-CoA dehydratase (HCD), and at last the enoyl-CoA is reduced again to a final saturated acyl-CoA two carbons longer than before by enoyl-CoA reductase (ECR) (Kunst and Samuels, 2003). Different from the other three enzymes, the condensing enzyme KCS is the rate-limiting enzyme that has strict substrate/product specificity and therefore determines the final chain length of the fatty acyl-CoA (Millar and Kunst, 1997). In *Arabidopsis thaliana*, 21 KCSs have been identified (Costaglioli *et al.*, 2005), but only a few have been described in detail (Reviewed in Joubes *et al*., 2008). Among them, only KCS6/CER6 has been shown to be specific for the biosynthesis of VLCFAs during wax production (Millar *et al.*, 1999).

It is to be noted that both KCS in VLCFA biosynthesis and ARS in alkylresorcinol biosynthesis are condensing enzymes, showing great similarity in their biosynthetic mechanisms. In both cases, elongation occurs by adding C_2 units derived from malonyl-CoAs into the linear fatty acyl-CoA chains. KCSs carry out only one condensation reaction to form diketides, with the further involvement of three other enzymes in every VLCFA elongation cycle. In contrast, ARSs catalyze three consecutive condensation reactions to tetraketides, without releasing the intermediate products for use by other enzymes, and further cyclize the products into the final structures.

After elongation by FAEs, the VLC fatty acyl-CoAs are modified by different biosynthetic machineries into diverse components needed in various contexts. The diversity of functional specialization is due to variation in chain lengths of VLCFAs in different organs and tissues in plants. In general, VLCFAs of prevailing chain lengths from C_{20} to C_{26} are the components of membrane lipids in all cells. VLCFAs of C_{20} and

C_{22} accumulate as triacylglycerols in storage lipids specifically in seeds. In contrast, VLCFAs from C_{20} to C_{32} are used for suberin synthesis, to form a hydrophobic barrier against pathogens, water and nutrient transport, both in roots and in wounded tissues. It should be noted that all three contexts described so far contain VLCFAs in relatively small concentration, as compared to normal fatty acids. Consequently, the pools of VLC fatty acyl-CoAs present in these cases are likely also relatively small. In contrast, epidermal cells of all aerial organs are highly specialized for the formation of cuticle components, including large quantities of VLC C_{20} to C_{36} wax compounds needed to limit uncontrolled water loss, protect against UV radiation and moderate pathogen and herbivore behavior. It has been estimated that, of the VLCFAs formed in epidermal cells during tissue growth, *ca.* 50% are dedicated to elongation and wax synthesis (Suh *et al.*, 2005). Thus, epidermal cells can be expected to contain the largest pools of VLC fatty acyl CoAs available in any plant tissue.

There are two downstream biosynthetic pathways occurring in epidermal cells, in which VLCFA precursors are further modified into different wax compound classes. The acyl reduction pathway produces even-numbered aliphatics including primary alcohols and alkyl esters, while the decarbonylation pathway yields odd-numbered alkanes, secondary alcohols and ketones (Kunst and Samuels, 2003). It seems plausible that the large pools of VLC fatty acyl-CoAs present in epidermal cells may also serve as substrates for other parallel pathways, including the formation of cuticular VLC alkylresorcinols.

1.5 Research objectives and questions

The current work was based on the central hypothesis that there is/are epidermis-specific ARS/s, differing from CHSs and STSs, involved in the biosynthesis of cuticular alkylresorcinols by preferentially accepting VLCFA precursors as substrates that are formed during wax production, and the biological function of the resulting VLC alkylresorcinols is to form a first line of defense at/near the plant surface. This model is synthesizing a number of smaller hypotheses, to each of which alternatives may be formulated: The enzymes involved may not be restricted to the epidermis;

they might have mainly CHS or STS activity, with minor side activity for alkylresorcinol formation; they may be ARSs that prefer short- to long-chain fatty acyl-CoA starter substrates, but have weak side activity on VLC substrates; the resulting alkylresorcinols may have roles in locations other than the cuticle; they may reach the cuticle only accidentally.

In order to test the above hypotheses, the current research was to study the biosynthesis of cuticular alkylresorcinols in selected plant species. Based on the previous evidence showing the surface accumulation of alkylresorcinols, *Secale cereale* was used for further studies here. Additionally, *Brachypodium distachyon*, which had recently emerged as a powerful and attractive experimental organism for grass studies with the accomplishment of genome sequencing (2010), was employed as another model system. Provided that cuticular alkylresorcinols are present in *B. distachyon*, this species may become a potential tool that facilitates the identification and characterization of enzymes involved in the biosynthesis of alkylresorcinols also in the closely related species *S. cereale*. To study alkylresorcinol biosynthesis in the two grass species, a set of parallel experiments were conducted using analytical techniques such as gas chromatography (GC) and thin layer chromatography (TLC) as well as biological techniques such as gene cloning, enzyme characterization and gene expression studies. In particular, the following seven questions were addressed.

In Chapter 3, wax analysis of *B. distachyon* leaves was performed to address:

a) Are alkylresorcinols present in *B. distachyon* total leaf wax mixture? How do they compare to those in *S. cereale*?

The finding of alkylresorcinols in *B. distachyon* total leaf wax in the current work together with the previous data on *S. cereale* provided the basic chemical information needed before proceeding with the following experiments. In their course, possible correlations between cuticular alkylresorcinol accumulation and the characteristics of candidate genes had to be studied.

To this end, in Chapter 4, cloning of putative ARS genes in *B. distachyon* and *S. cereale*

was carried out to answer:

b) Are ARS genes present in *B. distachyon* and *S. cereale*, respectively?

The activities of the enzymes encoded by the putative *ARS*s were tested in Chapter 5 by biochemical characterization in yeast *Saccharomyces cerevisiae* to reveal:

c) Do the putative enzymes indeed have ARS activities?

d) What are the substrate/product profiles correspondingly?

In order to explore the relatedness of functional ARSs and cuticular alkylresorcinol accumulation, biological characterization was performed as well in Chapter 5, including gene expression patterns and subcellular localization of the ARSs to answer the following questions:

e) What are the organ-specific expression patterns of the *ARS*s in *B. distachyon* and *S. cereale*, respectively?

f) What is the spatial expression pattern of *ScARS* in *S. cereale* along the leaf?

g) What is the subcellular localization of BdARS and ScARS?

Chapter 2 Materials and methods

2.1 Plant materials and growth conditions

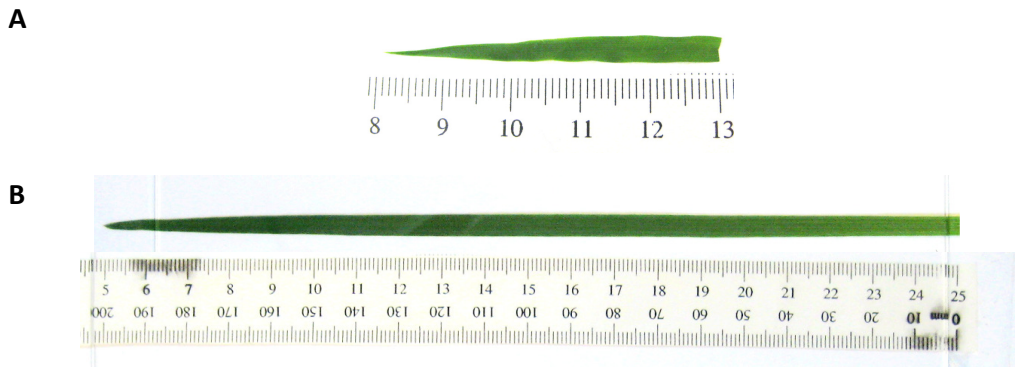
7.1.1 Plant growth conditions

Grains of *Brachypodium distachyon* (L.) P. Beauv. were ordered from National Germplasm Resources Laboratory, Beltsville, Maryland, United States. Grains of *Secale cereale* L. cv. Esprit were purchased from Capers, Vancouver, Canada. Seeds of *Nicotiana benthamiana* were from Dr. Etienne Grienenberger (Douglas lab, Department of Botany, the University of British Columbia). Grains of *B. distachyon* and *S. cereale* were soaked in water at room temperature overnight and grown either in pots containing moistened Sunshine Mix #4 (Sun Gro Horticulture Canada Ltd) to get aboveground materials for wax analyses, cloning and gene expression studies, or on Murashige and Skoog agar plates to get root material for root analysis and gene expression investigations. Seeds of *N. benthamiana* were sown directly in pots containing moistened Sunshine Mix #4 and grown for transient expression by agrobacterium-mediated infiltration. All plants described above were stratified for 3-4 days at 4°C and then moved to a growth chamber and grown at 22°C with a 20-h photoperiod (approximately 120 µmol m^{-2} s^{-1}) and a relatively humidity of 70%. One batch of *S. cereale* plants was grown at 22°C and 70% humidity, but in darkness for a comparative investigation of gene expression.

7.1.2 Plant materials

For wax compositional analysis of *B. distachyon*, intact leaves were used. The first fully expanded leaf typically around 5-cm long was harvested three weeks after germination under the growth condition described above. The entire leaf was subjected to wax extraction. Surface areas were calculated by digital photographs with ImageJ software and multiplied by 2 for total leaf surface areas (Figure 2.1A). Seven independent parallels were analyzed. The first fully expanded leaf of *S. cereale* around 20-cm long was harvested three weeks after germination and subjected to wax extraction for a comparative analysis (Figure 2.1B). Three independent parallels were analyzed. Roots of *S. cereale* were collected one week after grains had been

sown on agar plates, by removing agar, washing with water and air drying the material.



Chapter 8 Figure 2.1 **First fully expanded leaves of *B. distachyon*** (A) **and *S. cereale*** (B) **used for wax compositional analyses.**

For semi-quantitative RT-PCR analysis, leaves, stems, spikes and roots of *B. distachyon*, and cotyledons, sheaths, green leaves grown under normal light condition, etiolated leaves grown without light treatment and roots of *S. cereale* were harvested separately. For quantitative RT-PCR analysis, the first unfolded leaf of *S. cereale* grown under normal condition was used. Leaves exactly 20-cm long were harvested for sampling. Collected leaves were cut into 2-cm long segments, and the corresponding segments from 6-8 plants were pooled together into one sample.

Except for the stems and spikes harvested from *B. distachyon* and roots collected from *B. distachyon* and *S. cereale*, all other materials of *B. distachyon* and *S. cereale* used for wax analyses, gene cloning and expression studies were obtained consistently at an early point of the vegetative growth before the jointing stage, at which the stems start elongating from the leaves during grass development unless specified.

For subcellular localization studies, 4-week-old leaves of *N. benthamiana* were used for agrobacterium-mediated infiltration to initiate transient expression. Plants were then placed back into the growth chamber under the same conditions as described

earlier for 3-4 days, before the transformed leaves were examined using light microscopy and confocal microscopy.

2.2 Wax analyses

2.2.1 Wax extraction and derivatization

Total wax extraction was carried out by submerging an intact leaf of *B. distachyon* or *S. cereale* twice for 30 s in chloroform containing defined amounts of *n*-tetracosane and 5-*n*-tridecylresorcinol as internal standards. Roots of *S. cereale* were analyzed following the same extraction procedure. The resulting solutions were concentrated under a gentle stream of N_2 gas while heating to 50°C and transferred to sample vials. There, the solvent was removed under N_2 gas.

Samples were derivatized with bis-N,O-(trimethylsilyl) trifluoroacetamide (BSTFA; Sigma-Aldrich) in pyridine (1:1, v/v) at 70°C for 60 min to transform all hydroxyl-containing compounds into the corresponding trimethylsilyl (TMSi) derivatives. The solvents were evaporated, and chloroform was added to the samples again prior to quantitative analyses by gas chromatography-flame ionization detector (GC-FID) and identification by gas chromatography-mass spectrometry (GC-MS).

2.2.2 Chemical analyses using GC-FID and GC-MS

The qualitative wax composition was determined with capillary GC (5890N, Agilent, Avondale, PA; column 30 m HP-1, 0.32 mm i.d., df=0.1 µm, Agilent) using temperature-programmed on-column injection at 50°C, oven for 2 min at 50°C, raised by 40°C min^{-1} to 200°C, held for 2 min at 200°C, raised by 3°C min^{-1} to 320°C, and held for 30 min at 320°C. Qualitative analyses were carried out by GC with a mass spectrometric detector (5973N, Agilent) and He carrier gas inlet pressure programmed for a constant flow of 1.4 ml min^{-1}. Individual compounds were identified by comparison of characteristic fragments with those of authentic standards and literature data. Quantitative analyses were carried out by GC with a flame ionization detector (FID) and H_2 carrier gas inlet pressure programmed for a constant flow of 2.0 ml min^{-1}. Wax loads were determined by comparing GC-FID peak

areas against the internal standards and dividing by the surface area of the samples. Alkylresorcinol homologs were quantified against the synthesized internal standard 5-*n*-tridecylresorcinol according to the relative abundance of the characteristic fragment (*m/z* 268) in GC-MS runs. Other individual compounds were quantified against the other internal standard *n*-tetracosane by integrating peak areas in GC-FID runs. All quantitative data are given as means of parallel experiments and standard errors.

2.3 Genomic DNA extraction, RNA isolation and reverse transcription

2.3.1 Genomic DNA preparation and RNA isolation

Young leaves of *B. distachyon* and *S. cereale* were harvested for genomic DNA extraction using DNeasy Plant Mini Kit (Qiagen). Different organs of *B. distachyon* and *S. cereale* (see Chapter 2.1.2) were prepared for total RNA extraction using either Trizol Reagent (Invitrogen) or RNeasy Plant Mini Kit (Qiagen). For gene expression analysis using quantitative RT-PCR, RNA samples from 2 cm-long segments along the leaf of *S. cereale* were extracted using RNeasy Plant Mini Kit (Qiagen). During purification, RNA was treated with RNase-Free DNase (Qiagen) for DNase digestion to avoid any genomic DNA contamination.

2.3.2 Reverse transcription

Before reverse transcription, RNA concentration was measured using a NanoDrop 8000 Spectrophotometer (Thermo Scientific). After quantification, 2 µg of total RNA was subjected to reverse transcription with oligo (dT) primer for first-strand cDNA synthesis by SuperScript II Reverse Transcriptase (Invitrogen) at 42°C for 60 min. The resulting cDNA was subsequently used as template in PCR reactions.

2.4 Cloning of putative alkylresorcinol synthase (ARS) genes

2.4.1 Cloning of *BdARS* from *B. distachyon*

Taking advantage of the availability of the complete genome sequence information of *B. distachyon*, Brachypodium expressed sequence tag (EST) libraries were mined and

analyzed by BLASTN and TBLASTN programs to identify potential genes encoding ARSs belonging to the type III polyketide synthase (PKS) family in *B. distachyon* using the functionally characterized plant ARS sequences (two ARSs from *Sorghum bicolor* and three ARSs from *Oryza sativa*) as queries. The putative *BdARS* obtained from the EST libraries was PCR-amplified using leaf cDNA of *B. distachyon* as template with the gene-specific N-terminal primer 5'-CGCGGATCCATGACAAGAGCTAACGGTAAC-3' (*Bam*HI site underlined) and C-terminal primer 5'-AGACTCGAGCTAATTTCCCTTGAGAC CCGG-3' (*Xho*I site underlined). The PCR conditions were: 98°C for 30 s, 35 cycles of 98°C for 15 s, 62°C for 30 s, and 72°C for 60 s, and 72°C for 5 min with Phusion High-Fidelity DNA Polymerase (New England Biolabs). The resulting PCR product was gel-purified using QIAquick Gel Extraction Kit (Qiagen) and directly sent for sequencing to confirm the sequence information. The gene structure was investigated by comparing the cDNA sequence with the genomic DNA sequence from the Brachypodium Database.

2.4.2 Cloning of *ScARS* from *S. cereale*

Gene cloning from *S. cereale* was performed using a homology-based cloning approach (Figure 2.2). Two *ARS*s from *Sorghum bicolor* and three *ARS*s from *O. sativa* were used for sequence alignment. Different primer sets were designed accordingly from the consensus regions conserved between the functionally characterized *ARS*s to amplify the target *ARS* from *Secale cereale*. A core fragment of *ScARS* was amplified using leaf cDNA of *S. cereale* and the forward primer 5'-TCGCCATCGGCACGGCAAACC-3' as well as the reverse primer 5'-GTGATTCCAGGTCC GAAGGCCA-3' under the following PCR conditions: 94°C for 3 min, 35 cycles of 94°C for 15 s, 58°C for 30 s, and 72°C for 70 s, and 72°C for 10 min with *Taq* DNA Polymerase (Invitrogen). The resulting PCR product was gel-purified using QIAquick Gel Extraction Kit, cloned into pCR-Blunt vector using Zero Blunt PCR Cloning Kit (Invitrogen) and transformed into TOP10 chemically competent *E.coli* cells. Plasmid DNA was purified using GeneJET Plasmid Miniprep Kit (Fermentas) and sent for sequencing.

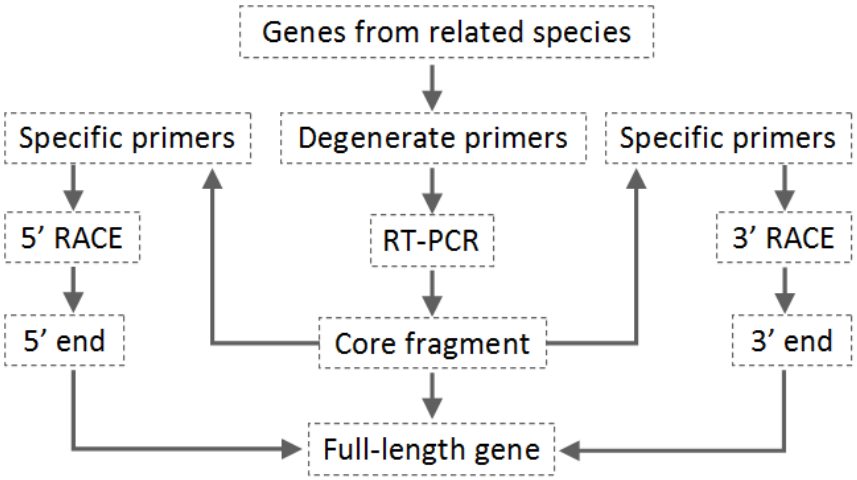


Figure 2.2 Homology-based cloning strategy of gene cloning from *S. cereale*.

Rapid Amplification of cDNA Ends (RACE) was used to amplify the 5'- and 3'-ends from the core fragment. For 3'-end amplification, cDNA was synthesized with Adapter Primer (AP) 5'-GGCCACGCGTCGACTAGTACTTTTTTTTTTTTTTTTT-3' using SuperScript II Reverse Transcriptase at 42°C for 60 min. The forward gene-specific primer 5'-TCGAGATGGTCCACGCCACGCAGAC-3' and the Abridged Universal Amplification Primer (AUAP) 5'-GGCCACGCGTCGACTAGTAC-3' were used in the first PCR reaction under the conditions: 98°C for 30 s, 35 cycles of 98°C for 15 s, 65°C for 30 s, and 72°C for 30 s, and 72°C for 5 min using Phusion High-Fidelity DNA Polymerase. A nested-PCR was performed directly using the first round PCR product as a template with the nested gene-specific forward primer 5'-AACGACCTCTTCTGGGCGGTGCAC-3' and the AUAP under the same conditions as in the first round of PCR. The resulting 3'-end PCR product was gel-purified, cloned into pCR-Blunt vector and sequenced as described above. For 5'-end amplification, cDNA was synthesized using a gene-specific reverse primer 5'-GAAGATCGGGCTCTGCTC-3' with SuperScript II Reverse Transcriptase at 42°C for 60 min. The gene-specific primer 5'-GAAGCAGACGA GCGTGAT-3' and Abridged Anchor Primer (AAP) 5'-GGCCACGCGTCGACTAGTACGGGIIG GGIIGGGIIG-3' were used in the first PCR reaction, following the amplification conditions: 94°C for 3 min, 35 cycles of 94°C for 15 s, 55°C for 30 s, and 72°C for 70 s, and 72°C for 10 min. Another round of PCR was performed for the nested

amplification with gene-specific primer 5'-GTAGGTGCTGAAGACGAGGTGGGTG-3' and AUAP under the conditions: 98°C for 30 s, 35 cycles of 98°C for 15 s, 60°C for 30 s, and 72°C for 30 s, and 72°C for 5 min. The resulting PCR product was sequenced as described above.

The corresponding full-length cDNA was amplified using the gene-specific N-terminal primer 5'-CGCGGATCCATGGGAAGCATAGGAACCACC-3' (*Bam*HI site underlined) and C-terminal primer 5'-AGACTCGAGCTAGCGTGGACAGCGGAGGAC-3' (*Xho*I site underlined). The PCR was carried out with Phusion High-Fidelity DNA Polymerase under the following conditions: 98°C for 30 s, 35 cycles of 98°C for 15 s, 60°C for 30 s, and 72°C for 70 s, and 72°C for 5 min. The resulting PCR product was gel-purified and cloned for sequencing to confirm the sequence information. In order to study the gene structure, N- and C-terminal primers designed for sequence confirmation in the previous PCR run were introduced again to amplify the gene from genomic DNA under the conditions: 98°C for 30 s, 35 cycles of 98°C for 15 s, 60°C for 30 s, and 72°C for 90 s, and 72°C for 5 min. The resulting PCR product was gel-purified and cloned for sequencing.

2.5 Phylogenetic analyses

Sequence alignments and phylogenetic analyses were performed with the ClustalX version 2.0 based on the neighbor-joining (N-J) method (Thompson *et al.*, 1997) using selected amino acid sequences of cloned and/or characterized type III PKSs. Phylogenetic trees were generated by MEGA4.0. The number of bootstrap replications was 1,000 and the indicated scale represents 0.05 amino acid substitutions per site.

Amino acid sequences of putative plant type III PKSs in *B. distachyon* were retrieved from Brachypodium ESTs (http://blast.brachypodium.org). The final candidate list contained nine full-length sequences of type III PKSs in Brachypodium EST libraries with the following gene identifiers: *B. distachyon* ARS (Bradi4g28070.1), *B. distachyon* chalcone synthase (CHS) (Bradi4g17230.1), *B. distachyon* chalcone synthase-like1

(CHSL1) (Bradi3g29230.1), *B. distachyon* CHSL2 (Bradi1g52580.1), *B. distachyon* CHSL3 (Bradi4g24780.1), *B. distachyon* CHSL4 (Bradi1g48200.1), *B. distachyon* CHSL5 (Bradi1g25920.1), *B. distachyon* CHSL6 (Bradi1g50710.1) and *B. distachyon* CHSL7 (Bradi1g12730.1). It is to be noted that Bradi4g28070.1 was biochemically characterized and designated as BdARS in this work. Bradi4g17230.1 was annotated as BdCHS based on the phylogenetic relationship in this work.

In addition to the nine type III PKSs from *B. distachyon* and the putative ARS from *S. cereale* obtained in this work, all other sequence data of plant type III PKSs and bacterial ARSs used in the phylogenetic analyses can be found in GenBank library under the following accession numbers: *Arabidopsis thaliana* CHS/TT4 (P13114), *A. thaliana* CHSL1/PKSA (NP_171707), *A. thaliana* CHSL2/PKSB (NP_567971), *A. thaliana* CHSL3 (NP_191915), *Arachis hypogaea* STS (P20178), *A. hypogaea* CHS (AAO32821), *Azotobacter vinelandii* ArsB (YP_002800096), *Cannabis sativa* OLS (BAG14339), *C. sativa* CHS (AAL92879), *Gerbera hybrida* 2PS (P48391), *G. hybrida* CHS1 (P48390), *Hordeum vulgare* CHS1 (P26018), *H. vulgare* CHS2 (Q96562), *Medicago sativa* CHS (P30074), *O. sativa* ARS1 (AAT44238), *O. sativa* ARS2 (NP_001064197), *O. sativa* ARS3 (AAN04188), *O. sativa* CHS1 (ABA94123), *O. sativa* CHS2 (NP_001059187), *O. sativa* CHSL1 (NP_001052003), *O. sativa* CHSL2 (BAD53112), *O. sativa* CHSL3 (BAD31062), *O. sativa* CHSL4 (NP_001068109), *O. sativa* CHSL5 (EEE60711), *O. sativa* CHSL6 (NP_001064891), *O. sativa* CHSL7 (NP_001059449), *O. sativa* CHSL8 (NP_001068007), *O. sativa* CHSL9 (NP_001068006), *O. sativa* CHSL10 (NP_001068008), *O. sativa* CHSL11 (NP_001059829), *O. sativa* CHSL12 (NP_001059830), *O. sativa* CHSL13 (NP_001059828), *O. sativa* CHSL14 (NP_001059345), *O. sativa* CHSL15 (NP_001059719), *O. sativa* CHSL16 (AAM01009), *O. sativa* CHSL17 (AAT44239), *O. sativa* CHSL18 (NP_001054922), *O. sativa* CHSL19 (AAT47098), *O. sativa* CHSL20 (AAL77133), *O. sativa* CHSL21 (EAZ24017), *O. sativa* CHSL22 (AAM01005), *Physcomitrella patens* ARS (EF593132), *P. patens* CHS (ABB84527), *Pinus sylvestris* CHS (P30079), *P. sylvestris* STS (AAB24341), *S. cereale* CHS1 (P53414), *S. cereale* CHS2 (P53415), *Sorghum bicolor* ARS1 (XP_002441839), *S. bicolor* ARS2 (XP_002449744), *S. bicolor* CHS1 (XP_002450874), *S. bicolor* CHS2 (Q9SBL7), *S. bicolor* CHS3 (XP_002450875), *S. bicolor* CHS4 (XP_002450870), *S. bicolor* CHS5 (Q9SBL4), *S. bicolor*

CHS6 (XP_002450877), *S. bicolor* CHS7 (XP_002450876), *S. bicolor* CHSL1 (XP_002450871), *S. bicolor* CHSL2 (XP_002449616), *S. bicolor* CHSL3 (XP_002445139), *S. bicolor* CHSL4 (XP_002461886), *S. bicolor* CHSL5 (XP_002467058), *S. bicolor* CHSL6 (XP_002450684), *S. bicolor* CHSL7 (XP_002450661), *S. bicolor* CHSL8 (XP_002450864), *S. bicolor* CHSL9 (XP_002449608), *S. bicolor* CHSL10 (XP_002449615), *S. bicolor* CHSL11 (XP_002449610), *S. bicolor* CHSL12 (XP_002449614), *S. bicolor* CHSL13 (XP_002449602), *S. bicolor* CHSL14 (XP_002450859), *S. bicolor* CHSL15 (XP_002450861), *S. bicolor* CHSL16 (XP_002457597), *S. bicolor* CHSL17 (XP_002441718), *S. bicolor* CHSL18 (XP_002459295), *S. bicolor* CHSL19 (XP_002462898), *S. bicolor* CHSL20 (XP_002445689), *S. bicolor* CHSL21 (XP_002462896), *S. bicolor* CHSL22 (XP_002462897), *S. bicolor* CHSL23 (XP_002452260), *S. bicolor* CHSL24 (XP_002454000), *Streptomyces griseus* SrsA (YP_001821984), *Triticum aestivum* CHS1 (AAQ19322), *T. aestivum* CHS2 (AAQ19323), *Vitis vinifera* CHS (CAA53583), *V. vinifera* STS (CAA54221), *Zea mays* CHS1/WHP1 (P24824), *Z. mays* CHS2/C2 (P24825), *Z. mays* CHSL1 (NP_001149022), *Z. mays* CHSL2/C2-Idf-III (AAW56963), *Z. mays* CHSL3 (ACF87981), *Z. mays* CHSL4 (NP_001131211), *Z. mays* CHSL5 (ACF85939), *Z. mays* CHSL6 (NP_001150611), *Z. mays* CHSL7 (NP_001149508), *Z. mays* CHSL8 (NP_001149157), and *Z. mays* CHSL9 (NP_001140848). It is to be noted that for *Arabidopsis thaliana*, *B. distachyon*, *O. sativa*, *S. bicolor* and *Z. mays*, all potential type III PKSs were included based on the search in the genome databases. Except for the ARS and CHS sequences that had been functionally characterized and reported previously, all other sequences were annotated as CHSL sequences. For other species including *Arachis hypogaea*, *C. sativa*, *G. hybrida*, *H. vulgare*, *M. sativa*, *P. patens*, *Pinus sylvestris*, *Secale cereale*, *T. aestivum* and *V. vinifera*, representative ARS, CHS, STS and 2PS sequences that had been characterized and reported earlier were used. However, more potential type III PKSs in these species may occur.

2.6 Heterologous expression of ARSs in yeast *Saccharomyces cerevisiae*

2.6.1 Yeast strains, transformation and expression

Different yeast strains used in this work were W303-1A (*MATα ade2-1 his3-11,15*

leu2-3,112 trp1-1 ura3-1 can1-100) served as wild type yeast, and BY4743 (*MATα/α his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 lys2Δ0/LYS2 MET15/met15Δ0 ura3Δ0/ura3Δ0 YLR372w::kanMX4/YLR372w::kanMX4*) served as the yeast mutant *elo3*Δ. The latter showed a defect in synthesizing C_{24} fatty acyl chain and beyond, and had elevated levels of C_{20} and C_{22} fatty acyl constituents instead. For heterologous expression in different yeast strains, the full-length cDNAs of *BdARS* and *ScARS* were digested with *Bam*HI and *Xho*I, and ligated into the yeast expression vector pESC-URA (Stratagene) under the control of inducible *GAL1* promoter. The resulting constructs were transformed into either wild type yeast or the yeast mutant *elo3*Δ using the LiAc/SS-DNA/PEG method (Gietz and Woods, 2002). Transformants were screened on synthetic dextrose minimal medium agar plates lacking uracil at 30°C for 2-3 days.

2.6.2 Yeast lipid extraction and analyses

The recombinant yeast cells were cultivated in synthetic dextrose minimal medium at 30°C overnight. The cells were harvested and resuspended in synthetic galactose minimal medium lacking uracil at 30°C and shaking at 200 rpm. After galactose induction for 2 days, yeast cells were harvested by centrifugation. Total lipids were extracted with chloroform-methanol (2:1, v/v; 20 volume) and washed with 0.9% NaCl (w/v; 0.2 volume) (Schneiter and Daum, 2006). Following phase separation, the chloroform phase was transferred to a new tube and evaporated to dryness under a gentle stream of N_2 gas while heating to 50°C. The lipid extracts were derivatized with BSTFA in pyridine at 70°C for 60 min. The solvents were evaporated, and chloroform was added to the samples again prior to chemical analyses using GC-FID and GC-MS as described above for wax analyses (Chapter 2.2.2).

2.6.3 Thin layer chromatography (TLC) analyses

Total yeast lipid extracts were separated on TLC plates (20 × 20 cm, silica gel 60 F_{254}, 0.25mm; Merck) using chloroform-ethyl acetate (7:3, v/v) as the mobile phase in the sandwich technique, stained with primuline (Sigma) and visualized under ultraviolet (UV) light. Synthetic 5-*n*-tridecylresorcinol was used as a standard. The bands corresponding to metabolic alkylresorcinols were scraped off from plates, extracted with chloroform, filtered, and prepared for GC analyses.

2.7 Semi-quantitative RT-PCR and quantitative RT-PCR analyses

2.7.1 Semi-quantitative RT-PCR

Gene-specific forward primer 5’-CGACCAGTTCTTCCGCGTGACC-3’ and reverse primer 5’-GATGGCTGGTCTCGGTCGAGGA-3’ were used to amplify a fragment of *BdARS* using cDNA templates derived from leaves, stems, spikes and roots of *B. distachyon*. Additionally, *18S rRNA* was introduced as a control using the gene-specific forward primer 5'-CCGTCCTAGTCTCAACCATAAAC-3' and reverse primer 5'-CCTTTAAGTTTCAGC CTTGCG-3’.

Information on primers targeting *Actin* sequences in Poaceae was obtained from a previous study (Paquet *et al.*, 2005). The forward primer 5’-AACTGGGATGATATGGAG AA-3' and reverse primer 5'-CCTCCAATCCAGACACTGTA-3' were used for amplification of a fragment of *Actin* in *S. cereale* using cDNA of green leaves with Phusion High-Fidelity DNA Polymerase under the following conditions: 98°C for 30 s, 30 cycles of 98°C for 15 s, 53°C for 30 s, and 72°C for 30 s, and 72°C for 5 min. The resulting PCR product was sequenced directly after gel purification. Gene-specific primers for the *ScActin* from *S. cereale* were subsequently designed and used in semi-quantitative RT-PCR. Green leaf, etiolated leaf, cotyledon, sheath and root cDNAs were used as templates.

ScARS gene-specific forward primer 5’-AAGCATAGGAACCACCAACGGCAA-3’ and reverse primer 5’-AAGACGAGGTGGGTGATCTCGCT-3’, as well as *ScCHS* gene-specific forward primer 5’-TCCGTGAAGCGCCTCATGATGTAT-3’ and reverse primer 5’- TCAGGT TTACCTTTGCCTCGACC A-3’ were designed to amplify fragments of *ScARS* and *ScCHS* in *S. cereale*. Also, *ScARS* and *ScCHS* primers were tested with plasmid templates harboring either *ScARS* or *ScCHS* cDNA to prove the specificity of designed primers (data not shown). In addition, the fragment of *ScActin* was amplified as a positive control using *ScActin* gene-specific forward primer 5’-ATGCTAGTGGACGCACAACAGGT A-3’ and reverse primer 5’-ATCTTCATGCTGCTTGGTGCAAGG-3’.

PCR cycle numbers and template amounts were optimized to yield products in the linear range of the reaction. In gene expression analysis of *BdARS*, RT-PCR was carried out under the conditions: 98°C for 30 s, 24 cycles of 98°C for 15 s, 60°C for 30 s, and 72°C for 30 s, and 72°C for 5 min with Phusion High-Fidelity DNA Polymerase. In gene expression analysis of *ScARS*, RT-PCR was carried out under the conditions: 98°C for 30 s, 26 cycles of 98°C for 15 s, 63°C for 30 s, and 72°C for 30 s, and 72°C for 5 min with Phusion High-Fidelity DNA Polymerase. The *ScCHS* expression profile was studied in parallel as a reference. PCR products were separated by electrophoresis on a 1% agarose gel.

2.7.2 Quantitative RT-PCR

In quantitative RT-PCR analysis, *ScARS* gene-specific forward primer 5'-CGAGGTGCCC CAGAACATCTTC-3' and reverse primer 5'-GTCAGGTGGAAGTGCCGCTTG-3' as well as *ScCHS* gene-specific forward primer 5'-AAGGAGAAGTTCAAGAGGATGTG-3' and reverse primer 5'-CTCCACGACGACGATATCCTG-3' were designed and employed to amplify the fragments of *ScARS* and *ScCHS* in *S. cereale*, respectively. The expression level of *18S rRNA* was determined as a reference using the forward primer 5'-CCGTCCTAGTCTCAACCATAAAC-3' and reverse primer 5'-CCTTTAAGTTTCAGCCTTGC G-3'. Different cDNAs from 2 cm-long segments of the first leaf of *S. cereale* were prepared as templates. Quantitative RT-PCR was performed using iQ SYBR Green Supermix (Bio-Rad) on a MJ MiniOpticon real-time PCR system using the program: 95°C for 3 min, 40 cycles of 95°C for 30 s, 60°C for 30 s, and 72°C for 30 s, with a fluorescence reading. A melting curve was generated ranging from 95 to 60°C. Threshold cycles (C_T) were adjusted manually and the C_T values for the reference gene *18S rRNA* amplified in parallel on each plate were subtracted from those of gene of interest to obtain the normalized ΔC_T values. The C_T values of an arbitrary calibrator (*e.g.* the 18-20 cm segment sample at the point of emergence of the leaf) were subtracted from ΔC_T values to generate the $\Delta\Delta C_T$ values. The relative expression levels were calculated using $2^{-\Delta\Delta C_T}$ method as described previously (Livak and Schmittgen, 2001).

2.8 Constructs and plant transformation

2.8.1 Plasmid constructs

To investigate the subcellular localization of BdARS and ScARS, the coding sequences of *BdARS* and *ScARS* and were fused to green fluorescent protein (GFP). *BdARS* gene-specific forward primer 5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATG ACAAGAGCTAACGGTAACGGT-3' and reverse primer 5'-GGGGACCACTTTGTACAA GAAAGCTGGGTCCTAATTTCCCTTGAGACCCGG-3', and *ScARS* gene-specific forward primer 5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGGAAGCATAGGAACCACCAA C-3' and reverse primer 5'-GGGGACCACTTTGTACAAGAAAGCTGGGTCCTAGCGTGGACA GCGGAGGAC-3' were used for PCR amplification of cDNAs from *B. distachyon* leaves and *S. cereale* leaves, respectively, using Phusion High-Fidelity DNA Polymerase. The resulting PCR products were cloned into the pDONR221 entry vector (Invitrogen) using BP Clonase II enzyme mix (Invitrogen) and transformed into TOP10 chemically competent *E.coli* cells. Individual clones were selected on Luria-Bertani (LB) plates containing kanamycin (50 µg/ml), confirmed by sequencing and introduced into the GATEWAY binary vectors pGWB6 (N-sGFP) and pGWB5 (C-sGFP) behind the constitutive cauliflower mosaic virus 35S promoter using LR Clonase II enzyme mix (Invitrogen), resulting in both N-terminal and C-terminal GFP fusions. Colonies were selected on LB plates containing both kanamycin (50 µg/ml) and hygromycin (50 µg/ml).

2.8.2 Agrobacterium-mediated transformation

The GFP fusion constructs were transformed into *Agrobacterium tumefaciens* strain GV3101 according to a previous protocol (Sparkes *et al.*, 2006). Agrobacteria with the plasmids containing 35S:sGFP-BdARS/ScARS or 35S:BdARS/ScARS-sGFP were grown at 28°C in LB medium with kanamycin (50 µg/ml), hygromycin (50 µg/ml), rifampicin (100 µg/ml) and gentamicin (50 µg/ml). Agrobacterium carrying p19 was grown at 28°C in LB medium with kanamycin (50 µg/ml) and rifampicin (100 µg/ml). Cells of *A. tumefaciens* were collected by centrifugation at 5000 rpm for 15 min at room temperature and resuspended in 10 mM $MgCl_2$ and 100 µM acetosyringone. *N. benthamiana* leaves of 4-week-old were used for infiltration of the agrobacterium

suspension into the abaxial air spaces. Plants were placed back into the growth chamber for 3-4 days before examining the fluorescence signals by light microscopy and confocal microscopy.

2.9 Light microscopy and laser scanning confocal microscopy

Leaf tissues of *N. benthamiana* transiently expressing 35S:sGFP-BdARS/ScARS or 35S:BdARS/ScARS-sGFP were prepared for light microscopy and confocal microscopy. Leaves were immersed in 1.6 mM hexyl rhodamine B solution for 10 to 30 min before confocal imaging. For confocal imaging, a Zeiss Pascal Excite laser scanning confocal microscope (http://www.zeiss.com) was used. GFP fluorescence was detected using excitation of 488 nm with a 505- to 530-nm emission filter. Hexyl rhodamin B was examined with a 543-nm argon ion laser line with a 560-nm long-pass emission filter. All confocal images obtained were processed using Zeiss LSM Image Browser and Adobe Photoshop.

Chapter 3 Accumulation of cuticular alkylresorcinols in *Brachypodium distachyon* leaves

3.1 Introduction

Brachypodium distachyon is a wild annual species belonging to the grass family Poaceae. With the release of the complete genome sequence in 2010, *B. distachyon* has recently become a potential and attractive experimental model system for research in grass species. In Poaceae there are three other species from which the genome has been completely sequenced, including *Oryza sativa* (rice), *Sorghum bicolor* (sorghum) and *Zea mays* (maize). *B. distachyon* stands out among them due to its small genome size of about 300 Mb, short lifecycle of 10-18 weeks, self-fertility and ease of growing (Draper *et al.*, 2001). Moreover, *B. distachyon* belongs to the subfamily Pooideae, which includes the majority of important temperate cereal species such as *Hordeum vulgare* (barley), *Secale cereale* (rye) and *Triticum aestivum* (wheat). In contrast, *O. sativa*, *Sorghum bicolor* and *Z. mays* are tropical cereal species belonging to two other subfamilies of Poaceae (Kellogg, 2001; Opanowicz *et al.*, 2008). Thus, according to the phylogenetic relationship, *B. distachyon* is closely related to the temperate cereal species such as *Secale cereale* and distant related to tropical cereal species, for instances, *Sorghum bicolor* and *O. sativa*. Therefore, *B. distachyon* can serve as a powerful tool for studies in gene identification and biochemical mechanism investigation, in particular, in temperate cereals. In this work, it was applied as a reference that facilitated gene cloning from *S. cereale*.

Little is currently known about the wax composition or genes involved in wax biosynthesis in *B. distachyon*. A comprehensive and reliable analysis of cuticular wax in general, and investigations into the presence of alkylresorcinols within the cuticle of *B. distachyon* is therefore needed as a basis for further studies on the biosynthesis of cuticular alkylresorcinols in this model system. Therefore, the first goal of the present work was to provide detailed chemical data on the cuticular wax on *B. distachyon* leaves, and to search for alkylresorcinols within them. Using gas chromatography-flame ionization detector (GC-FID) and gas chromatography-mass spectrometry (GC-MS), total leaf wax of *B. distachyon* was analyzed. In particular, it

was aimed at answering the following specific questions:

a) What is the wax composition in *B. distachyon* leaves?
b) Are alkylresorcinols present in the cuticular wax? If so, how do the chain length profiles and amounts of alkylresorcinols as well as other typical wax components compare to those in *Secale cereale* and other related species?

Provided that cuticular alkylresorcinols are indeed present in *B. distachyon* leaves, then a set of further experiments could be carried out in *B. distachyon* and in *S. cereale* in parallel, with the further goals of cloning and characterization of candidate alkylresorcinol synthases (ARSs) (see Chapter 4 and 5). Ultimately, data on cuticular alkylresorcinol accumulation can be correlated with gene expression profiles, to provide a better understanding on the biosynthesis of cuticular alkylresorcinols in the two selected grass species.

3.2 Results

The goal of the current investigations was to analyze the cuticular wax of *B. distachyon* leaves to allow a comparison with *S. cereale* and other literature data. Since the first unfolded leaf of *S. cereale* had been analyzed previously to minimize biological variation (Ji and Jetter, 2008), also the first fully expanded leaf of *B. distachyon* was analyzed in the present investigation. Analyses were carried out in seven independent parallels. The major emphasis was on finding alkylresorcinols globally in the wax mixture on *B. distachyon* leaves, rather than more detailed localization analyses distinguishing between adaxial and abaxial sides of the leaf, or between the epicuticular and intracuticular layers on each side of the leaf. Thus, the initial analyses were restricted to the total wax from both sides of the *B. distachyon* leaf. Roots of *S. cereale* were also extracted and analyzed following the same procedure for comparison.

3.2.1 Total leaf wax of *B. distachyon*

Total leaf wax was extracted from the first leaf of *B. distachyon* by submerging in chloroform, trimethylsilyl (TMSi)-derivatized and then analyzed on GC-FID and GC-MS.

The coverage of the wax mixture was 12.8 ± 0.8 µg/cm^2 (Figure 3.1). The wax was predominantly composed of primary alcohols (71%), with the rest as alkyl esters (11%), aldehydes (2%), alkanes (2%), trace amounts (these combined represented less than 0.03 µg/cm^2) of sterols (β-sitosterol; 0.1%), triterpenoids (β-amyrin; 0.1%) and alkyl benzoates (C_{26} benzoate; 0.03%), as well as unknowns (9%).

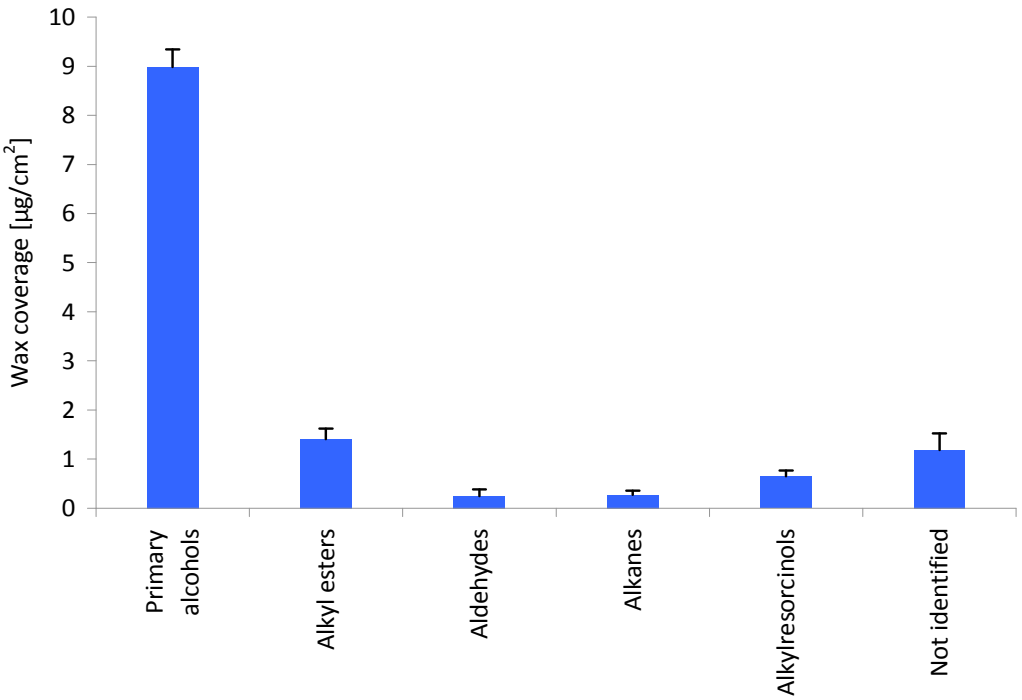
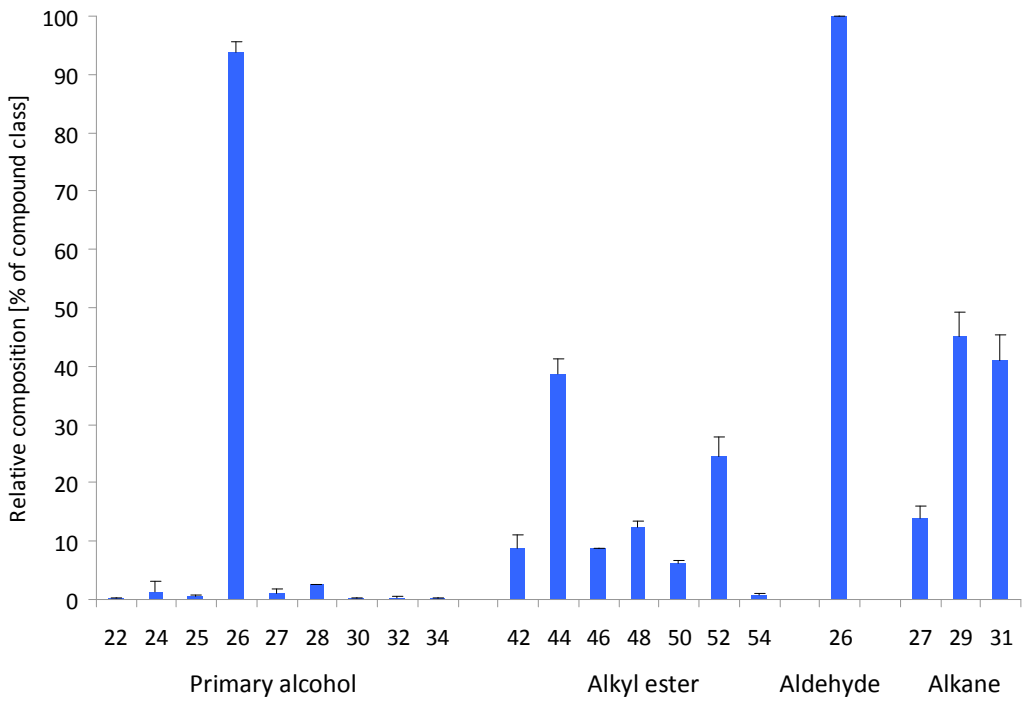


Figure 3.1 Wax compound classes from both sides of *B. distachyon* leaves. The coverages of compound classes in the total leaf wax are given as mean values (*n* = 7) ± SD.

Chapter 9

The chain length distributions of homologs within each of the compound classes showed patterns typical for plant cuticular wax mixtures (Figure 3.2). Primary alcohols ranged from C_{22} to C_{34}, dominated by even-numbered homologs and C_{26} alcohol accounting for 94% (8.4 ± 0.3 µg/cm^2) of the series. Odd-numbered homologs, C_{25} and C_{27} alcohols, were detected as well, albeit only at trace levels. Within the alkyl esters, even-numbered homologs ranging from C_{42} to C_{54} were identified, and found to be dominated by chain lengths C_{44} (0.5 ± 0.1 µg/cm^2) and C_{52} (0.4 ± 0.1 µg/cm^2). A detailed analysis using GC-MS showed that each ester homolog was composed of one

to three isomers, with primary alcohols ranging from C_{22} to C_{26} esterified to fatty acids ranging from C_{16} to C_{28} (Table 3.1). The predominant isomers of all alkyl esters had C_{26} alcohol esterified to respective fatty acids varying from C_{16} to C_{28}. The bimodal distribution of alkyl ester profile was due to an equally bimodal distribution of esterified fatty acids, with C_{18} and C_{26} fatty acids preferentially esterified to C_{26} alcohol. In the compound class of aldehydes, only the C_{26} homolog (0.2 ± 0.1 µg/cm^2) was detected. Alkanes were found as an odd-numbered series ranging from C_{27} to C_{31}, with a maximum at C_{29}. In summary, the chain length profiles of primary alcohols, alkyl esters and aldehydes were found to be dominated by the respective C_{26} homologs, whereas the C_{29} homolog was predominant in alkanes.



Chapter 10Figure 3.2 **Chain length distributions within compound classes in the total wax mixture from both sides of *B. distachyon* leaves. The percentages of individual homologs within the compound classes are shown as mean values (*n* = 7) ± SD.**

Chapter 11 Table 3.1 **Qualitative isomer composition of alkyl esters identified in *B. distachyon* total leaf wax.**

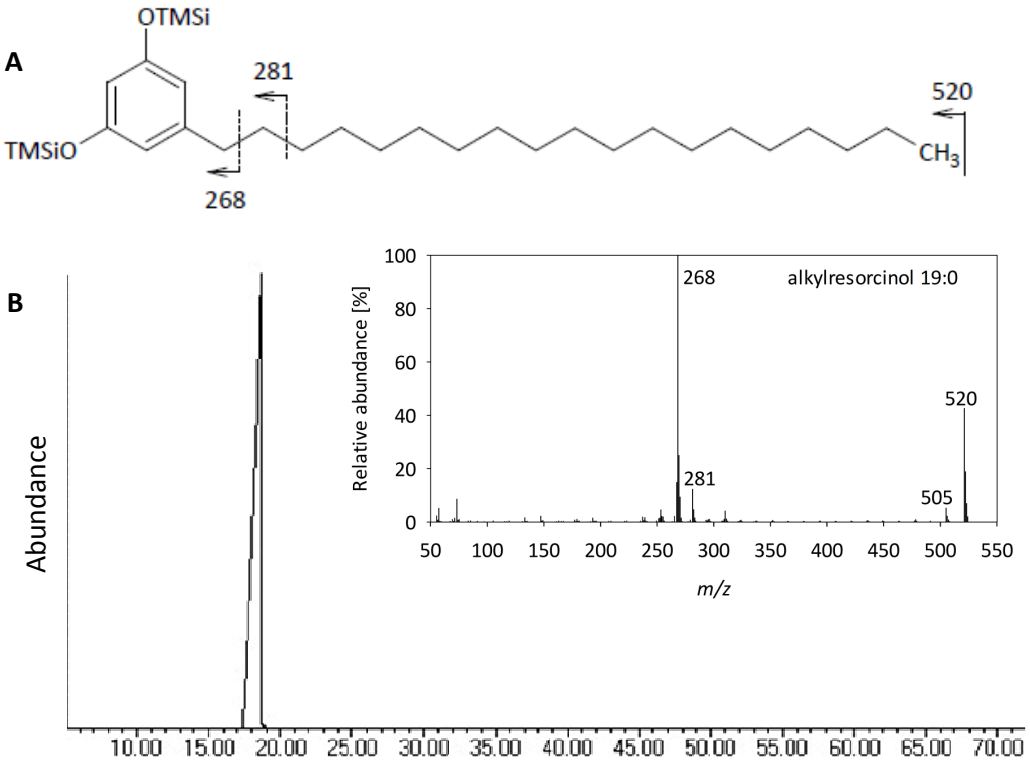
Ester chain length	Chain lengths of esterified acids						
C_{42}	**C_{16}**	C_{18}	C_{20}				
C_{44}		**C_{18}**	C_{20}				
C_{46}		C_{18}	**C_{20}**	C_{22}			
C_{48}			C_{20}	**C_{22}**			
C_{50}				C_{22}	**C_{24}**	C_{26}	
C_{52}						**C_{26}**	C_{28}
C_{54}							**C_{28}**

Note: Predominant (>80%) homologs are highlighted in bold.

3.2.2 Identification of alkylresorcinols in the total leaf wax of *B. distachyon*

In addition to those wax compounds described above, there were five even-spaced compounds that were not typically present in cuticular wax mixtures of other plant species. Their molecular ions differed by 28 mass units, indicating a homologous series differing by $-CH_2–CH_2-$ units. All of these compounds showed characteristic MS fragments of alkylresorcinols at *m/z* 73, 268, 281, together with molecular ions $[C_6H_3(OTMSi)_2(CH_2)_nCH_3]^+$ and corresponding fragments $[M-15]^+$ indicating the loss of a methyl group from the TMSi derivatives (Ji and Jetter, 2008).

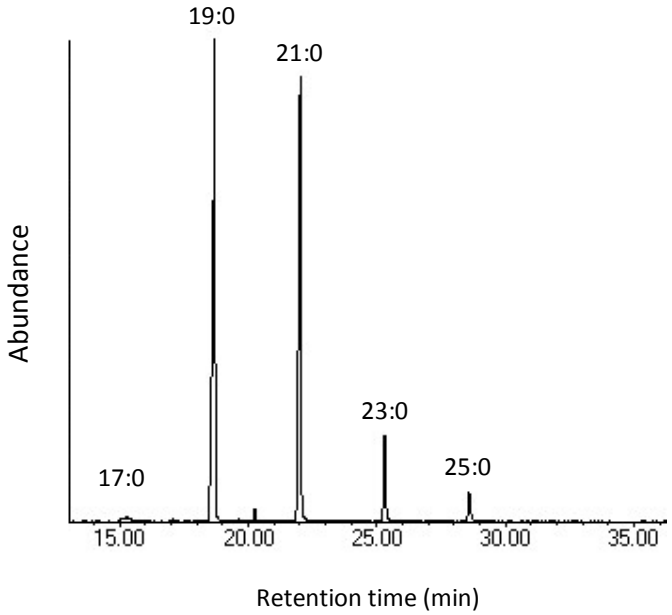
To further verify that the homologous series in the wax mixture indeed consisted of alkylresorcinols, a chemically synthesized standard of 5-*n*-nonadecylresorcinol was co-injected together with a wax sample on GC-MS for structural comparison. It was found to have identical MS characteristics and co-eluted with one of the putative alkylresorcinol homologs under the GC conditions used (Figure 3.3). Thus, the structure of alkylresorcinol 19:0 in *B. distachyon* total leaf wax mixture was unambiguously established, and other phenolic isomer structures were excluded.



Chapter 12 Figure 3.3 **GC-MS analysis of the TMSi-derivatized standard 5-*n*-nonadecylresorcinol.** A**, Structure and fragmentation pattern of 5-*n*-nonadecylresorcinol.** B**, Extracted chromatogram of ion *m/z* 268. The single GC peak has a retention time of 18.6 min, and the corresponding mass spectrum is shown as an inset. Abundance of the GC trace is not to scale.**

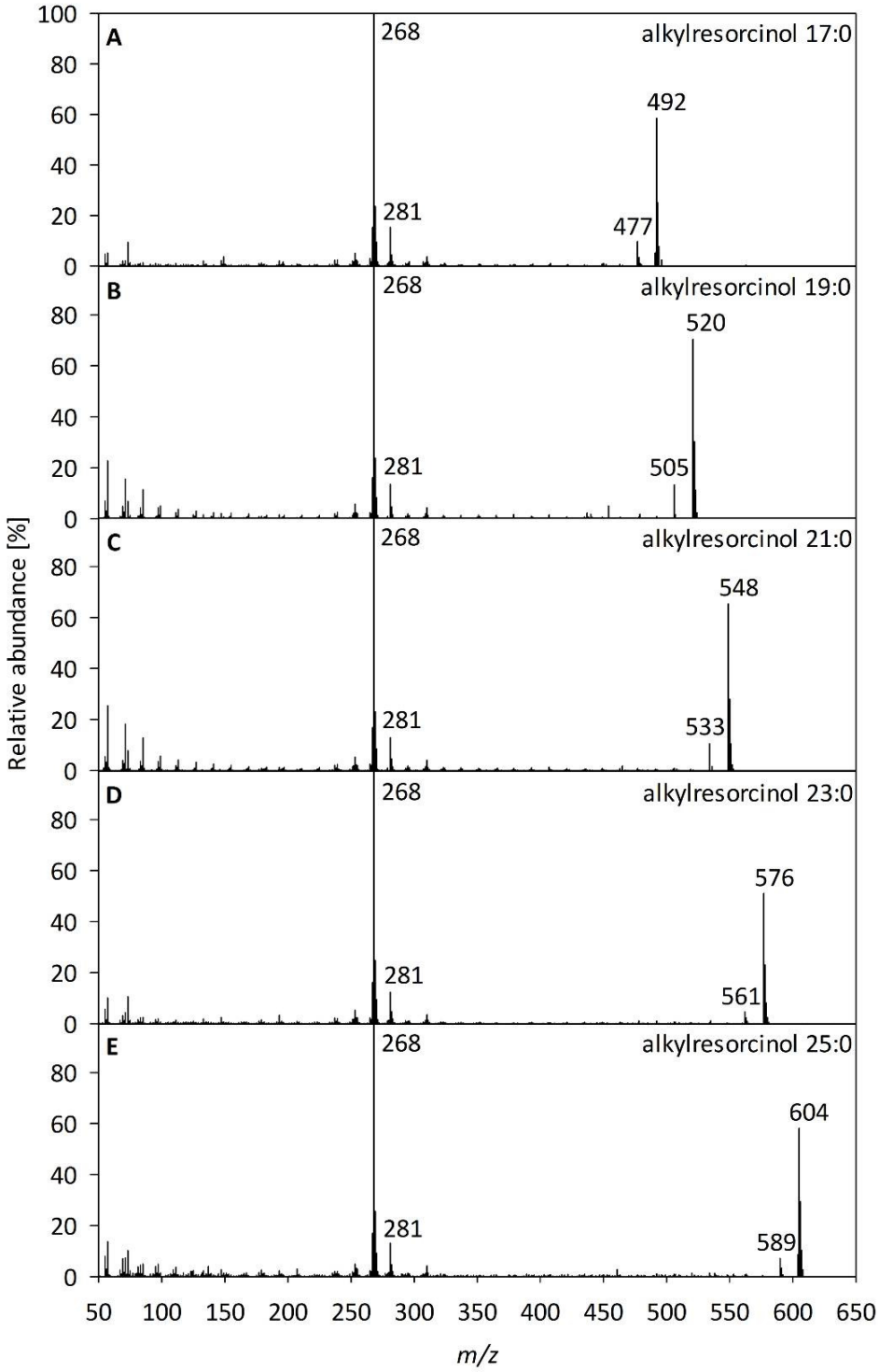
Based on the equal distances between GC peaks (Figure 3.4), all the wax constituents in *B. distachyon* leaves with identical MS characteristics (Figure 3.5) were determined as homologous alkylresorcinols with odd-numbered alkyl chains from C_{17} to C_{25}. In addition to the predominant odd-numbered alkylresorcinol homologs, trace amounts of alkylresorcinols with even-numbered side chains ranging from C_{18} to C_{24} were also found in the cuticular wax, but they only accounted for less than 1% of the total alkylresorcinol mixture. It is to be noted that signals in the low mass region *m/z* 50-73 slightly differed between spectra of alkylresorcinol homologs from the total leaf wax of *B. distachyon* (Figure 3.5), mainly due to subtraction of background signals from alkanes. Quantification of individual alkylresorcinol homologs with the alkyl side chains ranging from C_{17} to C_{25} in the total leaf wax of *B. distachyon* showed that the

homologous series was dominated by alkylresorcinols 19:0 and 21:0, followed by the homologs 23:0, 25:0 and 17:0. Overall, alkylresorcinols contributed 5% (0.6 ± 0.1 µg/cm^2) to the total wax mixture (Figures 3.1 and 3.6).

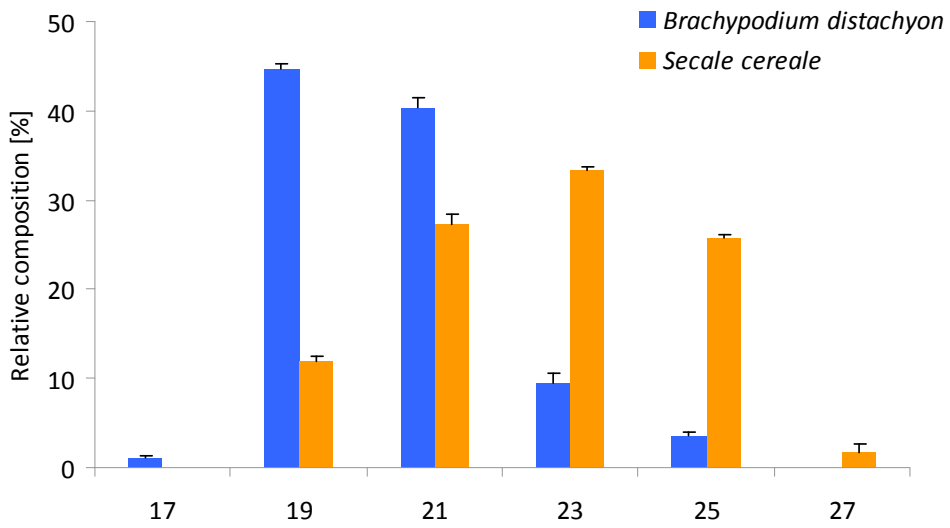


Chapter 13 Figure 3.4 **Extracted ion chromatogram (*m/z* 268) of TMSi-derivatized alkylresorcinols from the total leaf wax of *B. distachyon*. Abundance of the GC trace is not to scale.**

Chapter 14



Chapter 15 Figure 3.5 **Mass spectra of TMSi derivatives of individual alkylresorcinol homologs in *B. distachyon* total leaf wax.** A**, Alkylresorcinol 17:0 (molecular ion *m/z* 492).** B**, Alkylresorcinol 19:0 (molecular ion *m/z* 520).** C**, Alkylresorcinol 21:0 (molecular ion *m/z* 548).** D**, Alkylresorcinol 23:0 (molecular ion *m/z* 576).** E**, Alkylresorcinol 25:0 (molecular ion *m/z* 604).**



Chapter 16 Figure 3.6 **Chain length distributions of alkylresorcinols in the total leaf waxes of *B. distachyon* and *S. cereale*. The percentages of individual alkylresorcinol homologs in *B. distachyon* leaves are shown as mean values (*n* = 7) ± SD. The percentages of individual alkylresorcinol homologs in *S. cereale* leaves are shown as mean values (*n* = 3) ± SD.**

3.2.3 GC-MS analysis of alkylresorcinols in leaves and roots of *S. cereale*

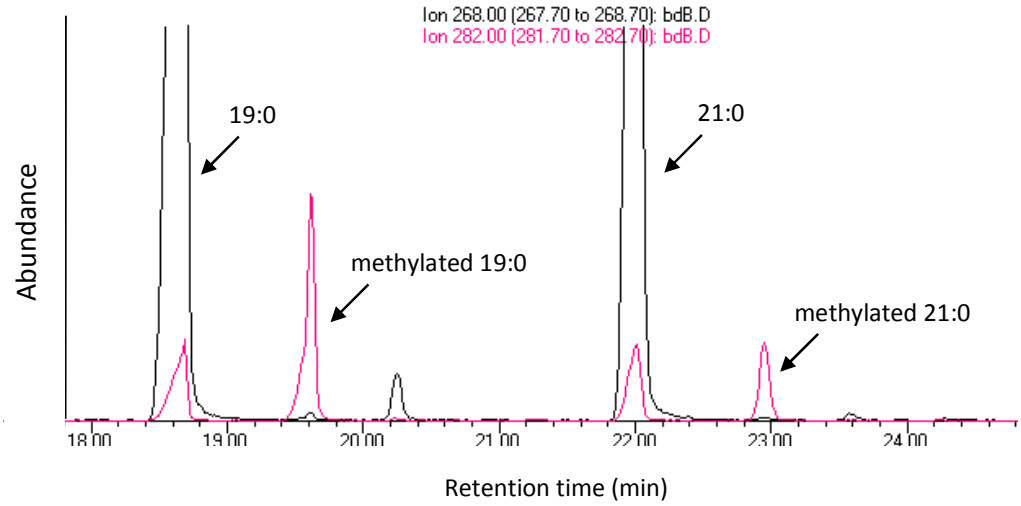
In order to obtain alkylresorcinol profiles for direct comparison between *B. distachyon* and *S. cereale*, an analysis of the total leaf wax of the second grass species was carried out. The data obtained from this work were consistent with previously published results (Ji and Jetter, 2008). They showed that the homologous alkylresorcinols in the total leaf wax of *S. cereale* were with the alkyl chains ranging from C_{19} to C_{27}. Among them, alkylresorcinols 21:0, 23:0 and 25:0 together accounted for 88%, with the homolog 23:0 slightly more abundant than the homologs 21:0 and

25:0. Thus, the alkylresorcinols in the total leaf waxes of both grass species were found to have slightly different chain length distributions, those in *B. distachyon* showing a prevalence of homologs with shorter side chains than in *S. cereale* (Figure 3.6).

Previous reports had shown that long-chain and very-long-chain (VLC) alkylresorcinols are present predominantly or exclusively at/near surfaces of aerial organs which are covered by a hydrophobic cuticle. To test the organ-specific accumulation of alkylresorcinols in *S. cereale*, the present findings on leaf wax alkylresorcinols had to be compared with the analyses of underground tissues. Therefore, *S. cereale* roots from 7-day-old seedlings grown on agar plates were extracted with chloroform and then analyzed by GC-MS. In root samples extracted without the presence of an internal standard, no alkylresorcinols could be detected. Spiking of further root samples with an internal standard of 5-*n*-tridecylresorcinol indicated that the detection limit for alkylresorcinols under the given conditions was at 0.02 µg/g (dry weight) (data not shown).

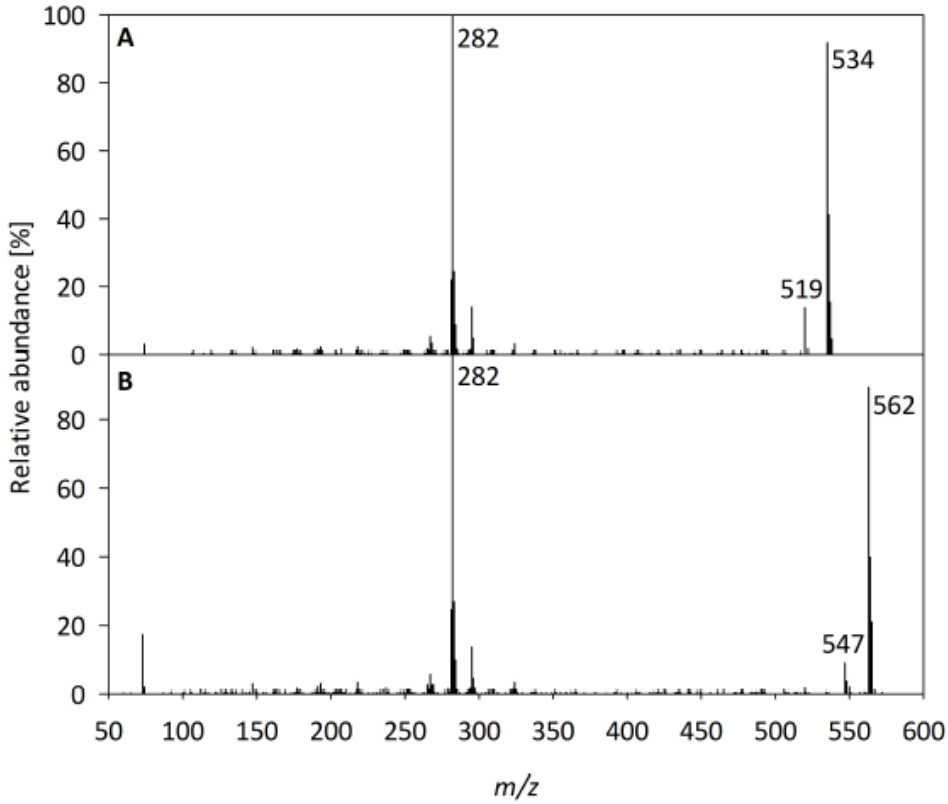
3.2.4 Two novel compounds identified as methylated alkylresorcinols in the total leaf wax from *B. distachyon*

In addition to the typical alkylresorcinol series identified as described above, two novel compounds present in the total leaf wax of *B. distachyon* were also noticed. Instead of having an MS fragment *m/z* 268 characteristic of alkylresorcinols, these two compounds both had an MS fragment *m/z* 282. Moreover, the equal distances between the GC peaks representing the elution time after alkylresorcinol 19:0 and 21:0, respectively, indicated they are homologous compounds.



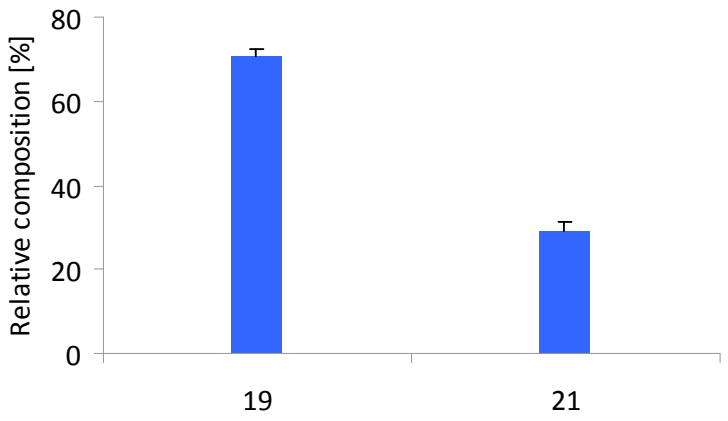
Chapter 17 Figure 3.7 **Extracted chromatograms of fragments *m/z* 268 and *m/z* 282 in the total leaf wax of *B. distachyon*. The trace for *m/z* 268 shows peaks of the TMSi-derivatized alkylresorcinols 19:0 and 21:0. In contrast, the trace for *m/z* 282 shows small peaks for the two alkylresorcinols and two novel compounds with the characteristic fragment *m/z* 282 that are tentatively identified as methylated alkylresorcinols 19:0 and 21:0. Abundances are shown in arbitrary units.**

The MS fragmentation patterns of the two novel compounds revealed that their molecular ions differed by 28 mass units, indicating a homologous series differing by $-CH_2–CH_2-$ units (Figure 3.8). The MS fragmentation patterns of the novel compounds had three major features all similar to those of alkylresorcinols, but differing from them by additional 14 mass units. Thus, the novel compounds had characteristic fragments at *m/z* 282, together with molecular ions $[C_6H_3(OTMSi)_2(CH_2)_nCH_3]^+$ and corresponding fragments $[M-15]^+$ indicating the loss of a methyl group from the TMSi derivatives. These findings suggested the presence of an additional methyl group in the novel compounds compared to their typical alkylresorcinols 19:0 and 21:0. Therefore, these two novel compounds were tentatively identified as methylated alkylresorcinols 19:0 and 21:0.



Chapter 18 Figure 3.8 Mass spectra of TMSi-derivatized methylated alkylresorcinol homologs**.** A**, Methylated alkylresorcinol 19:0 (molecular ion *m/z* 534).** B**, Methylated alkylresorcinol 21:0 (molecular ion *m/z* 562).**

Quantitative analysis revealed that the coverage of the methylated alkylresorcinol homologs 19:0 and 21:0 was 0.04 ± 0.01 µg/cm^2 in the total leaf wax of *B. distachyon*. Methylated alkylresorcinol 19:0 was the predominant homolog, accounting for 71% (Figure 3.9). However, the exact structures of methylated alkylresorcinol homologs 19:0 and 21:0 could not be determined in the current work due to very limited amounts of material present.



Chapter 19 Figure 3.9 **Relative abundance of methylated alkylresorcinols 19:0 and 21:0 in the total leaf wax of *B. distachyon*. The percentages of individual homologs are shown as mean values (*n* = 3) ± SD.**

The same MS fragments were used in search of methylated alkylresorcinols in the total leaf wax of *S. cereale* where the novel compounds had not been reported before. However, these compounds were not detectable (less than 0.0007 µg/cm^2) in the total wax mixture of *S. cereale* leaves.

Chapter 20

3.3 Discussion

There had been no chemical analyses of the leaf cuticles of *B. distachyon*, a new grass model system for studies in the majority of temperate cereal species. In the current work, the total leaf wax of *B. distachyon* was analyzed in order to obtain reliable data on wax composition for comparison with other related species, and more importantly, to search for cuticular alkylresorcinols along the leaves. In parallel, leaves of *S. cereale* were analyzed to confirm the profile of cuticular alkylresorcinols with the published data. Moreover, roots of *S. cereale* were also analyzed according to the same procedure as wax analysis. The root data supported the idea that alkylresorcinols were localized at/near the surfaces of aerial organs, but not in the underground organs in *S. cereale*.

The total leaf wax of *B. distachyon* had a coverage 12.8 ± 0.8 µg/cm^2, containing

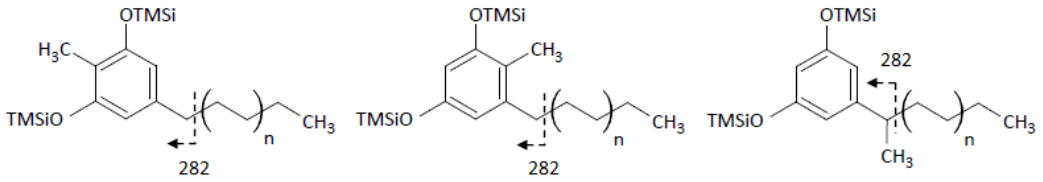
primary alcohols as principal components as well as alkyl esters, aldehydes and alkanes. Thus, the wax composition of *B. distachyon* was similar to that of related species of Poaceae, especially to the major temperate cereal species such as *H. vulgare*, *S. cereale* and *T. aestivum* (Tulloch, 1981) since primary alcohols had been found as the prevailing wax compound class. In *B. distachyon*, C_{26} alcohol was the dominant homolog within this compound class. Alcohols dominated by C_{26} homolog had also been reported for leaf wax of *H. vulgare* (Richardson *et al.*, 2005) and *S. cereale* (Ji and Jetter, 2008). However, the predominant chain length varied in a certain range between other grass species. For instance, in leaves of *T. aestivum* (Koch *et al.*, 2006) and *Sorghum* spp. (Bianchi *et al.*, 1978), the most abundant homolog was C_{28} alcohol, whereas in leaves of *O. sativa* it was C_{30} alcohol (Yu *et al.*, 2008), and in seedlings of *Z. mays* C_{32} alcohol (Bianchi *et al.*, 1989). Similar chain length distributions were also found in other common compound classes in grass waxes including alkyl esters, aldehydes and alkanes. For example, alkanes were usually identified as a mixture of C_{27}, C_{29} and/or C_{31} components with a peak at either the C_{29} or C_{31} homolog (Tulloch, 1981). Other wax constituents were also reported, varying between species, organs or even ages regardless of growth conditions and analytical procedures. For instance, β-amyrin was identified in *B. distachyon* leaves from the current work and *T. aestivum* leaves (Koch *et al.*, 2006), but not in *Secale cereale* (Ji and Jetter, 2008). In contrast, fatty acids were detected in *S. cereale* (Ji and Jetter, 2008) as well as other cereal species (Bianchi *et al.*, 1978; Tulloch, 1981), but were not detectable in *B. distachyon* in the current study. Furthermore, β-diketones had been reported in *Hordeum* spp., *Secale* spp., *Triticum* spp. (Tulloch *et al.*, 1980) and *Sorghum* spp. (Bianchi *et al.*, 1978) during anthesis, but they were not reported in *Secale cereale* leaves (Ji and Jetter, 2008). This discrepancy was explained by the variability between leaves harvested at different stages during plant development (Ji and Jetter, 2008).

Additionally, a homologous series not typically present in wax was found in *B. distachyon* leaves and determined to be alkylresorcinols, accounting for 5% of the total wax coverage. The alkylresorcinol side chains ranged from C_{17} to C_{25} with a maximum at homolog 19:0. Leaf cuticular alkylresorcinols had been reported once

from *S. cereale*, where they, accounting for 3% of the total wax, represented a homologous mixture with the chain lengths of C_{19} to C_{27} and peak at C_{23} (Ji and Jetter, 2008). Therefore, a comparison between two model systems in this work revealed that overall, *B. distachyon* and *S. cereale* had similar wax coverage and composition in leaves. However, alkylresorcinol homologs in *B. distachyon* had shorter chain lengths than those in *S. cereale* (Figure 3.6).

It is to be noted that in *S. cereale* the composition and abundance of cuticular alkylresorcinols were nearly the same on both sides of the leaves, and they were found to be exclusively within intracuticular wax while absent from epicuticular wax (Ji and Jetter, 2008). Although further wax analyses of *B. distachyon* leaves to distinguish wax composition between adaxial and abaxial sides, or between the epicuticular and intracuticular wax layers on each side of the leaves, were not conducted, similar results to those in *S. cereale* would be expected in leaves of *B. distachyon*. It is also believed that, as a particular type of compounds, the partitioning of alkylresorcinols in the wax layers is closely associated with their biological function to the cuticle.

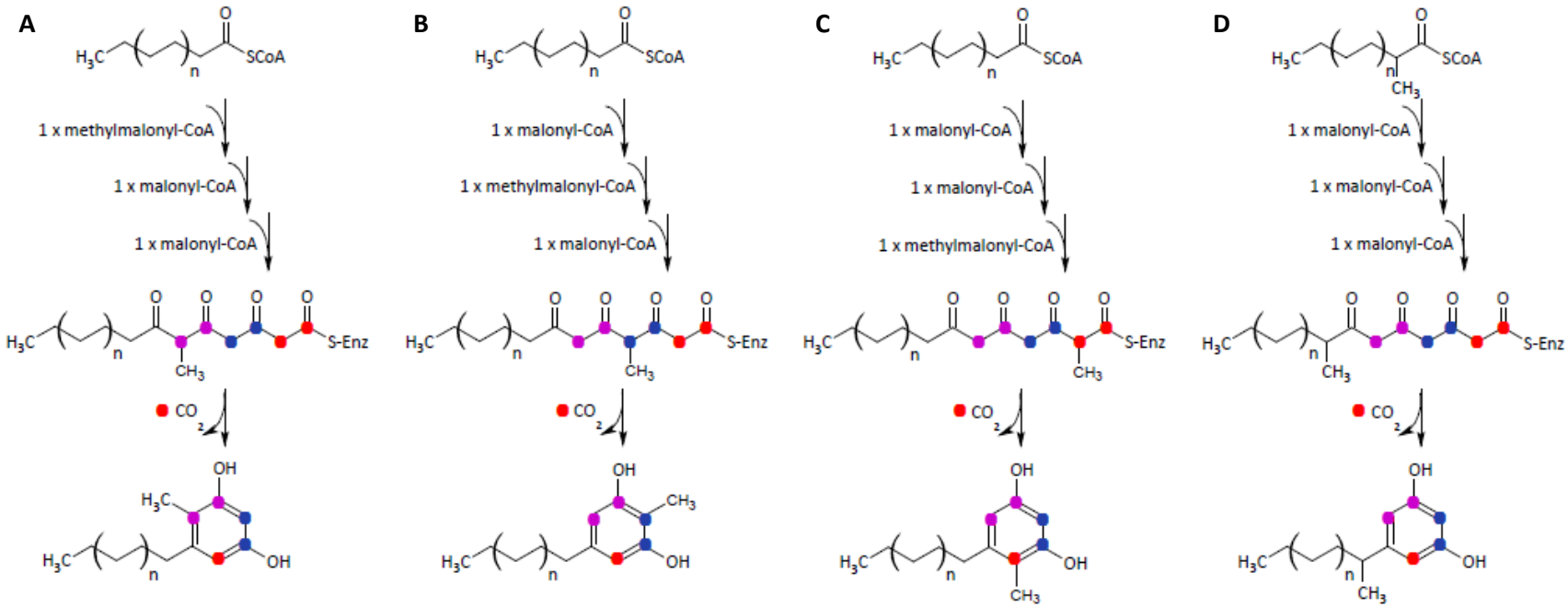
Interestingly, traces of two additional novel compounds were found in the total leaf wax of *B. distachyon* and they were identified as methlated alkylresorcinols 19:0 and 21:0. The exact structures of these compounds were not determined in this work; however, based on the available MS information, mainly the fragment *m/z* 282, three structures seem possible: The methyl group can be at either position 2 or 4 of the resorcinol ring, or at position 1 on the alkyl side chain (Figure 3.10). It seems very plausible that the possible methylated alkylresorcinol structures result from incorporation of either methylated malonyl-CoA extenders or fatty acyl-CoA starter substrates (Figure 3.11; see Chapter 1.3).



Chapter 21 Figure 3.10 **Possible positions of the methyl group in alkylresorcinols (n=8 or 9 as identified in *B. distachyon* total leaf wax).**

It is to be noted that the above proposed utilization of methylmalonyl-CoA as extender had been tested in the *in vivo* characterization of the bacterial SrsA from *Streptomyces griseus* as mentioned earlier in Chapter 1.3. Following a strictly controlled sequence, SrsA utilized one molecule of methylmalonyl-CoA as the extender in the first cycle of decarboxylative condensation reaction and two molecules of malonyl-CoA as extenders in the next two rounds of condensation to produce methylated alkylresorcinols as shown in Figure 3.11 A (Funabashi *et al.*, 2008).

Thus, it will be very interesting to determine the exact nature of methylated alkylresorcinols in the cuticular wax of *B. distachyon* leaves in future work, since the methyl group position will allow inferences on the biosynthesis of the compounds, and hence the characteristics of the ARS enzymes involved in their formation. Also, it will reveal whether the plants use the same biosynthetic pathway as their microbial counterparts.



Chapter 2 Figure 3.11 **The potential biosynthetic pathways leading to methylated alkylresorcinols. Biosynthetic route to 4-methyl-5-*n*-alkylresorcinols (**A **and** C**) and 2-methyl-5-*n*-alkylresorcinols (**B**) using one molecule of methylmalonyl-CoA and two molecules of malonyl-CoA as extenders in different sequences with straight chains of fatty acyl-CoA starter substrates. Biosynthetic route to 5-methyl-alkylresorcinols (**D**) using three molecules of malonyl-CoA as extenders with methyl-branched fatty acyl-CoA starter substrates (n=8 or 9 as identified in *B. distachyon* total leaf wax)**

The absence of alkylresorcinols of roots extracted with chloroform revealed the organ specificity of alkylresorcinol accumulation in *S. cereale*. The accumulation of alkylresorcinols in aboveground organs of *S. cereale* is contrary to *Sorghum bicolor*, where alkylresorcinols occur exclusively in roots (Cook *et al.*, 2010). The distinct accumulation specificities of alkylresorcinols suggest unique functions in different species of Poaceae.

Considering the difference between the chain length distribution profiles of cuticular alkylresorcinols in leaves of *B. distachyon* and *Secale cereale*, it now becomes interesting to compare the corresponding ARS enzymes involved in the biosynthesis of cuticular alkylresorcinols within these two grass species. In *B. distachyon*, alkylresorcinols with the alkyl chains ranging from C_{17} to C_{25} are hypothesized to be biosynthesized by the responsible condensing enzyme ARS via decarboxylative condensation from C_{18} to C_{26} fatty acyl-CoA starter substrates, whereas in *S. cereale* alkylresorcinols with the alkyl chains ranging from C_{19} to C_{27} are derived from C_{20} to C_{28} fatty acyl-CoAs. As is known, the VLC fatty acyl-CoAs present in epidermal cells are largely dedicated to wax biosynthesis (Suh *et al.*, 2005). Therefore, the responsible ARS enzymes should be present in the epidermis and possess the enzyme affinities to VLC fatty acyl-CoA substrates for the biosynthesis of cuticular alkylresorcinols. However, if the corresponding ARSs in *B. distachyon* and *S. cereale* both have access to all the VCL fatty acyl-CoAs with chain lengths varying from C_{18} to C_{28}, then difference in the resulting alkylresorcinol profiles must be due to different chain length selectivities of both enzymes. For this reason, gene cloning, biochemical and biological characterization of ARSs in *B. distachyon* and *S. cereale* were carried out, and will be described in the next two chapters.

Chapter 4 Cloning of putative alkylresorcinol synthases (ARSs) from *Brachypodium distachyon* and *Secale cereale*

4.1 Introduction

The research described in Chapter 3 showed that alkylresorcinols accumulate in the cuticular wax on leaves of *Brachypodium distachyon*. A homologous series of alkylresorcinols with the alkyl chains ranging from C_{17} to C_{25} was identified in the total leaf wax, accounting for 5% of the wax coverage. These findings are similar to previous reports on cuticular alkylresorcinols on leaves of *Secale cereale* (Ji and Jetter, 2008). Thus, both grass species can serve as models for further investigations into the biological function of alkylresorcinols. However, the biosynthesis of cuticular alkylresorcinols must be studied in order to provide tools for investigating the potential biological role of these compounds present in the cuticle. To this end, genes encoding alkylresorcinol synthases (ARSs) responsible for the biosynthesis of cuticular alkylresorcinols should be cloned and characterized first.

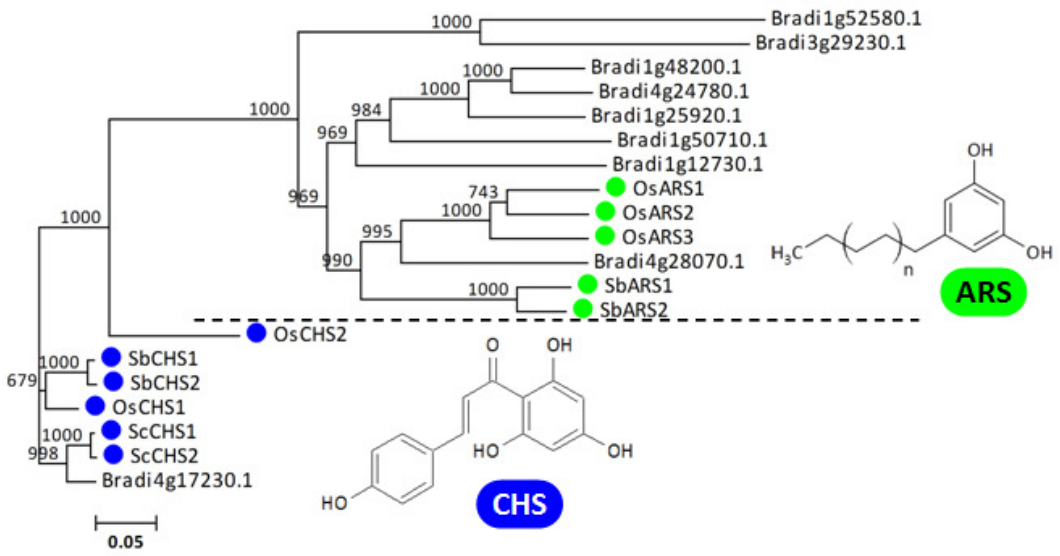
In this chapter, investigations aiming at the isolation of candidate ARS genes will be described. The accumulation of alkylresorcinol homologs in cuticular waxes on young, still expanding leaves of *B. distachyon* and *S. cereale* suggested that this material should be enriched in mRNAs encoding the ARS enzymes involved in the biosynthesis of cuticular alkylresorcinols. Therefore, the first true leaves of *B. distachyon* and *S. cereale* were harvested for cloning the target genes shortly before they reached full size. In *B. distachyon*, gene cloning was performed based on the mining of Brachypodium expressed sequence tag (EST) libraries. In *S. cereale*, PCR-based gene cloning was attempted using primer designs exploiting homology between the functionally characterized ARS sequences from *Sorghum bicolor* and *Oryza sativa* (Cook *et al.*, 2010). According to the different profiles of cuticular alkylresorcinols in the two selected grass species, the putative ARSs from *B. distachyon* and *Secale cereale* were expected to have subtle differences in product profiles. The following specific questions were addressed:

a) Are ARS genes present in *B. distachyon* and *S. cereale*? If so, what are the gene structures?

b) How do the sequences of putative ARSs from both species compare with each other and other related ARSs?
c) What is the phylogenetic relationship between the putative ARSs with other related type III polyketide synthases (PKSs)?

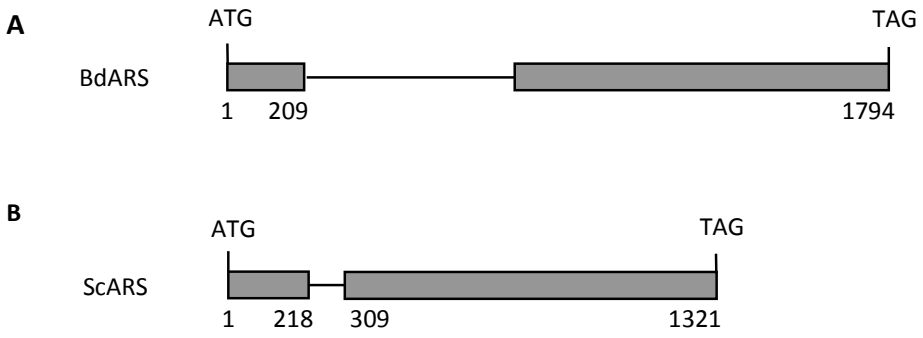
4.2 Results

In search for putative ARS sequences, the Brachypodium EST libraries were mined and nine full-length cDNAs encoding potential type III PKS sequences were identified. A sequence alignment and phylogenetic analysis were carried out using the corresponding amino acid sequences together with selected type III PKS sequences from *O. sativa*, *S. cereale*, and *Sorghum Bicolor*. In the resulting phylogenetic tree, the type III PKSs from *O. sativa*, *Secale cereale* and *Sorghum bicolor* that had been functionally characterized as either ARSs or chalcone synthases (CHSs) formed well-separated clades (Figure 4.1). While one type III PKS sequence from *B. distachyon* (gene identifier Bradi4g17230.1) fell within the CHSs, all other sequences from this species grouped together with the ARSs. Another type III PKS sequence (gene identifier Bradi4g28070.1) was found to have particularly high similarity with the ARSs, and it was therefore considered as the best candidate for the ARS enzyme responsible for alkylresorcinol biosynthesis in *B. distachyon.* It will be designated as BdARS in the following. The rest of the candidate sequences within the ARS clade are tentatively referred to as ARS-like enzymes.



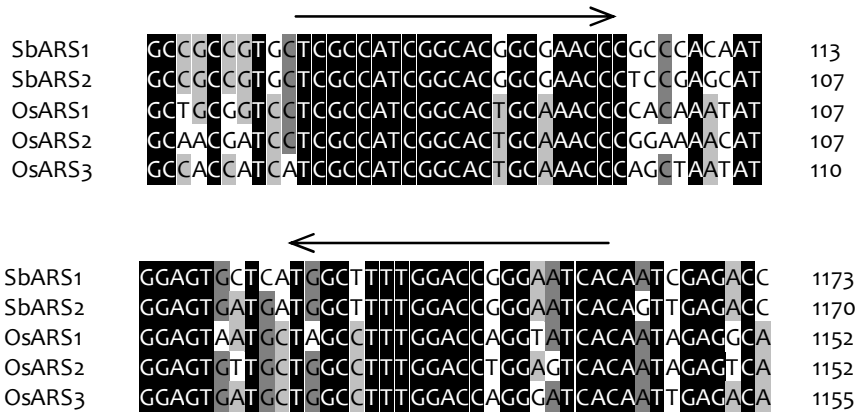
Chapter 22 Figure 4.1 **Phylogenetic analysis of *B. distachyon* candidate proteins related to selected type III PKSs. ARSs are alkylresorcinol synthases indicated by bright green circles and CHSs are chalcone synthases indicated by blue circles. OsARS1, OsARS2, OsARS3, OsCHS1 and OsCHS2 are from *O. sativa*. ScCHS1 and ScCHS2 are from *Secale cereale*. SbARS1, SbARS2, SbCHS1, and SbCHS2 are from *Sorghum bicolor*.**

The *BdARS* cDNA was cloned from 3-week-old *B. distachyon* leaves and the database sequence was confirmed. The corresponding gene was found to contain one intron of 567 bp and two exons of 209 and 1,018 bp. It was predicted to encode a 43.9 kD protein of 408 amino acids and with an isoelectric point of 5.96 (Figure 4.2A).



Chapter 23 Figure 4.2 **Gene structures of *BdARS* from *B. distachyon* (**A**) and *ScARS* from *Secale cereale* (**B**). Exons are shown in shaded boxes and introns are shown in lines. The positions of start codons and stop codons are indicated.**

In contrast to *B. distachyon*, no genome sequence information was available for *S. cereale*. Therefore, a homology-based cloning approach was used to clone putative ARS gene(s) from *S. cereale*. To this end, mRNA was isolated from 3-week-old leaves and subjected to reverse transcription. The resulting cDNA was used as the template for PCR with a set of gene-specific primers that had been designed based on conserved sequence regions that were characteristic of all the previously isolated ARSs but not for CHSs (Figure 4.3). The PCR resulted in a core fragment with the expected size of approximately 1,000 bp. 5' and 3' Rapid Amplification of cDNA Ends (RACE) reactions were employed to extend this sequence to full length, and the product was designated as *ScARS*. Sub-cloning and sequencing demonstrated that the *ScARS* cDNA represented an open reading frame of 1,230 bp, encoding a 43.3 kD protein of 409 amino acids with a predicted isoelectric point of 6.33. The corresponding gene was found to comprise two exons of 218 and 1,012 bp, separated by an intron of 91 bp (Figure 4.2B).



Chapter 24 Figure 4.3 **Conserved nucleotide sequences of *ARS*s used to design the forward and reverse primers for the homology-based cloning of a putative *ARS* from *S. cereale*.**

The full-length amino acid sequence of ScARS was 71% identical with BdARS and 58-62% identical with ARSs from *Sorghum bicolor* and *O. sativa*. In contrast, in a comparison of type III PKS enzymes within the same species, ScARS showed relatively low identity with the previously characterized CHSs from *Secale cereale*. This result suggested a difference of biochemical functions between ARS and CHS type III PKS enzymes. BdARS had the identities of *ca.* 62% with SbARSs and even higher identities *ca.* 68% with OsARSs (Table 4.1).

Table 4.1 Amino acid sequence identities between selected ARSs and CHSs.

	ScARS	BdARS	SbARS1	SbARS2	OsARS1	OsARS2	OsARS3	ScCHS1	ScCHS2
ScARS		70.7%	59.5%	58.4%	60.3%	62.3%	61.4%	46.3%	46.8%
BdARS			62.0%	60.8%	67.3%	67.7%	66.9%	49.6%	49.6%
SbARS1				89.6%	63.0%	63.6%	63.4%	50.6%	51.1%
SbARS2					62.0%	62.3%	63.7%	51.6%	52.4%
OsARS1						80.7%	84.2%	49.4%	49.4%
OsARS2							85.5%	48.1%	47.8%
OsARS3								49.1%	49.1%
ScCHS1									98.7%
ScCHS2									

In summary, the amino acid sequence analyses confirmed the hypothesis that BdARS and ScARS have biochemical functions differing from CHSs, and are possibly involved in alkylresorcinol biosynthesis. In order to elucidate the potential function of the two putative ARSs, a detailed characterization of BdARS and ScARS including biochemical and biological investigations will be described in the following chapter.

4.3 Discussion

The experiments described in this chapter aimed at the isolation of candidate ARS genes from *B. distachyon* and *S. cereale*. A candidate gene designated as BdARS was identified in *B. distachyon* by mining the Brachypodium EST libraries. A candidate designated as ScARS was identified in *S. cereale* using homology-based cloning approach. The sequence information on the resulting BdARS and ScARS serves as a basis for phylogenetic analysis and enzymatic function predictions discussed in the following.

4.3.1 Phylogenetic relationships of ARSs and other type III PKSs

Type III PKSs catalyze iterative decarboxylative condensation reactions over different cycles starting from a broad range of acyl-CoAs, to yield diverse products with antimicrobial activities and pharmaceutical uses. In this ancient enzyme family, CHS is the first characterized and most well-studied enzyme. Expressed ubiquitously in plants, CHS is involved in the biosynthesis of flavonoids that are important for flower pigmentation (Winkel-Shirley, 2001), plant defense (Cushnie and Lamb, 2005) and ultraviolet (UV) photoprotection (Winkel-Shirley, 2002). Therefore, type III PKSs are sometimes also called CHS superfamily enzymes, and other enzymes discovered afterwards have been annotated as non-CHS type PKS (a.k.a. CHS-like; CHSL) enzymes. For instance, stilbene synthase (STS) is a CHSL enzyme responsible for the biosynthesis of stilbenes which serve as phytoalexins in plant defense (Gorham, 1995; Austin *et al.*, 2004). Although it differs from CHS only in the mechanisms of the final cyclization of the linear tetraketide intermediate, STS is present only in a certain number of plant species, such as *Arachis hypogaea* (peanut), *Pinus sylvestris* (Scots pine) and *Vitis vinifera* (grape vine). A specialized type III PKS designated as ARS is responsible for the biosynthesis of alkylresorcinols, another class of polyketide-derived phenolic lipids. ARSs have been explored over the last five years, long after CHS and STS. In spite of the extensive occurrence of alkylresorcinols across the plant kingdom, fungi and bacteria, the responsible ARS enzymes have only been identified from a limited number of plant and microbe species.

In the model organism *Arabidopsis thaliana* there is one CHS (At5g13930) and three active CHSL enzymes (At1g02050, At4g00040 and At4g34850) (Wang *et al.*, 2007). Two of the CHSL enzymes (At1g02050 and At4g34850, a.k.a. PKSA and PKSB) are involved in the formation of alkylpyrones (Mizuuchi *et al.*, 2008), and the biochemical studies showed their significance in sporopollenin biosynthesis during pollen development (Grienenberger *et al.*, 2010; Kim *et al.*, 2010). Although the dicotyledonous species *A. thaliana* has only one true CHS, synthesizing naringenin chalcone, most angiosperm species have at least two CHSs (Huang *et al.*, 2004). In monocotyledonous species, the model organism *O. sativa* has 27 genes encoding

CHS/CHSL enzymes, including two true CHSs (Os11g32650 and Os07g11440) (Jiang *et al.*, 2008) and three ARSs (Cook *et al.*, 2010; Matsuzawa *et al.*, 2010). *Sorghum bicolor* has as many as seven CHSs (Contessotto *et al.*, 2001) and two ARSs (Cook *et al.*, 2010), among a total of 33 CHS/CHSL enzymes. *Zea mays* has eleven CHS/CHSL enzymes, two of which have been characterized as CHSs (C2 and Whp) (Franken *et al.*, 1991). The genome of *B. distachyon* has been completely sequenced very recently (2010). In this emerging model system, no type III PKSs have been characterized and reported so far. In the course of the investigations presented here nine CHS/CHSL enzymes were found in *B. distachyon* (Figure 4.1). One (Bradi4g17230.1) out of these nine sequences fell within the CHS clade, making it very likely that it is the only functional CHS in this species. Moreover, according to its position in the ARS sequence cluster, Bradi4g28070.1 was designated as BdARS and was cloned for further characterization. The remaining seven *B. distachyon* sequences are CHSL enzymes. Different from *O. sativa*, *S. bicolor*, *Z. mays* and *B. distachyon*, whose genome sequencing is complete, the whole genome sequencing of *Secale cereale* has not yet been achieved due to the large genome size of about 9,120 Mb. Nevertheless, two CHSs have already been identified in large scale EST collections from *S. cereale*.

A comprehensive phylogenetic analysis of CHS/CHSL representatives revealed three distinct clades in addition to the cluster of true CHSs (Figure 4.4). STSs showed no separate cluster, but grouped together with CHSs from the same plant species, which supported the idea that STSs have evolved repeatedly from CHSs after speciation (Tropf *et al.*, 1994; Tropf *et al.*, 1995). The most distant cluster to the true CHSs included three CHSLs from *A. thaliana*, where At1g02050/PKSA and At4g34850/PKSB were responsible for synthesizing alkylpyrones that were considered to be important for sporopollenin production during pollen development (Kim *et al.*, 2010). Several CHSL enzymes from *Sorghum bicolor*, *O. sativa* and *Z. mays* as well as Bradi3g29230.1/BdCHSL1 and Bradi1g52580.1/BdCHSL2 from *B. distachyon* belonged to this cluster, indicating that they could have a similar or identical biochemical function as the characterized Arabidopsis CHSL enzymes. A cluster more closely related to the true CHSs contained the ARSs from *S. bicolor* and *O. sativa* characterized in previous studies (Cook *et al.*, 2010), as well as the putative ARSs

cloned in this work from *B. distachyon* and *Secale cereale*. A few more CHSL enzymes from *Sorghum bicolor* and *O. sativa* belonged also to this ARS cluster. They probably act in the same manner as ARS enzymes, but their enzymatic functions need to be tested in the future. Surprisingly, the ARS enzyme from *Cannabis sativa* (olivetol synthase, OLS) stood out of the ARS cluster. The remaining CHSL enzymes from *S. bicolor*, *O. sativa*, *Z. mays* and *B. distachyon* fell between the ARS and the CHS clusters, making it interesting to characterize some of these enzymes using a variety of starter substrates from aromatic to aliphatic CoAs.

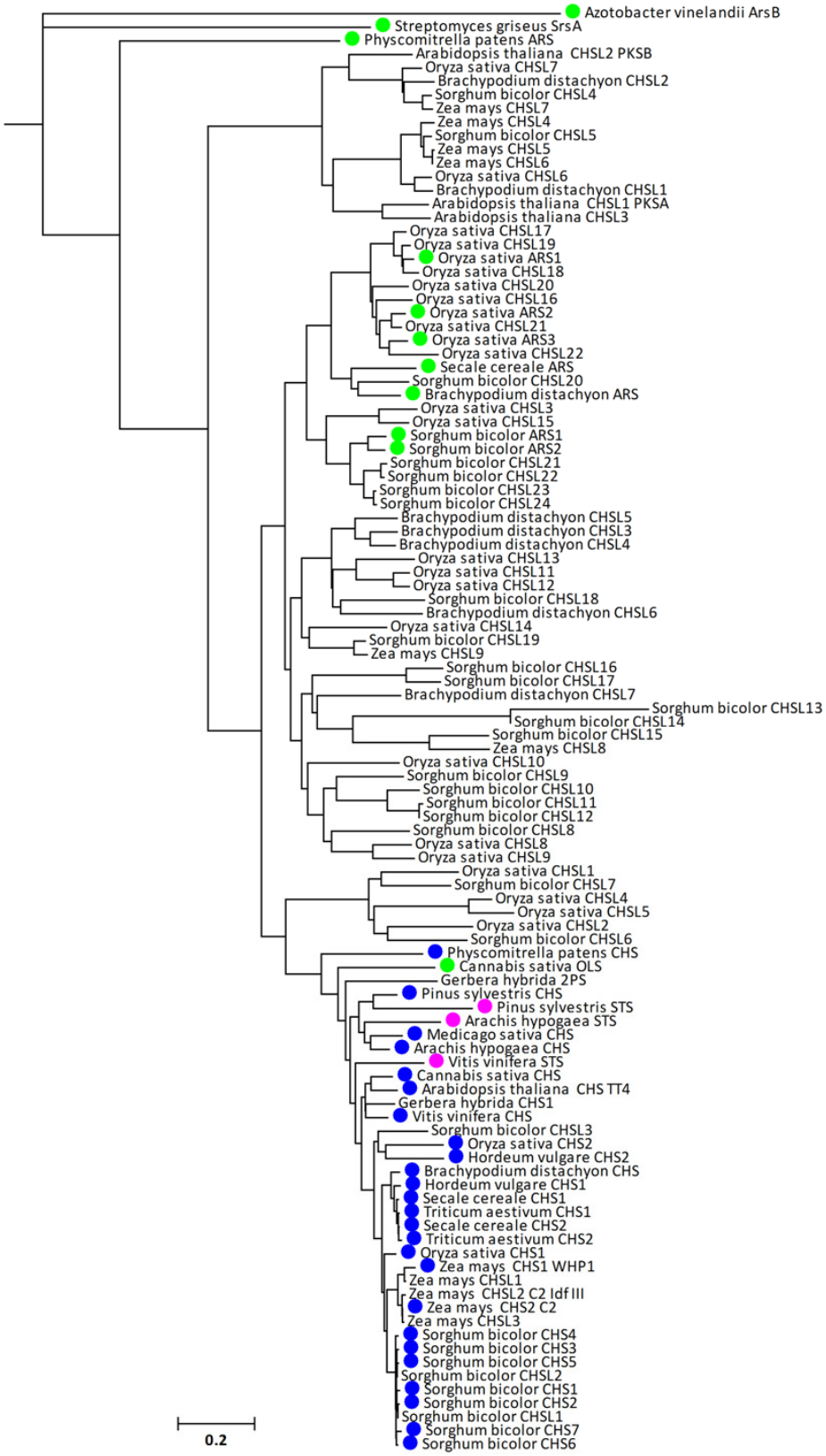


Figure 4.4 Phylogenetic relationships between ARSs and other related type III PKSs

(CHSs/CHSLs) in plants and bacteria. ARSs are indicated by bright green circles, CHSs are indicated by blue circles and STSs are indicated by pink circles.

Additionally, the existence of bacterial ARS enzymes ArsB and SrsA suggested that ARSs were ancient enzymes that had evolved a long time ago. Even though the bacterial ARSs have low identities to their plant counterparts, they still are functionally identical to the plant ARSs to some degree. The PpARS/PpCHS11 from *Physcomitrella patens* was closer to the bacterial ARSs compared to those in higher plants, which indicated the primitive origin of ARSs in the evolutionary process. It seems plausible that ARSs represent the original function in type III PKS family rather than CHSs (Baerson *et al.*, 2010). Eventually, the cluster patterns revealed by the phylogenetic studies of CHS/CHSL enzymes may be used to study the evolutionary origin and divergence of plant type III PKSs, and to predict the biochemical functions of the putative enzymes (Jiang *et al.*, 2008).

4.3.2 Biochemical function of ARSs

The functional difference between ARSs and other type III PKSs must be due to structural differences determined by amino acid sequences. In order to identify active site residues of BdARS and ScARS, their amino acid sequences were compared to two PKS templates, MsCHS from *Medicago sativa* (Ferrer *et al.*, 1999) and Gh2PS (2-pyrone synthase) from *Gerbera hybrida* (Jez *et al.*, 2000), whose crystal structures had been reported, as well as the ARSs from *S. bicolor* and *O. sativa* that had been functionally characterized recently (Cook *et al.*, 2010). The key residues putatively associated with the catalysis and CoA binding were predicted correspondingly. Overall, the residues Cys164, His303 and Asn336 (numbering in MsCHS) known to form the catalytic triad conserved in all plant type III PKSs were found in both BdARS and ScARS. Additionally, some other key residues putatively associated with CoA binding were also noticed (Figure 4.5). In contrast, distinct residues differing from CHS/CHSL representatives were observed to be consistent in all the ARS enzymes aligned. In particular, the substitutions of Thr132, Met137, Thr197 and Gly256 (numbering in MsCHS) by Tyr, Ala, Ala/Cys and Met residues were found in those

cereal ARSs that had been characterized as well as in the putative BdARS and ScARS in the current work. These critical amino acid residues had been predicted to play important roles in determining substrate specificities (Cook *et al.*, 2010). The presence of these residues in all the grass ARS enzymes strongly suggests that they play a significant role in substrate utilization and differentiate the enzymatic function of ARSs from CHSs.

```
MsCHS     ............MVSVAEIRQAQRAEGPATIMAIGTANPA.NCVEQSTYPDFYFKITNSEHKVELKEKFQRMCD..KSM   65
Gh2PS     .......MGSYSSDDVEVIREAGRAQGLATIIAIGTATPP.NCVAQADYADYYFRVTKSEHMVDLKEKFKRICE..KTA   70
ScCHS1    .........MAATMTVEEVRKAQRAEGPATVLAIGTATPA.NCVYQADYPDYYFKITKSDHMADLKEKFKRMCD..KSQ   68
ScCHS2    .........MAATMTVEEVRKAQRAEGPATVLAIGTATPA.NCVYQADYPDYYFKITKSDHMADLKEKFKRMCD..KSQ   68
ScARS     MGSIGTTNGNGIGHGSAAVARRQHAEGPAAMLGIGTANPTGVEVPQNIFAENLFRVTKSDHLTELQLKLTRICE..KTG   78
BdARS     ...MTRANGNGTVPVRDNRRSMQHAEGPAAVLAIGTANPTSTIAQQDQFADQFFRVTNSDDLTDLKAKFERICD..KTG   75
SbARS1    ...MGSMGKALP.ATVDEIRRAQRAEGPAAVLAIGTANPP.TIMPQDDYPDYYFRVTNSEHLTDLKAKLSRICNHNKSG   75
SbARS2    ......MGSAPPAATVQEMRRAQRADGPAAVLAIGTANPP.SIMPQDDYPDYYFRVTNSEHLTDLKAKLSRICNHNKSG   73
OsARS1    .....MPG.TATAAVVDSRPCTQHAEGPAAVLAIGTANPT.NIVYQDGFTDYYFGLTESEHLTELKDKMKRICH..RSG   71
OsARS2    .....MPG.ATTAAIVDSRRGTQHSEGPATILAIGTANPE.NIMFQDNFADYYFGLTKSEHLTELKEKMKRICH..KSG   71
OsARS3    .....MPGAATTAAVVDSRRSAQRAEGPATIIAIGTANPA.NIVPQDNFADYYFGLTKSEHLTELKDKMKRICK..KSG   72

MsCHS     KRRYMYLTEEILKDNPRVCEYMAPSLAARQDMAVVVVPRLGKEAAVKAIKEWGQPKSKITHLIFCTTSGVDMPGADYQLT  145
Gh2PS     KKRYLALTEDYLQENPTMCEFMAPSLNARQDLVVTGVPMLGKEAAVKAIDEWGLPKSKITHLIFCTTAGVDMPGADYQLV  150
ScCHS1    RKRYMHLTEEILQDNPNMCAYMAPSLDARQDIVVVEVPKLGKAAAQKAIKEWGQPRSKITHLVFCTTSGVDMPGADYQLT  148
ScCHS2    RKRYMHLTEEILQDNPNMCAYMAPSLDARQDIVVVEVPKLGKAAAQKAIKEWGQPRSKITHLVFCTTSGVDMPGADYQLT  148
ScARS     DKRHFHLTEETLAAHPELYDHEAPSLDNRIAMTVDAVPKLAQGAAKAIAEWGRPASEITHLVFSTYSAMGAPSADLRLA  158
BdARS     EKRHFHMTEEMLLAHPEFLDRDQPSLDARIEIVATAVPKLAESAARKAIAEWGRPASDITHLIFSTYSGCRAPAADLELA  155
SbARS1    RQRYLHLNEELLAANPGFIDPKRPSLDERVEMASAAVPELAAKAATKAIAEWGRPATDITHLIFSTYSGARAPSGDRRLA  155
SbARS2    RQRYLHLNEELLAANPGFIDPKRPSLDERVEMASAAVPELAAKAAAKAIAEWGRPATDITHLIFSTYSGARAPSGDRRLA  153
OsARS1    EKRYIHLDEKLIREHPEIIDKHMPSLETRVDIVTTEIPKLAESAARKAIAEWGRPAIDITHLIFSTYSGCSAPSADLKLA  151
OsARS2    EKRYIHLDAELISVHPEIIDKHLPSLETRVDIVATEVPKLAESAARKAIAEWGRPATDITHLIFSTYSGCRAPSADLQLA  151
OsARS3    EKRYIHLDEEIIRAHPEIIDKHQPSLEARVEIAAAEVPKLAESAARKAIAKWGRPATDITHLIFSTYSGCRAPSADLQLA  152

MsCHS     KLLGLRPYVKRYMYQQGCFAGGTVLRLAKDLAENNKGARVLVVCSEETPVTFRGPSDTHLDSLVGQALFGDGAAALIVG  225
Gh2PS     KLLGLSPSVKRYMLYQQGCAAGGTVLRLAKDLAENNKGSRVLIVCSEITAILFHGPNENHLDSLVAQALFGDGAAALIVG  230
ScCHS1    KMLGLRPSVKRLMYQQGCFAGGTVLRLAKDLAENNRGARVLVVCSEITAVTFRGPHE..FDSLVGQALFGDGAAAVIVG  226
ScCHS2    KMLGLRPSVKRLMYQQGCFAGGTVLRLAKDLAENNRGARVLVVCSEITAVTFRGPHESHLDSLVGQALFGDGAAAVIIG  228
ScARS     ALLGLRPTVSRTILSLHGGYGGGRALGLAREIAENNRGARVLVACAEITLVCFGGPDG...GNLVGHALFGDGACAVIVG  235
BdARS     TLLNLRPTVGRTILSLHGGYGGGRALHLAKEIAENNRGARVLVACSEITLVCFNGPDG...SNLVGHALFGDGACAAIVG  232
SbARS1    SLLGLRPTVSRTILNLHGGYGGGRSLQLAKEIAENNRGARVLVACSELTLIAFYGPEGGCVDNIIGQTLFGDGACAVVVG  235
SbARS2    SLLGLRPTVSRTILSLHGGYGGGRALQLAKEIAENNRGARVLVACSELTLIAFYGPEGGCVDNIIGQTLFGDGACAVIVG  233
OsARS1    SLLGLNPSVSRTILSLHGGSGGGRALQLAKEIAENNRDARVLIACAELTLICFSNPDE...SKIVGHGLFGDGACAIIVG  228
OsARS2    SLLGLRPSVSRTILSLHGGSGGGRALQLAKEIAENNRGARVLIACSELTLICFSTPDE...SKIIGHGLFGDGACAVIVG  228
OsARS3    SLLGLRPSVSRTILSLHGGSGGGRALQLAKEIAENNRGARVLVALSELTLVCFSTPDE...SKIVGHGLFGDGACAIIVG  229

MsCHS     SDPIPEIE..KPIFEMVWTAHTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNIDKALIEAFQ....PLNISDY....  295
Gh2PS     SGPHLAVE..RPIFEIVSTDQTILPDTEKAMKLHLREGGLTFQLHRDVPLMVAKNIENAAEKALS....PLGITDW....  300
ScCHS1    ADPDESVE..RPLFQLVSASQTILPDSEGAIDGHLREVGLTFHLLKDVPGLISKNIERALEDAFK....PLGIDDW....  296
ScCHS2    ADPDESIE..RPLFQLVSASQTILPDSEGAIDGHLREVGLTFHLLKDVPGLISKNIERALEDAFK....PLGIDDW....  298
ScARS     AGPFRDGE.QSPIFEMVHATQTTVPKTEHVLGMQVSGSGVDFHLAIQVPTLIGQNVERCLLDAFRGGDDGGGDDDGGAHL  314
BdARS     AGPFDAASGERPLFEMVAATQTTVPKTEHVLGMQVAGGGIDFHLAIQVPMLIGQNVEQCLRDAFR.ATLGDDDEDG...I  308
SbARS1    ADPD..AAVERPLFEMAFATQTTIPESEDAISMQYSKCGMEYHLSSKVPRLIGCNVERSLVDTFR....TLGVTAA....  305
SbARS2    ADPVG.APAERPLFEMVFASQTTIPETEDAISMQYSKCGMEYHLSSRVPRVLGSNVERCLVDTFR....TLGVSVA....  304
OsARS1    ADPLVDGE..RPLFEMVLASQTTIPGTEHALGMQTTSNGIDFHLSIQVPTLIKDNIRQCLLNTFR....SVGNMDP....  298
OsARS2    ADPSVDGE..CPLFEMVAASQTMIPGTEHALGMQATSSGIDFHLSIQVPTLIKDNIHQCLLNAFR....SVGNTDP....  298
OsARS3    AGPFSDGE..CPLFEMVAASQTMIPGTEHALGMQATSTGIDFHLSVQVPMLIKDNIQQSLLESFQ....SVGYTDP....  299

MsCHS     ...........NSIFWIAHPGGPAILDQVEEKLGLKPEKMKATREVLSEYGNMSSACVLFILDEMRKKSVQAGLKTTGEGL  365
Gh2PS     ...........NSVFWVHPGGRAILDQVERKLNLKEDKLRASRHVLSEYGNLISACVLFIIDEVRKRSMAEGKSTTGEGL  370
ScCHS1    ...........NSVFWIAHPGGPAILDMVEAKVNLNKERMRATRHVLSEYGNVSSACVLFIMDEMRKRSAEDGHTTTGEGM  366
ScCHS2    ...........NSVFWIAHPGGPAILDMVEAKVNLNKERMRATRHVLSEYGNVSSACVLFIMDEMRKRSAEDGHTTTGEGM  368
ScARS     PSPLSGNGKVNDLFWAVHPGGRPILDNIDKVLQLEPEKLGASRHVLREYGNVSGATIVFVLDELRRRRS........LLP  386
BdARS     PC.......VNDLFWAVHPGGRPILDNIDAVLKLEKGKLAASRQVLREYGNVSGATIVFVLDELRRRREKDGSGG.HQLP  380
SbARS1    .........VNDLFWAVHPGGRAILDNIEEVLGLEDDKLAASRHVLSEFGNVSGTTVIFVLDELRRRRAAAAKQG.GETP  375
SbARS2    .........VNDLFWAIHPGGRAILDNIEEVLRLEDGKLAASRHVLSEFGNVSGTTVIFVLDELRRRRAAAAKQG.GQAP  374
OsARS1    ........NVNDLFWAVHPGGRAILDNIEGELQLQPAKLAASRHVLSEYGNVSGTTIAFVLDELRCRREKEGDE..HQQP  368
OsARS2    ........NVNDLFWAVHPGGRAILDNIEDKLQLHPCKLAASRQVLSEYGNVSGATIAFVLDELRRRREKEQDI..QQQP  368
OsARS3    ........DVNNLFWAVHPGGRAILDNIEGKLQLQPWKLAASRQVLREFGNVSGATIAFVLDELCHRREKDEDE..SQQH  369

MsCHS     DVGVLFGFGPGLTIETVVLHSVAI.......  389
Gh2PS     DCGVLFGFGPGMTVETVVLRSVRVTAAVANG  401
ScCHS1    DVGVLFGFGPGLTVETVVLHSVPVTA.....  392
ScCHS2    DVGVLFGFGPGLTVETVVLHSVPVTA.....  394
ScARS     EVGAMLAFGPGVTIETMVLRCPR........  409
BdARS     EVGALLAFGPGITIETMVLRCPPG..LKGN.  408
SbARS1    EVGVLMAFGPGITIETIVLHTPSNPELEGN.  405
SbARS2    EVGVMMAFGPGITVETMVLHAPSNLELEGN.  404
OsARS1    EVGVMLAFGPGITIEAMVLRNPLS.......  392
OsARS2    EVGVLLAFGPGVTIESIVLRNPLSRGLKEN.  398
OsARS3    EVGVMLAFGPGITIETIVMRNPLARGLKQN.  399
```

Chapter 25 Figure 4.5 **Alignment of amino acid sequences of ARS and CHS/CHSL representatives. It includes BdARS from *B. distachyon*, Gh2PS from *G. hybrida*, MsCHS from *M. sativa*, OsARSs from *O. sativa*, ScARS and ScCHSs from *Secale cereale*, and SbARSs from *Sorghum Bicolor*. The residues (Cys164, His303 and Asn336, numbering in MsCHS) for the catalytic triad of all plant type III PKSs are highlighted by asterisks, the residues for the CoA binding are highlighted by triangles, and the residues for the functional diversity are highlighted by dots. The key residues associated with the C2→C7 aldol condensation mechanism are outlined.**

Chapter 26

To sum up, gene cloning of ARSs was carried out in this chapter, which revealed one putative ARS from *B. distachyon* (designated as BdARS) and one putative ARS from *Secale cereale* (designated as ScARS). After BdARS and ScARS genes were cloned, their enzymatic function had to be tested. Therefore, biochemical characterization and biological characterization of the two putative ARSs will be described in the chapter below. It is to be noted that there are two ARSs in *Sorghum bicolor* and three ARSs in *O. sativa* that have been functionally characterized so far. Additionally, genome searching showed that a few more members from *S. bicolor* and *O. sativa* in the type III PKS family pending characterization fell into the ARS cluster. They likely perform in the same manner as ARS does. On the other hand, *Secale cereale* has such a large genome size, but so far only one ARS was cloned from this species in this work. It is highly possible that more ARS enzymes exist in *S. cereale* which might be expressed in different tissues and/or at different developmental stages. Therefore, it turns out to be very interesting to identify more ARSs from *S. cereale* as well as related PKSs from related grass species as shown in Figure 4.4. With the increasing number of CHS/CHSL enzymes to be characterized, the divergent evolutionary relationships in the type III PKS family will become clear.

Chapter 5 Biochemical and biological characterization of alkylresorcinol synthases (ARSs) from *Brachypodium distachyon* and *Secale cereale*

5.1 Introduction

Alkylresorcinols are bioactive secondary metabolites formed by diverse plant species as well as fungi and bacteria. In plants, the accumulation of alkylresorcinols at/near tissue surfaces has raised the question whether these compounds play a role as first line of defense against pathogens and herbivores. However, direct evidence concerning the biological function of alkylresorcinols lining the surface of grass leaves is still lacking. In order to test this idea, molecular tools must be generated which will enable the manipulation of alkylresorcinol levels so that they can be correlated with plant defense performance. The overall goal of the present work was to provide such molecular tools for future work. To reach this goal, first model species had to be selected in which alkylresorcinols accumulate at/near the surface. Previous studies in our lab, together with the data presented in Chapter 3 in the current work, have shown that the grass species *Brachypodium distachyon* and *Secale cereale* produce very-long-chain (VLC) alkylresorcinol series which accumulate within the cuticular wax mixtures at/near the surface of their leaves. Based on these chemical data, the same plant species were then used to identify and clone genes encoding type III polyketide synthases (PKSs), as candidates for enzymes involved in the biosynthesis of cuticular alkylresorcinols. However, the mere sequence information on the putative alkylresorcinol synthases (ARSs) in Chapter 4 is not sufficient to prove the biochemical function and biological role of the enzymes. The objective of the work presented in this chapter was to provide experimental evidence that the two candidate enzymes, BdARS and ScARS, are responsible for formation of cuticular alkylresorcinols, and that they hence present the desired tools for testing the biological function of alkylresorcinols at the plant surface.

The potential enzymatic functions of BdARS and ScARS were first tested by *in vivo* biochemical characterization. Yeast *Saccharomyces cerevisiae* was chosen as the heterologous expression system for *in vivo* characterization, because it is alkylresorcinol-free and has the necessary substrates for ARS enzymes, in the form of

various fatty acyl-CoAs as starter substrates and malonyl-CoA as extenders. Yeast further offers the opportunity to test which fatty acyl-CoAs may serve as ARS substrates, since wild type yeast and the yeast mutant *elo3*Δ have different pools of VLC fatty acyl-CoA starters (Oh *et al.*, 1997). In particular, the following questions were addressed:

a) Do the two putative enzymes indeed have ARS activities?
b) What are their substrate/product profiles correspondingly?

After the biochemical characterization, the two ARS candidates were further characterized in terms of their biological properties, including gene expression studies and subcellular localization investigation. To this end, the spatial and temporal expression patterns, and the subcellular localization of the proteins were assessed in order to correlate them with product accumulation. It was hypothesized that, if the ARSs were involved in cuticular alkylresorcinol formation, then their expression patterns should parallel the patterns of product accumulation. The biological characterization aimed at answering the following questions:

c) What are the organ-specific expression patterns of *BdARS* and *ScARS*?
d) What is the spatial expression pattern of *ScARS* along the leaf of *S. cereale*? Does the expression pattern correlate with the deposition of alkylresorcinols within the leaf cuticle?
e) What is the subcellular localization of BdARS and ScARS?

5.2 Results

The characterization of the putative ARS genes was carried out in two sets of experiments, first to test their biochemical characteristics and then to further assess their biological properties, namely their expression patterns and subcellular localization. The two series of experiments will be described each in two sections below, first detailing the *in vivo* characterization of the enzymes in wild type yeast (Chapter 5.2.1) and the yeast mutant *elo3*Δ (Chapter 5.2.2), and then describing the gene expression (Chapter 5.2.3) and subcellular localization studies (Chapter 5.2.4).

5.2.1 Functional expression of BdARS and ScARS in wild type yeast

In order to test the biochemical functions of the two putative ARS enzymes, they were expressed in wild type yeast. The full-length cDNAs of *BdARS* and *ScARS* were amplified, cloned into the yeast expression vector pESC-URA and transformed into wild type yeast. Yeast transformants harboring pESC-URA:BdARS, pESC-URA:ScARS and pESC-URA empty vector were extracted after induction of expression with galactose. For rapid screening for the presence of alkylresorcinol homologs in recombinant yeast cells, thin layer chromatography (TLC) analysis was conducted first. Extracted lipids from yeast cells were separated by a solvent system of chloroform-ethyl acetate (7:3, v/v), using synthetic standard 5-*n*-tridecylresorcinol as a reference (Figure 5.1, lane 1). The lipid mixtures from yeast harboring pESC-URA:ScARS (Figure 5.1, lane 3) and pESC-URA:BdARS (Figure 5.1, lane 4) were both found to contain one fraction that was absent from the empty vector (Figure 5.1, lane 2). The fraction had retention similar to the alkylresorcinol standard, suggesting that the compounds might be alkylresorcinols.

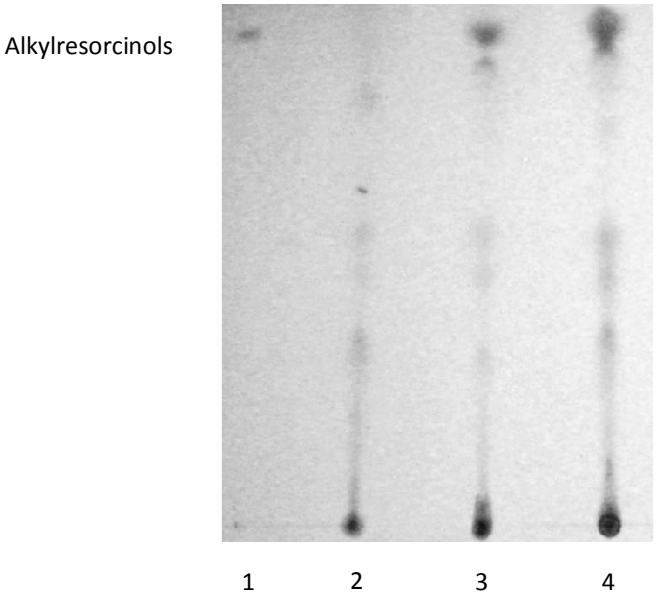
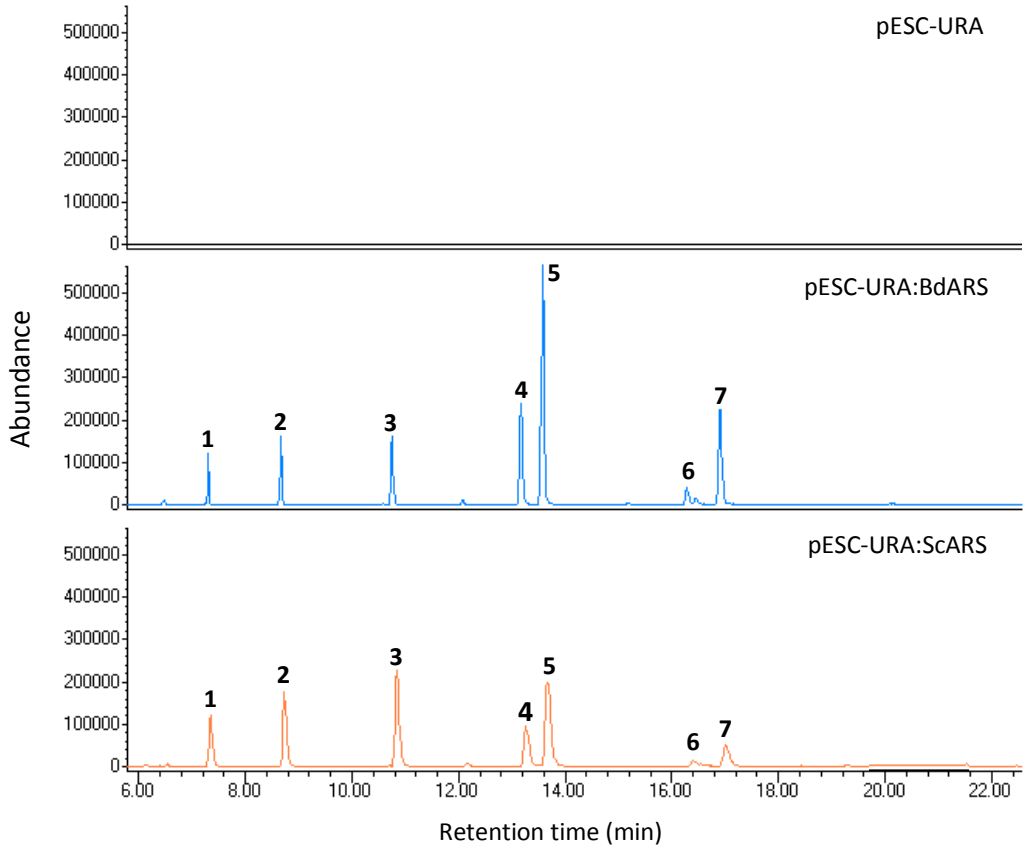
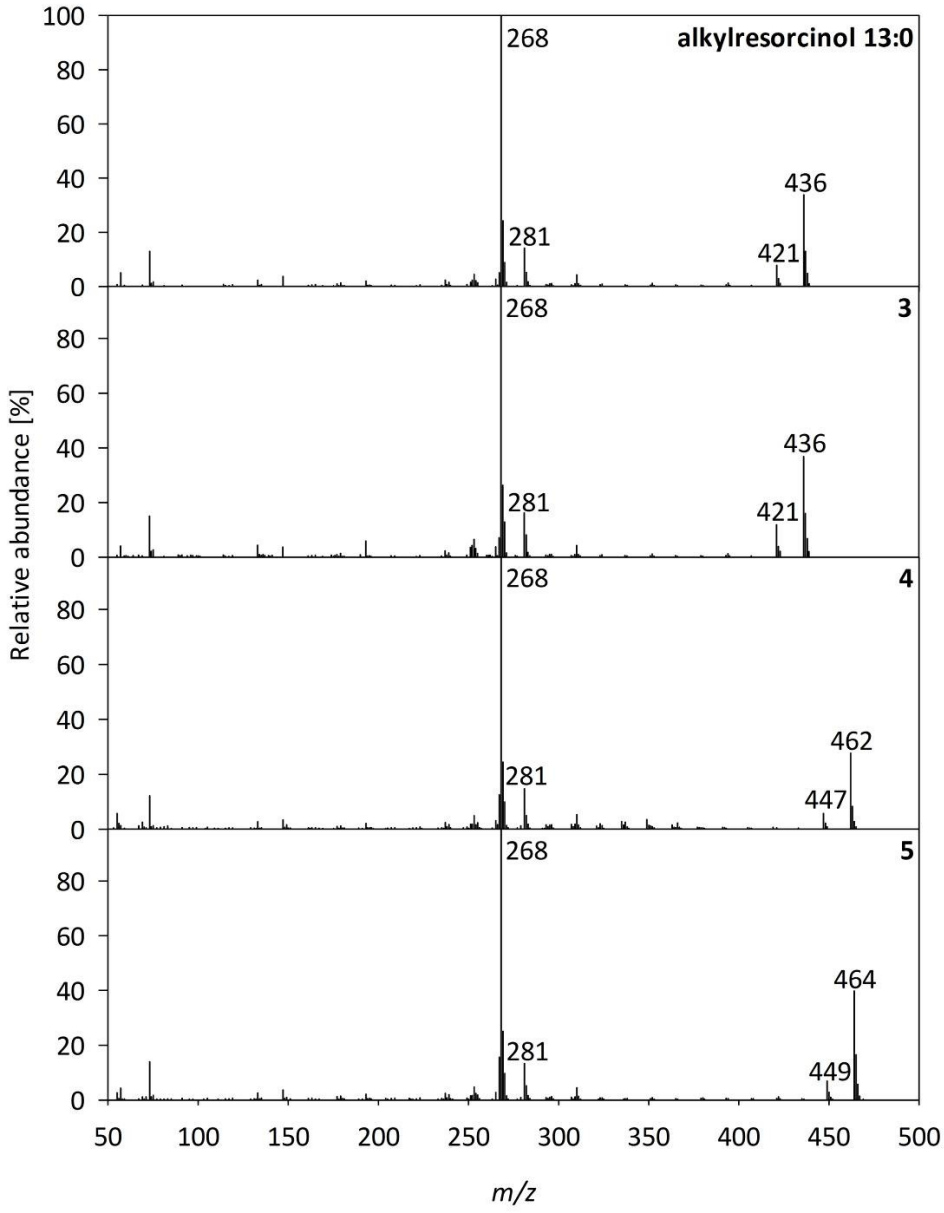


Figure 5.1 TLC analysis of yeast total lipids. **Lane 1**, Synthetic standard 5-*n*-tridecylresorcinol used as a reference. **Lane 2**, Wild type yeast expressing pESC-URA empty vector. **Lane 3**, Wild type yeast expressing pESC-URA:ScARS. **Lane 4**, Wild type yeast expressing pESC-URA:BdARS. The solvent system was chloroform-ethyl acetate (7:3, v/v).

In order to identify these novel compounds and reveal their detailed profile, the total lipids extracted from the three recombinant yeast lines were first trimethylsilyl (TMSi)-derivatized and then subjected to analysis by gas chromatography-mass spectrometry (GC-MS). The resulting data were analyzed using single ion monitoring at *m/z* 268, the fragment characteristic of all alkylresorcinols, to suppress background and better detect alkylresorcinol peaks. The recombinant yeast cells expressing BdARS and ScARS both showed a series of compounds numbered from **1** to **7** (Figure 5.2). Compounds **1** to **3**, **5** and **7** were found to have the characteristic MS fragments of alkylresorcinol homologs at *m/z* 73, 268 and 281, while differing in their molecular ions by *m/z* 28. This finding, together with the chromatographic behavior, suggested that the compounds belonged to a homologous series of alkylresorcinols differing by $-CH_2–CH_2-$ units. In order to further confirm the structure of these compounds, synthetic 5-*n*-tridecylresorcinol was co-injected. The standard had a MS fragmentation pattern (Figure 5.3) and GC retention behavior (data not shown) identical to the compound **3** under the current conditions. Thus, compound **3** was identified as alkylresorcinol 13:0, and compounds **1**, **2**, **5** and **7** were correspondingly determined as alkylresorcinol homologs 9:0, 11:0, 15:0 and 17:0 based on their respective molecular ions (see Chapter 3.2.2). Interestingly, two more compounds **4** and **6** eluted shortly before compounds **5** and **7**, and showed molecular ions $[C_6H_3(OTMSi)_2C_nH_{2n-1}]^+$ and corresponding fragments $[M-15]^+$ with two mass units less than **3** and **5**, suggesting that they were unsaturated alkylresorcinols containing one double-bond in the alkyl chains (Figure 5.3). Thus, compounds **4** and **6** were identified as alkylresorcinol 15:1 and 17:1. However, the exact position of the double bond cannot be determined based on the currently available information. The unsaturated alkylresorcinol homologs were novel compounds that had not been found in the total leaf wax of *B. distachyon* or *S. cereale*. In summary, all the compounds formed by transgenic yeast expressing BdARS and ScARS were determined as alkylresorcinols using GC-MS. Their formation must have been due to the presence of BdARS or ScARS, since corresponding products could not be detected in the empty vector control.

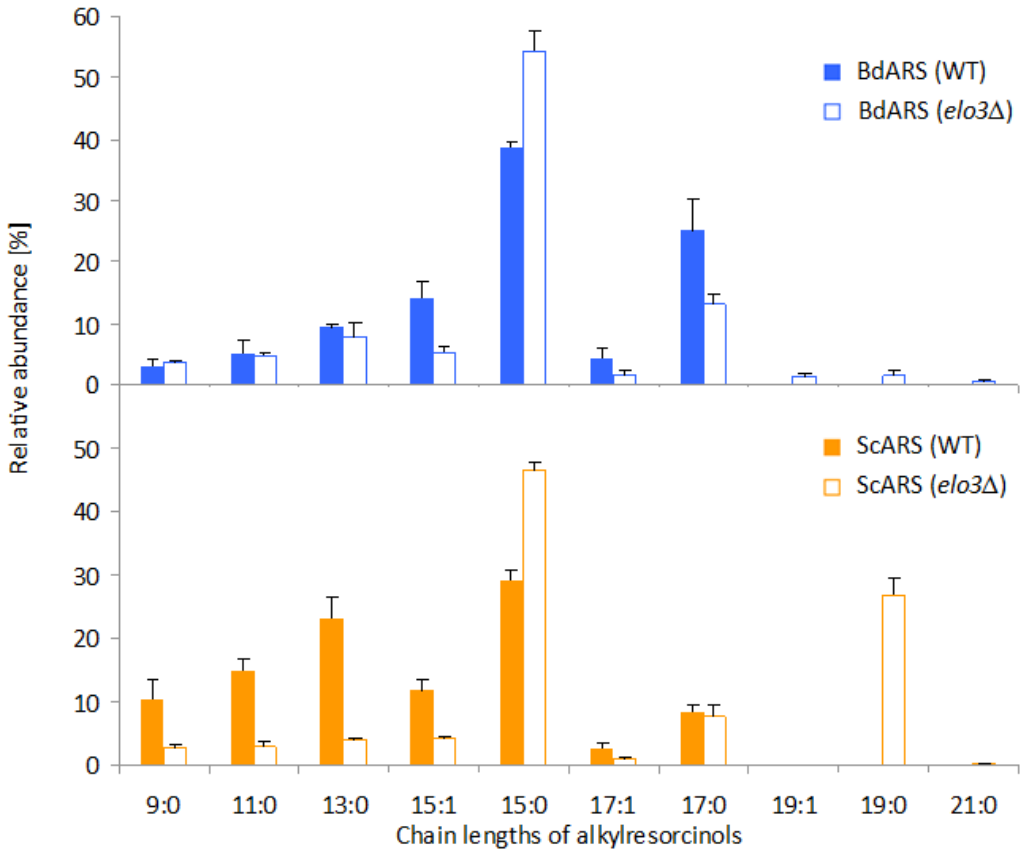


Chapter 27 Figure 5.2 **GC-MS analysis of alkylresorcinols produced by recombinant wild type yeast expressing BdARS and ScARS, respectively. Extracted chromatograms of ion *m/z* 268 are shown. In the empty vector control, no alkylresorcinol products were detected. In contrast, yeast expressing BdARS and ScARS were found to contain identical series of alkylresorcinols. Compounds were identified as** 1**, alkylresorcinol 9:0;** 2**, alkylresorcinol 11:0;** 3**, alkylresorcinol 13:0;** 4**, alkylresorcinol 15:1;** 5**, alkylresorcinol 15:0;** 6**, alkylresorcinol 17:1; and** 7**, alkylresorcinol 17:0.**



Chapter 28 Figure 5.3 **Mass spectra of TMSi derivatives of individual alkylresorcinol homologs. Peaks** 3, 4 **and** 5 **are shown as representatives from transgenic yeast lipids. Peak** 3**, formed by wild type yeast expressing ScARS, showing a fragmentation pattern identical to that of the authentic standard of alkylresorcinol 13:0 (**top**), is identified as alkylresorcinol 13:0 (molecular ion *m/z* 436). Compounds** 4 **and** 5**, formed by wild type yeast expressing BdARS, are identified as alkylresorcinol 15:1 (molecular ion *m/z* 462) and alkylresorcinol 15:0 (molecular ion *m/z* 464), correspondingly.**

The series of alkylresorcinol homologs in transgenic yeast expressing BdARS and ScARS both showed a broad range of alkyl side chains ranging from C_9 to C_{17}. However, the transgenic lines expressing BdARS and ScARS exhibited different chain length profiles of alkylresorcinol products (Figure 5.2). For BdARS enzyme, the most abundant homolog was alkylresorcinol 15:0, and other major homologs were alkylresorcinols 17:0 and 15:1. Expression of ScARS gave a homologous series with predominance of alkylresorcinol 15:0 as well, but for ScARS the next most abundant homologs were alkylresorcinols with shorter side chains. In particular, for BdARS the most abundant alkylresorcinols 15:0, 17:0 and 15:1 produced by wild type yeast accounted for 39%, 25% and 14%, respectively, comprising 78% of the total alkylresorcinol products, whereas for ScARS the major alkylresorcinol homologs 15:0, 13:0 and 11:0 constituted 29%, 23% and 15% in total alkylresorcinol products, making a contribution of 67% (Figure 5.4). In addition, trace amounts of alkylresorcinol homologs with even-numbered side chains were detected in the total lipids of the transgenic yeast as well (data not shown).



Chapter 29 Figure 5.4 **Profiles of alkylresorcinols from recombinant yeast lines (wild type, WT; and the *elo3*Δ) expressing BdARS and ScARS, respectively. The relative abundances of alkylresorcinol homologs are given as mean values (*n* = 3) ± SD.**

It should be noted that VLC alkylresorcinols (typically consisting of 19 and more carbons) were not detected in the yeast total lipids using the current GC conditions. In order to increase the sensitivity for detection in GC runs, TLC separation was employed for pre-purification of alkylresorcinols. When yeast lipids were first fractionated via TLC and then analyzed using GC-MS with single ion monitoring at *m/z* 268, no further alkylresorcinol homologs could be identified (data not shown). The failure to detect VLC alkylresorcinols might have been either due to a lack of enzyme activity with such substrates, or to the very low levels of VLC fatty acyl-CoAs that could serve as starter substrates in wild type yeast. To test these possibilities and elucidate whether BdARS and ScARS were able to accept VLC fatty acyl-CoA substrates, a yeast mutant with altered acyl chain length pools was used in the following experiments.

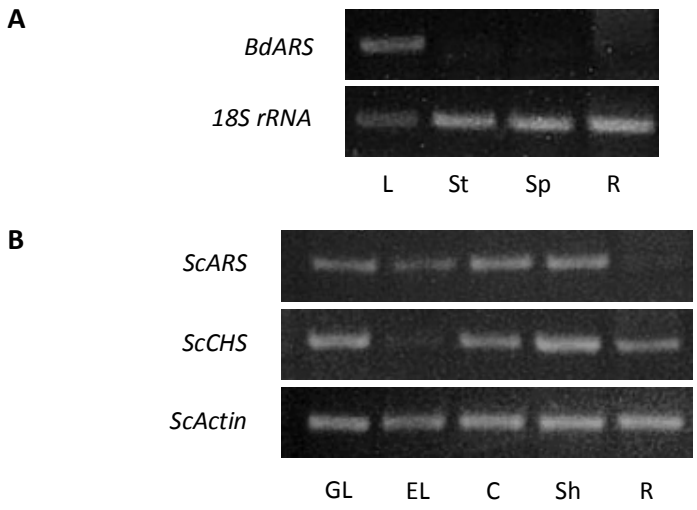
5.2.2 Functional expression of BdARS and ScARS in yeast mutant

To test the starter substrate specificity of BdARS and ScARS, the yeast mutant *elo3*Δ (ELO3, fatty acid elongase 3) was selected because it is known to have an increase of C_{20} and C_{22} fatty acyl substrates compared to the wild type yeast (Oh *et al.*, 1997). The yeast mutant was transformed, grown and analyzed as described above for heterologous expression experiments using the wild type yeast. In the yeast mutant *elo3*Δ transformed with empty vector as a negative control no alkylresorcinols could be detected. In contrast, the lines expressing BdARS and ScARS in the mutant background were found to contain homologous alkylresorcinols similar to the corresponding wild type yeast lines (Figure 5.4). However, expression of the ARSs in the yeast mutant background led to significant shifts in the chain length profiles of the alkylresorcinol products compared to that in the wild type yeast. In the line expressing BdARS in the mutant background, the relative abundances of alkylresorcinols 9:0, 11:0 and 13:0 were similar to those in the wild type, whereas

alkylresorcinols 15:1, 17:1 and 17:0 had decreased and alkylresorcinol 15:0 increased largely. Most importantly, new homologs of alkylresorcinols 19:1, 19:0 and 21:0, accounting for 3%, were found in the total alkylresorcinol mixture. Upon expression of ScARS in the yeast mutant, alkylresorcinol homologs with the alkyl chains ranging from C_9 to C_{17} were also reduced in comparison with those in the wild type, and alkylresorcinol 15:0 had increased. Alkylresorcinol 19:1 was not detected in this mutant line. However, alkylresorcinol 19:0 accumulated dramatically, constituting 27% of the alkylresorcinol products. Low amount of alkylresorcinol 21:0 was also detected in the line expressing ScARS in the mutant background.

5.2.3 Gene expression studies of *BdARS* and *ScARS*

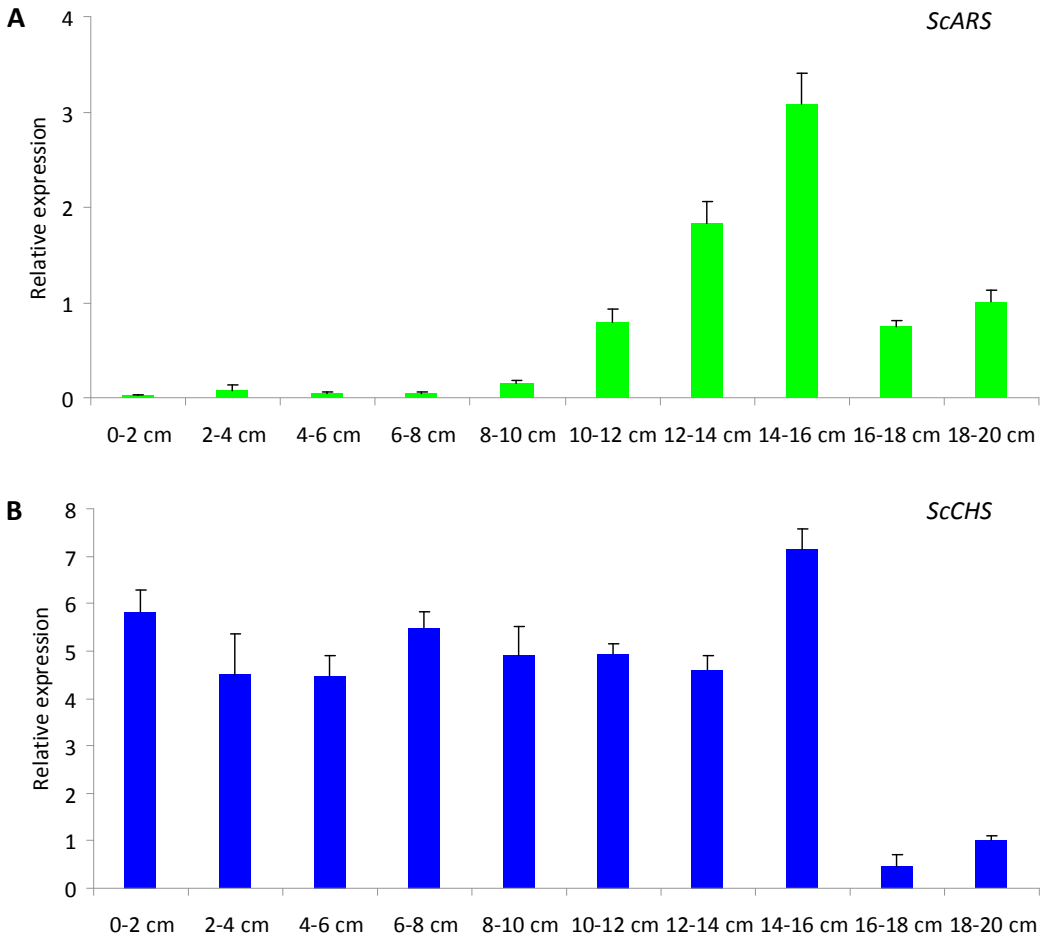
In order to assess the organ-specific expression of the gene encoding BdARS in *B. distachyon* and the gene encoding ScARS in *S. cereale*, semi-quantitative RT-PCR analyses were performed, respectively. Different organs of *B. distachyon* were examined, including expanding first leaf of 3-week-old seedlings, stem and spikes, as well as roots of 7-day-old seedlings. *BdARS* was found to be expressed only in leaves but not in stems, spikes or roots (Figure 5.5A). Due to the different developmental patterns of *S. cereale*, a different set of organs was selected for *ScARS* expression analysis, focusing on young leaves, cotyledons, sheaths and roots rather than stems or spikes. Another type III PKS gene, the ubiquitous chalcone synthase (CHS) gene, was employed as a reference that was expected to be expressed relatively broadly. Accordingly, *ScCHS* was found to be expressed in all the organs tested, with high expression levels in green leaves, cotyledons, sheaths as well as roots, and relatively low levels in etiolated leaves. *ScARS* was found to be expressed in all the organs tested except roots. The expression level of this gene was also lower in etiolated leaves than in other organs, consistent with that of *ScCHS* (Figure 5.5B).



Chapter 30 Figure 5.5 **RT-PCR analyses of gene expression patterns in *B. distachyon* (**A**) and *S. cereale* (**B**). Expression of *18S rRNA* was used as a constitutive control in the analysis of *BdARS*. Expression of *ScActin* was used as a constitutive control in the analysis of *ScARS*. Leaves of both *B. distachyon* and *S. cereale* were 3 weeks old, consistent with those used for wax analysis and gene cloning. Etiolated leaves of *S. cereale* were approximately 3 weeks old. Roots were harvested seven days after germination on agar plates. L, leaf; St, stem; Sp, spike; R, root; GL, green leaf; EL, etiolated leaf; C, cotyledon; and Sh, sheath.**

To further determine whether the ARS gene expression pattern was correlated with the accumulation of cuticular alkylresorcinols on different positions of the leaf of *S. cereale*, *ScARS* was subjected to quantitative RT-PCR analysis in different segments of the leaf. The sampling strategy used in this experiment was consistent with that described for the investigation of alkylresorcinol deposition in the cuticular wax during development of *S. cereale* leaves (Figure 1.3A). The first leaf of *S. cereale* was harvested when it was 20-cm long and thus just terminating expansive growth (growth stage IV), and cut into 2-cm long segments. The basal segment, designated as "segment 18-20 cm", thus represented the part of the leaf that was situated next to the point of emergence. It was chosen as a reference for normalizing data to enable comparison with all other segments. *ScARS* was found to be most strongly expressed in the leaf segments 10-20 cm away from the tip, and peaking in the segment 14-16 cm (Figure 5.6A). Expression was barely detectable in the distal regions within 10 cm

from the leaf tip. For comparison, *ScCHS* expression was profiled as well along the leaf of *S. cereale*. In contrast to *ScARS*, *ScCHS* was highly expressed in all leaf segments more than 4 cm away from the point of emergence (Figure 5.6B).

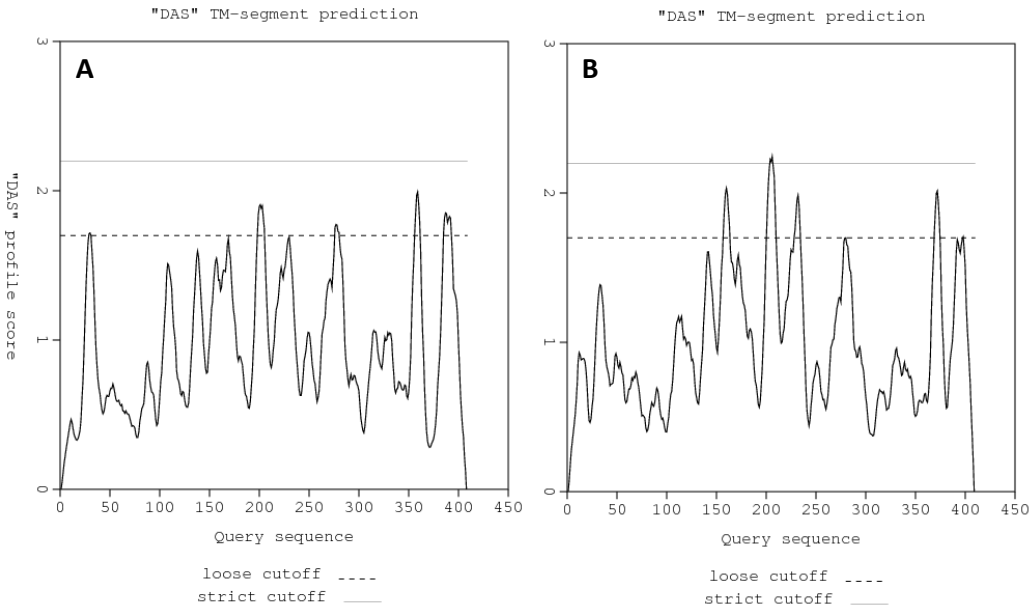


Chapter 31 Figure 5.6 **Quantitative RT-PCR analysis of relative gene expression levels of *ScARS* and *ScCHS* in *S. cereale* as a function of position along the first true leaf. The 20-cm long leaf was cut into 2-cm long segments that were used for the analysis. Segment labels represent the distance of the leaf pieces from the point of emergence. *18S rRNA* was used as a reference gene. Relative expression levels were calculated using the $2^{-\Delta\Delta C_T}$ method and the expression value of the 18-20 cm segment sample as the calibrator. The relative expression levels are given as mean values (*n* = 3) ± SD.**

Chapter 32

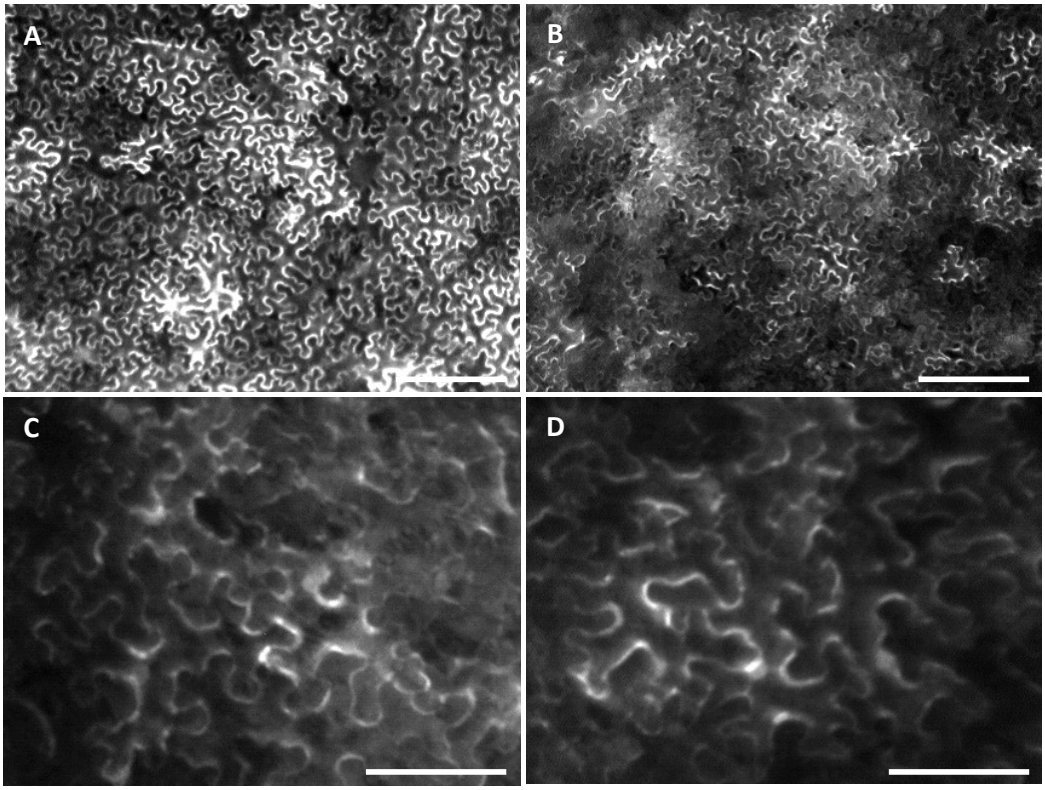
5.2.4 Subcellular localization of BdARS and ScARS

In the final experiment, the subcellular localization of the BdARS and ScARS enzymes was investigated. In preparation for this, the sequence information acquired in Chapter 4 was first used for *in silico* analyses to test possible membrane associations of the proteins. The amino acid sequences of BdARS and ScARS were employed for transmembrane domain prediction based on the "Dense Alignment Surface" (DAS) algorithm using the DAS transmembrane prediction server (http://www.sbc.su.se/~miklos/DAS) (Cserzo *et al.*, 1997). The results showed that both BdARS and ScARS might The results showed that both BdARS and ScARS might have two transmembrane segments, however the predicted hydrophobic domains are relatively short and their hydrophobicity is relatively small (Figure 5.7). The transmembrane domain prediction was therefore repeated with a second program, ARAMEMNON (http://aramemnon.uni-koeln.de/). Again, there was no strong evidence for a potential membrane association of the ARSs. To put this latter result into perspective, the ortholog PKSA from *Arabidopsis thaliana* was used as a reference for transmembrane prediction in ARAMEMNON. Even though this protein had been shown experimentally to have an ER localization (Kim *et al.*, 2010), the computational analysis did not predict any transmembrane domains.



Chapter 33 Figure 5.7 **Prediction of transmembrane domains for BdARS from *B. distachyon* (**A**) and ScARS from *S. cereale* (**B**).**

To verify the ambiguous results based on the computational prediction, the subcellular localization of BdARS and ScARS was investigated experimentally. To this end, both proteins were fused with the green fluorescent protein (GFP) at either their N- or C-terminus, and put under the control of the constitutive cauliflower mosaic virus 35S promoter. The resulting 35S:sGFP-BdARS/ScARS or 35S:BdARS/ScARS-sGFP were transiently expressed in *Nicotiana benthamiana* leaves after agrobacterium-mediated infiltration. The GFP fluorescence was examined using light microscopy first. In both cases of N- and C-terminal fusion with GFP, strong fluorescence was observed in *N. benthamiana* leaves transiently expressing BdARS and ScARS. Thus, the fusion proteins contained properly folded GFP and could be used for localization studies. Since the C-terminal fusions seemed to give stronger signals for both BdARS and ScARS, only the results for 35S:BdARS-sGFP and 35S:ScARS-sGFP will be illustrated here (Figure 5.8).



Chapter 34 Figure 5.8 **Light microscopy of transient expression of 35S:BdARS-sGFP (**A **and** C**) and 35S:ScARS-sGFP (**B **and** D**) in *N. benthamiana* leaves. Bars = 250 µm for** A **and** B**, and 100 µm for** C **and** D**.**

Furthermore, the fluorescence of GFP was examined by laser scanning confocal microscopy. Intense GFP fluorescence was observed in a reticulate pattern for both BdARS (Figure 5.9A) and ScARS (Figure 5.9B), seemingly characteristic of the endoplasmic reticulum (ER) membrane. Additionally, the reticulate structure seems to be connected with the nuclear envelope and the continuous observation showed the movement of reticulate structure over time, which is typical of the ER in vital cells (Figure 5.9B). These results taken together suggested that BdARS and ScARS are both associated with the ER membrane. Additionally, hexyl rhodamine B was used to counterstain the ER network in order to confirm the proposed ER localization of BdARS and ScARS by co-localization. However, the staining failed and the signal from hexyl rhodamine B could not be detected from the ER membrane (data not shown).

A

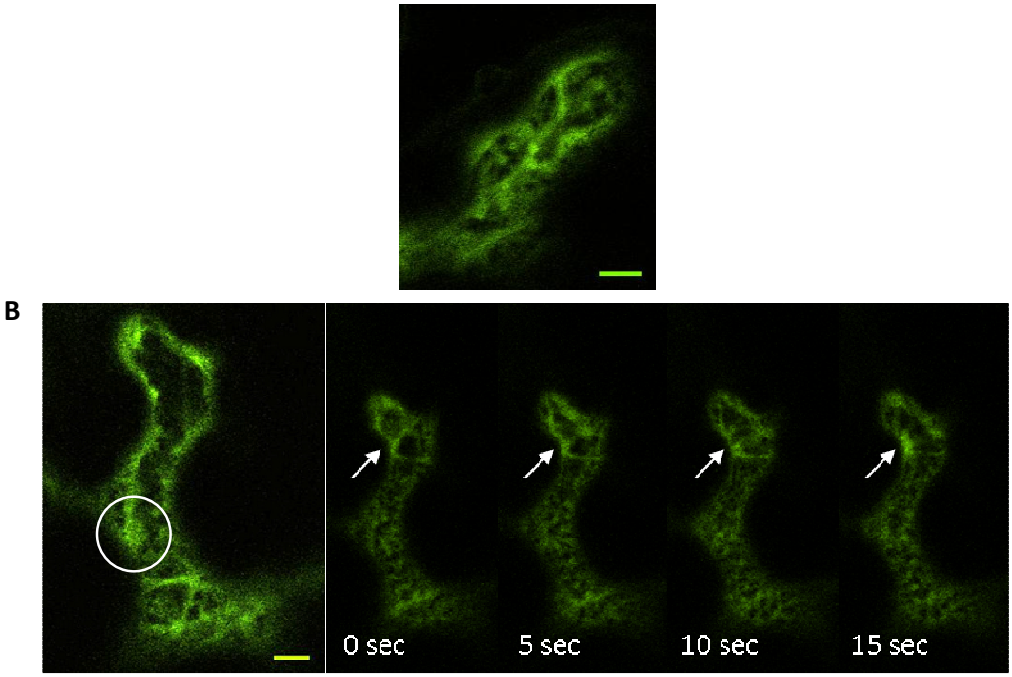


Figure 5.9 Subcellular localization of BdARS (**A**) and ScARS (**B**). 35S:BdARS-sGFP and 35S:ScARS-sGFP were transiently expressed in *N. benthamiana* leaves. A connection of the reticulate expression pattern of 35S:ScARS-sGFP with the nuclear envelope is indicated by a circle and the movement of the the reticulate expression pattern was tracked over time as indicated by an arrow. Bars = 10 µm.

5.3 Discussion

The experiments described in this chapter aimed at the biochemical and biological characterization of the two ARS candidate enzymes, BdARS and ScARS, to test their involvement in the biosynthesis of cuticular alkylresorcinols in *B. distachyon* and *S. cereale*, respectively.

5.3.1 Biochemical characterization in yeast

The first set of experiments was to elucidate the enzyme activities of the two proteins, to see whether they had the ARS activities predicted based on the amino acid sequences determined in Chapter 4. Overall, the results of *in vivo* heterologous expression in two different yeast strains confirmed that both BdARS and ScARS have ARS activities. Both enzymes are able to catalyze consecutive decarboxylative

condensation reactions utilizing malonyl-CoA as extenders together with a broad range of fatty acyl-CoA starter substrates to yield alkylresorcinols. In the wild type yeast, the heterologous expression of BdARS and ScARS led to the formation of alkylresorcinols with alkyl chains ranging from C_9 to C_{17}, suggesting that the enzymes accepted starter substrates of fatty acyl-CoAs varying from C_{10} to C_{18}. It seems plausible that the alkylresorcinol products 15:1 and 17:1 originated from 16:1 and 18:1 fatty acyl substrates, since both 16:1 and 18:1 substrates are abundant in yeast (Oh *et al.*, 1997).

In order to test whether BdARS and ScARS could also accept VLC fatty acyl-CoAs as substrates in addition to those with medium and long chains, the yeast mutant *elo3*Δ was used. This mutant is deficient in elongation to fatty acyl constituents of C_{24} and beyond, and therefore accumulates C_{20} and C_{22} intermediates. Overall, the mutant *elo3*Δ showed a 20% increase of VLC substrates compared to those in the wild type yeast (Oh *et al.*, 1997). The enhanced level of VLC fatty acyl substrates resulted in the formation of additional alkylresorcinols with VLCs from C_{19} to C_{21} in the transgenic lines expressing BdARS and ScARS in the mutant background. Thus, it can be concluded that both of the ARS enzymes can also use VLC fatty acyl-CoA substrates C_{20} and C_{22} to produce VLC alkylresorcinols C_{19} and C_{21}. It should be noted that, in the yeast lines expressing BdARS and ScARS, trace amounts of alkylresorcinols with even-numbered side chains were found, suggesting that the enzymes also accept odd-numbered acyl-CoA starters, and that small but not negligible amounts of these unusual fatty acid derivatives were present in yeast.

It is to be noted that the possibility of derailment product alkylpyrones was checked in both yeast strains expressing either BdARS or ScARS. The MS fragmentation information of alkylpyrones was refered to previous studies (Cook *et al.*, 2010; Kim *et al.*, 2010) and it revealed that these derailment products were not present in yeast lipids expressing BdARS and ScARS.

The alkylresorcinol product mixtures present in yeast lipids differed between the two ARS enzymes and between the two yeast strains used in this study (Figure 5.4). For

BdARS, alkylresorcinols 15:0 and 17:0 were the major products in the lipids extracted from both yeast strains, suggesting that C_{16} and C_{18} fatty acyl-CoAs were either the preferred substrates of this enzyme and/or the starter molecules most easily accessible to it. For ScARS, the dominant alkylresorcinols formed were 15:0 and 13:0 in the wild type yeast, likely formed from C_{16} and C_{14} fatty acyl substrates. This finding is in contrast to the typical acyl chain length profile of yeast lipids, and thus suggests that ScARS has substrate preferences towards starter substrates with shorter acyl chains compared to BdARS. When ScARS was expressed in the yeast mutant *elo3*Δ, a large accumulation of alkylresorcinol 19:0 was observed for ScARS, suggesting a significant preference of ScARS for C_{20} substrate. However, in *B. distachyon* and *S. cereale* leaves, the profiles of cuticular alkylresorcinols with the alkyl chains ranged from C_{17} to C_{25} and C_{19} to C_{27}, respectively, and thus included products with much longer chains than found even in the mutant. The *B. distachyon* cuticular alkylresorcinols were dominated by homologs 19:0 and 21:0, suggesting that the ARS forming them should have substrate preferences for C_{20} and C_{22} fatty acyl-CoAs. The *S. cereale* cuticular alkylresorcinols were dominated by homologs 23:0, 21:0 and 25:0, pointing to substrate preferences for even longer acyl-CoAs of C_{24}, C_{22} and C_{26} in this species. The apparent discrepancy between the wax data and the yeast characterization of the two ARS enzymes can be explained in two ways: a) BdARS and ScARS may not be involved in cuticular alkylresorcinol biosynthesis, but may be responsible for the formation of alkylresorcinols with shorter fatty acyl chains in other tissues, where they have so far eluded detection. b) Both enzymes may still have preferences for longer acyl-CoA substrates and be involved in cuticular alkylresorcinol biosynthesis, but in yeast they did not have access to sufficient quantities of these starter compounds. This might be due to differential localization of enzymes and substrates, to metabolic channeling of VLC fatty acyl-CoAs through yeast enzymes, or simply to the ratio of long-chain to VLC fatty acyl-CoAs that is much higher in the yeast environment than encountered by the same enzymes *in planta*.

On the other hand, both BdARS and ScARS showed activities of using fatty acyl substrates 16:1 and 18:1 to produce alkylresorcinols 15:1 and 17:1, whereas these alkylresorcinol homologs were not naturally present in the leaf cuticular waxes of

either species. It could be due to that the long chain alkylresorcinols cannot be exported outside the cells to the cuticle. However, according to a previous study by Ji and Jetter (2008), these homologs were absent in interior tissues either, indicating that long chain alkylresorcinols are very likely not present in grass leaves. If this is the case, the ARS enzymes must have no access to the long chain fatty acyl-CoA substrates which are then possibly used by other enzymes during other biosynthetic pathways, such as glycerolipid and phospholipid biosynthesis.

Further studies are needed in order to test these explanations and conclude on the fatty acyl-CoA chain length preferences of BdARS and ScARS. Moreover, an analysis of the supernatant after centrifugation of recombinant yeast cells could be done to test whether alkylresorcinol products, in addition to their presence within the cells ,can be exported to the outside of the cells. Overall, the results of the functional characterization above provided evidence showing that BdARS and ScARS are capable of accepting a variety of fatty acyl-CoAs to produce medium- to long-chain alkylresorcinols, and also C_{20} and C_{22} fatty acyl-CoAs to produce VLC alkylresorcinols.

5.3.2 Gene expression analyses

Investigation of gene expression patterns by semi-quantitative RT-PCR showed that both *ScARS* and *BdARS* were exclusively expressed in the aboveground organs that were examined, but not in underground organs (roots). Moreover, as investigated in *B. distachyon* in detail, *BdARS* was only expressed in young leaves representing an early vegetative stage but not in stems or spikes at a relatively late stage during plant development (Figure 5.5A). In *S. cereale* the expression level of *ScARS* was compared with *ScCHS*, a gene encoding the ubiquitous CHS enzyme responsible for flavonoid biosynthesis. As is shown in Figure 5.5B, *ScARS* expression was restricted to aerial organs, in contrast to the ubiquitous expression as seen for *ScCHS*. In addition, the expression levels were also tested between green leaves and etiolated leaves. Both *ScARS* and *ScCHS* showed significantly higher expression levels in green leaves, suggesting a correlation of gene activity with light. All these findings taken together showed that the two ARS genes were mostly expressed in organs that are covered by

a cuticle, thus supporting the idea that *BdARS* and *ScARS* are the potential genes involved in the biosynthesis of cuticular alkylresorcinols.

Quantitative RT-PCR analysis was further carried out in order to assess the expression levels of *ScARS* as a function of leaf development. It is to be noted that the leaf of *S. cereale* had been found to expand mainly in length with only small variation in width after emerging from the sheath of the older leaf. Previous studies revealed a steady growth rate of around 1.8 cm per day until the leaf reached a length of 20 cm (Ji and Jetter, unpublished data). The leaf grows exclusively in the basal region near the point of emergence, and consecutive sections along the blade thus have increasing ages. *ScARS* was preferentially expressed in the basal region of the leaf where relatively high expression levels were detected in 14-16 cm and 12-14 cm leaf segments (Figure 5.6A). The overall size of the leaf and the division into leaf segments used in this study were equal to those used previously in an investigation into the accumulation of waxes along the leaf blade, and the current expression data can thus be directly compared with the results describing the accumulation of cuticular alkylresorcinols (presented as $\mu g/cm^2$ in Figure 1.3B). Cuticular alkylresorcinols had been found to accumulate in a restricted time period at a relatively late stage during leaf development. Prior to growth stage III, no cuticular alkylresorcinols were detected, while at growth stage III the accumulation of cuticular alkylresorcinols started at 2 cm beyond the point of emergence, reached the highest coverage of 0.1 $\mu g/cm^2$ at 4 cm and remained at this level to the leaf tip (leaf segments 0-11 cm). The highest accumulation of alkylresorcinols was found at growth stage IV, with levels reaching 0.2-0.3 $\mu g/cm^2$ in leaf segments 10-18 cm (Figure 1.3B). Therefore, the comparison between growth stages III and IV clearly showed an increase in the abundance of cuticular alkylresorcinols within the zone from the point of emergence to 12 cm beyond, especially in the leaf segments of 12-14 cm and 14-16 cm. The region around 4 cm above the point of emergence showed the highest accumulation rate of cuticular alkylresorcinols, and this is also the region where the highest gene expression levels of *ScARS* were detected. Moreover, at growth stage IV the distal 10 cm near the tip of the leaf were covered with a constant low level of cuticular alkylresorcinols at around 0.1 $\mu g/cm^2$, and in this region *ScARS* expression was hardly

detected. Thus, the expression profile of *ScARS* was in good agreement with the profile of cuticular alkylresorcinol accumulation along the leaf overall, further confirming the notion that *ScARS* is involved in the formation of cuticular alkylresorcinols in *S. cereale* leaves.

As a gene encoding a specialized type III PKS, the expression profile of *ScARS* was also compared to that of *ScCHS*. Distinct from *ScARS*, the uniform expression levels of *ScCHS* along the leaf showed that it was ubiquitously expressed during leaf development (Figure 5.6B). It contrasted against the specialized expression pattern of *ScARS*, putting the correlation of the *ScARS* expression with alkylresorcinol accumulation further into perspective.

Assuming that ScARS is indeed involved in the formation of cuticular alkylresorcinols, the dynamics of the processes leading from the accumulation of mRNA transcripts of *ScARS* to the deposition of cuticular alkylresorcinols can be addressed based on the current data. Even though many of the mechanistic details involved remain to be elucidated, it is clear that the formation of cuticular alkylresorcinols involves at least the expression of biosynthetic protein(s), substrate accumulation, alkylresorcinol product formation, transport to the plasma membrane, export from the membrane through the cell wall towards the cuticle, and arrangement into the fine structure of the cuticle. All the currently available data for *S. cereale* taken together showed that the most rapid accumulation of cuticular alkylresorcinols occurred in the 3-day interval from growth stage III (13 days) to growth stage IV (16 days) in the basal zone of the leaf, while *ScARS* was most strongly expressed in the same area in the same time interval. It thus can be inferred that the overall process of wax production and export must occur on a time scale substantially shorter than three days, likely with rates of only a few hours.

5.3.3 Subcellular localization of BdARS and ScARS

The subcellular localization of BdARS and ScARS was investigated by transient expression in *N. benthamiana* leaves via agrobacterium-mediated infiltration. GFP

fluorescence was observed using light microscopy and confocal microscopy. It showed that both BdARS and ScARS exhibited patterns of a reticulate ER-like network, suggesting that they reside at the ER membrane (Figure 5.9). Moreover, the movement of the reticulate structure by observation over time also supported the idea that ScARS is associated with the ER membrane (Figure 5.9B). To confirm the subcellular localization, hexyl rhodamine B, a specific staining that labels the ER in plants, was introduced to check the signal to be co-localized with GFP (Boevink *et al.*, 1996). However, the co-localization was unsuccessful due to the failure of detecting the signal from hexyl rhodamine B at the ER. On one hand, this might have been due to the cuticle and trichomes on *N. benthamiana* leaves impeding hexyl rhodamine B from entering into the cells. Instead, most of the dye stayed at the cuticle outside the cells. On the other hand, in rare cases where the dye entered the cells, the staining patterns were inconclusive, without a clear staining of the ER, making it impossible to confirm the ER localization of the ARSs.

Overall, the results from localization studies of 35S:BdARS-sGFP and 35S:ScARS-sGFP revealed that both enzymes are likely localized to the ER, which is the site where the VLC fatty acyl-CoA substrates are present. Thus, the current experiments provided further evidence supporting the conclusion that BdARS and ScARS are the ARSs responsible for the biosynthesis of cuticular alkylresorcinols.

Chapter 6 Conclusion and future directions

Alkylresorcinols are phenolic lipids that are derived from polyketides. As natural products, alkylresorcinols occur in diverse plant species as well as fungi and bacteria. Interestingly, these compounds are relatively abundant in the grass family, in particular in cereal species such as *Secale cereale* (rye). Due to their antifungal and antibacterial activities, alkylresorcinols are of increasing interest. They usually accumulate at/near the plant surfaces and are present as series of homologs differing in carbon numbers of their alkyl side chains. Earlier studies in our lab had established that the localization of alkylresorcinols in *S. cereale* leaves was largely restricted to the cuticle, a protective waxy layer covering all primary aerial organs. It was shown that the cuticular alkylresorcinols accounted for 3% of the total wax coverage, and that the alkylresorcinols had very-long-chain (VLC) alkyls ranging from C_{19} to C_{27}. Additionally, the deposition of alkylresorcinols was found to be synchronized with the production of other wax compounds for a restricted time period relatively late during leaf ontogenesis. Based on all these findings, it was speculated that the alkylresorcinols are targeted to the cuticle to serve as a first line of defense in *S. cereale* leaves. However, the biological function of the cuticular alkylresorcinols could not be judged based on chemical data alone, but information on the mechanisms underlying the formation of these compounds was also needed. Therefore, the current work was aiming at a better understanding of the biosynthesis of cuticular alkylresorcinols.

For the investigations into the biosynthesis of cuticular alkylresorcinols, I selected *S. cereale*, since all the previous chemical studies had been performed on this species, as well as *Brachypodium distachyon*, a closely related genetic model. Chemical analysis of cuticular wax of *B. distachyon* leaves, gene cloning from the two selected species, and biochemical and biological characterization of the encoded enzymes were conducted to achieve the goal of this work. All data from the above experiments taken together provided evidence indicating that the alkylresorcinol synthases (ARSs) cloned from *B. distachyon* and *S. cereale* are the enzymes can account for the biosynthesis of cuticular alkylresorcinols. The results further suggested that cuticular

alkylresorcinols are biosynthesized for a function within the cuticle of *B. distachyon* and *S. cereale* leaves. Undoubtedly, the current ARSs can serve as fundamental tools in more detailed future studies into the biological function of these interesting cuticular compounds.

In the following sections, the major conclusions from the studies in each of the previous chapters will be summarized, and linked to further experiments that can now be carried out using the data and tools established here.

6.1 Identification of alkylresorcinols and other cuticular wax compounds in *B. distachyon* leaves

The genome of *B. distachyon*, belonging to the grass family (Poaceae), has recently been completely sequenced (2010) and, therefore, this species has become an attractive model system for studies of grass biology. It can serve as a powerful tool that facilitates research in related cereal species, such as gene identification and biochemical mechanism investigation. Investigations into the composition or formation of cuticular wax of *B. distachyon* have not been reported to date. Thus, the wax of *B. distachyon* leaves had to be analyzed at the beginning of the current work, in order to compare it with that of *S. cereale* as well as other literature data. The wax analysis was also undertaken in order to search for cuticular alkylresorcinols in *B. distachyon* leaves, the presence of which would be a prerequisite for further studies in cloning and characterization of the corresponding ARSs in parallel with *S. cereale*.

The chemical analyses carried out in this work revealed that the cuticular wax had a coverage of 12.8 ± 0.8 $\mu g/cm^2$ on *B. distachyon* leaves, and that it was mainly composed of primary alcohols (9.0 ± 0.4 $\mu g/cm^2$), alkyl esters (1.4 ± 0.2 $\mu g/cm^2$), aldehydes (0.2 ± 0.1 $\mu g/cm^2$) and alkanes (0.3 ± 0.1 $\mu g/cm^2$). Additionally, a series of homologous compounds that was not typically present in wax mixtures was noticed and identified as alkylresorcinols (0.6 ± 0.1 $\mu g/cm^2$) with alkyl side chains varying from C_{17} to C_{25}. The alkylresorcinols 19:0 (45%) and 21:0 (40%) were the prevalent homologs. Meanwhile, two novel compounds in the total wax (0.04 ± 0.01 $\mu g/cm^2$)

were tentatively identified as methylated alkylresorcinols with the side chain lengths of C_{19} (71%) and C_{21} (29%). However, the position of the methyl group was not determined at this stage. Three potential isomers may exist, and thin layer chromatography (TLC) analysis might help to further elucidate the true structures. Alternatively, feeding experiments with labeled precursor candidates, including methylmalonyl-CoA extenders and methyl-branched fatty acyl-CoA substrates, might help to resolve the structures. It would also reveal whether the plant ARSs accept the different extenders in a strict controlled sequence to generate methylated alkylresorcinols as observed for the bacterial SrsA (Funabashi *et al.*, 2008). Conversely, both the exact isomer structures and results from such feeding studies would add substantial information on the substrate specificities of the ARS enzymes from both *B. distachyon* and *S. cereale*. Besides, only less than one tenth of the wax (1.2 ± 0.3 $\mu g/cm^2$) remained unidentified (excluding trace amounts of the wax compound classes of sterols, triterpenoids and alkyl benzoates). TLC analysis might also help to identify the remaining unknowns in the cuticular wax mixtures of *B. distachyon* leaves.

Interestingly, only one homolog with chain length C_{26} was identified in the aldehyde compound class, whereas alkanes contained three homologs with chain lengths of C_{27}, C_{29} and C_{31}. According to the most widely accepted models, alkanes are formed from aldehydes via decarbonylation on one of the pathways of wax biosynthesis. Although much research effort has gone into the elucidation of wax alkane biosynthesis, the mechanism of plant alkane formation remains largely unknown. In the best-studied model plant *Arabidopsis thaliana*, several *eceriferum* (*CER*) genes have been identified that affect the alkane biosynthesis in stem wax, including *CER1* and *CER3*. However, the biochemical activities of the corresponding proteins remain unclear. Based on the simple composition of the aldehyde fraction in *B. distachyon* total leaf wax, a survey of gene candidates from the Brachypodium Database that are orthologs of the Arabidopsis *CER1* and *CER3* would definitely facilitate the identification and characterization of enzymes in the elusive alkane pathway.

6.2 Cloning and characterization of ARSs from *B. distachyon* and *S. cereale*

Relying on the presence of cuticular alkylresorcinols in *B. distachyon* leaves, cloning of one or more genes potentially encoding ARS(s) from this model system was attempted by mining the Brachypodium expressed sequence tag (EST) libraries with ARS sequence information from *Sorghum bicolor* and *Oryza sativa* as queries (Cook *et al.*, 2010). The mining resulted in nine candidates belonging to the chalcone synthase/chalcone synthase-like (CHS/CHSL) type III polyketide synthases (PKSs). Among them, one was defined as the potential ARS based on the close relationship with the functionally characterized orthologs in *S. bicolor* and *O. sativa* in the phylogenetic analysis. Then, using sequence information on the putative BdARS as well as the ARSs in *S. bicolor* and *O. sativa*, a homology-based cloning strategy was used for cloning potential ARS gene(s) from *Secale cereale*. Overall, this work revealed one ARS from *B. distachyon* and one from *S. cereale*. Designated as BdARS and ScARS, respectively, the two ARSs share high amino acid identity of 71% to each other, identities of 60-68% with the ARSs in *Sorghum bicolor* and *O. sativa*, and 47-50% with the CHSs in *Secale cereale*. Phylogenetic relationships between ARSs and other related CHS/CHSL suggested that the ARSs might represent an ancient and original enzyme subfamily within the type III PKSs. In addition to a conserved Cys-His-Asn catalytic triad characteristic of all type III PKSs, the alignment of amino acid sequences also revealed some other residues that are potentially determining a) the aldol condensation mechanism that is characteristic of ARSs as opposed to CHSs, b) the substrate selectivity, and c) the chain length specificity. Site-directed mutagenesis experiments should be attempted in the future to confirm the amino acids determining the ARS activity.

Biochemical characterization of the putative BdARS and ScARS by *in vivo* heterologous expression in different yeast strains confirmed the enzymatic function of the proteins as ARSs. Combining the profiles from two yeast strains, it can be concluded that both BdARS and ScARS have the ability to produce alkylresorcinols using a variety of C_{10} to C_{22} fatty acyl-CoAs as starter substrates with malonyl-CoAs as extenders. Thus, one of the most important findings of the current work is that there are indeed ARS enzymes

in both species that can account for the formation of the cuticular very-long-chain (VLC) alkylresorcinols.

However, the exact preferences of both enzymes for certain chain lengths of fatty acyl-CoA substrates can only be speculated based on the current data. It is well established that the most abundant substrates with fatty acyl chains in wild type yeast are 16:1, 18:1, 16:0 and 18:0, together comprising over 95% of the total fatty acyl constituents (Oh *et al.*, 1997). The finding that alkylresorcinols 15:0 and 17:0 were formed predominantly when expressing the ARS enzymes is probably due to the large abundance of the corresponding starter substrates 16:0 and 18:0 in yeast. Moreover, the formation of alkylresorcinols 15:1 and 17:1 in transgenic yeast clearly reflects the relatively high concentration of starter substrates of 16:1 and 18:1. Thus, the alkylresorcinol product profiles are at least in part due to substrate pool composition rather than enzyme specificities.

On the other hand, some of the subtle differences in the alkylresorcinol product profiles between yeast lines expressing either BdARS or ScARS, and between wild type yeast and yeast mutant *elo3*Δ backgrounds indicate also some substrate specificity. Overall, the results suggest that ScARS has substrate preferences towards shorter starter chain lengths compared to BdARS, and that ScARS also has significant affinity for C_{20} substrate. However, these findings leave the question open whether the enzymes actually have a preference for VLC fatty acyl-CoA substrates over medium- and long-chain starters. Although the yeast mutant *elo3*Δ is known to have an up to 10-fold increase in the levels of C_{20} and C_{22} VLC fatty acids (VLCFAs), those VLC fatty acyl constituents are still minor, with a proportion of 2% in the fatty acyl substrate pool (Oh *et al.*, 1997). Therefore, the current experiments with this yeast mutant cannot test the preference between relevant chain lengths in a truly competitive situation. Instead, the *in vivo* characterization in any of the available yeast systems is largely affected by the availability of fatty acyl-CoA constituents in the CoA pools in different strains. In order to overcome the limitation of yeast systems in defining the substrate specificity of the ARSs and, in particularly, in testing potential preferences for VLC substrates, feeding experiments with exogenous

VLCFAs (*e.g.* C_{20} to C_{26}) must be conducted. Ideally, *in vitro* assays with purified ARS enzymes should be performed to assess substrate selectivity quantitatively. Moreover, since methylated alkylresorcinols were not detected in the yeast lipids for BdARS as they were found in the total leaf wax of *B. distachyon*, the feeding experiment with labeled extenders/substrates as described earlier in Chapter 6.1 would also test whether BdARS is the enzyme responsible for the production of methlated alkylresorcinols in *B. distachyon* leaves. Concerning the *in vitro* assays, other factors such as protein stability, substrate solubility or the availability of enzyme cofactors need to be considered.

The investigation of gene expression patterns of *BdARS* and *ScARS* showed that both genes are exclusively expressed in the aerial organs but not in the roots. *BdARS* was expressed only in young leaves, rather than in late stages during plant development, *e.g.* in stems or spikes. It should be tested whether those gene expression patterns are paralleled by similar patterns in the accumulation of cuticular alkylresorcinols in various organs. If cuticular alkylresorcinols were detected in stems or spikes, this might suggest that one or more other *ARS*s apart from *BdARS* are present in *B. distachyon*, and that those homologs are specifically responsible for the biosynthesis of cuticular alkylresorcinols in those organs. On the other hand, although it is known that alkylresorcinols accumulate to high levels in *S. cereale* grains, expression of *ScARS* was not tested in spikes due to limited plant growth time. It would be interesting to examine the *ScARS* expression levels in spikes and stems. Since there are at least two *ARS*s in *Sorghum bicolor* and three in *O. sativa* (Cook *et al.*, 2010), it is well possible that more than one *ARS* exists in *Secale cereale*, since this species has a relatively large genome. If the current *ScARS* is leaf-specific, then there must be other candidate *ARS*s involved in alkylresorcinol biosynthesis particularly in other tissues/organs. Thus, cloning of other *ARS*s would become interesting and important for comparative studies with *ScARS* in *S. cereale*.

Furthermore, the profile of *ScARS* expression pattern was distinct from that of *ScCHS*, and was in agreement with the profile of cuticular alkylresorcinol accumulation along the leaf of *S. cereale*. This finding is in favor of the hypothesis that ScARS, different

from ScCHS, may be the enzyme responsible for biosynthesis of cuticular alkylresorcinols. To further underline this idea, it would be interesting to compare the *ScARS* expression profile with that of wax-associated genes. Such a comparison should focus, in particular, on the basal region of *S. cereale* leaves, in order to put the *ScARS* expression levels into perspective with wax production. In this case, ortholog(s) of *CER6*, *i.e.* a gene encoding a condensing enzyme exclusively expressed in epidermal cells during wax production in *A. thaliana*, should be identified from *S. cereale* and employed as a reference.

At last, the subcellular localization of BdARS and ScARS was examined using transient expression in *Nicotiana benthamiana* leaves. Both of the ARS enzymes were found to be associated with an endoplasmic reticulum (ER)-like network, supporting the idea that they are present at the ER membrane where the VLC fatty acyl-CoA substrates are thought to be localized. However, the green fluorescent protein (GFP) signal could not be co-localized with the signal from hexyl rhodamine B to confirm the ER localization. In order to confirm the ER localization of BdARS and ScARS, co-localization with an ER marker is still necessary. Instead of using chemical staining that may cause problems, ER protein markers such as HDEL and mCherry would be more reliable to use. Another protein for which the ER membrane association has been shown could be used as a positive control as well.

In summary, the following conclusions can be drawn from the current research. First, BdARS and ScARS, differing from CHS/STS type III PKS enzymes, have activities producing alkylresorcinols via decarboxylative condensation reactions of a broad range of fatty acyl-CoA starter substrates, including medium-, long- and some VLC substrates, to malonyl-CoA extenders. Second, the exclusive expression in aboveground organs and in the basal region of leaves at early stages during grass development, in particular, shows that the two ARSs are dedicated to the production of leaf alkylresorcinols, and that the production is synchronized with other compounds during wax biosynthesis. Third, their likely ER localization demonstrates that the ARSs reside at the site where the physiological VLC fatty acyl-CoA substrates are, supporting the idea that the ARSs are meant to synthesize VLC alkylresorcinols

for the cuticle. Taking all the above results together, it seems very plausible that BdARS and ScARS are the enzymes responsible for the biosynthesis of cuticular VLC alkylresorcinols in grass leaves, and that this is the sole biological function of the ARSs.

Wax biosynthesis occurs only in epidermal cells, and many wax biosynthetic genes are known to be expressed only in the (expanding) epidermis (Lessire *et al.*, 1982; Suh *et al.*, 2005). Therefore, if the *ARS*s were found to be expressed exclusively in the epidermis, then this would further argue in favor of the current overall conclusion that these ARSs are indeed responsible for synthesizing cuticular alkylresorcinols in leaves of *B. distachyon* and *S. cereale*. *In situ* hybridization analysis could be used to test tissue-specificity of *BdARS* and *ScARS* expression. Histochemical analysis using β-glucuronidase (GUS) staining would also be an option for *BdARS*, taking advantage of sequence information on the native promoter region of *BdARS* in the Brachypodium Database using promoter-GUS fusion and transformation in Brachypodium (Alves *et al.*, 2009) for expression studies. Alternatively, transformation of Arabidopsis might be attempted for GUS analysis.

BdARS and ScARS can serve as useful tools to identify ARSs in addition to the previously characterized orthologs in plant species. More importantly, these are the first potential ARSs responsible for the biosynthesis of cuticular alkylresorcinols. The current work permits further detailed investigations into the biological function(s) of alkylresorcinols in grass leaf cuticles. For example, cuticular alkylresorcinols could be synthesized or extracted from grass waxes, and applied in bioassays in order to study whether and how they affect pathogens and herbivores. Alkylresorcionols of different chain lengths could be compared in such assays, for example including those produced by heterologous expression in yeast. Moreover, ARS genes could be expressed in plants such as Arabidopsis, to create lines with varying levels of alkylresorcinols that could be used for comparative assessments of pathogen and/or herbivore performance. With a better understanding of the defensive mechanism that cuticular alkylresorcinols may perform, manipulation of the biosynthetic reactions generating these intriguing compounds could be accomplished for

biotechnological applications in the future.

References

(2010) Genome sequencing and analysis of the model grass *Brachypodium distachyon*. Nature, **463**, 763-768.

Alonso E, Ramn D J and Yus M (1997) Simple synthesis of 5-substituted resorcinols: a revisited family of interesting bioactive molecules. J Org Chem, **62**, 417-421.

Alves S C, Worland B, Thole V, Snape J W, Bevan M W and Vain P (2009) A protocol for Agrobacterium-mediated transformation of *Brachypodium distachyon* community standard line Bd21. Nat Protoc, **4**, 638-649.

Anderson H H, David N A and Leake C D (1931) Oral toxicity of certain alkylresorcinols in guinea pigs and rabbits. Proc Soc Exp Biol Med, **28**, 609-612.

Arisawa M, Ohmura K, Kobayashi A and Morita N (1989) A cytotoxic constituent of *Lysimachia japonica* Thunb. (Primulaceae) and the structure-activity relationships of related compounds. Chem Pharm Bull, **37**, 2431-2434.

Austin M B and Noel J P (2003) The chalcone synthase superfamily of type III polyketide synthases. Nat Prod Rep, **20**, 79-110.

Austin M B, Bowman M E, Ferrer J L, Schröder J and Noel J P (2004) An aldol switch discovered in stilbene synthases mediates cyclization specificity of type III polyketide synthases. Chem Biol, **11**, 1179-1194.

Baerson S R, Rimando A M and Pan Z (2008) Probing allelochemical biosynthesis in sorghum root hairs. Plant Signal Behav, **3**, 667-670.

Baerson S R, Schröder J, Cook D, Rimando A M, Pan Z, Dayan F E, Noonan B P and Duke S O (2010) Alkylresorcinol biosynthesis in plants: New insights from an ancient enzyme family? Plant Signal Behav, **5**.

Bandyopadhyay C, Gholap A S and Mamdapur V R (1985) Characterization of alkenylresorcinol in mango (*Mangifera indica* L.) latex. J Agric Food Chem, **33**, 377-379.

Bianchi G, Avato P, Bertorelli P and Mariani G (1978) Epicuticular waxes of two sorghum varieties. Phytochemistry, **17**, 999-1001.

Bianchi G, Avato P, Scarpa O, Murelli C, Audisio G and Rossini A (1989) Composition and structure of maize epicuticular wax esters. Phytochemistry, **28**, 165-171.

Boevink P, Santa Cruz S, Hawes C, Harris N and Oparka K J (1996) Virus-mediated delivery of the green fluorescent protein to the endoplasmic reticulum of plant cells. Plant J, **10**, 935-941.

Chaturvedula V S, Schilling J K, Miller J S, Andriantsiferana R, Rasamison V E and Kingston D G (2002) New cytotoxic bis 5-alkylresorcinol derivatives from the leaves of *Oncostemon bojerianum* from the Madagascar rainforest. J Nat Prod, **65**, 1627-1632.

Chen Y, Ross A B, Åman P and Kamal-Eldin A (2004) Alkylresorcinols as markers of whole grain wheat and rye in cereal products. J Agric Food Chem, **52**, 8242-8246.

Contessotto M G G, Monteiro-Vitorello C B, Mariani P D S C and Coutinho L L (2001) A new member of the chalcone synthase (CHS) family in sugarcane. Genet Mol Biol, **24**, 257-261.

Cook D, Rimando A M, Clemente T E, Schröder J, Dayan F E, Nanayakkara N P, Pan Z, Noonan B P, Fishbein M, Abe I, Duke S O and Baerson S R (2010) Alkylresorcinol synthases expressed in *Sorghum bicolor* root hairs play an essential role in the biosynthesis of the allelopathic benzoquinone sorgoleone. Plant Cell, **22**, 867-887.

Costaglioli P, Joubes J, Garcia C, Stef M, Arveiler B, Lessire R and Garbay B (2005) Profiling candidate genes involved in wax biosynthesis in *Arabidopsis thaliana* by microarray analysis. Biochim Biophys Acta, **1734**, 247-258.

Cserzo M, Wallin E, Simon I, von Heijne G and Elofsson A (1997) Prediction of transmembrane α-helices in prokaryotic membrane proteins: the dense alignment surface method. Protein Eng, **10**, 673-676.

Cushnie T P and Lamb A J (2005) Antimicrobial activity of flavonoids. Int J Antimicrob Agents, **26**, 343-356.

Deng J Z, Starck S R and Hecht S M (1999) Bis-5-alkylresorcinols from *Panopsis rubescens* that inhibit DNA polymerase β. J Nat Prod, **62**, 477-480.

Deszcz L and Kozubek A (2000) Higher cardol homologs (5-alkylresorcinols) in rye seedlings. Biochim Biophys Acta, **1483**, 241-250.

Draper J, Mur L A, Jenkins G, Ghosh-Biswas G C, Bablak P, Hasterok R and Routledge A P (2001) *Brachypodium distachyon*. A new model system for functional genomics in grasses. Plant Physiol, **127**, 1539-1555.

Droby S, Prusky D, Jacoby B and Goldman A (1986) Presence of antifungal compounds in the peel of mango fruits and their relation to latent infections of *Alternaria alternata*. Physiol Plant Pathol, **29**, 173-183.

Ferrer J L, Jez J M, Bowman M E, Dixon R A and Noel J P (1999) Structure of chalcone synthase and the molecular basis of plant polyketide biosynthesis. Nat Struct Biol, **6**, 775-784.

Franken P, Niesbach-Klosgen U, Weydemann U, Marechal-Drouard L, Saedler H and

Wienand U (1991) The duplicated chalcone synthase genes *C2* and *Whp* (white pollen) of *Zea mays* are independently regulated; evidence for translational control of *Whp* expression by the anthocyanin intensifying gene *in*. EMBO J, **10**, 2605-2612.

Funa N, Ozawa H, Hirata A and Horinouchi S (2006) Phenolic lipid synthesis by type III polyketide synthases is essential for cyst formation in *Azotobacter vinelandii*. Proc Natl Acad Sci USA, **103**, 6356-6361.

Funa N, Awakawa T and Horinouchi S (2007) Pentaketide resorcylic acid synthesis by type III polyketide synthase from *Neurospora crassa*. J Biol Chem, **282**, 14476-14481.

Funabashi M, Funa N and Horinouchi S (2008) Phenolic lipids synthesized by type III polyketide synthase confer penicillin resistance on *Streptomyces griseus*. J Biol Chem, **283**, 13983-13991.

García S, García C, Heinzen H and Moyna P (1997) Chemical basis of the resistance of barley seeds to pathogenic fungi. Phytochemistry, **44**, 415-418.

Gietz R D and Woods R A (2002) Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. Methods Enzymol, **350**, 87-96.

Gorham J (1995) The biochemistry of the stilbenoids. London: Chapman & Hall.

Goyal A, Saxena P, Rahman A, Singh P K, Kasbekar D P, Gokhale R S and Sankaranarayanan R (2008) Structural insights into biosynthesis of resorcinolic lipids by a type III polyketide synthase in *Neurospora crassa*. J Struct Biol, **162**, 411-421.

Grienenberger E, Kim S S, Lallemand B, Geoffroy P, Heintz D, Souza C A, Heitz T, Douglas C J and Legrand M (2010) Analysis of *TETRAKETIDE α-PYRONE REDUCTASE* function in *Arabidopsis thaliana* reveals a previously unknown, but conserved, biochemical pathway in sporopollenin monomer biosynthesis. Plant Cell, **22**, 4067-4083.

Guhling O, Hobl B, Yeats T and Jetter R (2006) Cloning and characterization of a lupeol synthase involved in the synthesis of epicuticular wax crystals on stem and hypocotyl surfaces of *Ricinus communis*. Arch Biochem Biophys, **448**, 60-72.

Heredia A (2003) Biophysical and biochemical characteristics of cutin, a plant barrier biopolymer. Biochim Biophys Acta, **1620**, 1-7.

Huang J X, Qu L J, Yang J, Yin H and Gu H Y (2004) A preliminary study on the origin and evolution of chlacone synthase (*CHS*) gene in angiosperms. Acta Bot Sin, **46**, 10-19.

Itokawa H, Totsuka N, Nakahara K, Maezuru M, Takeya K, Kondo M, Inamatsu M and Morita H (1989) A quantitative structure-activity relationship for antitumor activity of long-chain phenols from *Ginkgo biloba* L. Chem Pharm Bull, **37**, 1619-1621.

Jetter R, Schäffer S and Riederer M (2000) Leaf cuticular waxes are arranged in chemically and mechanically distinct layers : evidence from Prunus laurocerasus L. Plant, Cell Environ., **23**, 619-628.

Jetter R, Kunst L and Samuels L (2006) Composition of plant cuticular waxes. In Biology of the Plant Cuticle. (Riederer M and Müller C, eds). Oxford: Blackwell Publishing, pp. 145-181.

Jez J M, Austin M B, Ferrer J, Bowman M E, Schröder J and Noel J P (2000) Structural control of polyketide formation in plant-specific polyketide synthases. Chem Biol, **7**, 919-930.

Ji X and Jetter R (2008) Very long chain alkylresorcinols accumulate in the intracuticular wax of rye (*Secale cereale* L.) leaves near the tissue surface. Phytochemistry, **69**, 1197-1207.

Jiang C, Kim S Y and Suh D Y (2008) Divergent evolution of the thiolase superfamily and chalcone synthase family. Mol Phylogenet Evol, **49**, 691-701.

Jin W and Zjawiony J K (2006) 5-alkylresorcinols from *Merulius incarnatus*. J Nat Prod, **69**, 704-706.

Kellogg E A (2001) Evolutionary history of the grasses. Plant Physiol, **125**, 1198-1205.

Kim S S, Grienenberger E, Lallemand B, Colpitts C C, Kim S Y, Souza C A, Geoffroy P, Heintz D, Krahn D, Kaiser M, Kombrink E, Heitz T, Suh D Y, Legrand M and Douglas C J (2010) *LAP6/POLYKETIDE SYNTHASE A* and *LAP5/POLYKETIDE SYNTHASE B* encode hydroxyalkyl α-pyrone synthases required for pollen development and sporopollenin biosynthesis in *Arabidopsis thaliana*. Plant Cell, **22**, 4045-4066.

Knödler M, Berardini N, Kammerer D R, Carle R and Schieber A (2007) Characterization of major and minor alk(en)ylresorcinols from mango (*Mangifera indica* L.) peels by high-performance liquid chromatography/atmospheric pressure chemical ionization mass spectrometry. Rapid Commun Mass Spectrom, **21**, 945-951.

Koch K, Barthlott W, Koch S, Hommes A, Wandelt K, Mamdouh W, De-Feyter S and Broekmann P (2006) Structural analysis of wheat wax (*Triticum aestivum*, *c.v.* 'Naturastar' L.): from the molecular level to three dimensional crystals. Planta, **223**, 258-270.

Kozubek A and Tyman J H P (1995) Cereal grain resorcinolic lipids: mono and dienoic homologues are present in rye grains. Chem Phys Lipids, **78**, 29-35.

Kozubek A, Pietr S and Czerwonka A (1996) Alkylresorcinols are abundant lipid components in different strains of *Azotobacter chroococcum* and *Pseudomonas* spp. J Bacteriol, **178**, 4027-4030.

Kozubek A and Tyman J H P (1999) Resorcinolic lipids, the natural non-isoprenoid phenolic amphiphiles and their biological activity. Chem Rev, **99**, 1-26.

Kulawinek M and Kozubek A (2008) Quantitative determination of alkylresorcinols in cereal grains: independence of the length of the aliphatic side chain. J Food Lipids, **15**, 251-262.

Kunst L and Samuels A L (2003) Biosynthesis and secretion of plant cuticular wax. Prog Lipid Res, **42**, 51-80.

Lessire R, Hartmann-Bouillon M and Cassagne C (1982) Very long chain fatty acids: Occurrence and biosynthesis in membrane fractions from etiolated maize coleoptiles. Phytochemistry, **21**, 55-59.

Livak K J and Schmittgen T D (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. Methods, **25**, 402-408.

Lytollis W, Scannell R T, An H, Murty V S, Reddy K S, Barr J R and Hecht S M (1995) 5-Alkylresorcinols from *Hakea trifurcata* that cleave DNA. J Am Chem Soc, **117**, 12683-12690.

Magnucka E, Żarnowski R, Suzuki Y, Yamaguchi I, Pietr S J and Kozubek A (2001) The influence of herbicides on biosynthesis of antifungal resorcinols by rye seedlings. Bull Pol Acad Sci: Biol Sci, **49**, 361-369.

Matsuzawa M, Katsuyama Y, Funa N and Horinouchi S (2010) Alkylresorcylic acid synthesis by type III polyketide synthases from rice *Oryza sativa*. Phytochemistry.

Millar A A and Kunst L (1997) Very-long-chain fatty acid biosynthesis is controlled through the expression and specificity of the condensing enzyme. Plant J, **12**, 121-131.

Millar A A, Clemens S, Zachgo S, Giblin E M, Taylor D C and Kunst L (1999) *CUT1*, an Arabidopsis gene required for cuticular wax biosynthesis and pollen fertility, encodes a very-long-chain fatty acid condensing enzyme. Plant Cell, **11**, 825-838.

Miyanaga A and Horinouchi S (2009) Enzymatic synthesis of bis-5-alkylresorcinols by resorcinol-producing type III polyketide synthases. J Antibiot (Tokyo), **62**, 371-376.

Mizuuchi Y, Shimokawa Y, Wanibuchi K, Noguchi H and Abe I (2008) Structure

function analysis of novel type III polyketide synthases from *Arabidopsis thaliana*. Biol Pharm Bull, **31**, 2205-2210.

Oh C S, Toke D A, Mandala S and Martin C E (1997) ELO2 and ELO3, homologues of the *Saccharomyces cerevisiae* ELO1 gene, function in fatty acid elongation and are required for sphingolipid formation. J Biol Chem, **272**, 17376-17384.

Opanowicz M, Vain P, Draper J, Parker D and Doonan J H (2008) *Brachypodium distachyon*: making hay with a wild grass. Trends Plant Sci, **13**, 172-177.

Paquet N, Bernadet M, Morin H, Traas J, Dron M and Charon C (2005) Expression patterns of TEL genes in Poaceae suggest a conserved association with cell differentiation. J Exp Bot, **56**, 1605-1614.

Reiss J (1989) Influence of alkylresorcinols from rye and related compounds on the growth of food-borne molds. Cereal Chem, **66**, 491-493.

Richardson A, Franke R, Kerstiens G, Jarvis M, Schreiber L and Fricke W (2005) Cuticular wax deposition in growing barley (*Hordeum vulgare*) leaves commences in relation to the point of emergence of epidermal cells from the sheaths of older leaves. Planta, **222**, 472-483.

Riederer M (2006) Introduction: biology of the plant cuticle. In Biology of the Plant Cuticle. (Riederer M and Müller C, eds). Oxford: Blackwell Publishing, p. 3.

Ross A B, Shepherd M J, Schupphaus M, Sinclair V, Alfaro B, Kamal-Eldin A and Åman P (2003) Alkylresorcinols in cereals and cereal products. J Agric Food Chem, **51**, 4111-4118.

Ross A B, Kamal-Eldin A and Åman P (2004) Dietary alkylresorcinols: absorption, bioactivities, and possible use as biomarkers of whole-grain wheat- and rye-rich foods. Nutr Rev, **62**, 81-95.

Samuels A L, Kunst L and Jetter R (2008) Sealing plant surfaces: cuticular wax formation by epidermal cells. In Annu Rev Plant Biol, Vol. 59.

Schneiter R and Daum G (2006) Extraction of yeast lipids. Methods Mol Biol, **313**, 41-45.

Segura D, Cruz T and Espin G (2003) Encystment and alkylresorcinol production by *Azotobacter vinelandii* strains impaired in poly-β-hydroxybutyrate synthesis. Arch Microbiol, **179**, 437-443.

Sparkes I A, Runions J, Kearns A and Hawes C (2006) Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. Nat Protoc, **1**, 2019-2025.

Suh M C, Samuels A L, Jetter R, Kunst L, Pollard M, Ohlrogge J and Beisson F (2005) Cuticular lipid composition, surface structure, and gene expression in Arabidopsis stem epidermis. Plant Physiol, **139**, 1649-1665.

Suzuki Y, Esumi Y, Hyakutake H, Kono Y and Sakurai A (1996) Isolation of 5-(8'*Z*-heptadecenyl)-resorcinol from etiolated rice seedlings as an antifungal agent. Phytochemistry, **41**, 1485-1489.

Suzuki Y and Yamaguchi I (1998) Antimicrobial agents (phytoanticipins) in Gramineae crops, produced specifically during seedling stage. J Pest Sci, **23**, 316-321.

Suzuki Y, Esumi Y and Yamaguchi I (1999) Structures of 5-alkylresorcinol-related analogues in rye. Phytochemistry, **52**, 281-289.

Suzuki Y, Kurano M, Esumi Y, Yamaguchi I and Doi Y (2003) Biosynthesis of 5-alkylresorcinol in rice: incorporation of a putative fatty acid unit in the 5-alkylresorcinol carbon chain. Bioorg Chem, **31**, 437-452.

Taura F, Sirikantaramas S, Shoyama Y, Shoyama Y and Morimoto S (2007) Phytocannabinoids in *Cannabis sativa*: recent studies on biosynthetic enzymes. Chem Biodivers, **4**, 1649-1663.

Taura F, Tanaka S, Taguchi C, Fukamizu T, Tanaka H, Shoyama Y and Morimoto S (2009) Characterization of olivetol synthase, a polyketide synthase putatively involved in cannabinoid biosynthetic pathway. FEBS Lett, **583**, 2061-2066.

Thompson J D, Gibson T J, Plewniak F, Jeanmougin F and Higgins D G (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res, **25**, 4876-4882.

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.

Tropf S, Kärcher B, Schröder G and Schröder J (1995) Reaction mechanisms of homodimeric plant polyketide synthase (stilbenes and chalcone synthase). A single active site for the condensing reaction is sufficient for synthesis of stilbenes, chalcones, and 6'-deoxychalcones. J Biol Chem, **270**, 7922-7928.

Tulloch A P, Baum B R and Hoffman L L (1980) A survey of epicuticular waxes among genera of Triticeae. 2. Chemistry. Can J Bot, **58**.

Tulloch A P (1981) Composition of epicuticular waxes from 28 genera of Gramineae: differences between subfamilies. Can J Bot, **59**, 1213-1221.

Verdeal K and Lorenz K (1977) Alkylresorcinols in wheat, rye, and triticale. Cereal Chem, **54**, 475-483.

Wang W K, Schaal B A, Chiou Y M, Murakami N, Ge X J, Huang C C and Chiang T Y (2007) Diverse selective modes among orthologs/paralogs of the chalcone synthase (*Chs*) gene family of *Arabidopsis thaliana* and its relative *A. halleri* ssp. *gemmifera*. Mol Phylogenet Evol, **44**, 503-520.

Wasserman D and Dawson C R (1948) Cashew nut shell liquid. III. The cardol component of Indian cashew nut shell liquid with reference to the liquid's vesicant activity. J Am Chem Soc, **70**, 3675-3679.

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

Winkel-Shirley B (2002) Biosynthesis of flavonoids and effects of stress. Curr Opin Plant Biol, **5**, 218-223.

Yamashita Y, Matsunami K, Otsuka H, Shinzato T and Takeda Y (2008) Grevillosides A-F: glucosides of 5-alkylresorcinol derivatives from leaves of *Grevillea robusta*. Phytochemistry, **69**, 2749-2752.

Yamashita Y, Matsunami K, Otsuka H, Shinzato T and Takeda Y (2010) 5-Alkylresorcinol glucosides from the leaves of *Grevillea robusta* Allan Cunningham. J Nat Med, **64**, 474-477.

Yu D, Ranathunge K, Huang H, Pei Z, Franke R, Schreiber L and He C (2008) *Wax Crystal-Sparse Leaf1* encodes a β-ketoacyl CoA synthase involved in biosynthesis of cuticular waxes on rice leaf. Planta, **228**, 675-685.

Żarnowska E D, Żarnowski R and Kozubek A (2000) Alkylresorcinols in fruit pulp and leaves of *Ginkgo biloba* L. Z Naturforsch [C], **55**, 881-885.

Żarnowski R, Suzuki Y, Yamaguchi I and Pietr S J (2002) Alkylresorcinols in barley (*Hordeum vulgare* L. *distichon*) grains. Z Naturforsch [C], **57**, 57-62.