Molecular Cloning and Characterization of the Three Genes Coding for Desulfoglucosinolate Sulfotransferase in *Arabidopsis thaliana*

Hui Zhang

A Thesis in the

Department of Biology

Presented in Partial Fulfillment of the Requirements

For the Degree of Master of Science at

Concordia University

Montreal, Quebec, Canada

August 2003

Your file Votre référence
ISBN: 0-612-94794-7
Our file Notre référence
ISBN: 0-612-94794-7

ABSTRACT

Molecular Cloning and Characterization of the Three Genes Coding for Desulfo-glucosinolate Sulfotransferases in *Arabidopsis thaliana*

Hui Zhang

Three genes designated *AtST5a, AtST5b* and *AtST5c* from Arabidopsis thaliana, were characterized at the biochemical and molecular levels. They were found to code for desulfoglucosinolate sulfotransferases. Amino acid alignment of the three deduced amino acid sequences showed that they share 70-95% amino acid identity and 81-95% similarity. The results of enzymology studies showed that the three enzymes exhibit strict requirement for the thioglucose moiety of the desulfoglucosinolate substrate and for 3'-phosphoadenosine 5'-phosphosulfate as sulfate donor. The recombinant AtST5b and AtST5c enzymes exhibit broad substrate specificity with preference for 3-methylthiobutyl- and 2-phenylethyl desulfoglucosinolates. In contrast, the recombinant AtST5a enzyme is more specific and accepts only 3-methylthiobutyl desulfoglucosinolate to a significant level.

Promoter sequences of the *AtST5a, 5b* and *5c* genes (*Pro5a, Pro5b* and *Pro5c*) were cloned into β-glucuronidase-containing vectors and transformed into Arabidopsis thaliana. The results of histochemical studies revealed that the three genes have similar expression patterns. Their expression is restricted to the vascular bundles of hypocotyls, leaves, stems, mature siliques and roots. No activity was found in developing seeds. At the flowering stage, expression was found in the vascular bundles of receptacles, sepals, petals, pollen grains, stamens and stigmas.

The glucosinolate profile of *AtST5a* T-DNA insertion knock-out plants was

found to exhibit significantly different than the one of wild type plants. In contrast, no significant difference was found between the profiles of *AtST5c* T-DNA insertion knock-out and wild type plants. These results are consistent with the results of the biochemical studies of the three *AtST5* genes which showed redundancy between the biochemical function of the AtST5b and 5c enzymes.

The expression of the *AtST5a, 5b* and *5c* genes is not affected by treatments with indole-3-acetic acid, methyl jasmonate and salicylic acid. These results plus the ones of previous expression studies of other genes involved in glucosinolate biosynthesis suggest that the genes involved in the fist step of glucosinolate biosynthesis pathway may be inducible, while those genes involved in the final steps, such as the ones encoding desulfoglucosinolate sulfotransferases, are constitutively expressed.

ACKNOWLEDGMENTS

I would like to express my appreciation to my thesis supervisor, Dr. Luc Varin, for giving me the opportunity to work in his lab and providing me the flexible and challenging environment to study.

I thank the members of my committee, Dr. P. Gulick and Dr. R. Storms, for their support and patience. I appreciate the collaboration of Dr. Bill Zerges as external examiner. Thanks to Dr. Grant who acted as chair during my defense. I am obligated to the Graduate Program Directors at the Department of Biology, Dr. Narinder Kapoor, and Dr. Paul Albert, for their warm reception and encouragement.

I am so grateful to my IRBV colleagues for their kind and friendly encouragement for my thesis defense. Especially I would like to express my deep gratitude to Dr. Jean Rivoal for his scientific advice and support.

Thanks to the members of Dr. Luc Varin's laboratory, Diego Spertini, Frederic Savard, Anastassia Levitin and other undergraduate students for their friendly and warm company during several years. Special thanks to Diego Spetini and Frederic Savard for their supervision and help for my experiments.

I would like to thank all my family members, specially my mom and my sisters, who encouraged me through the most difficult times. Finally I would like to express my deep gratitude to the most important person of my life, my husband, for his support, encouragement and understanding during those difficult years.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SCHEMES

LIST OF ABREVIATIONS

AtST#	*Arabidopsis thaliana* sulfotransferase #
BLAST	Basic local alignment search tool
BR	Brassinosteroid
cDNA	Complementary deoxyribonucleic acid
ChoST	Choline sulfotransferase
ClustalW1.8	Multiple sequence alignment program version 1.8
ddH2O	Double distilled, ultra-purified water
DEPC	Diethyl pyrocarbonate
Desulfoglucosinolate	DSG
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphates
DPI	diphenylene iodonium
DTT	DL-Dithiothreitol
EDTA	Ethylenediamine tetraacetic acid
EtOH	Ethanol
ExPASy	Expert protein analysis system
FST	Flavonol sulfotransferase
GI	Glucosinolate
GUS	β-glucuronidase
HCL	Hydrochloric acid

HPLC	High performance liquid chromatography
IAA	Indoleacetic acid
IPTG	Isopropyl β-D-thioglucoside
Kbp	Kilo base pair
MeOH	Methanol
MeJA	Methyljasmonate
MES	2-[N-morpholino] ethanesulfonic acid
$MgSO_4$	Magnesium sulfate
$MnCl_2$	Magnesium chloride
mRNA	Messenger ribonucleic acid
MS	Murashige and Skoog salt mixture
NaCl	Sodium chloride
NaOAc	Sodium acetate
NaOH	Sodium hydroxide
Ni-NTA	Nickel-nitrilotriacetic acid
OD	Optical density
PAGE	Polyactylamide gel electrophoresis
PAP	3`-phosphoadenosine 5`-phosphate
PAPS	3`-phosphoadenosine 5`-phosphosulfate
PCR	Polymerase chain reaction
P-OHBG	p-hydroxybenzylglucosinolate
pkatal	picomoles of product of formed /second of reaction time
PLMF-1	Periodic-leaf-movement-factor-1

A. INTRODUCTION

Glucosinolates (GIs) (mustard oil glycosides) are plant secondary metabolites found almost exclusively in plants of the order Capparales (Rodman, et al., 1996; Selmar, 1999). These compounds are sulfated thioglucosides, which are stored in the vacuole and co-exist with endogenous thioglucosidases known as myrosinases (Bones and Rossiter, 1996). Tissue damage brings GIs into contact with myrosinase, resulting in the release of glucose and numerous compounds such as isothiocyanates, organic nitriles and thiocyanates having diverse biological activities. These enzymatic degradation products mainly account for the distinctive flavors of cabbage (*Brassica oleracea*) and condiments, e.g. in white mustard (*Sinapis alba*) and black mustard (*B. nigra*) (Fenwich, et al., 1983). Isothiocyanate, a hydrolytic degradation product, was shown to have anticarcinogenic properties (Zhang et al., 1992). Furthermore, GIs are thought to serve in plant defense due to the toxic and deterrent properities of their degradation product to many pathogens and herbivores (Louda and Mole, 1991). Because of the toxic properties and pungent taste of their degradation products, GIs are often classified as anti-nutritional compounds. For instance, rapeseed (*B. napus*) is one of the main sources of edible vegetable oil. After oil extraction, the by-product seed meal is an excellent feed concentrate containing about 40% protein with balanced amino acid composition. However, such high nutritional value is impaired by high concentrations of GIs due to the pungency and goitrogenic effect of their breaking products (Bone et al., 1996; Fenwich et al., 1983).

As consequence of the wide-range biological activities of GIs, their biosynthesis has been studied intensively in recent years. However, the characteristics of some of the

enzymes involved are still imperfectly known. Isotopic tracer studies have clearly outlined how GIs are derived from corresponding amino acids, with aldoximes, thiohydroximes and desulfoglucosinolates (DSGs) as intermediates (Underhill et al., 1973). It is the goal of our lab to isolate and characterize the DSG sulfotransferase (s) (STs) which catalyzes the final step of GI biosynthesis from the model plant *Arabidopsis thaliana*. *A. thaliana* is a small flowering plant that is widely used as a model organism in plant biology. It offers many advantages for detailed genetic and molecular studies because of its small size, short life cycle, small genome, and ability to be transformed. However, the greatest advantage to work with Arabidopsis is the fact that its genome has been totally sequenced. Besides, it contains about 30 different GIs (Hogge et al., 1988; Haughn et al., 1991). Therefore, this species is well suited for the examination of GI biosynthesis and their physiological roles. Three genes, *AtST5a/5b/5c*, from Arabidopsis have been shown to code for DSG-ST in our lab. Amino acid alignment of the three enzymes indicates that they share 70-95% amino acid sequence identity and 81-95% similarity. AtST5a and AtST5b have been partially characterized previously (Diego Spertini, unpublished result). They were found to sulfonate DSGs with high specificity. The present study covers the following aspects:

- Characterization of the biochemical function of AtST5c and subsequent comparison of its biochemical function with the other two DSG -STs, AtST5a and AtST5b.
- In vivo studies on the expression of the three *AtST5* genes using histochemical assays.
- Phenotypic and GI content analysis of *AtSt5a/5c* knock-out plants by HPLC.
- Analysis of the expression pattern of the three genes following induction by plant

signal molecules, such as indole acetic acid (IAA), methyljasmonate (MeJA), salicylic acid (SA) by reverse transcriptase-PCR (RT-PCR).

B. LITERATURE REVIEW

In this section, I will present the GIs structure and nomenclature, their distribution and bioynthesis. In addition, GI degradation and transport as well as the regulation of their biosynthesis and degradation will be presented. Other aspects covered by this review are the mechanism of sulfate activation, sulfonate transfer and the sulfotransferase (ST) enzymes that catalyze this reaction in plants.

B.1. GIs

B.1.1. Structure and nomenclature

GIs are a class of naturally occurring anions which have so far, been found exclusively in the plant kingdom. They have a well-established basic chemical structure (Figure 1) where the sugar moiety is attached through β-thioglucosidic linkage. Prior to 1948, only nine GIs were known, by 1980 this number had increased to seventy, and to date, more than 110 GIs have been identified (http://robig.pmf.ukim.edu.mk/bbogdanov/GIs/webG2.htm).

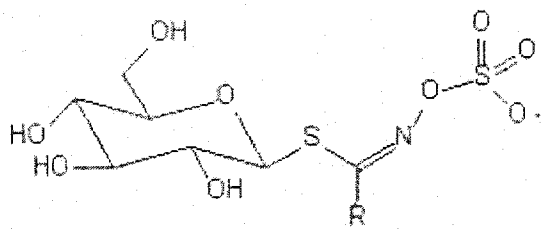


Figure 1. The basic structure of GIs

Because the number of GIs identified has increased, the nomenclature has shifted from classical names such as sinigrin and sinalbin, or trivial names such as

glucotropeolin to the current systematic method, in which the anion of the salt (I; R=H) is designated by the term "GI" and the chemical name of the radical "R" (the GI side chain) is used as a prefix. For example, sinigrin and glucotropaecolin are renamed potassium ally-GI and potassium benzyl-GI, respectively. Generally they are grouped into aliphatic, aromatic and indole GIs depending on whether they originate from methionine, phenylalanine or tryptophan. The structural diversity of GIs is largely due to chain elongation resulting from the inclusion of amino acids such as methionine and phenylalanine taking place before the formation of the GI core structure and by secondary modifications of the GI side chain (e.g. thiol oxidation, desaturation, hydroxylation and esterification) and /or of the glucose moiety (esterification).

B.1.2. Distribution in the plant kingdom

GIs have only been found in 15 families of dicotyledonous plants with an uneven distribution pattern (Rodman, 1991). GIs seem to be present in all species of three families, the Capparidaceae, the Cruciferae and the Resedaceae, with a great variability in side chain structures. In the remaining families only GIs either derived directly from protein amino acids or with additional hydroxyl groups have been found. This simple pattern is similar to that known for cyanogenic glycosides. Since a number of important vegetables, herbs, and agricultural crops belong to Cruciferae family, the large variety of GIs found in this family became of special interest.

Generally, a plant species contains more than one GIs and often there are both qualitative and quantitative differences between the different organs and during the life cycle of the plant. For instance, in the Brassicaceae, an extensive variation of GIs has

been observed. Over 30 different GIs have been found in *B. napus* (Sorensen H, et al., 1990) and 23 in *Arabidopsis thaliana* (Hogge, et al., 1988). Studies on the concentration of GIs in different vegetables and crop plants have been reported. Levels of 0.1% or less in fresh weight (Cole, 1976; VanEtten *et al.*, 1976) and of up to 10% in dry seed weight (Josefsson, 1972; VanEtten et al., 1974; Wetter and Dyck, 1973) were observed. GIs always occur together with myrosinase, an enzyme involved in the break down of GIs into several harmful degradation products. A detailed review of the myrosinase system is presented in section B.1.4.

B.1.3. Biosynthesis

Biosynthetic studies of GIs have revealed that they are all derived from only seven amino acids (alanine, leucine, isoleucine, valine, phenylalanine, tyrosine and tryptophan) and a number of chain-elongated homologues. In Arabidopsis, GIs are derived from only three amino acids (methionine, phenylalanine and tryptophan) (Hogge et al., 1988). It has been demonstrated that GIs are derived from amino acids by isotopic labelling studies. Their origin from amino acids was first demonstrated by incorporation of DL- [β –^{14}C] tryptophan into indol-3-ylmethylGI in *Brassica oleraceae* (Kuráček et al., 1962) and by the incorporation of [^{14}C] from phenylalanine into benzylGI in *Tropaeplum mahjus* (Underhill et al., 1962; Benn, 1962). GIs are formed by a common biosynthetic pathway, which comprises chain elongation of protein amino acids, oxidative decarboxylation of the amino acid to its corresponding aldoxime, conversion of the oxime into the basic GI structure followed by secondary modifications. The synthesis of the basic GI structures is presented in figure 2.

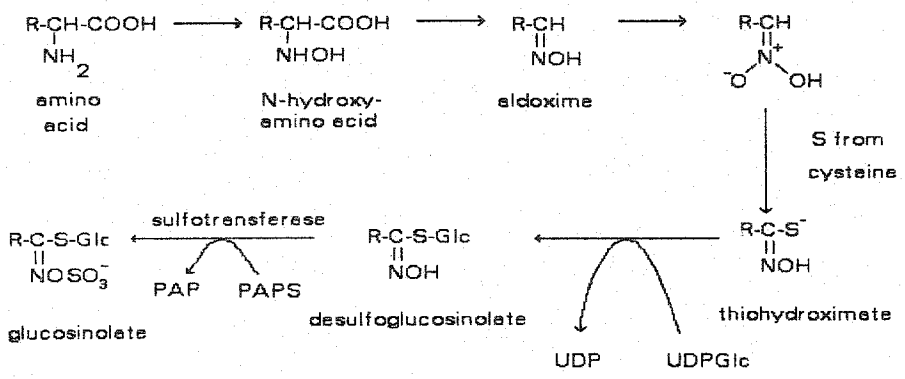


Figure 2. Proposed GI biosynthesis pathway

B.1.3.1. The intermediates between amino acids and GIs

A number of aldoximes have been demonstrated to be intermediates in GI biosynthesis (Kjaer and Olesen Larsen, 1973; Underhill and Wetter, 1973). Three enzyme systems, cytochrome P450 dependent monooxygenases, flavin-containing monooxygenases and plasma membrane peroxidases, have been shown independently to be involved in the conversion of amino acids to aldoximes. Microsomes isolated from *S. alba, Tropaeolum majus* and *Carica papya* were able to convert the aromatic amino acids tyrosine and phenylalanine to their corresponding oximes. The activities were subjected to photoreversible carbon monoxide inhibition as well as being inhibited by cytochrome P450 inhibitors (Bennett et al., 1997; Du et al., 1995), indicating cytochrome P450 dependent oxime formation. The role of flavin-containing monooxygenases in the formation of aldoximes from homophenylalanine and chain-elongated methionine derivatives was established by measuring $^{14}CO_2$ release from [1- ^{14}C]-homophenylalanine and dihomophenylalanine catalyzed by microsomes isolated from *B. napus* leaves. The

reaction is inhibited by the flavoprotein-specific inhibitor diphenylene iodonium (DPI) and copper salts (Bennett, et al., 1996; Bennett, et al., 1995; Burmeister, et al., 1997). The third system, the plasma membrane peroxidase, has been shown to be able to catalyze the conversion of tryptophan to indole acetaldoxime which can be used for both indoleGI and possibly IAA biosynthesis (Ludwig-Mueller, et al., 1990). However, recent studies suggest that the majority of the enzymes catalyzing oxime formation in GI biosynthesis are cytochrome P450 monooxygenases. To date, five out of the seven cytochrome P450 *CYP79* homologues have been characterized and shown to be involved in GI biosynthesis in *A. thaliana* (Halkier et al, 1999). It has become evident that CYP79A2 catalyze the conversion of phenylalanine to its corresponding aldoximes (Wittstock et al., 2000), CYP79B2 and CYP79B3 convert tryptophan to indole-3-acetaldoxime (Mikkelsen et al., 2000), and CYP79F1 and CYP79F2 metabolize chain-elongated methionine derivatives (Hansen et al., 2001; Reintanz et al., 2001). CYP79F1 and CYP79F2 differ in their substrate specificity. *CYP79F1* knockout mutants are free of GIs derived from short-chain methionine homologues (with one to four additional methylene groups) but contain wild-type levels of pentahomo- and hexahomomethionine-derived GIs. Characterization of these mutants suggested that CYP79F1 only converts short-chain methionine derivatives to the corresponding aldoximes (Reintanz et al., 2001). However, recent investigations have shown that CYP79F1 is actually able to metabolize not only short-chain methionine homologues but also pentahomo- and hexahomomethionine, whereas the catalytic activity of CYP79F2 is restricted to pentahomo- and hexahomomethionine (Wittstock et al., 2002).

The search for intermediates between the amino acids and the aldoximes has been

done in a series of experiments. N-hydroxyamino acids are proposed to be the intermediates as deduced from experiments with ^{14}C-labeled N-hydroxyamino acids (Kindle and Underhill, 1968). This conclusion is further supported by the definite establishment of N-hydroxytyrosine as an intermediate between tyrosine and p-hydroxyphenylacetaldoxime in the biosynthesis of the cyanogenic glucoside dhurrin in *Sorghum* seedlings.

Aldoximes are transformed into GIs via thiohydroximic acids and DSGs. This conclusion was drawn from the results of the incorporation studies with phenylacetothiohydroximic acids labelled with both ^{14}C and ^{35}S, and with benzyl-DSG labelled with ^{14}C both in the glucose moiety and the aglycone (Underhill and Wetter, 1969). The conversion of oximes to the thiohydroximates is a poorly understood step. The reduced sulfur donor and the oxime-derived sulfur acceptor ultimately leading to the formation of the thiohydroximate remain unclear. It has been proposed that the aldoxime is N-hydroxylated to give rise to a reactive intermediate (proposed to be an aci-nitro compound or nitrile oxide), which is then subsequently conjugated with a sulfur donor (presumably cysteine) to produce a cysteine conjugate (S-alkylthiohydroximates). Two cytochrome P450, CYP83A1 and CYP83B1 from Arabidopsis have been identified as the aldoxime-metabolizing enzymes involved in this intermediate step by two different approaches: a systematic reverse-genetic approach, combined with a screen for auxin phenotypes (Bak et al., 2001; Bak and Feyereisen, 2001), and a biochemical approach (Hansen et al., 2001). Biochemical characterization of the two enzymes shows that they can accept all the aldoximes tested as substrates. However, CYP83A1 has high affinity for the aliphatic aldoximes, whereas CYP83B1 has high affinity for indole-3-

acetaldoximes and the aromatic aldoximes (Bak and Feyereisen, 2001). The conversion of s-alkylthiohydroxime to thiohydroximate is thought to be catalyzed by a C-S lyase. The conversion of thiohydroximic acid into DSG is catalyzed by a soluble thiohydroximate glucosyltransferase (S-GT). Several S-GTs have been partially purified and characterized (Halkier, 1999). It has been shown that S-GTs are present in all GI containing plants tested, whereas they are absent in non-GI containing plants (GrootWassink et al., 1997). Furthermore, *S-GTs* have been cloned in several species (GrootWassink et al. 1997; Marillia et al., 2001) and their deduced amino acid sequence share a significant degree of homology with other glucosyltransferases characterized in other species, including a highly conserved motif within this family of enzymes corresponding to the UDP-glucose-binding domain (Marillia et al., 2001). The final step, the conversion of DSG to GI, is catalyzed by DSG-ST (s) using PAPS as sulfate donor. Both enzymes, S-GT and DSG-ST, display a high specificity for the GI basic structure respectively but no specificity for the nature of the side chain (Halkier, 1999). The lack of side-chain specificity appears to be a common feature for the enzymes involved in the last two steps in the pathway, and contrasts with the high specificity of upstream enzymes involved in the conversion of amino acids to oximes.

B.1.4. Degradation of GIs

B.1.4.1. Myrosinase

Myrosinase is an enzyme, which cleaves the thioglucosidic linkage of GIs yielding D-glucose and an unstable sulfate compound. The latter rearrange to form either an isothiocyanate, a nitrile, an amine, an epithionitrile, a thiocyanate, an oxazolidine-3-

thione or other less prevalent products depending upon the nature of the GI aglycone, the pH, the presence of metal ions and epithiospecifier proteins (Foo et al, 2000). Several reports have described the isolation and physico-chemical characterization of the myrosinase enzyme. The most extensive studies have been made in Brassicaceae species, mainly *Lepidium sativum* seedlings (Durham, et al., 1990), *Sinapi. alba* seeds (Bjorkman, et al., 1972, and Bjorkman, et al., 1973.) and *Brassica. napus* seeds (Bjorkman, et al., 1973, and James, et al., 1991).

Myrosinases are generally characterized as soluble proteins. But more recent analysis has shown that certain myrosinases also exist in complex with other proteins and are insoluble (Rask, et al., 2000). All plant myrosinases characterized so far are highly glycosylated, usually have pH optima ranging from 4 to 7, and have GIs as substrates (Bones, et al., 1996, and Chen, et al., 1999). The genes encoding for myrosinases in *B. napus* have been divided into three subfamilies, MA, MB, MC, which contain five, ten and fifteen members, respectively (Rask, et al., 2000). The embryo is the only organ where all subfamilies are expressed.

B.1.5. GI transport

GIs have been detected in all organs of Brassica plants, but the capacity of de novo biosynthesis varies between organs at different development stages (Porter et al, 1991). GIs are known to be transported within the plant, e.g., from the silique walls to the seeds. For example, in *Tropeolum majus* high amount of benzylGI, primarily synthesized in the leaves, is also found to accumulate in other tissue, such as developing seeds (Lykkesfeldt et al., 1993). Additionally, studies of GI profile in seeds and leaf

tissue of *Brassica. napus* F1 hybrids from reciprocal crosses between cv. Cobra and a synthetic line showed that the profile of the aliphatic GIs in the seeds was identical to the profile in the leaves of the maternal parent (Magrath et al., 1993), suggesting that fully formed GIs are transferred from maternal tissue into the developing seeds. Furthermore, in vivo feeding of labelled tyrosine to isolated seeds and intact siliques of *Sinapis alba* showed that although a low rate of *de novo* biosynthesis of p-hydroxybenzylGI (p-OHBG) occurred in the seeds, the majority of p-OHBG was synthesized de novo in the silique wall and subsequently transported to the seeds (Du et al., 1998). A recent detailed quantitative study of GI variation among 39 *A.thaliana* ecotypes also showed a significant positive correlation between the levels of aliphatic GIs in leaves and seeds (Kliebenstein et al., 2001). Moreover, DSG-ST activity was found in seeds of *B. napus* (Toroser et al., 1995), suggesting that DSGs were the transport form of GIs.

Recently, studies on phloem mobility of radiolabelled GIs and DSGs in *B. napus* plants showed that both compounds had the suitable physicochemical properties to allow mobility (Brudenell et al., 1999). This data suggests the possibility of phloem transport of GIs. In support of phloem transport, aphid feeding experiments on the black mustard *B. nigra,* have shown that there was more than 10 mM sinigrin in the phloem sap of young leaves, whereas there was only about 1-2 mM GIs in mature, presenescent and senescent leaves (Merritt, 1996). These data shed light on the possibility that long-distance transport of GIs follows the principle of assimilating translocation, i.e. mass flow from source to sink. Thus, GIs synthesized in the mature leaves are supposed to be loaded into and transported by the phloem, together with photoassimilates in the

osmotically driven translocation stream from source to sink. In addition to possible nutritional function in case of environmental stress, the phloem transport of GIs may help the plant to defend itself against phloem-feeding organisms and gain the ability to co-ordinate the synthesis and use of protective resources between various organs.

B.1.6. Regulation of GI metabolism

It has been demonstrated that GI profiles vary greatly in different plant species and at different ages within a single species (Clossais-Besnard et al, 1991; Porter et al, 1991). In addition, the level of GI is affected by environmental factors such as light (Hasegawa et al, 2000), fungal infection (Li et al, 1999), wounding (Koritsas et al, 1991; Bartlet et al, 1999), insect damage (Koritsas et al, 1991; Ludwig-Mueller et al, 1999) and other forms of stress (Bartlet et al, 1999; Blake-Kalff et al, 1998).

B.1.6.1. Regulation of biosynthesis

Generally speaking, GI biosynthetic activities are high in young tissues and decrease as the tissues mature (Bennett et al, 1995; Lykkesfeldt et al, 1993; Poter et al, 1991; Clossais-Besnard et al, 1991). For example, the concentration of aliphatic GIs is decreased dramatically during the first 7 days following imbibition of *B. napus* seeds. Meanwhile, there is de novo synthesis of indole GIs and gluconasturtin. During the rapid growth period, additional GIs may accumulate (Clossais-Besnard et al, 1991). However some GI concentration increases only slightly and may even decrease (Poter et al, 1991). At the flowering stage, both vegetative tissues and inflorescences exhibit a decrease of GI concentration (Poter et al, 1991; Clossais-Besnard et al, 1991). These phenomena are

consistent with the decrease of the enzymatic activities involved in the formation of oximes from both chain-elongated amino acids and from tryptophan with the increasing maturity of the tissues (Bennett et al, 1995).

Several biologically active molecules are implicated in the induction of GI biosynthesis. For instance, treatment of plants with methyl jasmonate (MeJA), a signal molecule associated with reactions pertaining to plant wounding and herbivory, led to a systematic accumulation of indole GIs, suggesting a role for indole GIs in the plant defence response (Barlet et al, 1999; Doughty et al, 1995; Li et al, 1999; Porter et al, 1991). Similarly, the concentration of 2-phenylethyl GI is specifically increased by the treatment of plants with salicylic acid (SA), a phenolic compound involved in pathogen related responses (Kiddle et al, 1994). In contrast, treatments with abscisic acid result in low levels of GIs (Mollers et al, 1999). Interestingly, nitrogen starvation could induce the synthesis of GIs, which may imply a 'fine-tuning' short-term nutritional storage/balance function. It has been reported that the *CYP79B2* and *CYP79B3* genes are involved in indole-3-acetic acid (IAA) and indole GI biosynthesis (Hull et al, 2000; Mikkelsen et al, 2000) in Arabidopsis. After MeJA treatment, both genes were highly induced and the amount of their corresponding metabolites, indole GIs, increased 3- to 4-folds (Mikkelsen et al., 2003). Meanwhile, it has been reported that infection by a virulent bacterial pathogen induces the expression of *CYP79B2* (Hull et al., 2000). Further study indicates that the level of IAA is regulated by the flux of indole-3-acetaldoxime (IAOx) through a cytochrome P450, CYP83B1, to the GI pathway (Soren et al, 2001). Under stress conditions, IAOx could be preferentially channelled into indole GI biosynthesis. Since large variations in GI profiles exist between individual tissues at

different developmental stages, it would be interesting to determine which step of biosynthesis account for these variations and how the signals cross-talk in the regulation of GI biosynthesis.

B.1.6.2. Regulation of degradation

Studies of the myrosinase genes have clearly shown that their expression is regulated during development in a tissue specific manner (Lenman et al, 1993). This pattern of expression correlates with the dramatic decrease in GI content during the early developmental stage of *B.napus* (Lenman et al, 1993), and the reduction in GI levels in leaves upon maturation (Porter et al, 1991) could be due to the hydrolysis of GIs by myrosinase. It has been reported that the level of myrosinase activity in plants is modulated by nutrition (Blake-kalff et al, 1998; Visvakubgan et al, 1998), induced by blue illumination (Hasegawa et al, 2000) and down regulated by SA treatment (Taipalensuu et al, 1997). Although regulatory mechanisms involved in myrosinase expression and GI degradation in vivo still remain unclear, enhanced degradation of GI under abiotic and biotic stresses could lead to efficient response to stress conditions.

B.1.7. GI and plant defence

The GI/myrosinase system is thought to be a dynamic system, responding to environmental changes and to plant damage. It has been reported that the plants pre-treated with certain elicitor compounds, which selectively stimulate GI accumulation (Kiddle et al, 1994; Doughty et al, 1995), can enhance resistance of the plant to subsequent infection by pathogens (Doughty et al, 1995). It has been proposed that the

role of GIs in plant defence is due to their breakdown allelochemical products, such as isothiothiocyanates (Rask et al, 2000). For instance, naturally occurring isothiocyanates possess a range of antifungal, antibacterial and antimicrobial activities, and thus repel microorganisms, insects and molluscs (Fenwich et al, 1983), but very few examples have been studied in vivo.

Although numerous studies attempting to correlate the levels of GIs with resistance to specific pathogens have been done so far, none were successful. This may reflect a complex interaction between numerous different defence compounds in plants and their interacting organisms. Successful pathogens are able to circumvent the protective effects either by limiting the extent of tissue damage and thus minimizing the release of toxic constituents.

B.2. Sulfonation reaction

B.1.1. Sulfoconjugation reaction

The sulfoconjugation reaction is the transfer of a sulfonate group from 3'-phosphoadenosine 5'-phosphosulfate (PAPS) to an acceptor substrate. This reaction is catalyzed by a family of sulfotransferase enzymes. Many different endogenous and xenobiotic molecules containing OH group are substrates for the sulfotransferases. The general sulfoconjugation reaction is as follows:

$$PAPS + R\text{-}OH \xrightarrow{\text{Sulfotransferase}} PAP + R\text{-}OSO_3^-$$

Scheme 1. General sulfoconjugation reaction

In higher organisms, PAPS is strictly required for this reaction. This strict requirement has been demonstrated in a number of biochemical experiments (Rens Domiano et al, 1987Komatsu et al, 1994; Chiba et al, 1995). PAPS is the "activated" form of sulfate and is synthesized from endogenous sulfate. Its synthesis is the result of a two-step, coupled reaction. The first step combines ATP and inorganic sulfate (SO_4^{2-}) to form APS and pyrophosphate (PPi). The reaction is catalyzed by the enzyme ATP-sulfurylase in the presence of Mg^{2+}.

$$ATP + SO_4^{2-} \xrightarrow{\text{ATP-sulfurylase}} APS + PPi$$

Scheme 2. Step 1 of PAPS biosynthesis

The subsequent step combines the APS formed with an additional ATP molecule to form PAPS and ADP. This reaction is catalyzed by APS-kinase in the presence of Mg^{2+}.

$$ATP + APS \xrightarrow{\text{APS-kinase}} PAPS + ADP$$

Scheme 3. Step 2 of PAPS biosynthesis

B.2.2. Plant sulfotransferases

A number of sufotransferase enzymes have been isolated from animal and plant species. Based on sequence homology, these enzymes are classified into three groups: the phenol STs (PST), the hydroxysteroid STs (HSST) and the flavonol STs (FST) (Weinshilboum et al., 1994). The plant STs are grouped under the FST family, also known as the SULT3 family. Although the study of plant STs have lagged considerably

compared to those of their mammalian homologs, several plant STs have been characterized at the biochemical and molecular levels.

Flavonoid Sulfotransferase

The flavonoids are synthesized by plants in response to UV irradiation and microbial attack (Hahlbrock et al., 1989). More recent reports indicate that a number of flavonoids act as signal molecules that regulate the expression of early nodulation genes in *Rhizobium* (Long, 1989) and polar auxin transport (Faulkner et al., 1992).

To date, four position-specific flavonol sulfotransferases from *Flaveria chloraefolia* and *Flaveria bidentis* have been isolated and characterized (Varin et al., 1989; Varin et al., 1991). These enzymes, specific for the hydroxyl groups at position 3 of flavonol aglycones, 3' and 4' of flavonol 3-sulfate, and 7 of flavonol 3, 3'- or 3, 4'-disulfate, exhibit no requirement for divalent cations, are not sensitive to SH group reagents and have similar molecular masses (35 kDa). Their Km values for substrate and cosubstrate (PAPS) vary between 0.2 and 0.4 μM. Furthermore, the four enzymes from *Flaveria* share 69% amino acid identity and have 30% similarity when compared with STs characterized from animal tissues.

Comparative sequence studies of the cytosolic STs defined four conserved regions, suggesting a common evolutionary origin (Varin et al., 1992; Weinshilboum et al., 1994). The conserved region I, which is located in the amino-terminal portion, and the conserved region IV, which is located in the carboxyl-terminal portion, represent almost uninterrupted blocks of sequence identity and have been proposed to participate in shared functions of these enzymes such as co-substrate (PAPS) binding or specifying the

proper protein folding for catalysis (Varin et al., 1992; Weinshilboum et al., 1994). Further structure-function studies of cytosolic STs identified two strictly conserved residues important for catalysis and substrate binding. Residue lys^{59} within conserved region I was suggested to be involved in catalysis of the flavonol 3-ST rather than involved in co-substrate binding (Marsolais et al., 1995). The second residue Arg^{276} within conserved region IV is thought to be critical for the binding of PAPS (Marsolais et al., 1995). Studies with the flavonol 3- and 4'-STs from *F. chloraefolis* have identified a segment, designated as domain II, determining their substrate specificity (Varin et al., 1995). This domain, encompassing amino acids 92 to 194 of the flavonol 3-ST, can be subdivided into a highly conserved region and two regions of high divergence.

Choline Sulfotransferase

Choline sulfate has been found to accumulate in all species of the salt stress-tolerant Plumbaginaceae family (Hanson et al., 1994). It is believed to act as an osmoprotectant in response to salinity or drought stress. In 1991, in vivo radiotracer experiments with *Limonium* species showed that [^{14}C] choline was converted to [^{14}C] choline sulfate (Hanson et al., 1991), suggesting the presence of a choline ST activity. Later, the enzyme was partially purified from *L. sinuatum* roots (Rivoal et al, 1994). Furthermore, the choline ST strictly requires PAPS as sulfate donor and has apparent Km values of 5.5 and 25 μM for PAPS and choline, respectively.

Gallic Acid Glucoside Sulfotransferase

Some plants have the ability to change the orientation of their leaves in response to external and internal stimuli. Examples of these movements include nyctinastism-leaf closure at night, heliotropism-solar tracking leaves and seismonasty-leaf closure in response to touch. Both nyctinastic and seismonastic movements can be observed in *Mimosa pudica.* Gallic acid 4-*O*- (β-D-glucopyranosyl-6'-sulfate), which was named the periodic leaf movement factor 1 (PLMF-1), was shown to induce leaf closure movement in *Mimosa* (Schildknecht et al., 1990). Further studies of structure-activity relationship of PLMF-1 and other natural and synthetic derivatives demonstrated that the sulfate group, the carboxyl group on the aromatic ring, and the specific substitution pattern of the phenolic ring were required for biological activity (Kallas et al., 1989).

The last step of PLMF-1 biosynthesis is catalyzed by a sulfotransferase, which transfers the sulfonate group from PAPS to gallic acid glucoside. This enzyme was partially purified from the plasma membrane of Mimosa and was found to exhibit strict specificity and high affinity for both gallic acid glucoside and the cosubstrate PAPS with Km values of 3.0 μM and 0.5 μM, respectively (Varin et al., 1997a).

DSG Sulfotransferase

Up to date, two DSG-STs (see section B.1.3.2) have been purified from *Lepidium sativun* (cress) (Koritsas et al., 1991) and *Brassica juncea* (Jain et al., 1990). The purified enzyme from *Lepidium* was found to be extremely unstable and to require bovine serum albumin for stabilization. It has an apparent molecular weight of 31±5 kDa

and an isoelectric point of 5.2; whereas the enzyme from *Brassica* had a molecular mass of 44 kDa and an isoelectric point of 5.2. Both enzymes display strict requirement for PAPS as a sulfate donor and for the thioglucose moiety of DSGs as a sulfate acceptor. Among all the substrates tested, the cress ST exhibits strict specificity for DSGs with an apparent km of 82 μM for benzyl-DSG, 670 μM for p-hydroxylbenzyl DSG and 6.5 mM for allyl DSG. The enzyme from *B. juncea* shows the highest affinity for benzyl-DSG with a km value of 2.3 μM. This strict substrate specificity is a common feature of the plant STs and has also been observed with the flavonol sulfating enzymes. The two DSG-ST enzymes have many features in common, such as no requirement for divalent cations and no inhibition by SH group reagents.

The Sulfotransferase of Arabidopsis thaliana and Brassica napus

18 ST sequences have been found to be present in the *A. thaliana* genome. Recently, a cDNA clone encoding a putative ST (RaRo47) has been isolated from Arabidopsis cell suspensions (Lacomme et al., 1996). Amino acid sequence comparison revealed that RaR047 is highly related to the plant flavonol STs with an overall identity of 50-52.5% and also to STs from animal tissues. The four regions of sequence conservation among all cytosolic STs are also present in the predicted protein sequence of RaR047.

The expression of the Arabidopsis ST gene corresponding to RaR047 was examined during different developmental stages. It was found that, at the level of steady-state mRNA, expression of the gene encoding this ST was rapidly induced in the aerial parts of young seedlings and during growth of Arabidopsis cell cultures. No expression could be detected in roots. The study of the expression of the RaR047 gene showed that

SA and MeJA efficiently induced its expression. SA is a signal molecule in the plant defence response to pathogen infection (Raskin, 1992), and MeJA is known as a regulator of plant defence genes (Xu et al. 199). Furthermore, this gene was preferentially expressed in response to avirulent pathogens causing the hypersensitive reaction, thus suggesting that the RaR047 ST protein might participate in establishing the systemic acquired resistance.

Recently, an ST (BnST3) involved in the sulfonation of brassinolsteroids (BRs) was isolated from *Brassica napus* using RaR047 as a probe (Rouleau et al., 1999). The deduced amino acid sequence of BnST3 shares 87% sequence identity with that of the RaR047 protein. The study of ST activity revealed that BnST3 preferentially catalyze the *O*-sulfonation of brassinosteroids and of mammalian estrogenic steroids. The enzyme is specific for the hydroxyl group at position 22 of 24-epicathasterone, an intermediate in the biosynthesis of 24-epibrassinolide. The sulfonation of 24-brassinolide by BnST3 abolished its biological activity in the bean second internode bioassay (Rouleau et al., 1999). This mechanism of hormone inactivation by sulfonation is similar to the modulation of estrogen biological activity observed in mammals (Hobkirk, 1993). Furthermore, the expression of the BnST3 gene was also induced by SA, suggesting that plants respond to pathogen infection by modulating steroid-dependent growth and development process.

C MATERIAL AND METHODS

C.1. Molecular cloning

C.1.1 Plasmid DNA preparation

Plasmid DNA was prepared by the centrifugation of 4 ml of overnight-grown *E. coli* bacterial culture for 5 minutes at 9000 rpm. The supernatant was removed and the bacterial pellet was resuspended in 100µl of lysis buffer containing 25 mM Tris-HCl, 10 mM EDTA, 1% glucose (w/v), and 5 µl RNase A (10mg/ml), followed by 5 minutes incubation at room temperature. 200 µl of a fresh solution containing 0.2N NaOH and 1% SDS (w/v) was added and chilled on iced for 5 minutes, followed by the addition of 150 µl of 3M NaOAC for 5 minutes. After centrifugation of the mixed solution at 9000 rpm for 5 minutes, the yielding supernatant was mixed with 1 volume of 7.5 M cold ammonium acetate. The plasmid DNA was precipitated by the addition of 0.8 volume of isopropanol, washed with 70% EtOH and resuspended in sterile ddH_2O.

Plasmid DNA prepared for sequencing was prepared using the Midi Prep kit (Qiagen) as recommended by the supplier.

C.1.2 *E. coli* bacterial transformation

5 µl of ligation mix was added to 100 µl of competent cells (*E. coli* XL-1 blue strain) and chilled on ice for 40 minutes. Following a heat shock at 42^0C for 1 minute, the cells were incubated on ice for 2 minutes. The cell mixture was transferred to 1 ml of LB medium and incubated at 37^0C for 1 hour with gentle shaking. Aliquots of cell culture were plated on petri dishes containing LB medium supplemented with the required antibiotic.

C.1.3. Restriction enzyme digestion and agarose gel electrophoresis

Restriction enzymes were used as recommended by the suppliers (New England Biolabs or MBI Fermentas). One unit of Restriction Enzyme was used per microgram of DNA and the reaction was allowed to proceed for 2 hours at the recommended temperature. The digested DNA was loaded on 1% agarose gel for electrophoresis. The electrophoresis was run at 70-85 volts and 1kb Mass Ruler^{TM} DNA ladder was loaded on the gel as molecular weight marker (MBI Fermentas).

DNA fragments for cloning were excised and isolated from agarose gels using the Qiagen II gel extraction kit (Qiagen).

C.1.4. Molecular cloning of *AtST5c*

The nucleotide sequence of *AtST5c* (F25I16.7) was identified through a BLAST search of the Arabidopsis Information Resource's (TAIR) database using *AtST5a* (F2P9.3) as a query sequence. The intronless coding sequence was amplified by PCR using an annealing temperature of 60^0C, an elongation time of 40 seconds and Taq DNA polymerase (New England Biolabs). The following oligonulceotides were used for amplification.

AtST5c-KpnI3' 5' ACGGGGTACCTCACATTCATGATTTTGAAGA3'

AtST5c-BamHI5' 5'ACGCGGATCCATGGAATCCAAAACCATAAA3'

The amplified DNA product was digested with the restriction enzymes KpnI and BamHI (New England and Biolab) and ligated into the corresponding sites of the bacterial expression vector pQE30 (Qiagen) using T4 DNA ligase (New England Biolabs) to produce the recombinant vector *pQE30-AtST5c.*

C.1.5. Cloning of the *AtST5a/5b* promoters (Pro5a and Pro5b)

The promoter sequences of *AtST5a* (F2P9.3) and *AtST5b* (F2P9.4) can be retrieved from the BAC clone F2P9 under the accession number AC016662. Their nucleotide sequences were amplified by PCR using the Taq DNA polymerase (New England Biolabs) and the following oligonucleotides:

Pro5a (5')-HindIII 5'ACCCAAGCTTAGTAGGGTTTCACAGACTATC3'

Pro5a (3')-Xbal 5'CTAGTCTAGACATTGGGGATGAGATC3'

Pro5b (5')-HindIII 5'ACCCAAGCTTTATTGTGGATATTTTTGTCA3'

Pro5b (3')-BamHI 5'ACGCGGATCCCATTGAAGATTAGAAAGAGAA3'

The sequences were amplified by PCR using an annealing temperature of 60^0C and an elongation time of 60 seconds. The PCR products were digested with the appropriate restriction enzymes and cloned into the corresponding sites of the *GUS* -containing vector pBI101 (Clontech).

C.1.6. Cloning of the *AtST5c* promoter (Pro5c)

The promoter sequence of *AtST5c* can be retrieved from chromosome1 (BAC F25I16) under the accession number AC026238. The nucleotide sequence of *Pro5c* was amplified by PCR using Taq polymerase and the following oligonucleotides as primers:

Pro5c (5')-HindIII 5'ACCAAGCTTAGGAACCCAAAGTACTTGTCATTACT3'

Pro5c (3')-BamHI 5'ACGGGATCCCATTATAGAGAGTGTTTTCTTTTTGGCT3'

The sequences were amplified by PCR using an annealing temperature of 60^0C and an elongation time of 40 seconds. The PCR product was digested with appropriate

restriction enzymes and cloned into the corresponding sites of the *GUS*-containing vector pBI 101 (Clontech).

C.1.7. Alignment of deduced sequences

The DNA sequence of *AtST5a, AtST5b* and *AtST5c* were translated with the Expert Protein Analysis System (ExPASy, http://www.expasy.ch/tools/dna.html). Multiple sequence alignments were performed with the ClustalW1.8 program (http://dot.imgen.bcm.tmc.edu:9331/mutialign/multi-align.htmL) provided by the Baylor College of Medicine. Genomic distribution, map locations and the number of *Arabidopsis ST* genes were determined with resources available at the TAIR site (http://www.arabidopsis.org).

C.2. Enzymology

C.2.1. Protein extraction and purification

An overnight pre-culture (2-3mL) from single fresh *E. coli* XLI Blue colony containing the recombinant vector pQE30-AtST5c was grown at 37^{0}C in LB supplemented with tetracycline (100μg/mL) and ampiciline (50 μg/mL). The preculture was used to inoculate a 200ml culture under the same conditions as described above. After 4 hours of incubation at 37 ^{0}C, the cultures were supplemented with isopropyl β-D-thiogalactopyranoside (IPTG) at a final concentration of 1mM and grown overnight at room temperature. The cells were centrifuged at 5000 rpm for 15 minutes, and the pellet resuspended in 10 mL of 50 mM phosphate buffer pH 8.0 (buffer A) containing 0.3 M NaCl and 14 mM β-mercaptoethanol. The resuspended cells were lysed by sonication.

The cell lysate was centrifuged at 10000 rpm for 15 minutes. The supernatant was applied to a nickel-nitrilotriacetic acid (Ni-NTA) agarose column (Qiagen) preequilibriated with buffer A. The column was washed with phosphate buffer pH 6.0 (buffer B) containing 0.3 M NaCl and 14 mM β-mercaptoethanol. The enzyme was eluted from the column with 500 mM imidazole in buffer B. The purity and solubility of the recombinant protein was tested using SDS –PAGE.

C.2.2. Glucosinolate extraction

To obtain samples of 3-indolylmethyl and 4-benzyloxylbutyl DSG, seeds of *Arabdopsis thaliana* (L.) Heynh ecotype Columbia (Col-0) were ground in liquid nitrogen and extracted in boiling 80% ethanol for 10 minutes. The extract was filtered under vacuum with a Whatman no.1 filter paper, followed by filtration on a 0.45 μm nylon mesh. The ethanol was evaporated by flash evaporation at 40 or 45^0C. The remaining water fraction was extracted with four volumes of chloroform and the aqueous phase was applied to a DEAE-Sephadex column pre-equilibrated with HPLC grade water. A sulphatase solution (2%, w/v) from Helix pomatia was added to the column for desulfation of the GIs. The DSG mixture was eluted with water and concentrated using a Speed-Vac. DSG were purified on a Waters Nova-Pak C 18 column (3.9 x150 mm, 4 μm particle size) using a Waters 625 LC HPLC system equipped with a Waters 486 tuneable absorbance detector as described previously (Hogge et al, 1988). The final identification of the purified compounds was carried out by mass spectrometry using a Quattro II triple quadrupole (Micromass, Manchester, UK) equipped with an atmospheric source and a nanoflow probe. The MS/MS fragmentation

pattern of DSGs was characterized by a fragment at m/z 145 and the loss of a neutral of 162 Da (glucose moiety). Allyl GI (sinigrin) was purchased from sigma (S-0875). 3-Methylsulfonylpropyl-, 3-Methylsulfinylpropyl-, 4-methylthiobutyl, 2-hydroxy 3-butenyl-, benzyl- and 2-phenylethyl-DSG were kindly provided by Dr. D. W. Reed (Plant Biotechnology Insistitute, Saskatoon, Canada). The DSGs were prepared from their GIs as described above. The concentration of individual compounds was estimated from absorbance at 227 nm using sinigrin as reference compound.

C.2.3. Protein determination, DSG-ST assays and identification of products

Protein concentration was determined by the method of Bradford using the Bio-Rad dye reagent and bovine serum albumin as standard protein (Bradford, 1976). Enzyme assays were performed in a 50 µl reaction mixture, containing 5µl of 10.0 µM [^{35} S] PAPS (New England nuclear), 5µl of DSG at various concentrations and 40 µl of enzyme diluted in Tris-HCl pH 8.0. The optimal enzyme concentrations were determined using 5 µl of 10.0 µM Benzyl-DSG as acceptor substrate. Following a 10-minute incubation, 2.5% acetic acid and 250µl of water saturated n-butanol was added to extract the product of the reaction. 100 µl of the butanol fraction was counted for radioactivity using liquid scintillation counting. For kinetic analysis, substrate concentrations were selected according to the apparent Km values obtained in preliminary experiments. Reaction mixtures were incubated for 10 minutes at room temperature and were stopped by freezing at –80 ^{0}C. 10 µl of the assay mixture was chromatographed on cellulose TLC plates (Alltech) using n-butanol-acetic acid-water (12:3:5) as solvent system. TLC plates were autoradiographed and the labelled products

were cut from the plates and counted for radioactivity using liquid scintillation counting. Km values for PAPS were measured with benzyl-DSG as substrate and the kinetic parameters were calculated from double-reciprocal Lineweaver-Burk plots.

C.2.4. Western blot analysis

Protein extracts were separated by SDS-PAGE on a 12% polyacrylamide gel and transferred to nitrocellulose. AtST5c was detected using anti-AtST2 (*AtST2* Genbank accession No.T43254) polyclonal antibodies (dilution 1:1000). Goat Anti-Rabbit IgG (H+L)- AP (alkaline phosphatase) conjugate (dilution 1:1000; Bio-Rad) was used to detect the primary antibodies. The AP conjugate substrate kit (Bio-Rad) was used for colour detection. Anti-AtST2 antibodies were raised in rabbits using purified recombinant AtST2 expressed in *E. coli* and were found to cross-react with several recombinant *A.thaliana* STs including AtST5c.

C.3. Histochemical localization of DSG-STs

C.3.1. *Agarobacterium tumefaciens* transformation

1μg of recombinant plasmid DNA was added to 100μl of competent *Agarobacterium tumefaciens* (Strain GV3101pmp 90) cells. Following a freezing cycle in liquid N2, the cell mixture was immediately incubated at 37^0C for 5 minutes. 1 ml of LB or YT was added and cells were grown at 28^0C for 2 hours with shaking. The concentrated cell culture was plated on petri dishes with AB medium supplemented with kanamycin and gentimycin.

C.3.2. Vaccum infiltration of *Arabidopsis thaliana* ecotype Col-0 and selection of transformants

Arabidopsis thaliana (Col-0) inflorescences were allowed to grow to approximately 5 cm and were decapitated at their base (Benchtold et al. 1993). After 4-6 days of growth, the plants were carefully removed from soil and their secondary inflorescences were submerged in an infiltration medium (2.15g/l MS salts, 1x B5 vitamins, 50g/l sucrose in 0.5g/l MES, pH 5.7) containing the Agrobacterium strain, GV3101pmp 90 transformed with the proper recombinant vector (pBI101-p5a, pRD400-p5b and pBI101-p5c) for 1 minutes under 25 inches HG of vacuum pressure. The infiltrated plants were potted in newly prepared soil and covered with plastic wrap for 2 days under optimal growth conditions. The seeds were collected from dry plants, sterilized and stratified at 4^0C for 4 days. The sterile seeds were plated on AB media (containing 3% sucrose, 1x vitamin B, 0.43% MS salts, 0.7% phytoagar and 50µg/ml kanamycin at pH 5.7) for 2 weeks, to select for transformants. Resistant individuals were transferred to newly prepared soil and grown under optimal condition until seed set (T2 seeds).

C.3.3. Histochemical assays

Sterile T1 seeds of each transformants were sowed in AB medium supplemented with kanamycin and grown at 22^0C in a growth chamber to reselect for transformants. 7 day-old seedlings, leaves, cauline leaves, primary flowers, secondary flowers and stem from several transformants were used to perform the histochemical assays. The plant tissues were transferred to 0.5-1ml 1.5mM X-GLUC (5-bromo-4-chloro-3-indolyl glucuronide) containing 50 mM NaPO4 pH 7.0 and incubated for one hour to overnight

at 37^{0}C. Following staining treatment, the plant tissues were rinsed with 70% ethanol to remove the chlorophyll and improve the contrast. GUS-stained tissues were examined under a Nikon dissecting microscope equipped with a digital camera.

C.4. Analysis of knock out *AtST5a* and *AtST5c* plants

C.4.1. Selection of homozygous knockout plants by PCR

The knockout *AtST5a* and *AtST5c* seed stocks were searched from the SIGnAL website (http://signal.salk.edu/tdna_protocols.html) using At1g74100 for *AtST5a* and At1g18590 for *AtST5c* as the query gene names, and their T3 generation seeds were ordered from Salk Institute. The homozygous knockout lines of *AtST5a* were selected by two sets of PCR. PCR1 used the following primers:

AtST5a-Sac1: 5'AGGCGAGCTCATGGAATCAAAGACAACCCA3'

AtST5a-xbal: 5'CTAGTCTAGATCAGTTATCATGTTGAAGCAAG3'

and genomic DNA extracted from putative knockout plants to amplify the coding sequence of the *AtST5a* gene with an elongation time of 40 seconds and an annealing temperature of 62^{0}C

PCR 2 used the following primers:

AtST5a-Xbal: 5'CTAGTCTAGATCAGTTATCATGTTGAAGCAAG3'

LBa1: 5' TGGTTCACGTAGTGGGCCATCG3'

to amplify a DNA fragment between 5' of *AtST5a* and the left border of the T-DNA with an elongation time of 40 seconds and an annealing temperature of 60 ^{0}C.

The homozygous lines of AtST5c were also selected by two sets of PCR.

PCR 1 used the following gene specific primers:

AtST5c-Kpn (3'): 5'ACGGGGTACCTCACATTCATGATTTTGAAGA3'

AtST5c-BamHI (5'): 5'ACGCGGATCCATGGAATCCAAAACCATAAA3'

and genomic DNAs extracted from putative knockout plants to amplify the coding region of AtST5c with an elongation time of 40 seconds and an annealing temperature of 60^0C.

PCR 2 used the following primers:

AtST5c-BamHI (5'): 5'ACGCGGATCCATGGAATCCAAAACCATAAA3'

LBa1: 5'TGGTTCACGTAGTGGGCCATCG3'

to amplify a DNA fragment between 5' of *AtST5c* and the left border of the T-DNA with an elongation time of 40 seconds and an annealing temperature of 60^0C.

C.4.2. Genomic DNA extraction

T3 seeds of putative *AtST5a* and *AtST5c* knock out plants, ordered from the Salk institute, were sowed in soil and grown for 20 days under optimal conditions. 1-2 leaves were cut from each plant and ground in liquid N2. The cold plant powder was immediately mixed with 500 µl of CTAB buffer (containing 1% PVP40, 100mM Tris.HCl ph9.0, 20 mM EDTA, 1.4 M NaCl, 2%CTAB and 0.2% ß-mercaptoethenol) preheated at 65^0C and incubated at 65^0C for 2x 20 minutes. The solution was extracted with 1 volume of phenol: chloroform (1:1) and then re-extracted with 1 volume of chloroform. The supernatant was precipitated with 1/10 volume of 3M of NaoAc pH 5.2 and 0.6 volume of isopropanol. The DNA pellet was dissolved in water.

C.4.3. DSG extraction and analysis of DSG variation by HPLC

100 plants for Col-0 and knockout mutants were grown for 28 days. DSGs extracts were prepared from 10 individual Col-0 plants and 10 individual knockout plants. The DSG extraction procedure was described in section C.2.3. Extracted DSGs from each sample were injected in a Waters 625 LC HPLC system equipped with a Water 486 tunable absorbance detector as described previously (Hogge et al, 1988) GI profiles of the mutant plants were compared with those of wild type col-0 plants. Students T-test (ttp://www.bio.miami.edu/rob/perls_t) was used to evaluate the significance of the variations in DSG content.

C.5. Studies on gene expression pattern following treatments with signal molecules

C.5.1. Signal molecules induction of Col-0 plants

Seeds of *Arabidopsis thaliana* (L.) Heyne ecotype Columbia (Col-0) were surface sterilized, plated on germination medium and grown under long day conditions (16hr light, 22^0C: 8 hr dark 20^0C) for 12 days. For auxin response analysis, 12-day-old seedlings were treated with increasing concentration (0, 0.1 μm, 1 μm, 2.5 μm and 5 μm, 10 μm) of IAA and harvested after 8 hours; or treated with a final concentration of 5 μm IAA and harvested after different incubation times (0, 2, 4, and 8 hours). For methyl-jasmonate and salicylic acid treatments, 12-day-old seedlings were grown in the present of 10 μm, 50 μm and100 μm MeJA or SA and harvested after 8 hours. Alternatively, plants were treated with a final concentration of 100 μm of MeJA or SA and harvested after different incubation times (0, 2, 4 and 8 hours).

C. 5.2. RNA extraction

Total RNA was extracted by the following procedure. The whole plants were ground to powder with liquid N_2 and transferred to a solution containing the extraction buffer (200mM Tris-HCl pH 9.0, 400mM NaCl, 20mM MgOAC and 500mM sucrose, 5mM DTT), 1/10 volume of 10% SDS, 1/20 volume of 400mM EDTA and mixed vigourously. One volume of phenol-chloroform-isoamyalcool (50:48:2) was added to the mixture shaked vigorously for 20 minutes, followed by centrifugation at 4000 rpm for 15 mimutes. The aqueous phase was extracted with phenol-chloroform-isoamyalcohol one more time. Finally RNA was precipitated with LiCl, resuspended in water and precipitated with NaOAC (3M pH 5.2) and 95% ETOH. The RNA pellet was dissolved in 0.1% diethyl pyrocarbonate (DEPC) water.

C.5.3. Reverse-transcription PCR analysis of *AtST5a, AtST5b* and *AtST5c* expression

The synthesis of cDNA was performed according to the recommendations of New England Biolabs. To produce reverse transcription (RT) product, three steps are required. Step 1: 2.0 µg total RNA was treated with 2.0 units of DNAse I (Roche Molecular Biochemicals Molecular Biochemicals) in 0.5x Expand RT buffer (Roche Molecular Biochemicals Molecular Biochemicals) for 15 minutes at room temperature. DNAse I was inactivated at 65^0C for 10 minutes. The RNA solution was put on ice for 2 minutes. Step 2: 100 pmoles OligodT was added to 1µg of DNAse-treated RNA solution. The mixture was incubated at 65^0C for 10 minutes and put on ice for 2 min for Oligo dT annealing. A reaction cocktail was added to a final concentration of 1x Expand RT Buffer, 10 mM of DTT, 1mM dNTP, 25U RNAse inhibitor and 62.5U Expand RT

(ROCHE). The reaction was allowed to proceed for 60 min at 43^0C. Step 3: PCR reaction: PCR analysis was performed using the gene specific primers, Ex-Taq DNA polymerase (TaKaRa), and 2μL of reverse-transcription (RT) products. The gene specific primers for *AtST5a* are:

AtST9 (527): 5'-CTTTCTTACACAAGGAGAAGTCTCAAGA-3'

OBO-3' (SamI): 5'-TCCTCCCCCGGGTCAGTTATCATGTTGAAGCAAG-3'

The gene specific primers for *AtST5b* are:

AtST-12 (568): 5'-CTTCACAAGGAAAGGACAGAGCT-3'

AtST-12 kpn: 5'-ACGGGGTACCTCATTTTACCATGTTCAAGCA-3'

The gene specific primers for *AtST5c* are:

AtST5c (534): 5'-GTTTGTGTCGATGTGGACTTTTGCT-3'

AtST5c Kpn 3': 5'-ACGGGGTACCTCACATTCATGATTTTGAAGA-3'

Control PCR reactions were performed with 100ng of *A. thaliana* ecotype Col-0 genomic DNA using the actin primers (ACT1, GeneBank accession number U39449), 5'-GCTGATGGTGAAGACATTCAAC-3' and 5'-CATACATAGCAGGGGCATAAG-3'. The primers were designed within regions flanking a 100 bp single intron of the actin gene and could be used to monitor genomic DNA contamination. Actin controls also confirmed that similar quantities of RT-products were present in each reaction.

D. RESULTS

D.1 Enzymology of the sulfonation reaction

D.1.1 Molecular characterization of *AtST5a/5b/5c*

Eighteen sulfotransferase-coding sequences were retrieved from the Arabidopsis Information Resource (TAIR) database using the flavonol 3-sulfotransferase (ST) of *Flaveria chloraefolia* as query sequence. The coding sequences of the putative STs were amplified by PCR from Col-0 genomic DNA and cloned in the bacterial expression vector pQE30. Out of the eighteen recombinant proteins, AtST5a (Figue 3) and AtST5b (Figure 4) were the only members to exhibit specificity for DSGs. AtST5c, another protein to exhibit specificity for DSGs, was identified through BLAST searches of TAIR database using *AtST5a* as a query sequence (Figure 5). *AtST5a*, *AtST5b* and *AtST5c* are presented on chromosome I. *AtST5a* (F2P9.3) spans nucleotides 7673 to 8689 of BAC clone; *AtST5b* (T2P9.4) spans nucleotides 9123 to 10175 of BAC clone and *AtST5c* (F25I16.7) spans nucleotides 21629 to 22669 of BAC clone. *AtST5a* and *AtST5b* are clustered on one arm of chromosome I with an intergenic distance of 434 nucleotides. *AtST5c* is located on the other arm of the chromosome I (Figure 6).

Amino acid alignment of AtST5a, AtST5b and AtST5c indicates that they share 70-95% of amino acid identity and 81-95% of amino acid similarity (Figure 7). All three deduced protein sequences contain the four regions that are conserved among all cytosolic STs (Marsolais and Varin, 1997). Conserved region I, located in the amino terminal portion with consensus sequence as PKSGTxW, and region IV, located in the carboxyl-terminal portion with consensus sequence as RKGxxGDWKxxFT are thought to be involved in the binding of the sulfate donor 3'-phosphoadenosine5'-phosphosulfate

(PAPS) and in catalysis (Marsolais and Varin, 1997). The variation observed among the three sequences is mostly clustered in the two subdomains known to be involved in ST acceptor substrate recognition (Varin et al, 1995). The three deduced proteins have 340 to 350 amino acids and predicted molecular weights of around 40KDa. The AtST5a, AtST5b and AtST5c proteins share 45% amino acid sequence identities with the flavonol STs of *Flaveria* species, 40% identities with the other *A.thaliana* STs and approximately 25% identities with the mammalian cytosolic STs.

D.1.2. Molecular cloning and protein expression of *AtST5c*

The availability of the *AtST5c* coding sequence allowed us to amplify the DNA by PCR from Col-0 genomic DNA for cloning in the bacterial expression vector pQE30. The amplified PCR product was verified on ethidium bromide-staining agarose gel (Figure 8A). One band of approximately 1 kb was observed on the gel. This result is consistent with the predicted size of the product. The amplified DNA and the pQE30 plasmid were digested with KpnI and BamHI. The *AtST5c* coding sequence was then ligated in pQE30. The recombinant DNA was transformed in *E. coli* XLI blue strain and plated on petri dishes containing LB medium supplemented with ampicilin. Several colonies were screened and the plasmids were digested with BamHI and KpnI.

```
1       atggaatcaaagacaacccaaaacggatccgaagtcgtcgaactcacagagttcgagaaa
        M  E  S  K  T  T  Q  N  G  S  E  V  V  E  L  T  E  F  E  K
61      acccagaagaagtatcaagatttcatcgctacacttccaaagagcaaaggctggagacca
        T  Q  K  K  Y  Q  D  F  I  A  T  L  P  K  S  K  G  W  R  P
121     gatgagatcttaacccaatacggtggacactggtggcaagaatgtctcctcgaaggtctt
        D  E  I  L  T  Q  Y  G  G  H  W  W  Q  E  C  L  L  E  G  L
181     tttcacgctaaagaccatttcgaagcacgaccaactgatttcctcgtctgtagctaccca
        F  H  A  K  D  H  F  E  A  R  P  T  D  F  L  V  C  S  Y  P
241     aaaaccggtacaacttggctcaaagcactaacttacgcaatcgtcaatcgttctcgttac
        K  T  G  T  T  W  L  K  A  L  T  Y  A  I  V  N  R  S  R  Y
301     gacgacgccgcaaacccactcctcaaacgaaaccctcacgagtttgtcccttacgttgag
        D  D  A  A  N  P  L  L  K  R  N  P  H  E  F  V  P  Y  V  E
361     atcgacttcgcgttttaccccaccgttgatgttcttcaagacagaaagaacccacttttc
        I  D  F  A  F  Y  P  T  V  D  V  L  Q  D  R  K  N  P  L  F
421     tctactcatatcccaaacgggttattacccgattcgattgtgaactctggttgtaagatg
        S  T  H  I  P  N  G  L  L  P  D  S  I  V  N  S  G  C  K  M
481     gtgtacatatggagagacccgaaagatactttcatctccatgtggactttcttacacaag
        V  Y  I  W  R  D  P  K  D  T  F  I  S  M  W  T  F  L  H  K
541     gagaagtctcaagaaggtcaattagcgagtcttgaggacagctttgatatgttttgtaaa
        E  K  S  Q  E  G  Q  L  A  S  L  E  D  S  F  D  M  F  C  K
601     ggtttatctgtgtatggtccttatctggatcatgttttgggttattggaaagcttaccaa
        G  L  S  V  Y  G  P  Y  L  D  H  V  L  G  Y  W  K  A  Y  Q
661     gagaatccagataggattttgttccttaggtacgagaccatgagggccaatcctttgcct
        E  N  P  D  R  I  L  F  L  R  Y  E  T  M  R  A  N  P  L  P
721     tttgtgaagagattggctgagttcatgggttatggattcactgatgaggaagaggagaat
        F  V  K  R  L  A  E  F  M  G  Y  G  F  T  D  E  E  E  E  N
781     ggtgttgctgagaaagtggtgaagctttgtagctttgagacgttgaagaatcttgaagct
        G  V  A  E  K  V  V  K  L  C  S  F  E  T  L  K  N  L  E  A
841     aacaaaggtgataaagaaagagaggatcgtcctgctgtttatgcgaatagcgcgtatttt
        N  K  G  D  K  E  R  E  D  R  P  A  V  Y  A  N  S  A  Y  F
901     aggaaaggaaaggttggagattgggctaattatttgactcctgagatggctgctcgtatt
        R  K  G  K  V  G  D  W  A  N  Y  L  T  P  E  M  A  A  R  I
961     gatggcttagtggaggagaaattcaaagatactggcttgcttcaacatgataactga
        D  G  L  V  E  E  K  F  K  D  T  G  L  L  Q  H  D  N  -
```

Figure 3. Nucleotide and deduced amino acid sequence of AtST5a. Start and stop codons are bolded

```
1       atggaatcagaaaccctaaccgccaaggctacgatcacgaccacgaccctaccgagtcac
        M  E  S  E  T  L  T  A  K  A  T  I  T  T  T  T  L  P  S  H
61      gacgagaccaagacagaatcaacagagttcgagaaaaatcaaaaacggtatcaagacctc
        D  E  T  K  T  E  S  T  E  F  E  K  N  Q  K  R  Y  Q  D  L
121     atctccacgtttcctcacgagaaaggctggagaccgaaagagcccctgatcgagtatggt
        I  S  T  F  P  H  E  K  G  W  R  P  K  E  P  L  I  E  Y  G
181     ggttactggtggctaccgtctctcctcgaaggttgtattcacgcgcaagagttctttcaa
        G  Y  W  W  L  P  S  L  L  E  G  C  I  H  A  Q  E  F  F  Q
241     gcacgacccagtgacttcctcgtctgtagctacccaaagacaggcaccacttggctcaaa
        A  R  P  S  D  F  L  V  C  S  Y  P  K  T  G  T  T  W  L  K
301     gccctgactttcgccatcgcaaatcgttcccgcttcgatgattcctccaaccctctcctg
        A  L  T  F  A  I  A  N  R  S  R  F  D  D  S  S  N  P  L  L
361     aaacgtaaccctcacgagtttgttccttacattgagatagatttccctttcttccctgaa
        K  R  N  P  H  E  F  V  P  Y  I  E  I  D  F  P  F  F  P  E
421     gttgatgttctcaaagacaaagggaacactctgttttcgactcatatcccatacgagtta
        V  D  V  L  K  D  K  G  N  T  L  F  S  T  H  I  P  Y  E  L
481     ttacctgattcggttgtgaaatccggttgtaagatggtttacatatggagagaaccaaag
        L  P  D  S  V  V  K  S  G  C  K  M  V  Y  I  W  R  E  P  K
541     gacactttcatctccatgtggactttccttcacaaggaaaggacagagcttggacctgtc
        D  T  F  I  S  M  W  T  F  L  H  K  E  R  T  E  L  G  P  V
601     agcaatcttgaggagtcttttgatatgttctgtcgtggtctgtctgggtatggtccttat
        S  N  L  E  E  S  F  D  M  F  C  R  G  L  S  G  Y  G  P  Y
661     cttaatcatatcctggcgtattggaaagcataccaagagaatccagataggatcttgttc
        L  N  H  I  L  A  Y  W  K  A  Y  Q  E  N  P  D  R  I  L  F
721     ctcaagtacgagacgatgagagctgatcctttaccgtacgtgaagagtctggctgagttt
        L  K  Y  E  T  M  R  A  D  P  L  P  Y  V  K  S  L  A  E  F
781     atgggtcatggattcacagccgaggaagaggagaaaggtgttgttgagaaagtggtgaat
        M  G  H  G  F  T  A  E  E  E  E  K  G  V  V  E  K  V  V  N
841     ctttgcagcttcgagacgttgaagaatcttgaagctaacaaaggggagaaagacagagag
        L  C  S  F  E  T  L  K  N  L  E  A  N  K  G  E  K  D  R  E
901     gatcgtcctggtgtttacgcgaatagcgcgtatttcaggaaaggaaaggtgggagattgg
        D  R  P  G  V  Y  A  N  S  A  Y  F  R  K  G  K  V  G  D  W
961     tcgaactatctgactccggagatggctgctcgtatagatgggttaatggaagagaaattt
        S  N  Y  L  T  P  E  M  A  A  R  I  D  G  L  M  E  E  K  F
1021    aagggcaccggcttgcttgaacatggtaaatga
        K  G  T  G  L  L  E  H  G  K  -
```

Figure 4. Nucleotide and deduced amino acid sequence of AtST5b. Start and stop codons are bolded.

```
1       atggaatccaaaaccataaacgacgtcgtcgtatcagagtcaaatcacgagttagcatct
        M  E  S  K  T  I  N  D  V  V  V  S  E  S  N  H  E  L  A  S
61      tcgtcgccatcagaatttgagaagaaccagaaacattaccaagaaatcatcgcaactctt
        S  S  P  S  E  F  E  K  N  Q  K  H  Y  Q  E  I  I  A  T  L
121     cctcacaaagatggctggagaccaaaagatccgttcgttgagtacggtggtcactggtgg
        P  H  K  D  G  W  R  P  K  D  P  F  V  E  Y  G  G  H  W  W
181     ctacaacctcttcttgaaggtttacttcacgcccagaagttcttcaaggcacgtcccaat
        L  Q  P  L  L  E  G  L  L  H  A  Q  K  F  F  K  A  R  P  N
241     gatttcttcgtctgcagctacccaaaaaccggcacgacttggctcaaagccttaactttc
        D  F  F  V  C  S  Y  P  K  T  G  T  T  W  L  K  A  L  T  F
301     gcaatcgcaaatcgctccaagttcgacgtttcaacaaaccctcttctcaaacgtaatcca
        A  I  A  N  R  S  K  F  D  V  S  T  N  P  L  L  K  R  N  P
361     cacgagtttgtgccttacatcgaaatcgacttcccgtttttcccaagcgttgatgttctt
        H  E  F  V  P  Y  I  E  I  D  F  P  F  F  P  S  V  D  V  L
421     aaagacgaaggaaacacgcttttctcgactcatatcccttatgatctcttacctgaatca
        K  D  E  G  N  T  L  F  S  T  H  I  P  Y  D  L  L  P  E  S
481     gttgtgaaatctggttgcaagattgtttacatatggagagacccaaaggacacgtttgtg
        V  V  K  S  G  C  K  I  V  Y  I  W  R  D  P  K  D  T  F  V
541     tcgatgtggacttttgctcacaaggagagatcacaacaaggaccagtggttagcattgag
        S  M  W  T  F  A  H  K  E  R  S  Q  Q  G  P  V  V  S  I  E
601     gaagcttttgataagtactgtcaaggtttatcagcttacggtccttatcttgatcatgtt
        E  A  F  D  K  Y  C  Q  G  L  S  A  Y  G  P  Y  L  D  H  V
661     ttagggtattggaaagcttaccaagcaaacccggatcagattttgtttctcaagtatgag
        L  G  Y  W  K  A  Y  Q  A  N  P  D  Q  I  L  F  L  K  Y  E
721     acaatgagagctgatccattgccttatgtgaagagattagctgagtttatgggttatggg
        T  M  R  A  D  P  L  P  Y  V  K  R  L  A  E  F  M  G  Y  G
781     ttcacgaaggaggaagaggagggaaatgttgttgagaaagtagtgaagctttgtagcttt
        F  T  K  E  E  E  E  G  N  V  V  E  K  V  V  K  L  C  S  F
841     gagactttgaagaatcttgaagctaacaaaggtgagaaagatagagaagatcgtcctgct
        E  T  L  K  N  L  E  A  N  K  G  E  K  D  R  E  D  R  P  A
901     gtttatgccaacagtgcgtatttcaggaaagggaaagttggggattggcagaattatctg
        V  Y  A  N  S  A  Y  F  R  K  G  K  V  G  D  W  Q  N  Y  L
961     actccggagatggtggctcgtattgatggtttgatggaagagaagtttaaaggaactggc
        T  P  E  M  V  A  R  I  D  G  L  M  E  E  K  F  K  G  T  G
1021    tttctttcttcaaaatcatga
        F  L  S  S  K  S  -
```

Figure 5. Nucleotide and deduced amino acid sequence of AtST5c. Start and stop codons are bolded.

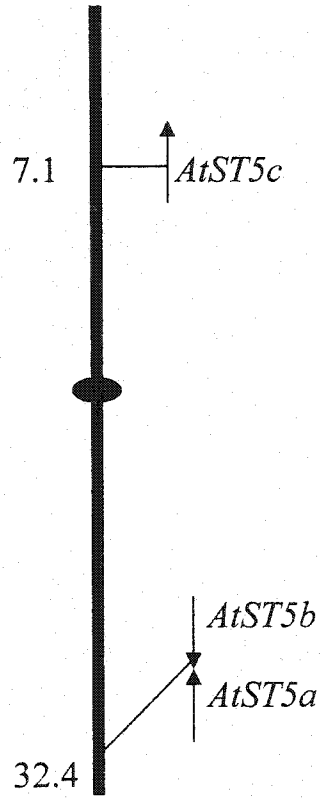


Figure 6. Schematic representation of *Arabidopsis thaliana* chromosome I (the centromere is depicted by a black dot). Numbers indicate the approximate chromosomal position of the three genes *AtST5a, AtST5b* and *AtSt5c* in millions of base pairs. The arrows indicate the orientation of the three genes.

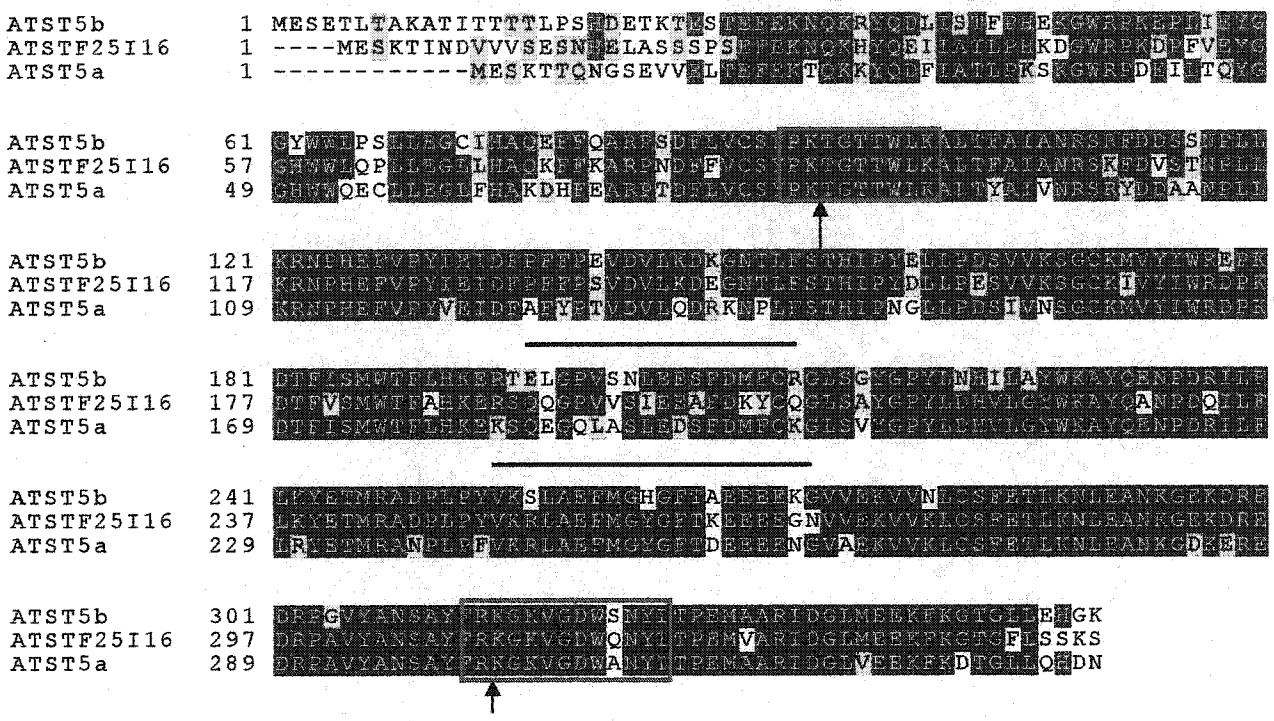


Figure 7. AtST5a, AtST5b and AtST5c amino acid alignment.
Deduced amino acid sequences of AtST5a, AtST5b and AtST5c were aligned using the CLUSTAL W 1.8 program (multiple sequence alignments, Baylor College of Medicine search Launcher, Houston, TX). Identical amino acid residues are boxed in black and conservative changes are boxed in gray. The conserved region I and IV are boxed in red. Residues critical for catalysis and PAPS binding are indicated by arrows. The domain II involved in the determination of substrate specificity is underlined

The resulting digests were size fractionated on an ethidium bromide-stained agarose gel and exposed to UV light. A plasmid containing the insert was identified (Figure 8B). The purified recombinant plasmid was partially sequenced at the Genomic Centre of Laval University. The sequence results demonstrated that the DNA inserted in the plasmid was in fact *AtST5c*. Recombinant AtST5c was expressed in *E. coli* XL1 blue and purified by Ni-agarose affinity chromatography. The high purity and solubility of the recombinant protein was confirmed by SDS-PAGE (Figure 9A). Western blot analysis of the lysate from the AtST5c transformed *E. coli* strain, utilizing anti-AtST2 polyclonal antibodies as probe, confirmed the presence of the ST protein (Figure 9B). Antibody reacted with a protein band from the lysate with an estimated molecular mass of approximately 47 KDa, consistent with the predicted molecular mass of the recombinant protein, which includes 12 additional amino acid residues contributed by the vector. An additional band with a molecular mass of approximately 35 kDa is also observed, which might be the result of proteolytic cleavage of the recombinant protein.

D.1.3. Substrate specificity of AtST5c

The purified AtST5c recombinant protein was tested for enzymatic activity using DSGs as well as a wide range of substrates, including steroids, flavonoids and phenolic acids. AtST5c was found to exhibit strict specificity for DSGs. It did not accept structurally related oxime precursors, such as 3-indolyacetaldoxime and 5-hydroxypenanaldoxime, indicating a strict requirement for the

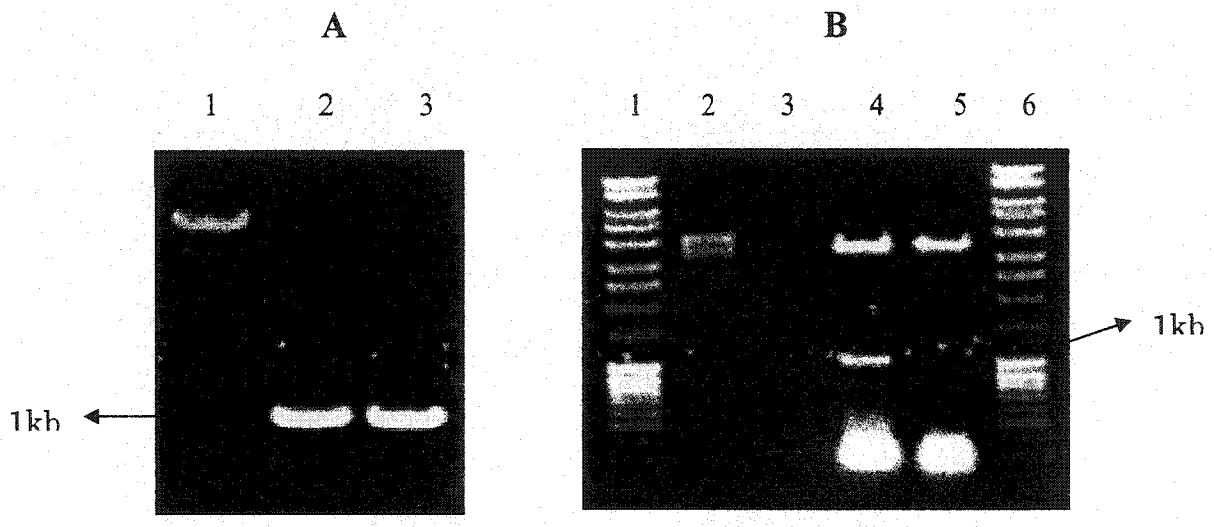


Figure 8A) PCR amplified DNA products of *AtST5c*. PCR experiment was performed using a set of primers specific for *AtST5c* and Taq DNA polymerase. The amplified DNA product was verified on ethidium bromide-stained agarose gel. Lane 1 is Gene Ruler™ 1-kb DNA ladder, lane 2 and 3 are amplified DNA products. Amplified DNA is approximately 1 kb.

B) Screening for Recombinant pQE30. The PCR DNA product was cloned in pQE30 and transformed in *E. coli* and plated on petri dishes containing LB medium supplemented with ampicilin. Several colonies were screened and purified using the Qiagen miniprep kit. The purified plasmids were verified on ethidium bromide-stained agarose gels after digestion with KpnI and BamHI. Lane1 and 6 are Gene Ruler™ 1-kb DNA ladder, lane 2 is undigested pQE30, lane 3 is pQE30 digested with KpnI and BamHI as a negative control, lane 4 is the plasmid containing insert, lane 5 is a plasmid without insert.

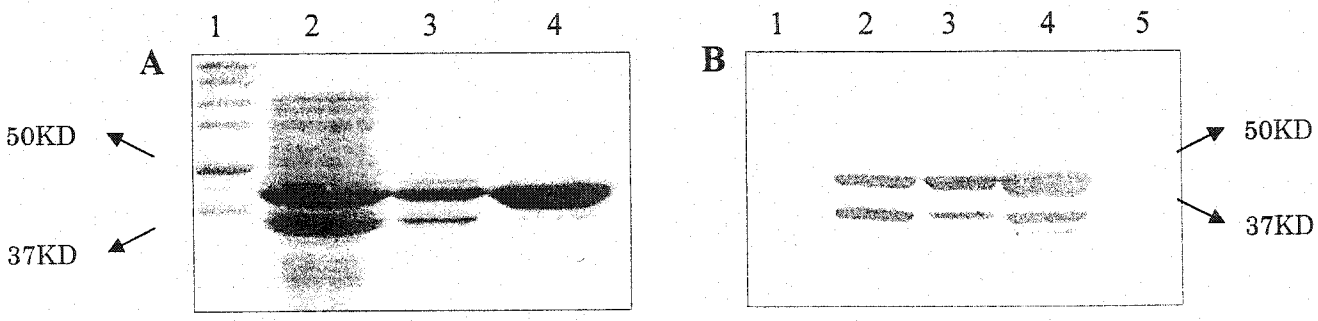


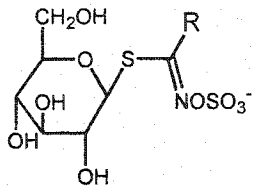
Figure 9A) Identification of the solubility and purity of recombinant AtST5c by SDS-PAGE. The expression of AtST5c was induced by IPTG and purified by Ni-agarose affinity chromatography. The solubility and purity of the recombinant protein was confirmed by SDS-PAGE. Lane 1 is protein standard (Bio-Rad); Lane 2 is total *E. coli* protein extraction. Lane 3 is cell lysate following centrifugation. Lane 4 is the nickel agarose purified recombinant protein.

B) Western blot of the target protein.
Proteins separated by SDS-PAGE (See A) were transferred to a nitrocellulose membrane. AtST5c was detected using anti-AtST2 (Genbank accessionNo. T43254) polyclonal antibodies (dilution 1:1000). Goat anti-rabbit secondary antibodies conjugated to alkaline phosphatase were used to detect the primary antibodies. One band estimated at about 47kDa was observed in all three extracts, confirming the presence of a sulfotransferase (ST) protein. The presence of another band less than 37 kDa may result from the degradation of the recombinant protein. Lane 1) Prestained-protein standard (Bio-Rad). Land 2) Total *E. coli* protein extract. Lane 3) Cell lysate following centrifugation. Lane 4) The nickel agarose purified recombinant protein. Lane 5) Protein extract from pQE30 in XL1-blue strain used as negative control.

presence of the thioglucose moiety. DSGs with different side chain structures were tested to determine the substrate preference of AtST5c (Figure 8). Two final concentrations of substrates (1μM and 5μM) were used to determine the substrate preference.

The results showed that AtST5c accepts DSGs with different side chains including those that are not naturally occurring in *A. thaliana*, such as benzyl- and methylsulfonylpropyl DSGs (Table 1). AtST5c accepts all substrates tested at 5μM, with preference for benzyl- > 4-methylthiobutyl > 3-indolylmethyl- DSG. At 1μM substrate concentration, AtST5c exhibits preference for benzyl->2-phenylethyl-> 4-methylthiobutyl DSGs and low activity for both ally and 2-hydroxyl 3-butenyl DSG. When considering only the naturally occurring Arabidopsis GIs, 4-methylthiobutyl- and 2-phenylethyl DSGs are the preferred substrates among the aliphatic and aromatic ones respectively.

Kinetic analyses were performed with the desulfated form of the naturally occurring GIs from *A.thaliana* and the non-occurring benzyl-DSG (Table II). The Km and Vmax values of the enzyme for the different substrates were obtained from double-reciprocal Lineweaver-Burk plots. Enzyme kinetic studies reveal that AtST5c has high affinity for all the substrates tested with the exception of indolymethyl-DSG. Therefore further experimental studies will be required to determine the real affinity of the enzyme for indolymethyl-DSG. The results reveal that AtST5c shows no apparent



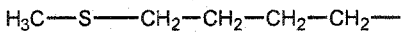
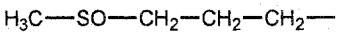
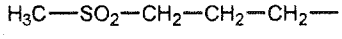
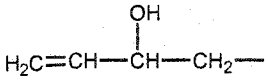
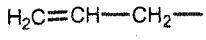
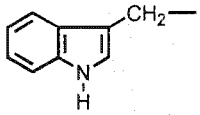
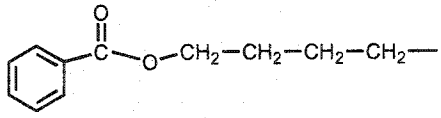
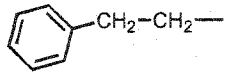
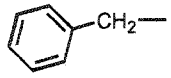
Aliphatic GIs	Structure of the R group
4-methylthiobutyl	
3-methylsulfinylpropyl	
3-methylsulfonylpropyl	
2-hydroxy 3-butenyl	
allyl	
Indole GI	
3-indolylmethyl	
Aromatic GIs	
4-benzoyloxybutyl	
2-phenylethyl	
benzyl	

Figure 10. The chemical structures of selected GI

Table 1. Substrate specificity of recombinant AtST5c

DS-GI	1μM AtST5c	5 μM AtST5c
4-methylthiobutyl*	84.3	95.4
3-methylsulfinylpropyl*	13.4	37.9
3-methysulfonylpropyl	9.3	25.2
2-hydroxyl 3-butenyl	5.2	23
Allyl	19.8	45
3-indolylmethyl*	31.2	86.4
4-benzoyloxybutyl*	11.7	23.8
2-phenylethyl*	99.6	68.2
Benzyl	100[a]	100[b]

* GIs detected in Arabidopsis (Hogge et al, 1988)
a. 34 pkatal/mg protein
b. 41 pkatal/mg protein

Table 2. Kinetic parameters of *AtST5c* for DSGs with different side chain structures

DS GI	Km (μM)	Vmax (pkatal/mg)	Vmax/Km (pkatal/mg. μM)
Benzyl	0.97	124	127.8
2-phenylethyl	0.62	35.8	57.7
4-methythiobutyl	1.30	130.9	100.7
PAPS	0.17		

difference of catalytic efficiency for the aromatic and aliphatic substrates. Although AtST5c exhibits similar affinity for three substrates, a 4-fold lower catalytic activity is observed with 2-phenylethyl DSG compared to 4-methylthiobutyl DSG. The apparent Km value for the sulfonate donor, PAPS, was found to be 0.17μM.

D.2. Studies on the expression of *AtST5a, 5b* and *5c*

D.2.1. Cloning of the promoters of *AtST5a, 5b* and *5c* (*pro5a, pro5b* and *pro5c*)

Due to the great similarity of *AtST5a, AtST5b* and *AtST5c*, it was not possible either to generate the specific DNA probes required for northern blot analysis and in situ hybridization studies or to synthesize the specific peptides needed for antibody production and subsequent western blot analysis. Therefore we performed promoter analysis for each gene in transgenic Arabidopsis plants. All three promoters were amplified by PCR using gene specific primers and fused with the β-glucuronidase (*GUS*) reporter gene. *Pro5a* corresponds to the entire intergenic region between gene *F2P9.2* and *F2P9.3* (*AtST5a*), whereas *Pro5b* spans a region of 2053 bp between the gene *F2P9.4 (AtST5b*) and an AT-rich region on *A.thaliana* chromosome I BAC clone F2P9 (Figure 11). *Pro5c* comprises the entire intergenic region (953 bp) between gene *F25I16.7 (AtST5c)* and *F25I16.6* on *A.thaliana* chromosome I BAC clone F25I16 (Figure 11). PCR amplified DNA products were analyzed on ethidium bromide-stained agarose gels and were found to have the predicted sizes (Figure 12). The recombinant vectors containing the expected inserts were screened for on agarose gel. The purified recombinant plasmids were partially sequenced at the Genomic Centre of

Laval University. The sequence results demonstrated the proper fusion of the promoters with the *GUS* reporter gene.

D.2.2. Expression patterns of the *AtST5a, 5b and 5c* genes

The GUS reporter gene constructs were transformed into Arabidopsis. The tissue-specific expression pattern of 10 independent lines was analyzed over three generations. Histochemical staining revealed variations in the amount of GUS activity in different transgenic lines of *Pro5a, Pro5b* and *Pro5c*. Such variability is common and probably reflects the positional effects of the insertion sites on transgene expression. Nevertheless, all lines exhibited the same qualitative staining patterns.

No difference in the staining pattern was observed between apical plant parts of transformants harbouring from the three GUS reporter gene constructs. Microscopic analyses of whole organs revealed promoter activities restricted in the vascular bundles of hypocotyls, leaves, stems, petioles, mature silique walls and seed funicles (Figure 13). Similar results were observed in studies on the gene expression pattern of *CYP79 F1* and *CYP79F2* (Reintanz, et al. 2001), both of which are involved in the oxidation of amino acids to aldoximes, the initial step in GI biosynthesis. CYP79F1 is thought to be restricted to use short-chain methionine derivatives as substrates, whereas CYP79F2 uses long-chain methionine derivatives as substrates.

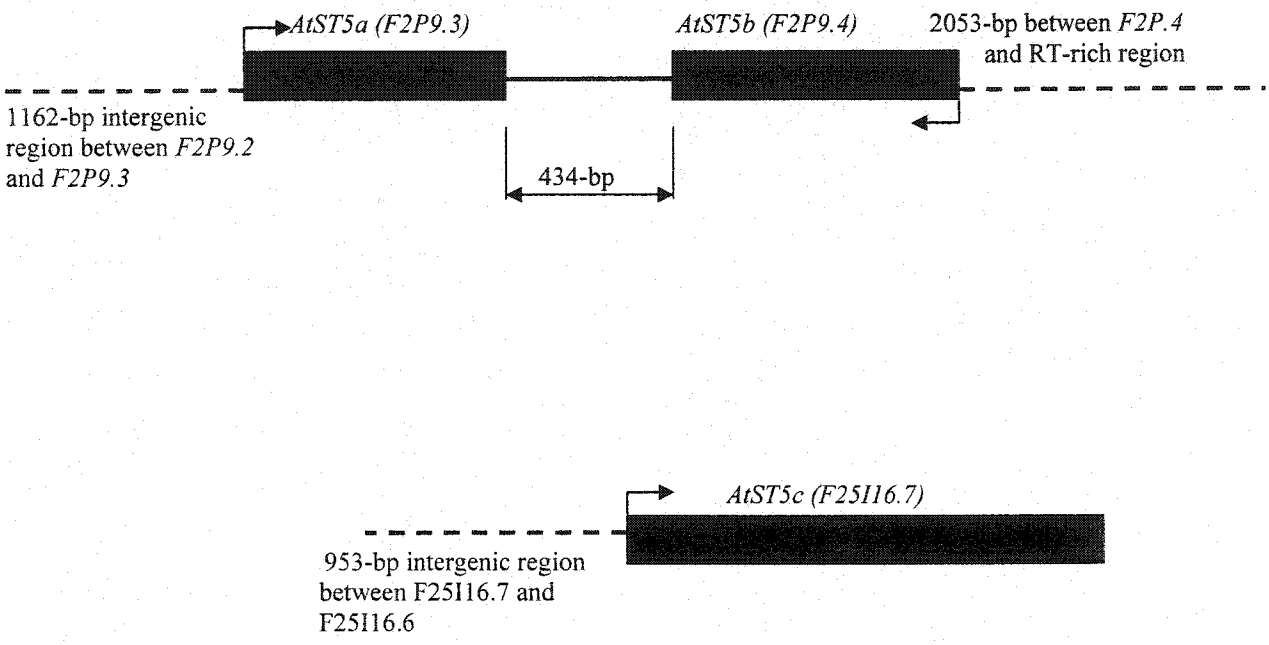


Figure 11. Genomic organization of *AtST5a, AtST5b* and *AtST5c*
The coding regions of all three genes contain one exon. Initiation sites of translations are indicated by arrows, and intergenic regions that have been used for GUS expression studies are indicated by dashed lines. Intergenic region between AtST5a and AtST5b is 434-bp.

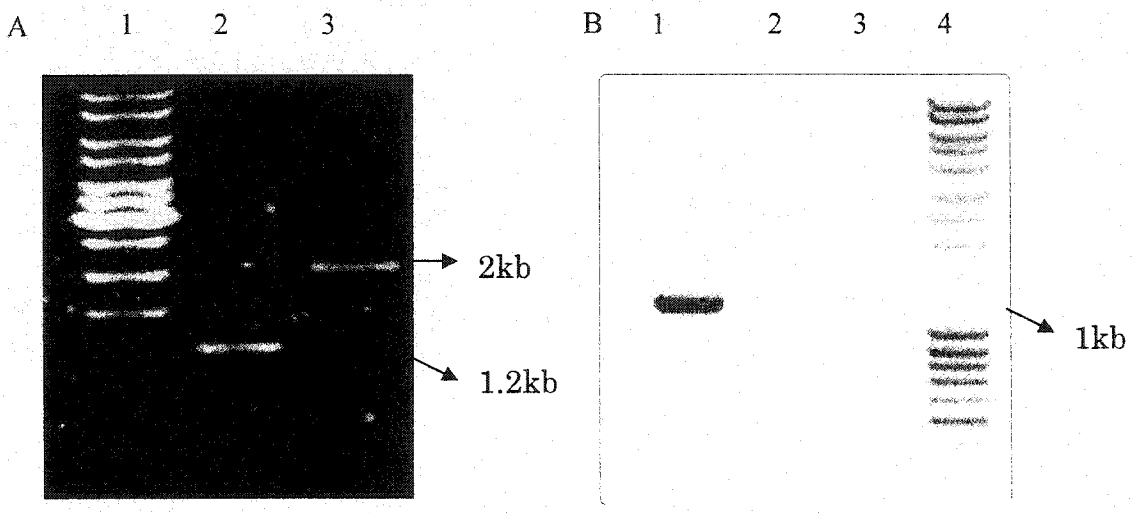


Figure 12.

A) PCR amplified products of *Pro5a* and *Pro5b*

PCR amplifications were performed using sets of primers specific for *Pro5a* and *Pro5b* respectively and Taq DNA polymerase. The amplified DNA products were verified on ethidium bromide-stained agarose gels. Lane 1) Gene Ruler^{TM} 1kb DNA ladder; Lane 2) Amplified *pro5a* with a length of approximately 1.2 kb; Lane 3) Amplified *Pro5b* with a length around 2 kb.

B) PCR amplified products of *Pro5c*

Lane 1) Amplified *Pro5c* with a length of approximately 1kb. Lane 2) 5x diluted PCR product of *Pro5c*; Lane 3) 10x diluted PCR product of *Pro5c*; Lane 4) Mass Ruler^{TM} 1kb DNA ladder.

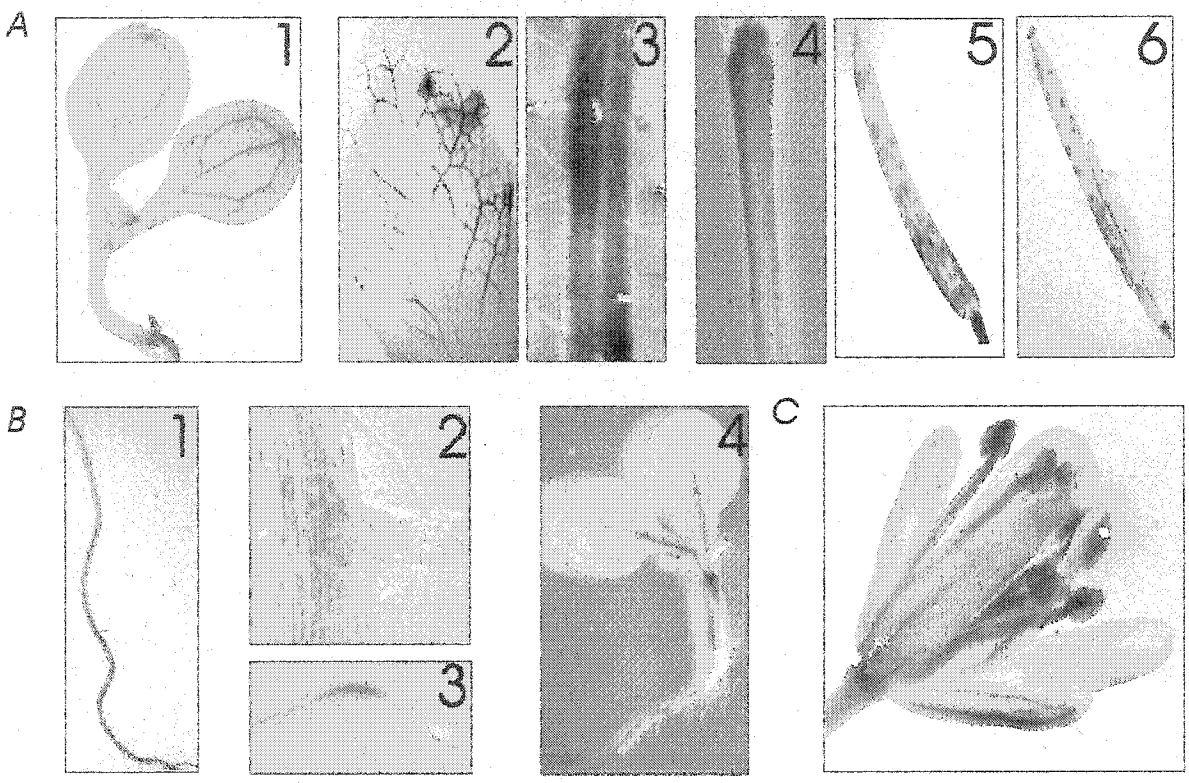


Figure 13. Expression studies of *AtST5a, 5b* and *5c* by GUS staining
(*A*) Promoter activities of *Pro5a* in apical plant parts. Expressions in hypocotyls, leaves, stems, petioles and silique walls (1 to 6) are restricted to the vascular tissue. However, no activity was found in seeds except seeds funicles. *Pro5b* and *Pro5c* exhibited the same expression pattern (not shown).
(*B*) *Pro5a* is active in the vascular tissue throughout the entire roots (1 and 2) except young main root tip (e.g. 6 days) (3) and is first detectable in 3-day-old seedlings at the basal end and the vascular tissues of cotyledons (4). However, the older root tips were also found to have *Pro5a* activity (after 10 days). The same expression patterns were also found for *Pro5b* and *Pro5c* (not shown).
(*C*). In flowers, *Pro5a* is active in the vascular tissue of the receptacle, sepals and petals, stamen mainly in pollen grains and stigma. The same expression patterns were found for *Pro5b* and *Pro5c* (not shown).

GUS staining reveals that the three promoters were expressed throughout the roots except for young main root tips (younger than 7 days). The staining was restricted to the vascular system of both main and lateral roots as well as young lateral root primordia, and was first detectable 3 days after germination at the basal end of the main roots and the vascular tissues of leaves. Furthermore, according to GUS-staining time, the activities of the three promoters were found to vary between organs at different stages. For example, promoter activity is increased in the mature silique walls and decreased in the roots and leaves during plant development. However, the overall promoter activities are high in young leaves, shoots and roots, decrease as the tissues mature and are absent in seeds. At the flowering stage, the promoter activity of the three genes was observed in all parts of the flower organ mainly in pollen grains and stigma, but was also found in the vascular bundle of sepals, petals and receptacles.

D.3. Analysis of GI patterns

AtST5a, AtST5b and AtST5c were found to catalyze the conversion of DSGs to GIs, the final step of GI biosynthesis. Based on the difference in substrate preference, we expected that AtST5a, AtST5b and AtST5c might be involved in the synthesis of different GIs in vivo. The knockout T3 seeds of the *AtST5a* and *AtST5c* genes were found to be available and were ordered from the ABRC collection. The homozygous knockout lines were identified by PCR using gene specific primers. Three out of 12 plants were found to be homozygous for the insertion in *AtST5a* (*KO5a*) (Figure 14), whereas four out of fourteen plants were found to be homozygous for the insertion in *AtST5c* (*KO5c*) (Figure 15). 10 knockout and 10 wild-type plants were prepared for the

extraction of GIs for both *KO5a* and *KO5c*. The GI extract from each plant was injected into an HPLC to analyze GI patterns. GI profiles of *KO5a* and *KO5c* were compared with those of wild-type plants using students T-test. After the analysis of 6 peaks in the *KO5a* and wild type plant profile, we found that peak 4' was below detectable limits in *KO5a*. Peak 17.5' was also lower with a P value of 0.0001, which means that there is significant difference between *KO5a* and wild type plants for this peak (Figure 16). However, 7 peaks were analyzed in *KO5c* and its corresponding wild type plant. None of them exhibits significant difference with wild-type plant profiles (Figure17).

D.4. Expression of *AtST5a, 5b* and *5c* following treatment with signal molecules

The primary auxin, indole-3-acetic acid (IAA) and indole GI have been shown to share the same precursor (tryptophan) and the same metabolic branch point (indole-3-acetaldoxime) (Bak and Feyereisen, 2001; Normanly and Bartel, 1999; Hull et al., 2000; Mikkelsen et al., 2000). GIs are believed to play a role in plant protection against herbivores and pathogen infection. In support of this hypothesis, experiments with the treatment of signal molecules in *Brassica napus* have shown that the application of SA leads to an increase of GIs in leaves, in particular 2-phenylethyl-GI (Kiddle, et al., 1994), whereas MeJA induces mainly indole GIs (Doughty et al, 1995). Furthermore, enzymology studies have shown that three enzymes display the different substrate preference. Therefore 14-days old wild-type Arabidopsis plants were treated with IAA, SA and MeJA to study *AtST5a, AtST5b* and *AtST5c* expression using RT-PCR. RT-PCT results showed the absence of induction by the different treatments (Figure19 I.II, III).

However, RT-PCR results are not quantitative and the experiments should be repeated using other approaches (e.g. real time PCR).

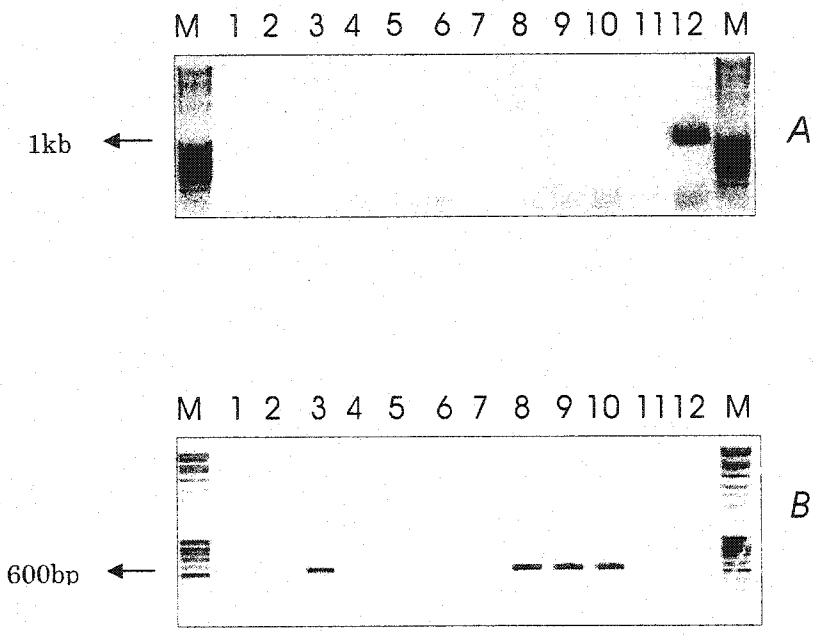


Figure 14 Selection of homozygous line of KO5c using PCR.
11 putative KO5c plants plus one wild-type plant were used for genomic DNA extraction. The first set of primers was used to amplify the wild-type allele of *AtST5c* (around 1 kb), whereas the second set of primers was used to amplify the region (about 600 bp) between the left border of the T-DNA and the 3' end of *AtST5c*. A) PCR results using the first set of primers. Lane1 to 11 corresponds to 11 putative KO5c plants. Lane12 corresponds to wild-type plants. M is Mass Ruler™ 1-kb DNA marker. B) PCR results using the second set of primers.

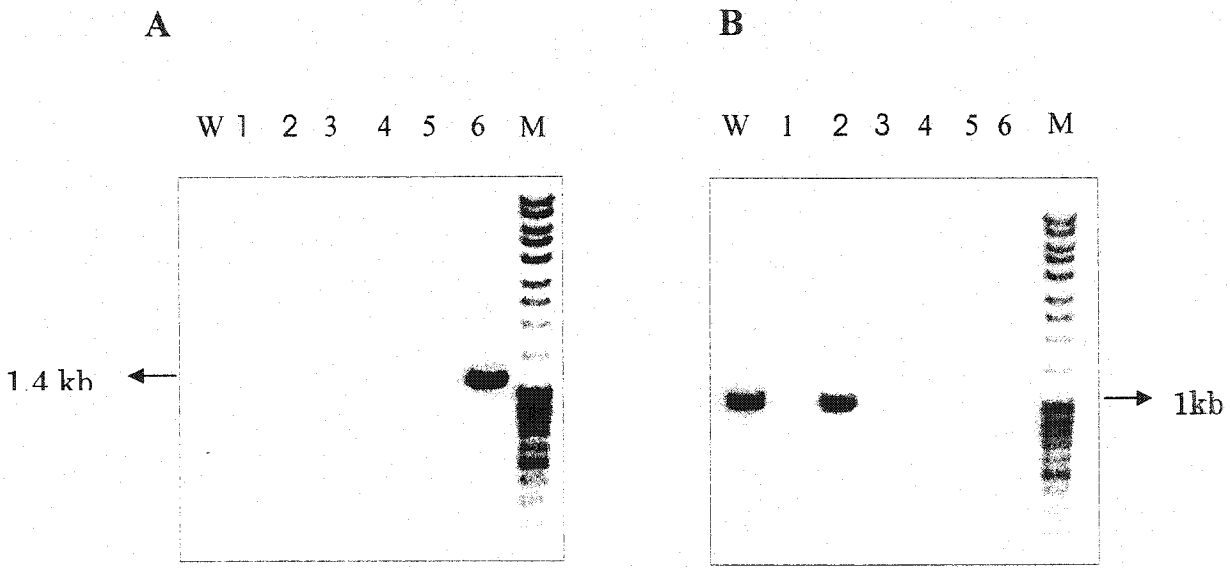


Figure 15 Selection of homozygous *KO5a* using PCR
Genomic DNA was extracted from 6 putative KO5c plants and from 1 wild-type plant. The first set of primers used in the PCR reaction was expected to amplify the region (about 1.4kb) between the left border of the T-DNA and the 5' end of *AtST5a*. The second set of primers used in the PCR reaction was expected to amplify the wild-type allele of *AtST5c* (around 1 kb). A) PCR results using the first set of primers. B) PCR results using the second set of primers. Lane 1 to 6 corresponds to 6 putative knock out plants. W corresponds to wild-type plant. M corresponds to mass Ruler™ 1 kb DNA markers. .

A.

Ret time	Average area of WT	Std.dev. of WT	Average area of KO5a	Std.dev of KO5a	P value
3'	33142652	24740224.43	22057264	23134494.14	0.095
4'	22140383.7	-	-	-	-
15'	34144565	19980002.41	33855079	19229069.94	0.48
17'	19829443	6540560.39	21654428	6830862.34	0.2747
17.5'	6792100	3047798.59	716017.3	940569.48	0.0001
20'	25336446	10232681.87	19804502	7926673.47	0.0966

B.

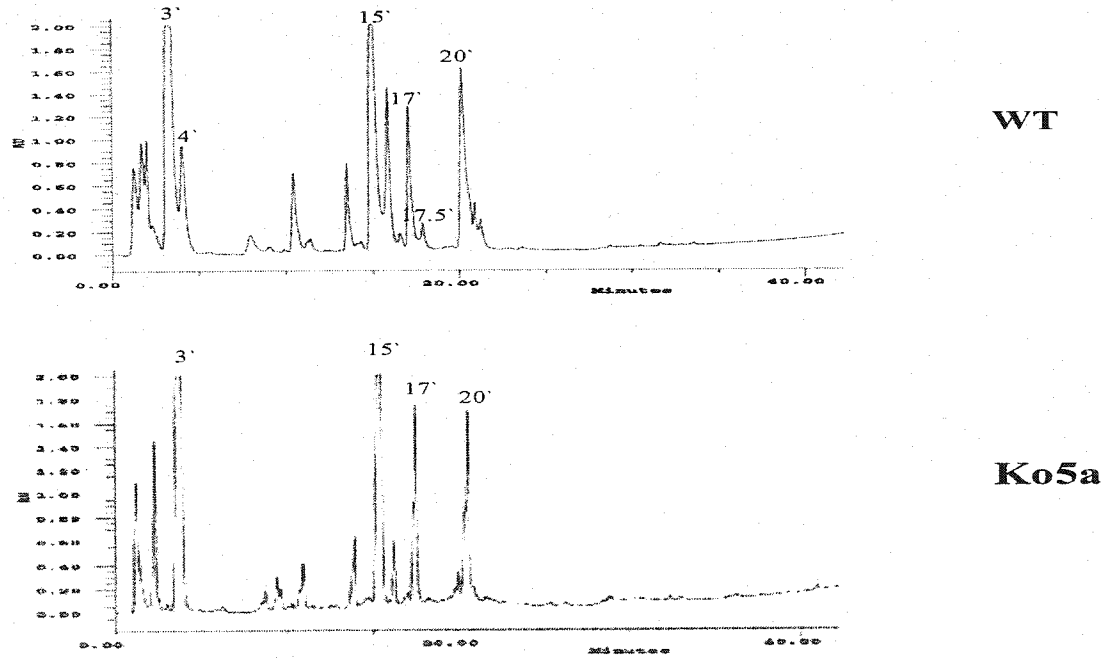


Figure16. HPLC profile variation in KO5a plants
(A) GI contents from wild-type (WT) and knock out *AtST5a* (KO5a) plants. 10 samples for each were analyzed (see methods). An equal quantity of extract from each sample was injected for peak quantification. HPLC results from 10 wild- type samples were compared with those from 10 KO5a samples using students T-test. (-), not detected.
(B). Representative HPLC chromatograms from plant extract of one wild-type sample and one KO5a sample. Signals are given in arbitrary units (mAU). Peaks that were used for the students T test are marked in minutes

A

Ret time	Average of WT	std.dev.	Average of KO5c	std.dev.	P value
1'	20149885	7460026.18	23411730	12670891	0.246
3'	66378995	52021357	54778855	29530871	0.3105
13'	14038892	7390101.6	12001851	5677336.19	0.2606
14'	54043786	32885752.69	27548358	22950954.17	0.0256
15'	12714274	4.62068E+13	7914201	2680239.74	0.1356
17'	32659506	12491066.22	32199092	16210582.23	0.473
20'	40621659	16997846.15	47436174	24177786.67	0.2377

B

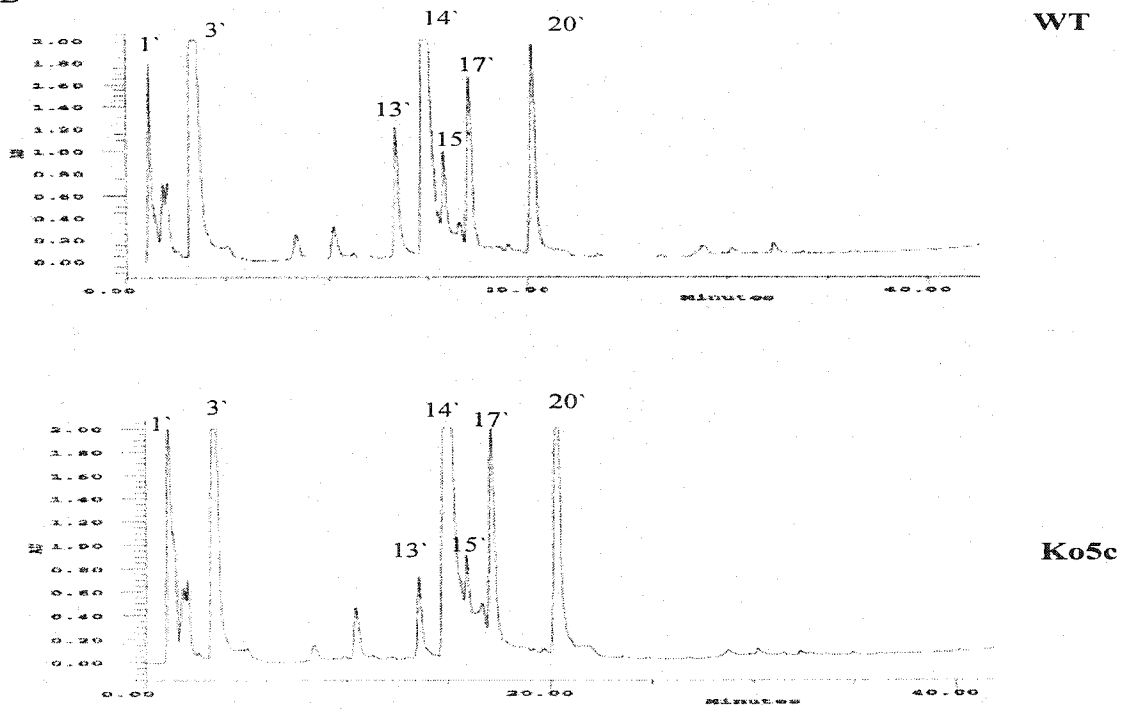


Figure 17. HPLC profile variation in KO5c plants
(A) GI contents from wild-type (WT) and knock out *AtST5c* (KO5c) plants. 10 samples for each were analyzed (see method). An equal quantity of extract from each sample was injected for peak quantification. HPLC results from 10 wild- type samples were compared with those from 10 *KO5c* samples using student T-test.
(B) Represent HPLC chromatograms from plant extract of one wild-type sample and one KO5a sample. Signals are given in arbitrary units (mAU). Peaks that were used for the students T test are marked in minutes.

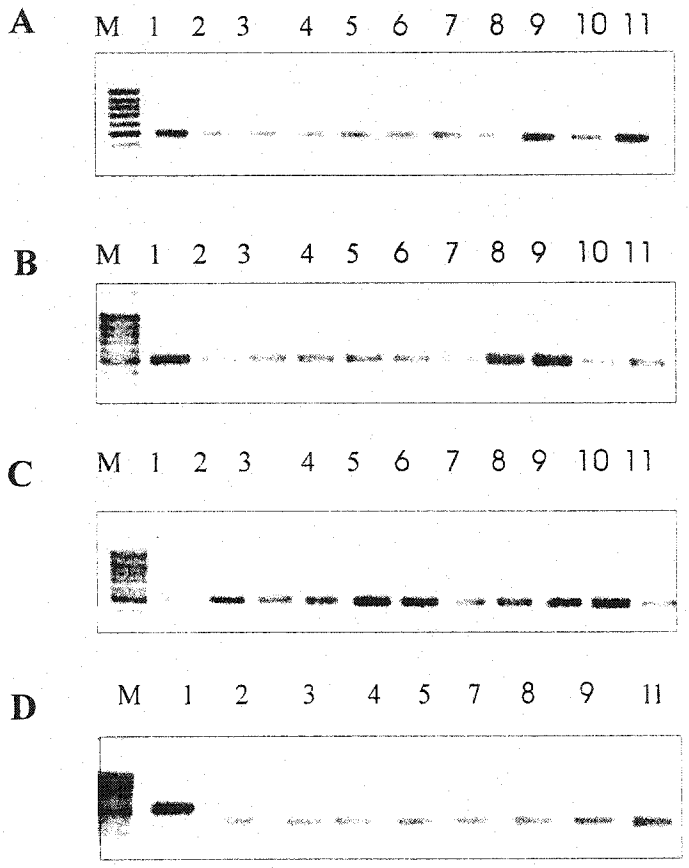


Figure 18I) RT-PCR reaction using mRNA from *Arabidopsis thaliana* 14 day-seedlings treated with IAA for 0h (2, 6); 2h, 5μM (3); 4h, 5μM (4); 8h, 5μM (5,10); 8h, 0.10μM (7); 8h, 1.0 μM (8); 8h, 2.5 μM (9); and 8h, 10μM (11) using gene specific primers for *AtST5a* (*A*), *AtST5b (B)*, *AtST5c (C)* and *actin (D)*. M correspond to Mass Ruler™ 1-kb DNA marker. Lane 1 corresponds to genomic DNA.

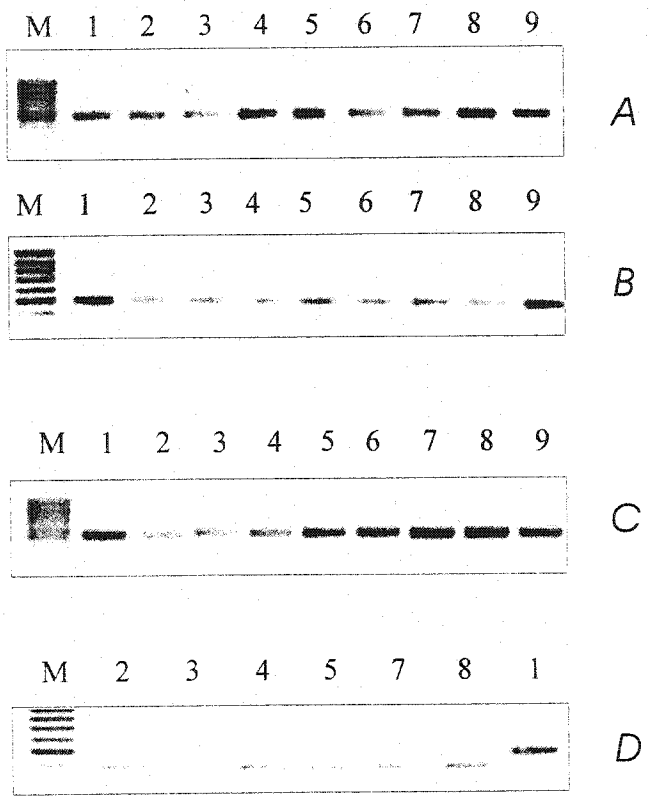


Figure 18II) RT-PCR reaction using mRNA from *Arabidopsis thaliana* 14 - days old seedlings treated with MeJA for 0h (2, 6); 2h, 100μM (3); 4h, 100μM (4); 8h, 100μM (5, 9); 8h, 10μM (7); and 8h 50 μM (8) using gene specific primers for *AtST5a* (*A*), *AtST5b (B), AtST5c (C)* and *actin (D)*. M correspond to Mass Ruler™ 1-kb DNA marker. Lane 1 corresponds to genomic DNA.

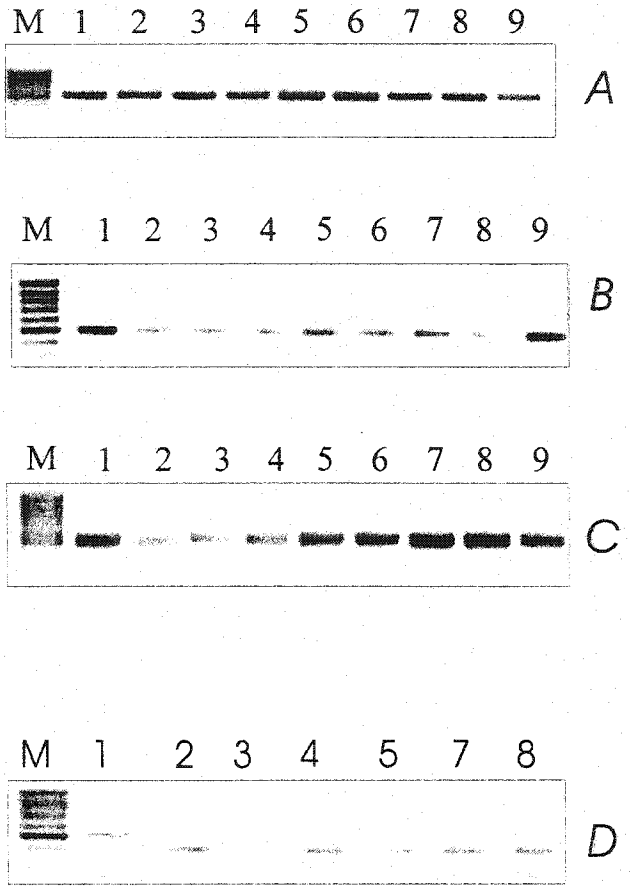


Figure 18III) RT-PCR reaction using mRNA from *Arabidopsis thaliana* 14- days old seedlings treated with SA for 0h (2, 6); 2h, 100μM (3); 4h, 100μM (4); 8h, 100μM (5, 9); 8h, 10μM (7); and 8h 50 μM (8) using gene specific primers for *AtST5a (A)*, *AtST5b (B)*, *AtST5c (C)*, and *actin (D)*. M correspond to Mass Ruler™ 1-kb DNA marker. Lane 1 corresponds to genomic DNA.

E. DISCUSSION

E.1. Substrate specificity of Arabidopsis DSG-STs

One of the Arabidopsis genes encoding DSG-ST, *AtST5c*, has been cloned and characterized at the biochemical level. Its encoded protein shares a high level of amino acid sequence identity and similarity with the other two previously characterized Arabidopsis DSG-STs, encoded by *AtST5a* and *AtST5b* (Diego Spertini, unpublished result). Amino acid alignment of the three deduced amino acid sequences reveals that they possess all the conserved domains known to be involved in cosubstrate binding (PAPS) and in catalysis, suggesting that their catalytic mechanism is similar to those of previously characterized plants and animal cytosolic STs. A significant proportion of the variation observed between the three protein sequences is located in the regions which determine substrate preference, suggesting that they might have specialized functions in vivo. This hypothesis is partially supported by the results of the substrate preference studies, which reveal that the three enzymes have different substrate preferences. However, AtST5c like the two previously characterized enzymes exhibit a strict requirement for the thioglucose moiety. The same structural requirement was also observed for the *L. sativum* DSG-STs (Glendening et al., 1990). Previous studies indicated that the *Brassica* and *Lepidium* DSG-STs have many features in common: such as the absence of requirement for divalent cations and inhibition by the addition of SH group reagents. These features are also observed for the three Arabidopsis DSG-STs studied here, suggesting the importance of one or more thiol groups for enzyme activity.

The enzymes involved in the last two steps of GI biosynthesis were thought to exhibit no specificity for the side-chain structures (Poulton, et al., 1993 and Underhill, 1980). More recently, Bak et al (2001) further suggested that post oxime enzymes involved in the GI biosynthesis pathway have low substrate specificity. These hypotheses are supported by the fact that the three Arabidopsis DSG-STs enzymes were found to accept all GIs tested, including those that do not occur naturally in Arabidopsis, such as benzyl-GI. Similar results were reported for the thiohydroxymate glucosyltransferase from T. *majus* (Matsuo et al., 1971) and *B. napus* (Reed et al, 1993). The two enzymes were found to accept precursors with different side-chain structures. Furthermore, CYP83A1 and CYP83B1, two enzymes downstream of the oxime formation, were reported to have overlapping substrate specificity in vitro (Bak et al., 2001). This lack of substrate specificity is further supported by the fact that cell suspension cultures of *Brassica juncea* can produce artificial and novel GIs from *p*-nitrobenzaldoxime (Grootwassink et al., 1990). Likewise, the ectopic expression of the sorghum CYP79A1 encoding the enzyme involved in the conversion of tyrosine to *p*-hydroxyphenyl-acetaldoxime resulted in Arabidopsis plants with high levels of tyrosine-derived *p*-hydroxybenzyl-GI, which is not naturally occurring in this species (Bak et al., 1999)

Although the three Arabidopsis DSG-STs exhibit broad substrate specificities, they display distinct substrate preferences. AtST5c accepts all of the DSGs tested but with preference for 3-methylthiobutyl-, 3-indolylmethyl- and 2-

phenylethyl-DSGs (table I). AtST5b exhibits also a wide substrate specificity but with preference for 3-methylthiobutyl-, 3-methylsulfonylpropyl- and 2-phenylethyl-DSGs (Diego Spertini, unpublished results). In contrast, AtST5a is not as flexible as the other two enzymes, and only accepts 3-indolylmethyl-DSG to a significant level (Diego spertini, unpublished results). It has been proposed that secondary modifications of the GI side chain might occur following the GI core structure formation (Halkier et al., 1997). For instance, the methionine-derived aliphatic side chains are found to be modified extensively. The initial products after core structure formation are likely to be methythioalkyl-GIs. Both methylsulfinylalkyl- and methylsulfonylalkyl- GIs are believed to be oxidation products of methylthioalkyl-GIs (Chisholm, 1972). Furthermore, the modification of the methylthio group is thought to be an early step in the biosynthesis of the benzyloxyalkyl GIs (Haughn et al., 1991). The three Arabidopsis DSG-STs exhibit preference for methylthiobutyl-DSG and considerably lower activity for 2-hydroxyl 3-butenyl- and 4-benzyloxybutyl-DSG, which are derivatives of methylthiobutyl-DSG, supporting the hypothesis that secondary modifications of the side chains occur after GI core structure formation. Although benzyl-GI is not a naturally occurring metabolite in Arabidopsis, kinetic studies reveal that benzyl-DSG was the best substrate for AtST5c with the highest Vmax/km value (table II). The other two enzymes, AtST5a and AtST5b were also found to prefer benzyl-DSG as their best substrate (Diego Spertini, unpublished result). The preference for benyzyl-DSG as opposed to allyl DSG was also observed for the *L. Sativum* DSG-ST (Glendening et al., 1990). These results demonstrate the flexibility of the active site of the DSG-ST enzymes

for different side-chain structures. When we consider only the naturally occurring Arabidopsis GIs, 4-methylthiobutyl- DSG is the best substrate for AtST5c (Table II); whereas AtST5b and AtST5a display preference for 2-phenylethyl- and 3-indolylmethyl-DSG respectively (Diego Spertini, unpublished results). The fact that the three enzymes display distinct substrate preference suggests that each enzyme might have a specific function in vivo. To test this hypothesis, analyses of *AtST5a, AtST5b* and *AtST5c* T-DNA insertion mutants might reveal the contribution of the individual enzymes to the synthesis of *A.thaliana* GIs.

The GI pattern of *AtST5a* and *AtST5c* knock out mutants was analyzed by HPLC. Statistical analyses reveal that the GI pattern of the AtST5a knock out mutant exhibited a detectable difference when compared with wild type plants. This includes the absence of one peak eluting at 4 minutes and variation in a peak eluting at 17.5 minutes. In contrast, no apparent difference in the GI profile was observed for the AtST5c knock out mutant. Enzymology studies have revealed that the three enzymes display distinctive substrate preference such as AtST5a shows preference for 3-indolmethyl DSG, while AtST5b and AtST5c prefer 2-phenylethyl and 4-methylthiobutyl-DSG (Diego Spertini, unpublished) (Table II). The changes observed for the GI pattern of the AtST5a knock out plant are consistent with the result of the enzymology studies and further support the hypothesis that each DSG-ST enzyme might have specialized functions in vivo. The peak that is absent in AtSt5a knock out plant might correspond to 3-indolymethylDSG. Structural studies of the eluting GI at this peak using LC-MS/MS will be required to confirm this hypothesis.

Amino acid sequence alignment of the three DSG-STs shows that although they are very similar, AtST5b and AtST5c are more closely related to each other as compared with AtST5a especially in the domain involved in substrate binding. This observation supports the results of the substrate preference studies, which show that AtST5b and AtST5c can accept all the substrates tested, and have overlapping substrate preferences for 3-methylthiotyl- and 2-phenylethyl-DSG in vitro. The close relationship between AtST5b and AtST5c might explain the absence of changes in the GI profile of the AtST5c knock out mutant plants since AtST5b might compensate for the loss of the AtST5c enzyme in vivo. To confirm this hypothesis, analysis of the GI patterns of *AtST5b* T-DNA insertion mutants and the construction of the *AtST5b* and *AtST5c* double mutant will be required.

E.2. Histochemical analysis of the expression patterns of the DSG-ST genes

Studies of the expression pattern of the three Arabidopsis DSG-ST genes, *AtST5a/AtST5b/AtST5c*, provide valuable information regarding the site of GI biosynthesis in Arabidopsis. Previous reports indicated that GIs in Arabidopsis accumulate in all tissues that have been analysed, including roots, leaves, and seeds (Haughn et al., 1991). In this study, we found that *AtST5a, 5b* and *5c* are expressed in roots and leaves. *CYP79F1,* which encodes an enzyme involved in the conversion of methionine to the corresponding aldoxime, was also found to be expressed in these tissues (Reintanz et al., 2001).

Plant secondary metabolites are often transported in plants, e.g. nicotine, caffeine and cytokinin are produced in roots and translocated to the leaves. GIs are also

believed to be actively transported in plants. For instance, in *T. majus,* benzyl GI was found to be primarily synthesized in leaves, but was found to accumulate in other tissues such as developing seeds, suggesting the existence of translocation (Lykkesfeldt et al., 1993). Furthermore, in white mustard and rapeseed, it has been demonstrated that one of the major sites of GI biosynthesis is the silique wall (Toroser et al., 1995; Du et al, 1998) and that GIs are transferred from the walls to the seeds, where they accumulate during seed maturation. In addition, there was a correlation between the level of aliphatic GIs in leaves and seeds of Arabidopsis suggesting that they are transported from the leaves where they are synthesized to the seeds (Kliebenstein et al., 2001). In our study, none of the three DSG-ST genes is expressed in the seeds whereas all of them are expressed in the mature silique walls and leaves, suggesting that GIs in Arabidopsis are transported from mature silique walls or leaves to the developing seeds.

Previous studies with radiolabelled precursors revealed that the sulfonation reaction was extremely active in the pod walls of *B. napus* (Toroser et al., 1995). Our studies demonstrate that the three DSG-ST genes are highly expressed in the mature silique walls of *A. thaliana*. The fact that the three DSG-ST genes, *CYP79F1* and *CYP79F2* are not expressed in the seeds of Arabidopsis (Reintanz et al., 2001) suggests that the seeds may not take part in GI biosynthesis. The seeds in oil seed rape were also proposed not to take part in GIs biosynthesis (Magrath and Mithen, 1993).

The phloem tissue is responsible for the translocation of carbohydrates and other metabolites in plants. Both GIs and DSGs were found to have suitable

physicochemical properties to allow phloem mobility (Brudenell et al, 1999). GI phloem transport was demonstrated using aphid feeding experiments on *B. nigra*, where it was shown that more than 10 mM sinigrin is present in the phloem sap of young leaves whereas only 1-2 mM GIs were found in mature, presenescent and senescent leaves (Merritt et al., 1996). In this study, we demonstrate that *AtST5a, AtST5b and AtST5c* expression seems to be associated with the vascular bundle of leaves, roots, stems and silique walls. These results support the possibility that GIs are synthesized and transported in the vascular tissue.

It has been demonstrated that the GI profile of different species vary greatly and that within a single species there are important age-dependent fluctuations (Clossais-Besnard et al., 1991; Porter et al., 1991). The dynamic changes of GI levels in a plant are dependent on the regulation of de novo biosynthesis, degradation and mobilization of glucosinolates. For instance, in oilseed rape and Chinese cabbage, aldoxime formation was found to be high in young leaves and low in mature leaves (Bennett et al., 1995). *In A. thaliana,* the expression of *CYP79F1 and CYPF2* was found to be the highest in fully expanded rosette leaves and lowest in stems and young stem leaves (Reintanz et al., 2001). In this study, the expression of *AtST5a, AtST5b and AtST5*c was found to be high in young leaves, shoots and roots, to decrease as the tissue mature and was absent in seeds, suggesting a similar regulation for each step of GI biosynthesis. Although the three genes have the same expression pattern in the plants, they might have a different level of expression in certain tissues and at certain development stages.

Interestingly, the three DSG-ST genes were found to be expressed in

pollen, stamen, stigma, and in the vascular bundles of sepals, petals and receptacles. Since the sepals develop into siliques, and the petals and receptacles are derived from leaves, it is not surprising to see the expression of the three genes in these tissues. However, it is quite interesting to see the expression of the three genes in pollen grains, anthers and stigmas but not in the seeds since seed formation is due to interactions between these plant sex organs. Furthermore, the three DSG-ST genes were found to be expressed in the pollen grains of anthers and stigmas, whereas none of other genes involved in GI biosynthesis were found to be expressed in the flower organs of Arabidopsis (Reintanz et al., 2001). This indicates that the conversion of DSG to GI may be the only step of GI biosynthesis occurring in the flower organ. These results suggest that DSG might be transported from the leaves to the flowers in the plants. Similarly, DSGs in *B. napus* were also suggested to be the transport form of GIs since DSG-ST activity was found in the seeds of this species (Toroser et al., 1995).

E.3. Regulation of three *DSG-ST* gene expressions

The expression of *AtST5a, 5b* and *5c* seems not affected by treatments with signal molecules. Previous reports have shown that the treatment of *B.nupus* with jasmonate lead to a systematic accumulation of indole GIs (Barlet et al, 1999; Doughty et al, 1995; Li et al, 1999; Porter et al, 1991), whereas 2-phenylethylGI levels were found to increase following salicylic acid treatment (Kiddle et al, 1999). In contrast, the treatment with abscisic acid resulted in lower levels of GIs (Mollers et al, 1999). Furthermore, certain signal molecules were found to induce some of the

genes involved in the GI biosynthesis. For instance, *CYP79B2* and *CYPB3,* which encode enzymes catalyzing the first step of indole GI synthesis, are highly induced by treatments with MeJA (Mikkelsen et al., 2003). It has also been reported that the infection by a virulent bacterial pathogen induces the expression of *CYP79B2* (Hull et al., 2000). These observations, as well as our results suggest that the genes involved in the first steps of GI biosynthesis may tend to be regulated, whereas those genes involved in later steps, such as *DSG-STs,* are constitutively expressed.

E.4. GI in crop improvement and food design

The ultimate applied aspect of GI research is to specifically alter the quality and quantity of GIs, e.g. to decrease the amount of unwanted GIs and increase the amount of desirable GIs to boost crop protection and nutritional values.

GIs are often classified as antinutritional compound due to their toxic degradation products. Thus, the nutritional value of rapeseed meal as animal feed is impaired by high GI content. For this reason, over 40 years ago, plant breeders initiated a search for rapeseed that is low in GIs. To efficiently reduce the GI content in rapeseed meal, a biosynthetic block of GI biosynthesis might provide an efficient alternative way to breeding. To date, several genes involved in GI biosynthesis have been identified. The identification of the genes encoding DSG-STs provides additional tools to modulate GI profiles in vivo.

GIs are thought to play a role in plant defence (Rask et al., 2000) and act as powerful antibacterial, antifungal and antimicrobial agents (Louda and Mole, 1991), which therefore can influence the behavior of insect, avian, and molluscan

herbivores. It has been shown that different GIs and their degradation products play different roles in plant defence. For example, indole GIs and their hydrolysis products play a role in the control of club root disease, caused by the obligate biotroph *Plasmodiophora brassicae* (Ludwig-Mueller et al., 1999); whereas the length of the side chain of methionine-derived aliphatic GIs have a special function in the defence against flea beetles (Giamoustaris and Mithen, 1995). Since each *DSG-ST* may play a unique role in the synthesis of GIs in vivo, altering or enhancing their expression might allow to tailor plants with distinct resistance profiles.

REFERENCES

Bak, S., Olsen, C.E., Petersen, B.I., Moller, B.L. and Halkier, B.A. (1999) Metabolic engineering of p-hydroxybenzyl-glucosinolate in *Arabidopsis* by expression of the cyanogenic CYP79A1 from *Sorghum* bicolor. *Plant J.* **20,** 663-672.

Bak, S. and Feyereisen, R. (2001) The involvement of two P450 enzymes, CYP83B1 and CYP83A1, in auxin homeostasis and glucosinolate biosynthesis. *Plant Physiology.* **127**, 108-118.

Bartlet, E., Kiddle, G., Willians, I., Wallsgrove, R. (1999) Wound-induced increases in the glucosinolate content of oilseed rape and their effect on subsequent herbivory by a crucifer specialist. *Entomol. Exp. Apple.* **91,** 163-167.

Benchtold, N., Ellis, J. and Pelletier, G. (1993) In planta Agrobacterium mediated gene transfer by infiltration of adult *Arabidopsis thaliana* plants. *C.R. Acad. Sci Paris, Life Sciences.* **316**, 1194-1199.

Benn, M.H. (1962) Biosynthesis of mustard oils. *Chem. Ind.* (London) 1907.

Bennett, R.N., Ludwig-Muller, J., Kiddle, G., Hilgenberg, W. and Wallsgrove, R.M. (1995) Developmental regulation of aldoxime formation in seedlings and mature plants of Chinese cabbage (*Brassica campestris ssp. pekinensis*) and oilseed rape (*Brassica napus*): Glucosinolate and IAA biosynthetic enzymes. *Planta.* **196,** 239-244.

Bennett, R.N., Kiddle, G., Hick, A.J., Dawson, G.W. and Wallsgrove, R.M. (1996) Distribution and activity of microsomal NADPH-dependent monooxygenases and amino acid decarboxylases in cruciferous and non-cruciferous plants, and their relationship to foliar glucosinolate content. *Plant Cell Environ.* **19,** 801-812.

Bennett, R.N., Kiddle, G. and Wallsgrove, R.M. (1997) Biosynthesis of benzylglucosinolate, cyanogenic glucosides and phenylpropanoids in *Carica papaya. Phytochemistry.* **45,** 59-66.

Bjorkman, R. and Janson, J.C. (1972) Studies on myrosinases. I. Purification and characterization of a myrosinase from white mustard seed (*Sinapis alba L.*). *Biochim. Biophys. Acta.* 276, 508-518.

Bjorkman, R. and Lonnerdal, B. (1973) Studies on myrosinases. III. Enzymatic properties of myrosinases from *Sinapis alba* and *Brassica napus* seeds. *Biochim. Biophys. Acta.* **327**, 121-131.

Blake-Kalff, M.M., Harrison, K.R., Hawkesford, M.J., Zhao, F.J. and McGrath, S.P. (1998) Distribution of sulfur within oilseed rape leaves in response to sulfur

deficiency during vegetative growth. *Plant Physiol.* **118**, 1337-1344.

Bones, A.M. and Rossiter, J.T. (1996) The myrosinase-glucosinolate system, its organization and biochemistry. *Physiol. Plant.* **97,** 194-208.

Brudenell, A.J.P., Griffiths, H., Rossiter, J.T. and Baker, D.A. (1999) The phloem mobility of Glucosinolates. *J. Exp. Bot.* **50,** 745-756.

Burmeister, W.P., Cottaz, S., Driguez, H., Iori, R., Palmieri, S. and Henrissat, B. (1997) The crystal structures of *Sinapis alba* myrosinase and a covalent glycosyl-enzyme intermediate provide insights into the substrate recognition and active-site machinery of an S-glycosidase. *Structure.* **5,** 663-675.

Chen, S. and Halkier, B.A. (1999) Functional expression and characterization of the myrosinase MYR1 from *B. napus* in *S. cerevisiae*. *Protein Expr. Purif.* **17,** 414-420.

Chiba, H., Komatsu, K., Lee, Y.C., Tomizuka, T. and Strott, C.A. (1995) The 3'-terminal exon of the family of steroid and phenol sulfotransferase genes is spliced at the N-terminal glycine of the universally conserved GXXGXXK motif that forms the sulfonate donor binding site. *Proc. Natl. Acad. Sci. USA.* **92,** 8176-8179.

Chisholm, M.D. (1972) Biosynthesis of 3-methylthiopropylglucosinolate and 3-methylsulfinylpropyl glucosinolate in wallflower *Cheiranthus kewensis, Phytochemistry.* **11,** 197-202.

Clossais-Besnard, N. and Larher, F. (1991) Physiological role of glucosinoaltes in *Brassica napus.* Concentration and distribution pattern of Glucosinolates among plant organs during a complete life cycle. *J. Sci. Food Agric.* **56,** 25-38.

Daughty, K.H., Kiddle, G., Pye, B.J. and wallsgrove, R.M. (1995) Selective induction of Glucosinolates in oilseed rape leaves by methyl jasmonate. *Phytochemistry.* **38**, 347-350

Du, L. and Halkier, B.A. (1998) Biosynthesis of Glucosinolates in the developing silique walls and seeds of *Sinapis alba*. *Phytochemistry* **48,** 1145-1150

Du, L., Lykkesfeldt, J., Losen, C.E. and Halkier, B.A. (1995) Involvement of cytochrome P450 in oxime production in glucosinolate biosynthesis as demonstrated by an in vitro microsomal enzyme system isolated from jasmonic acid-induced seedlings of *Sinapis alba* L. *Proc. Natl. Acad. Sci. USA.* **92**, 12505-12509

Durham, P.L. and Poulton, J.E. (1990) Enzymatic properties of purified myrosinase from *Lepidium sativum* seedlings. *Z. Naturforsch.* **45,** 173-178

Faulkner, I.J., Rubery, P.H. (1992) Flavonoids and flavonoid sulfates as probes of auxin transport regulation in *Cucurbila pepo* hypocotyls segments and vescles. *Planta*. **186,** 618-625.

Fenwich, R., Heaney, R.K. and Mullin, W.J. (1983) Glucosinolates and their breakdown products in food and food plants. *CRC Crit. Rev. Food Sci. Nutri.* **18,** 123-201.

Foo, H.L., Gronning, L.M., Goodenough, L. Bones, A.M., Danielsen, B.E., Whiting, D.A., Rossiter, J.T. (2000) Purification and characterization of epithiospecifier protein from *Brassica napus*: enzymic intramolecular sulphur addition within alkenyl thiohydroximates derived from alkenyl glucosinolate hydrolysis. *FEBS Lett.* **468,** 243-246.

Giamoustaris, A., and Mithen, R. (1995) The effect of modifying the glucosinolate content of leaves of oilseed rape *(Brassica napus* ssp. *Oleifera*) on its interaction with specialist and generalist pests. *Ann. Appl. Biol.* **126**, 347-363.

Glendening, T.M. and Poulton, J.E. (1990) Partial purification and characterization of a 3'-phosphoadenosine 5'-phosphosulfate: Desulfoglucosinolate sulfotransferase from cress *(Lepidium sativum)*. *Plant Physiol.* **94,** 811-818.

Grootwassink, J.W.D., Balsevich, J.J., Kolenovsky, A.D. (1990) Formation of sulfateglucosides from exogenous aldoximes in plant cell cultures and organs. *Plant Sci.* **66**, 11-20.

Grootwassink, J.W.D., Underhill, E.W., Hemmingsen, S.M., Reed, D.W. and Kolenovsky, A.D. (1997) Plants with reduced Glucosinolate content. *European Patent,* EP 0 771 878 A1.

Hahlbrock, K., and Scheel, D. (1989) Physiology and molecular biology of phenylpropanoid metabolism. *Annu. Rev. plant Physiol. Plant Mol. Biol.* **41**, 339-367.

Halkier, B.A. (1999) GIs, in: Ikan R. (Ed.), Naturally Occuring Glycosides: Chemistry, Distribution and Biological Properties. *Jon Wiley and Sons Ltd, London.* pp. 193-223.

Halkier, B.A., and Du, L. (1997) The biosynthesis of glucosinolates. *Trends in Plant Sci.* **2,** 425-431.

Hasegawa, T., Yamada, K., Kosemura, S., Yamamura, S. and Hasegawa, K. (2000) Phototropic stimulation induces the conversion of glucosinolate to phototropism-regulating substances of radish hypocotyls. *Phytochemistry.* **54,** 275-279.

Hanson, A.D, Rathinasabapathi, B., Chamberlin, B., and Gage, D.A. (1991)

Comparative physiological evidence that β-alanine betaine and choline-*O*-sulfate act as compatible osmolytes in halophytic *Limonium* species. *Plant Physiol.* **97,** 1199-1205.

Hanson, A.D., Ratgubasabaoathi, B., Rivoal, J., Burnet, M., Dillon, M.O. and Gage, D.A. (1994). Osmoprotective compounds in the Plumbaginaceae: a natureal experiment in metabolic engineering of stress tolerence. *Proc. Natl. Acad. Sci. USA.* **91,** 306-310.

Hansen, D.H., Wittstock, U., Olsen, C.K., Hick, A.J., Pickett, J.A., Halkier, B.A. (2001) Cytochrome P450 CYP79 F1 from Arabidosis catalyzes the conversion of dihomomethionine and trihomomethionine to the corresponding aldoximes in the biosynthesis of aliphatic GIs. *J. Biol. Chem.* **276,** 11078-11085.

Haughn, G.W., Davin, L., Giblin, M. and Underhill, E.W. (1991) Biochemical genetics of plant secondary metabolites in *Arabidopsis thaliana*: The Glucosinolates. *Plant Physiol.* **97,** 217-226.

Hobkirk, R (1993) Steroid sulfation: current concepts. *Trends Endocrinol. Metab.* **4,** 69-74.

Hogge, L.R., Reed, D.W., Underhill, E.W. and Haughn GW (1988) HPLC separation of Glucosinolates from leaves and seeds of *Arabidopsis thaliana* and their identification using thermospray liquid chromatography/mass spectrometry. *J. Chromatogr. Sci.* **26,** 551-556.

Hull, A.K., Vij, R. and Celenza, J.I. (2000) Arabidopsis cytochrome P450s that catalyze the first step of tryptophan-dependent indole-3-acetic acid biosynthesis. *Proc. Natl. Acad. Sci. USA.* **97,** 2379-2384.

Jain, J.C., Grootwassink, J.W., Reed, D.W., Underhill, E.W. (1990) Purification and properties of 3'-phosphoadenosine 5'-phosphosulfate: Desulfoglucosinolate sulfotransferase from *Brassica juncae* cell cultures. *J. Plant Physiol.* **136,** 356-361.

James, D.C. and Rossiter, J.T. (1991) Development and characteristics of myrosinase in *B. napus* during early seedling growth. *Physiol Plant.* **82,** 163-170.

Kallas, P., Meier-Augenstein, W., Schildknecht, H. (1989) Turgorine-Neue phytochormone. *Naturwiss. Rundsch.* **42,** 309-317.

Kiddle, G.A., Bennett, R.N., Hick, A.J., Wallsgrove, R.M. (1999) C-S lysase activities in leaves of crucifers and non-crucifers, and the characterization of three classes of C-S lysase activities from oilseed rape (*Brassica napus* L.). *Plant Cell Environ.* **22,** 433-445.

Kiddle, G.A., Doughty, K.L., Wallsgrove, R.M. (1994) Salicylic acid-induced

accumulation of Glucosinolates in oilseed rape leaves. *J. Exp. Biol.* **45,** 1343-1346.

Kindle, H. and Underhill, E. W. (1968) Biosynthesis of mustard oil glucosides: N-hydroxyphenylalanine, a precursor of glucotropaeoline and a substrate for the enzymatic and nonenzymatic formation of phenylacetaldehyde oxime. *Phytochemistry.* **7,** 745-756

Kjaer, A. and Larsen, P. (1973) Nonprotein amio acids. Cyanogenic glycosides and GIs. In: biosynthesis. Geissman, T. A. (ed). Vol. 2, pp. 71-105. Specialist periodical reports. *London: The Chemical Society.*

Kliebenstein, D.J., Lambrix, V.M., Reichelt, M., Gershenzon, J. and Mitchell-Olds, T. (2001) Gene duplication in diversification of secondary metabolism: tandem 2-oxoglutarate-dependent dioxygenases control Glucosinolate biosynthesis in *Arabidopsis thaliana. Plant Cell.* **13,** 681-693.

Koritsas, V.M., Lewis, J.A. and Fenwich, G.R. (1991). Glucosinolate response of oilseed rape, mustard and kale to mechanical wounding and infestation by cabbage stem flea beetle (*Psyliodes chysocephala*). *Ann. Appl. Biol.* **118,** 209-221.

Komatsu, K., Driscoll, W.J., Koh, Y.C. and Strott, C.A. (1994) A P-loop related motif (GxxGxxK) highly conserved in sulfotransferases is required for binding the activated sulfate donor. *Biochem. Biophys. Res. Commun.* **204,** 1178-1185.

Kuracek, M., Prochazka, Z. and Veres, K. (1962) Biogenesis of glucobrassicin, the in vitro precusor of ascorbigen. *Nature.* **194,** 393-394.

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

Lenman, M., Falk, A., Rodin, J., Hoglund, A-S., Ek, B. and Rask, L. (1993) Differential expression of myrosinase gene families. *Plant Physiol.* **103,** 703-711.

Li, Y.C., Kiddle, G., Bennett, R., Wallsgrove, R.M. (1999) Local and systemic changes in Glucosinolates in Chinese and European cultivars of oilseed rape (*Brassica napus* L.) after inoculation with *Sclerotinia sclerotiorum* (stem rot). *Ann. Appl. Biol.* **134,** 45-58.

Long, S. (1989) *Rhizobium*-legume nodulation: Life together in the underground. *Cell.* **56,** 203-214.

Lykkesfeldt, J. and Moller, B.L. (1993) Synthesis of benzylGlucosinolate in *Tropaeolum majus* L.: isothiocyanates as potent enzyme inhibitors. *Plant Physiol.* **102,** 609-613.

Luda, S., and Mole, S (1991) Glucosinolates: Chemistry and ecology. In Herbivores: Their interaction with Secondary Plant Metabolites. Vol.1. The Chemical Paticipants, 2nd ed, G.A. Rosenthal and M.R. Berenbaum, eds (San Diego, CA: Academic Press), pp. 123-164.

Ludwig-Mueller, J., Bennett, R., Kiddle, G., Ihmig, S., Ruppel, M. and Hilgenberg, W. (1999) The host range of *Plasmodiophora brassicae* and its relationship to endogenous. Glucosinolate content. *New phytol.* **141**, 443-458.

Ludwig-Mueller, J., Rausch, T., Lang, S. and Hilgenberg, W. (1990) Plasma membrane bound high plant isoenzymes convert tryptophan to indole-3-acetaldoxime. *Phytochemisty* **29**, 1397-1400.

Magrath, R. and Mithen, R. (1993) Maternal effects on the expression of individual aliphatic Glucosinolates in seeds and seedlings of *Brassica napus. Plant breeding.* **111**, 249-252.

Marsolais, F., and Varin, L. (1995) Identification of amino acid residues critical for catalysis and cosubstrate binding in the flavonol 3-sulfotransferase. *J. Biol. Chem.* **270**, 30458-30463.

Marsolais, F. and Varin, L. (1997) Mutational analysis of domain II of flavonol 3-sulfotransferase. *Eur. J. Biochem.* **247**, 1056-62.

Marillia, E.F., MacPheson, J.M., Tsang, E.W., Van Audenhove, K., Keller, W.A. and GrootWassink, J.W. (2001) Molecular cloning of a Brassica napus thiohydroximate S-glucosyltransferase gene and its expression in *Escherichia coli. Physiol Plant.* **113(2)**, 176-184.

Matsuo, M. and Underhill, E.W. (1971) Purification and properties of a UDP glucose: thiohydroximate glucosyltransferase form higher plants. *Phytochemisty.* **10**, 2279-2288.

Merritt, S.Z. (1996) Within-plant variation in concentrations of amino acids, sugar, and sinigrin in phloem sap of black mustard, *Brassica nigra* L. *Koch. J. Chem. Ecol.* **22**, 1133-45.

Mikkelsen, M.D., Hansen, C.H., Wittstock, U. and Halkier, B.A. (2000) Cytochrome P450 CYP79B2 from Arabidopsis catalyzes the conversion of tryptophan to indole-3-acetaldoxime, a precusor of indoleglucosinolates and indole-3-acetic acid. *J. Biol. Chem.* **275**, 33712-33717.

Mikkelsen, M.D., Petersen, B.L., Glawischnig, E., Jensen, A.B., Andreasson, E. and Halkier, B.A. (2003) Modulation of *CYP79* Genes and Glucosinolate Profiles in Arabidopsis by Defense Signaling Pathways. *Plant Physiol.* **131(1)**, 298-308

Mollers, C., Nehlin, L., Glimelius, K. and Iqbal, M.C.M. (1999) Influence of in vitro culture conditions on Glucosinolate composition of microsopore-derived embryos of *Brassica napu. Physiol. Plant.* **107,** 441-446.

Normally, J. and Bartel, B. (1993) Redundancy as a way of life-IAA metabolism. *Curr. Opin. Plant Biol.* **2**, 207-213.

Poulton, J.E. and Moller, B.L. (1993) GLucosinolates. In *Methods in plant biochemistry* (Acadamic Press, San Diego, CA), pp. 209-237.

Porter, A.J.R., Mortor, A.M., Kiddle, G., Doughty, K.H. and Wallsgrove, R.M. (1991) Variation in the GI content of oilseed rape (*Brassica napus* L.), I. Effects of leaf age and position. *Ann. Appl. Biol.* **118,** 461-467.

Rask, L., Andreasson, E., Ekbom, B., Erikkson, S., Pontoppidan, B. and Meijer, J. (2000) Myrosinase: Gene family evolution and herbivore defense in *Brassicaceae. Plant Mol. Bio.* **42**, 93-113.

Raskin, I. (1992) Role of salicylic acid in plants. *Annu Rev. Plant Physiol. Plant Mol. Biol.* **43**, 439-463.

Reed, D.W., Davin, L., Jain, J.C., Deluca, V., Nelson, L. and Underhill, E.W. (1993) Purification and properties of UDP-glucose:thiohydroximate glucosyltransferase from *Brassica napus* L. seedlings. *Arch. Biochem. Biophys.* **305,** 526-532.

Rens Domiano, S.S. and Toth, J.A. (1987). Inhibition of M and P phenol sulfotransferase by analogues of 3'-phosphoadenosine-5'-phosphosulfate. J. *Neurochem.* **48**, 1411-1415.

Reintanz, B., Lehnen, M., Reichelt, M., Gershenzon, J., Kowalczyk, M., Sandberg, G., Godde, M., Uhl, R. and Palme, K. (2001) Bus, a bushy Arabidopsis *CYP79F1* knockout mutant with abolished synthesis of short-chain aliphatic Glucosinolates. *The plant cell.* **13**: 351-367

Rivoal, J. and Hanson, A.D. (1994) Choline-*O*-sulfate biosynthesis in plants. *Plant Physiol.* **106,** 1187-1193.

Rodman, J.E. (1991) A taxonomic analysis of Glucosinolate producing plants. part 1, Phenetics. *Syst. Bot.* **1,** 598-618

Rodman, J.Z., Karol, K.G., Price, R.A. and Sytsma, K.J. (1996) Molecular, morphology, and Dahlgren's expanded order Capparales. *Syst. Bot.* **21,** 289-307

Rouleau, M., Marsolais, F., Richard, M., Nicolle, L., Voigt, B., Adam, G. and Varin, L. (1999) Inactivation of Brassinosteroid biological activity by a salicylate-

inducible steroid sulfotransferase from *Brassica napus*. *J. Biol. Chem.*. **274,** 20925-20930.

Schildknecht, H., and Meler-Augenstein, W 91990) Role of turgorins in leaf movement in *The puloinus: Moto Organ For leaf Movement* (Satter, R.L., Gorton, H.L., and Vogelman, T.C. eds) pp. 101-129. American Society of Plant Physiologists. Rockville, Md.

Selmar, D. (1999) Biosynthesis of cyanogenic glycosides, Glucosinolates and non protein amino acids. In *Biochemistry of Plant Secondary Metabolism, M. Wink, ed* (Boca Raton, Fl: CRC Press), pp. 79-150.

Soren, B, et al. (2001). *CYP83B1*, a Cytochrome P450 at the Metabolic Branch Point in Auxin and Indole glucosinolate Biosynthesis in Arabidopsis. *The plant cell.* **1,** 101-111.

Taipalensuu, J., Andreasson, E., Eriksson, S. and Rask, L. (1997) Regulation of the wound-induced myrosinase-associated protein transcript in *Brassica napus* plant. *Eur. J. Biochem.* **247**, 963-971.

Toroser, D., Grifiths, H., Wood, C. and Thomas, D.R. (1995) GI biosynthesis in oilseed rape (*Brassica napus* L): studies with ^{35}SO and gllucosinolate precusors using oilseed rape pods and seeds. *J. Exp. Bot.* **46,** 787-794.

Underhill, E.W. (1980) In *Encyclopedia* of plant biochemistry (Springer-Verlag, New York), Vol. 8, pp. 439-511.

Underhill, E. W., Chisholm, M. D. and Wetter, L. R. (1962) Biosynthesis of mustard oil glucosides. I. Administratin of C^{14}-labelled compounds to horseradish, nasturtium and watercress. *Can. J. Biochem. Physiol.* **40,** 1505-1514.

Underhill, E.W. and Weter, L.R. (1969) Biosynthesis of mustard oil glusides: sodium phenylacetothiohydroximate and desulfobenzylglucosinolate, precursors of benzylglucosinolate in *Tropaeolum majus* L. *Plant Physiol.* **44**, 584-590.

Underhill, E.W., Wetter, L.R. and Chisholm, M.D. (1973) Biosynthesis of Glucosinolates. *Biochem. Soc. Symp.* **38,** 303-326.

Varin, L., Chamberland, H., Lafontaine, J.G. and Richard, M. (1997) The enzyme involved in the sulfation of the turgorin, gallic acid 4-O-(β-D-glucopyranosyl-6'-sulfate) is pulvini-localized in *Mimosa pudica*. *Plant Journal.* **12**, 831-837.

Varin, L., Deluca, V., Ibrahim, R.K. and Brisson, N. (1992) Molucular characterization of flavonol sulfotransferases. *Proc. Natl. Acad. Sci. USA.* **89,** 1286-1290.

Varin, L. and Ibrahim, R.K. (1989) Partial purification and characterization of three flavonol specific sulfotransferase from *Flaveria chloraefolia. Plant physiol.* **90,** 977-981.

Varin, L. and Ibrahim, R.K. (1991) Partial purification and some properties of flavonol 7-sulfotransferases from Flaveria bidentis. *Plant physiol.* **95,** 1254-1258.

Varin, L., Marsolais, F., and Brisson, N. (1995) Chimeric flavonol sulfotransfrases define a domain responsible for substrate and position specificities. *J. Biol. Chem.* **270,** 12498-12502.

Varin, L., Marsolais, F., Richard, M. and Rouleau, M. (1997) Biochemistry and molecular biology of plant sulfotransferase. *FASEB.* **11,** 517-525.

Visvalingam, S., Hinsi, T.G. and Bones, A.M. (1998) Sulphate and micronutrients can modulate the expression levels of myrosinase in *Sinapus alba* plants. *Physiol. Plant.* **104,** 30-37.

Weinshilboum, R. and Otterness, D.B.S. (1994) Sulfotransferase enzymes. In: FC Kauffman (ed) Handbook of Experimental Pharmacology: conjugation-Deconjugation Reactions in Drug Metabolism and Toxicity. Vol 112, pp. 45-78. *Springer-Verlag, New York.*

Wittstick, U., and Halkier, B.A. (2002) Glucosinolate research in the Arabidopsis era. *Trends in Plant Sci.* 7, 263-270

Wittstock, U., and Halkier, B.A. (2000) Cytochrome P450 CYP79A2 from *Arabidopsis thaliana* L. catalyzes the conversion of L-phenylalanine to phenylacetaldoxime in the biosynthesis of benzylglucosinlate. *J. Biol. Chem.* **275,** 14659-14666.

Xu, Y., Chang, P.F.L., Narasimhan, M.L., Raghothama, K.G., Hasegawa, P.M. and Bressan, R.A. (1994) Plant defense genes are synergistically induced by ethylene and methyl jamonate. *Plant Cell.* **6,** 1077-1085.

Zhang, Y., Talalay, P., Cho, C-G and Posner, G.H. (1992) A major inducer of anticarcinogenic protective enzymes from brocholi: isolation and elucidation of structure. *Proc. Natl. Acad. Sci. USA* **89,** 2399-2403.