# Chapter 6

# Sulfotransferases from Plants, Algae and Phototrophic Bacteria

Cinta Hernández-Sebastià and Luc Varin*
*Centre for Structural and Functional Genomics, Department of Biology, Concordia University, 7141 Sherbrooke West, Montréal, Canada H4A 1R6*

Frédéric Marsolais
*Agriculture and Agri-Food Canada, Southern Crop Protection and Food Research Centre, 1391 Sandford St., London, Ontario, Canada N5V 4T3*



## Summary

Sulfotransferases (SULTs) catalyze the transfer of a sulfuryl group ($SO_3$) from the universal donor 3′-phosphoadenosine 5′-phosphosulfate to a hydroxyl group of various substrates in a process called the sulfonation reaction. These enzymes share highly conserved sequence regions across all kingdoms;

* Corresponding author, Fax: 514 848 2881, E-mail: varinl@alcor.concordia.ca

*Rüdiger Hell et al. (eds.), Sulfur Metabolism in Phototrophic Organisms,* 111–130.

however, their substrates and physiological function are predicted to be very diverse. In mammals, the sulfonation reaction is involved in mechanisms of cellular detoxification and in the modulation of the biological activity of steroid hormones and neurotransmitters. In plants, few SULTs have been characterized despite the large number of sequences found in databases. Understanding of their role in plant biology is still relatively speculative. Since the cloning of the first plant SULT cDNA more than 10 years ago from *Flaveria chloraefolia*, functional genomic approaches, particularly in *Arabidopsis thaliana*, have led to the characterization of brassinosteroid, hydroxyjasmonate, desulfoglucosinolate and flavonoid SULTs. However, these efforts are hindered by a limited knowledge of plant sulfated metabolites. A detailed analysis of plant SULT sequences revealed that at present, phylogeny is of limited value to predict biochemical function. For instance, *A. thaliana* desulfoglucosinolate SULTs form a clade with the *Flaveria* flavonol SULTs. Ten of the 17 *A. thaliana* SULTs belong to a single homogeneous clade, suggesting that most of the divergence occurred after the diversification of this plant lineage. We present here a detailed overview of the molecular phylogeny, characterization, and biological roles of SULTs in plants. We also describe the current state of knowledge of sulfonation in algae, as well as in phototrophic bacteria, where the SULT domain can be present in multidomain proteins.

## I. Introduction

Sulfotransferases (SULTs) catalyze the transfer of a sulfuryl group ($SO_3$) from the universal donor 3′-phosphoadenosine 5′-phosphosulfate (see Chapter 3 in this book) to a hydroxyl group of various substrates in a process called the sulfonation reaction (Fig. 1). There are two classes of SULTs; cytosolic and membrane-associated SULTs. Cytosolic SULTs sulfonate small organic molecules such as steroids, flavonoids, glucosinolates and hydroxyjasmonates, to mention a few. Membrane-associated SULTs sulfonate larger biomolecules such as complex carbohydrates, peptides and proteins. Figure 2 illustrates the structural diversity of selected sulfated metabolites found in plants.

The modification of metabolites by the addition of a sulfonate group can have a profound influence on their biological properties. Initially, the cytosolic SULTs were thought to be primarily involved in detoxification of endogenous and exogenous metabolites. They were considered as part of the detoxification arsenal, which includes hydroxylases and glucuronidases. It is now clear that this is not their only function. For example, the sulfonation of steroids has been shown to be required for the modulation of their biological activity (Clarke et al., 1982). The presence of the sulfate group on some molecules can also be a prerequisite for their biological activity. For example, it is well known that in mammals the sulfate groups are essential for the molecular interaction between heparan sulfate and antithrombin III accounting for its anticoagulation property (Atha et al., 1985). In another example pertinent to human health, the sulfonation of tyrosine residues on the chemokine receptor CCR5 is a modification that is required for its biological activity, and indirectly for the efficient binding and entry of HIV-1 (Farzan et al., 1999).

The roles of sulfated compounds in bacteria and plants are less clear. However, in a well-documented example, the sulfonation of a glycolipid secondary messenger was shown to modulate the specificity of the interaction between the bacterium *Sinorhizobium meliloti* and alfalfa (Lerouge et al., 1990). *S. meliloti* mutants lacking the SULT responsible for the sulfonation of the glycolipid are unable to induce root nodulation in alfalfa but gain the ability to colonize the roots of vetch (Roche et al., 1991). In this biological system, the sulfonation of a single glycolipid is determinant for the proper cell to cell communication between the bacterium and its natural host plant.

In plants, the sulfonation of brassinosteroids by SULTs from *Brassica napus* was shown to produce a sulfated compound lacking biological activity in a manner similar to the inactivation of estrogens in mammals (Rouleau et al., 1999). The presence of a

*Abbreviations*: DSG – desulfoglucosinolate; PAP – 3′-phosphoadenosine-5′-phosphate; PAPS – 3′-phosphoadenosine-5′-phospho sulfate; PB – 3′-phosphate binding loop; PLMF1 – periodic leaf movement factor 1; PSB – 5′-phosphosulfate binding loop; PSK – phytosulfokine; SULT – sulfotransferase; TPR – tetratricopeptide; TPSULT – tyrosylprotein sulfotransferaseI

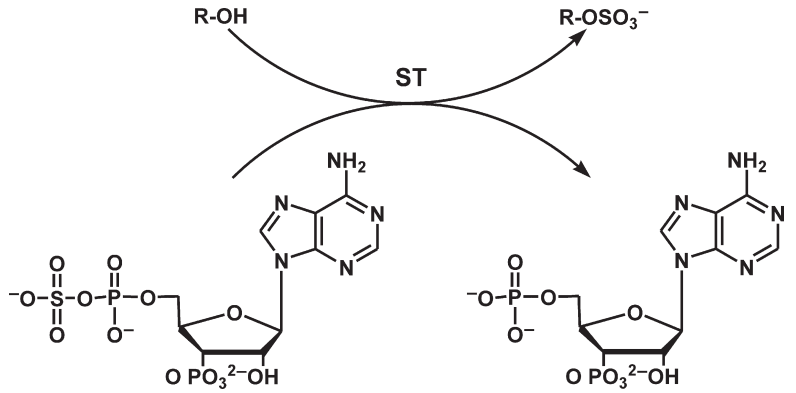
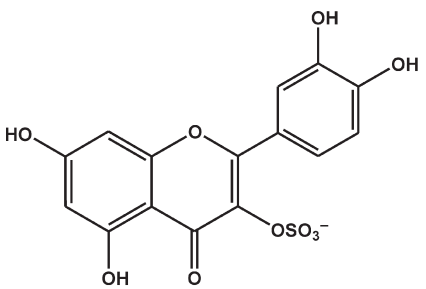
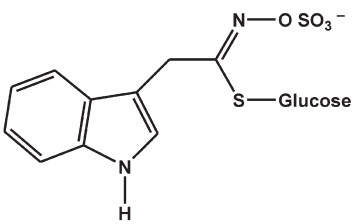


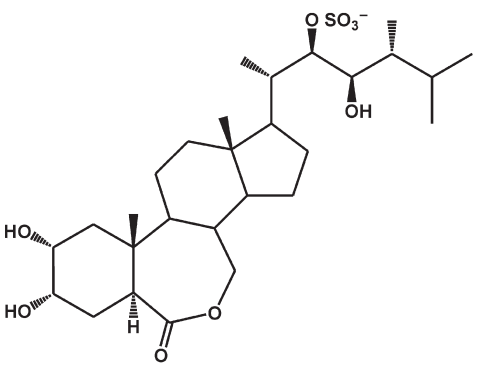

*Fig. 1*. Enzymatic reaction catalyzed by sulfotransferases.



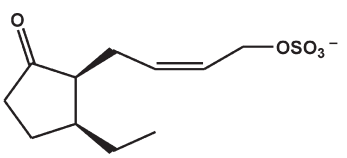

**Quercetin 3-sulfate**



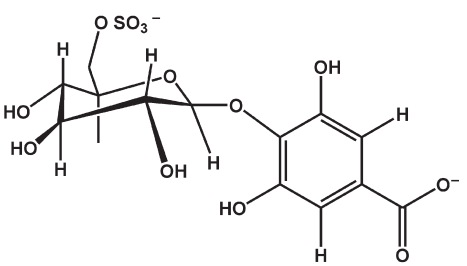

**Glucobrassicin**



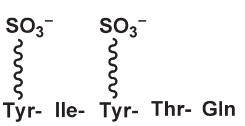

**24- epibrassinolide sulfate**




**12-hydroxyjasmonate sulfate**




**Gallic acid glucoside sulfate**




**Phytosulfokine -α**

*Fig. 2*. Molecular structure of selected plant sulfated metabolites.

sulfate group on some molecules can also be a prerequisite for their function. For example, a sulfated pentapeptide exhibiting mitogenic activity was isolated from the conditioned medium of asparagus cell culture (Matsubayashi and Sakagami, 1996). Structure–activity relationship studies revealed that the presence of two sulfated tyrosine residues was essential for the activity of the peptide hormone. Herein we present the current state of our knowledge on the molecular phylogeny, structure and biological activity of plant, algal and phototrophic bacterial SULTs.

## II. Nomenclature

Despite the fact that the literature on plant SULTs is quite limited, the nomenclature of these enzymes is suffering from the same confusion that was initially encountered with their mammalian homologs. Presently, the designations ST and SOT have been used to refer to the same enzymes (Piotrowski et al., 2004; Klein et al., 2006). Furthermore, the numbering of the SULTs in one species is leading to confusion since the same protein can get two different numbers in two different publications (e.g. AtST2a and AtSOT14).

In an effort to rationalize the nomenclature of SULTs across kingdoms, an international committee set up rules to assign names to SULTs (Blanchard et al., 2004). The committee used a system similar to the one applied successfully for the cytochrome P450 monooxygenases. An international curator is responsible to assign final names and can be contacted at rl_blanchard@fccc.edu. The designation SULT was chosen by the international committee and genes having more than 45% amino acid sequence identity are members of the same family and are assigned the same number. The committee assigned the numbers from 201 to 400 to designate the plant enzymes. SULT401 and higher are reserved for prokaryotes. Members of the same family sharing 60% or more amino acid sequence identity are classified in the same subfamily and are assigned the same letter following the family number. For example, the *Flaveria chloraefolia* flavonol 4′-sulfotransferase was assigned the designation SULT201A2. This enzyme belongs to family 201 and subfamily A, which comprises two other members, the flavonol 3-SULT from the same species and the flavonol 3-SULT from *F. bidentis*. In order to differentiate between orthologs, the family number is preceded by the abbreviation of the species name, for example, (FLACH) for *F. chloraefolia* and (FLABI) for *F. bidentis*. The family number 202 was assigned to the *B. napus* brassinosteroid SULTs, previously called BnST1 to -4. Since the four genes share more than 60% identity at the amino acid level, they were assigned to the same subfamily A. One protein from *A. thaliana* encoded by the locus *At2g03760* (previously called *RaR047* or *AtST1*) also belongs to this subfamily. The hydroxyjasmonate SULT from *A. thaliana* (initially called AtST2a) encoded by the locus *At5g07010* was assigned to family 203. Table 1 shows the result of the application of this nomenclature to the *A. thaliana* SULTs for which the biochemical function is known and to a few related sequences from other plant species.

## III. Molecular Phylogeny of Plant Sulfotransferases

The availability of complete or near complete genome sequences for an increasing number of species renders possible to carry out thorough and detailed large-scale analysis of the plant SULT superfamily. The fully sequenced *A. thaliana* genome contains 18 SULT-coding genes including 1 apparent pseudogene (www.arabidopsis.org) (Table 2). For other plants with genomic information (e.g. rice and poplar), SULTs also seem to be encoded by comparatively small gene families. This is contrasting with the size of the glycosyltransferase family, which comprises 120 members in *A. thaliana* including eight apparent pseudogenes (Gachon et al., 2005). These results indicate that the modification of metabolites by sulfonation is a relatively rare event as compared to glycosylation. However, the number of SULTs in plants is significantly higher than in mammals (Table 2).

In spite of the rapid progress in the acquisition of genomic information from plants, only 13 genes encoding plant SULTs have been characterized at the biochemical level and only a handful have been characterized *in planta*. To date, more than 200 plant SULT sequences can be retrieved from public databases. However, their alignment reveals that a smaller number appear

*Table 1*. Nomenclature of biochemically characterized *Arabidopsis thaliana* sulfotransferases and few selected sequences from other plant species.

| Sulfotransferase family | Sulfotransferase subfamily and number | Accession number | Biochemical function |
|---|---|---|---|
| SULT201 | A1 | Fc_M84135 | Flavonol 3-SULT |
| | A2 | Fc_M84136 | Flavonol 4′-SULT |
| | A1 | Fb_10277 | Flavonol 3-SULT |
| | B1 | At1G74090 | Desulfoglucosinolate SULT |
| | B2 | At1G18590 | Desulfoglucosinolate SULT |
| | B3 | At1G74100 | Desulfoglucosinolate SULT |
| SULT202 | A1 | Bn_AF000305 | Brassinosteroid SULT |
| | A2 | Bn_AF000306 | Brassinosteroid SULT |
| | A3 | Bn_AF000307 | Brassinosteroid SULT |
| | A4 | At2G03760 | Brassinosteroid SULT |
| | A5 | Bn_AY442306 | Brassinosteroid SULT |
| SULT203 | A1 | At5G07010 | Hydroxyjasmonate SULT |
| | A2 | At5G07000 | Unknown |
| | A3 | Pt_LG11001725 | Unknown |

*Table 2*. Number of sulfotransferase sequences found in a selected number of species.

| Species name | Total number of unique sequences | Putative full-length sequences |
|---|---|---|
| *Arabidopsis thaliana* | 18 | 17 |
| *Oryza sativa* | 35 | 31 |
| *Populus trichocarpa* | 42 | 32 |
| *Homo sapiens* | 11 | 11 |

to contain a full-length open reading frame. In an attempt to predict biochemical function based on sequence homology, we constructed an unrooted cladogram from 78 plant SULT sequences (Fig. 3). The sequences used to generate the cladogram presented in Fig. 3 were selected according to the following criteria. First, the regions I to IV conserved in all cytosolic SULTs had to be present in the sequence (Varin et al., 1992). Second, only SULTs that could be assembled into full-length sequences were included in the alignment and in the phylogenetic tree. Finally, sequences containing large gaps that might have originated from sequencing errors were eliminated from the dataset. We also generated cladograms from the alignment of the conserved region I or conserved region IV. In all cases, the trees looked similar to the one presented in Fig. 3 (data not shown).

Ten of the 17 *A. thaliana* SULT proteins are present in the same homogeneous clade (Fig. 3). In contrast, the rice and poplar SULTs are distributed in several clades containing sequences from at least two other species. The presence of the majority of the SULT proteins from *A. thaliana* in a homogeneous clade suggests that most of the divergence occurred after the diversification of this plant lineage. The remaining seven sequences are distributed in three clades containing proteins from at least three other species.

As expected, the brassinosteroid SULT enzymes from *B. napus* are present in the clade which also comprises *A. thaliana* At2g03760. The latter has recently been shown to be a brassinosteroid SULT (Marsolais et al., 2007) (see section V.A.4.a). The uncharacterized protein At2g03770 from *A. thaliana* is also present in this clade. The *At2g03770* locus is located adjacent to *At2g03760* on chromosome 2 and probably originates from a recent duplication event. Its position in the brassinosteroid SULT clade suggests that it might also be a brassinosteroid SULT. A number of SULTs from rice and one from barley are present in one branch of this clade. Their involvement in brassinosteroid sulfonation remains to be demonstrated. The absence of a protein from poplar in this clade is surprising considering that its genome is almost fully sequenced and that a conservation of the brassinosteroid SULT function is expected. However, a functional homolog from poplar might be represented in the few genes that had to be excluded following the application of the sequence selection criteria.

To date, all monocotyledonous and dicotyledonous plants that were tested were found to accumulate sulfated hydroxyjasmonates. This

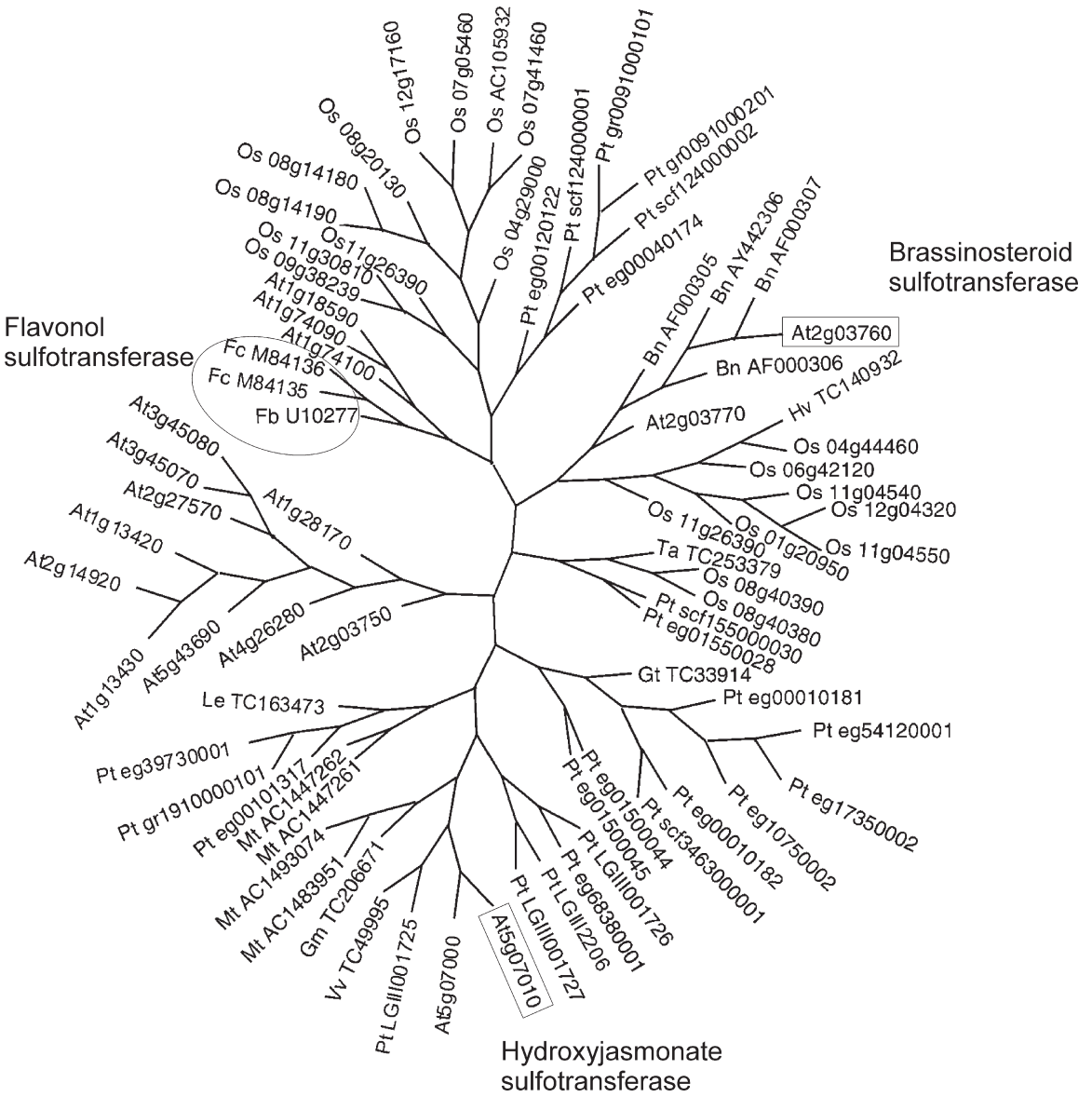



*Fig. 3*. Cladogram showing the relationship between 78 plant SULT sequences. The accession number is preceded by the initial of the plant species in which they were identified. Abbreviations for species: At, *Arabidopsis thaliana*; Bn, *Brassica napus*; Fb, *Flaveria bidentis*; Fc, *Flaveria chloraefolia*; Gm, *Glycine max*; Gt, *Gentiana triflora*; Hv, *Hordeum vulgare*; Le, *Lycopersicon esculentum*; Mt, *Medicago truncatula*; Os, *Oryza sativa*; Pt, *Populus trichocarpa*; Ta, *Triticum aestivum*; Vv *Vitis vinifera*. The *Arabidopsis* and *Flaveria* genes which were characterized at the biochemical level are highlighted. The sequences were aligned with ClustalW and the cladogram was generated using PAUP 4.0b10 (Ativec) by a neighbor joining method.

suggests that the SULT enzyme involved in their synthesis is conserved among species. Recently, the *A. thaliana* hydroxyjasmonate SULT (AtST2a) was shown to be encoded by the locus *At5g07010* (Gidda et al., 2003). Interestingly, this enzyme is located in a clade comprising members from poplar, grape, alfalfa and soybean (Fig. 3). Its closest relative is encoded by the adjacent locus *At5g07000* on chromosome 5. Despite an 85% sequence identity with At5g07010, the At5g07000 protein did not accept hydroxyjasmonates as substrate suggesting that high sequence identity alone is not a good predictor of substrate preference. Alternatively, the absence of

activity of At5g07000 with hydroxyjasmonates may have been due to the improper folding of the recombinant enzyme. The second and third closest relatives of At5g07010 are from poplar and grape with 60% and 58% identity, respectively. It will be interesting to test their enzymatic activity with hydroxyjasmonates.

Surprisingly, the closest relatives to the three SULTs involved in glucosinolate biosynthesis in *A. thaliana* are those involved in flavonoid sulfonation in *Flaveria* species (Fig. 3). This result illustrates one of the limitations of molecular phylogeny studies to predict protein function. It is well known that few amino acid changes in SULT sequences can lead to profound changes in substrate specificity (Varin et al., 1995; Sakakibara et al., 1998). However, it is possible that the nature of the desulfoglucosinolate and flavonoid substrates impose similar constraints on the architecture of the protein. It is also possible that a common ancestral protein diverged recently to give rise to the two enzyme groups.

The utility of the cladogram to predict protein function is severely limited by the small number of SULTs that have been thoroughly characterized and by the limited number of sequences that were aligned. The biochemical characterization of more SULTs is required to better understand the evolutionary relationship between SULT genes.

## IV. Enzymatic Mechanism and Structural Requirements

The amino acid alignment of the first characterized plant SULT with animal SULTs allowed defining four conserved regions (Varin et al., 1992). The stability of these conserved sequences during evolution suggested that they are essential for the catalytic function of these enzymes. To date, the structures of six SULTs have been solved including one plant enzyme from *A. thaliana* (At2g03760) (Yoshinari et al., 2001; Smith et al., 2004). They share a common globular structure and contain a single $\alpha/\beta$ fold with a central four- or five-stranded parallel $\beta$ sheet surrounded by $\alpha$-helices. The SULT structures are similar to those of nucleotide kinases suggesting a common ancestral protein. The most important structural features of SULTs are briefly described in the following section.

### *A. 3´-Phosphoadenosine 5´-Phosphosulfate Binding Site*

Amino acids from conserved regions I and IV contribute to the PAPS binding site (Fig. 4a and d). Region I forms the 5′-phosphosulfate loop (PSB loop). Hydrogen bonds are formed between the 5′-phosphate and the N6 atom of the strictly conserved lysine residue and the oxygen of the two threonine residues located at the extremity of the PSB loop (Fig. 4a). The 3′-phosphate binding loop (PB loop) is made up of the three amino acids RKG located in region IV (Fig. 4d). Hydrogen bonds are formed between the arginine and glycine residues and the 3′-phosphate of the cosubstrate. The motif GXXGXXK located after the PB loop was shown to be essential for the binding of PAPS and estradiol in the estrogen SULT from mammals (Driscoll et al., 1995). Finally, aromatic amino acids in region II (Trp) and region III (Phe) form a parallel stack with the adenine group of PAPS (Fig. 4b and c).

### *B. Substrate Binding Site*

The flavonol 3- and 4′-SULTs catalyze the transfer of the sulfuryl group of PAPS to position 3 of flavonol aglycones and 4′- of flavonol 3-sulfates (see section V.A.1). To elucidate the structural aspects underlying the difference in substrate specificity of these enzymes, a number of hybrid proteins were constructed by the manipulation of their cloned cDNA sequences (Varin et al., 1995). Analysis of the substrate preference of the chimeric proteins indicated that a segment, designated domain II, contains all the amino acid residues responsible for their substrate and position specificities. Within this domain, two regions of high amino acid divergence corresponding to amino acids 98 to 110 and 153 to 170 were identified. A similar study conducted with the human phenol (SULT1A1) and catecholamine (SULT1A3) SULTs identified the same two variable regions as being responsible for the specificity of the two enzymes (Sakakibara et al., 1998). To refine the analysis of the determinants of specificity, a number of divergent amino acids located in the variable regions of the flavonol 3-SULT were replaced with the corresponding amino acids of the flavonol 4′-SULT (Marsolais and Varin, 1997). No reversal of the specificity was observed after the individual

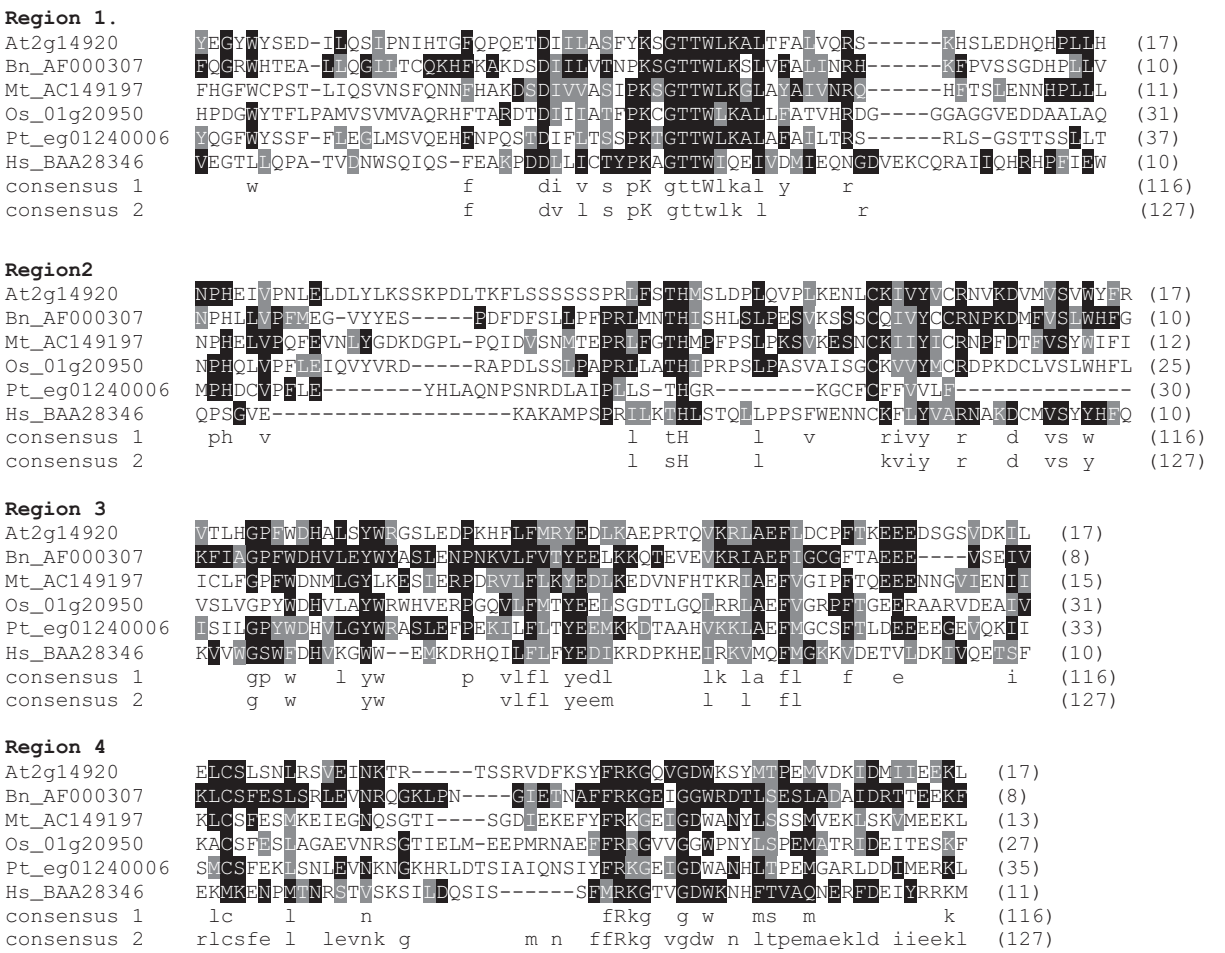



*Fig. 4*. Conservation of amino acid residues in region I to IV of plant SULTs. Sequences from individual species were aligned separately using ClustalW. The level of conservation is defined by Boxshade 3.21 with a threshold value of 0.9 for black boxes. The number of genes from each species to generate the alignment is shown in the parentheses. Only one gene from each species is illustrated on the figure. Species: *A. thaliana* (At), *Brassica napus* (Bn), *Medicago trunculata* (Mt), *Oryza sativa* (Os) and *Populus trichocarpa* (Pt). Residues conserved in more than 90% of the sequences are shown in black while those conserved in more than 50% of the sequences are shown in gray. Consensus 1 is derived from the alignment of 116 plant sequences. Consensus 2 is derived from the alignment of 116 plant and 11 human SULT sequences.

mutation. However, replacement of leucine 95 of the flavonol 3-SULT by the corresponding tyrosine of the 4′-SULT had different effects on the kinetic constants depending on the flavonoid ring B structure suggesting that the tyrosine side chain may be in direct contact with this part of the molecule. Modeling studies of the flavonol 3- and 4′-SULTs on the structure of the estrogen sulfotransferase showed that the leucine residue of the flavonol 3-SULT and the tyrosine residue of the 4′-SULT are in close proximity with the catalytic site supporting their role in substrate binding (data not shown).

## *C. Catalytic Mechanism*

A great deal of information on the catalytic mechanism of SULTs has been derived from the solved crystal structures and from the results of numerous site-directed mutagenesis studies (Marsolais and Varin, 1998; Chapman et al., 2004). The sulfonation reaction proceeds by an in-line attack of the sulfate group of PAPS. The conserved histidine residue in region II probably assists in the reaction by deprotonating the attacking substrate. Alternatively, the histidine at the active site might act as a nucleophile to form

an unstable protein–sulfate complex. The lysine residue located in region I is also clearly involved in catalysis. Initially, this residue is interacting with a conserved serine located in region II. Upon catalysis, the lysine side chain probably interacts with the leaving sulfate group of PAPS (Kakuta et al., 1998). Even a conservative replacement of the lysine with arginine led to a significant reduction in catalytic efficiency without affecting the binding of PAPS supporting its role in catalysis (Marsolais and Varin, 1995).

## V. Functional Characterization of Plant Sulfotransferases

### A. Cytosolic Plant Sulfotransferases

Cytosolic SULTs sulfonate small organic molecules such as flavonoids, brassinosteroids, glucosinolates and hydroxyjasmonates. The term cytosolic can be misleading since in most cases, the subcellular localization of theses enzymes is unknown. The term cytosolic is in fact referring to the ability of the enzymes to be extracted from cells in a soluble form. Although cytosolic SULTs were initially thought to be primarily involved in detoxification, it is now clear that this is not their main function. The next section describes our current knowledge of the plant SULTs and focuses on the biological significance of the formation of sulfated conjugates.

#### 1. Flavonoid Sulfotransferases from Flaveria Species

The position-specific flavonol SULTs from *F. chloraefolia* were the first plant SULTs to be characterized. These enzymes exhibit strict specificity for position 3 of flavonol aglycones, 3′ and 4′ of flavonol 3-sulfates and 7 of flavonol 3,3′- and 3,4′-disulfates (Fig. 2). Together, they participate in the sequential sulfonation of flavonol polysulfates (Varin and Ibrahim, 1989, 1991). The highly purified enzymes were active as monomers, did not require divalent cations for activity and had a similar molecular mass of 35 kDa. The $K_m$ values of the four enzymes for PAPS and the flavonol acceptors varied between 0.2 and 0.4 μM.

The flavonol 3-SULT was purified to apparent homogeneity and its kinetic properties were extensively studied (Varin and Ibrahim, 1992). The results of substrate interaction kinetics and product inhibition studies are consistent with an ordered Bi Bi mechanism, where PAPS is the first substrate to bind to the enzyme and PAP the last product to be released. These results are similar to those obtained for several mammalian SULT enzymes (Chapman et al., 2004).

Despite the early discovery and abundance of flavonol sulfates in *Flaveria* species, their biological significance remains to be elucidated. One report suggests that sulfated flavonols might play a role in the regulation of polar auxin transport (Jacobs and Rubery, 1988). Flavonol aglycones, such as quercetin and kaempferol, were found to bind to the naphthylphthalamic acid receptor and thus inhibit polar auxin transport from the basal end of stem cells. In contrast, sulfated flavonols act as antagonists of quercetin and thus allow auxin efflux from tissues where it is produced (Faulkner and Rubery, 1992).

#### 2. Choline Sulfotransferase from Limonium sativum

Choline sulfate accumulates in all the species of the salt stress-tolerant Plumbaginaceae family investigated to date (Hanson et al., 1994) (Fig. 2). It has been proposed that choline sulfate and other betaines act as osmoprotectants in response to salinity or drought stress. *In vivo* feeding experiments with *Limonium* species showed that [$^{14}$C] choline was converted to the sulfated derivative suggesting the presence of a choline SULT (Hanson et al., 1991). The enzyme was identified and characterized *in vitro* from root extracts of *Limonium sinuatum* (Rivoal and Hanson, 1994). The fact that more than 98% of the choline SULT activity could be recovered in the supernatant following high speed centrifugation suggests that it is cytosolic. The choline SULT activity was found to be strictly dependent on PAPS and showed a pH optimum of 9.0. The apparent $K_m$ values were of 5.5 μM for PAPS and 25 μM for choline. The presence of the choline SULT was investigated in a number of species including non-accumulators of choline sulfate such as barley, maize, sunflower and *Brassica* spp., and was found to be restricted to members of the genus *Limonium*. In spite of the potential of this enzyme for the engineering of plants with increased salt tolerance, no attempts have been made to isolate a cDNA clone encoding this activity.

### 3. Brassinosteroid Sulfotransferases from Brassica napus

Two members of the plant sulfotransferase superfamily, SULT202A3 and SULT202A5 from *B. napus* (initially called BnST3-4) were found to catalyze the *in vitro* sulfonation of brassinosteroids, as well as mammalian estrogenic steroids and hydroxysteroids (Rouleau et al., 1999; Marsolais et al., 2004) (Fig. 2). SULT202A3 and SULT202A5 are stereospecific for 24-epibrassinosteroids, with a substrate preference for the metabolic precursor 24-epicathasterone. Based on the lack of biological activity of 24-epibrassinolide sulfate, and the induction of *B.napus SULT* genes by salicylic acid, a function in brassinosteroid inactivation was initially hypothesized for these sulfotransferases (Rouleau et al., 1999). However, SULT202A3 and SULT202A5 exhibit a relatively broad specificity towards steroids *in vitro*. Besides salicylic acid, the *B. napus SULT* genes are inducible by ethanol and other xenobiotics, including the herbicide safener naphthalic anhydride (Marsolais et al., 2004), which would be compatible with a general function of the sulfotransferases in detoxication. The observation that constitutive expression of SULT202A3 in transgenic *A. thaliana* did not lead to a brassinosteroid-deficient phenotype suggested that the steroid sulfotransferase is not directly involved in brassinosteroid inactivation (Marsolais et al., 2004).

### 4. Cytosolic Sulfotransferases from Arabidopsis thaliana

The genome of *A. thaliana* contains a total of 18 sulfotransferase-coding genes based on sequence similarity with previously characterized SULTs from animals and plants. Of these, one (*At3g51210*) is probably a pseudogene since it encodes a truncated protein lacking the strictly conserved region I. To date, seven *A. thalina* SULTs have been characterized at the biochemical level.

#### a. Brassinosteroid SULTs (At2g03760 and At2g14920)

##### At2g03760

*SULT202A4* (*At2g03760, RaR047*) was first characterized as a gene up-regulated in response to pathogens and the pathogen-related signals, methyl jasmonate and salicylic acid (Lacomme and Roby, 1996). As can be seen in Fig. 3, *At2g03760* is an ortholog of the *B. napus SULT* genes in *A. thaliana*. *At2g03760* shares 87% of amino acid sequence identity with its closest relative SULT202A5. Together, At2g03760 and the *B. napus* SULT proteins belong to the SULT202 family of plant soluble SULTs, according to the proposed guidelines for SULT nomenclature (Table 1) (http://www.fccc.edu/research/labs/blanchard/sult/) (Marsolais et al., 2000; Blanchard et al., 2004). *At2g03760* is located on chromosome 1 and clustered with two other SULT genes of unknown function, *At2g03770*, which is also part of the SULT202 family (53% identity with *At2g03760* in amino acid sequence), and the more distant *At2g03750* (Fig. 3).

Recombinant SULT202A4 displayed similar substrate specificity towards brassinosteroids as the previously characterized SULT202A3 and SULT202A5 (Rouleau et al., 1999; Marsolais et al., 2007). Among naturally occurring brassinosteroids, the enzyme was stereospecific for 24-epibrassinosteroids, since no activity was observed with the 24-epimers castasterone and brassinolide. The preferred substrates of SULT202A4 were 24-epicathasterone ($V_{max}/K_m$ 8.3 pkatal mg$^{-1}$ μM$^{-1}$), followed by 6-deoxo-24-epicathasterone ($V_{max}/K_m$ 2.2 pkatal mg$^{-1}$ μM$^{-1}$) (Marsolais et al., 2007). SULT202A3 was previously hypothesized to transfer the sulfonate group at position 22 of 24-epibrassinosteroids, based on its lack of catalytic activity with the synthetic 22-deoxy-24-epiteasterone, and the fact that estrogens are sulfonated at position 17, since 17β-estradiol 3-methyl ether was conjugated but not estrone (Rouleau et al., 1999). Like SULT202A5 (Marsolais et al., 2004), SULT202A4 displayed a significant but low catalytic activity with the synthetic 22-deoxy-24-epiteasterone. Like SULT202A3, SULT202A4 catalyzed the sulfonation of both hydroxysteroids and estrogens (Rouleau et al., 1999), whereas SULT202A5 only accepted hydroxysteroids (Marsolais et al., 2004). The apparent $K_m$ for PAPS of SULT202A4 was 3 μM.

##### At2g14920

At2g14920 was recently identified and characterized as a novel enzyme which sulfonates brassinosteroids *in vitro* (Marsolais et al., 2007). Recombinant At2g14920 was screened against

a wide array of potential substrates, including phenolic acids, desulfoglucosinolates, flavonoids, steroids, gibberellic acids, cytokinins, phenylpropanoids, hydroxyjasmonates and coumarins. Catalytic activity was detected exclusively with brassinosteroids. In contrast with SULT202A4, no enzymatic activity was observed towards hydroxysteroids or estrogens. Among brassinosteroids tested, At2g14920 was specific for biologically active end-products of the biosynthetic pathway, including castasterone, brassinolide, related 24-epimers, and the naturally occurring (22*R*, 23*R*)-28-homobrassinosteroids. The preferred substrates of the enzyme were (22*R*, 23*R*)-28-homocastasterone ($V_{max}/K_m$ 2.7 pkatal $mg^{-1}$ $\mu M^{-1}$), followed by (22*R*, 23*R*)-28-homobrassinolide ($V_{max}/K_m$ 1.8 pkatal $mg^{-1}$ $\mu M^{-1}$). The apparent $K_m$ of the enzyme for PAPS was 0.4 µM (Marsolais et al., 2007).

The *A. thaliana* genome contains two closely related genes, *At1g13420* and *At1g13430*, located in tandem on chromosome 1. At1g13430 is more closely related to At2g14920 (84% identical in amino acid sequence) than At1g13420 (69%). It is possible that *At1g13420* and *At1g13430* also encode brassinosteroid SULTs, given the high degree of sequence relatedness.

### b. Hydroxyjasmonate SULT (At5g07010)

12-Hydroxyjasmonate, also known as tuberonic acid, was first isolated from *Solanum tuberosum* and was shown to have tuber-inducing properties (Yoshihara et al., 1989) (Fig. 2). It is derived from the ubiquitously occurring jasmonic acid, an important signaling molecule mediating diverse developmental processes and plant defense responses. It has recently been shown that 12-hydroxyjasmonate and its sulfated derivative occur naturally in *A. thaliana* (Gidda et al., 2003). Treatments with methyl jasmonate led to an increase in the amount of 12-hydroxyjasmonate and 12-hydroxyjasmonate sulfate in this plant, and it has been proposed that the metabolism of jasmonic acid to 12-hydroxyjasmonate sulfate might be a route leading to its inactivation.

The enzyme catalyzing the sulfonation of 12-hydroxyjasmonate is encoded by the gene *SULT203A1* (At5g07010, previously *AtST2a*) (Gidda et al., 2003) (Table 1). The recombinant SULT203A1 protein was found to exhibit strict specificity for 11- and 12-hydroxyjasmonate with $K_m$ values of 50 and 10 µM, respectively. The $K_m$ value for PAPS was found to be 1 µM. *SULT203A1* expression was induced following treatment with methyljasmonate and 12-hydroxyjasmonate. In contrast, the expression of the methyljasmonate-responsive gene *Thi2.1*, a marker gene in plant defense responses, is not induced upon treatment with 12-hydroxyjasmonate indicating the existence of independent signaling pathways. The presence of two independent response pathways suggests that the function of SULT203A1 might be to directly control the biological activity of 12-hydroxyjasmonate.

*At5g07010* is localized on chromosome 5 and is clustered with the locus *At5g07000*. The latter encodes SULT203A2 that shares 85% amino acid identity with SULT203A1. SULT203A2 was also expressed in *Escherichia coli* and could be recovered in a soluble form (Gidda et al., 2003). However, no activity was observed with any of the substrates tested including 11- and 12-hydroxyjasmonate. Other substrates structurally related to 11- and 12-hydroxyjasmonate will need to be tested to elucidate the biochemical function of this enzyme.

### c. Desulfoglucosinolate SULTs (At1g74100, At1g74090 and At1g18590)

Glucosinolates are sulfated, non-volatile thioglucosides derived from aliphatic, indolyl or aromatic amino acids (Fig. 2). They are found throughout the order Capparales (Cronquist, 1988), particularly in the Brassicaceae family, which includes important crop plants such as *B. napus* and the model species *A. thaliana*. Upon mechanical damage, infection or pest attack, cellular breakdown exposes the stored glucosinolates to catabolic enzymes (e.g. myrosinases), yielding a variety of reactive products such as isothiocyanates, organic nitriles, thiocyanates and oxazolidine-2-thiones (Bones and Rossiter, 1996). These catabolites contribute to the distinctive flavor and aroma of cruciferous plants and are believed to play an important role in plant protection against herbivores and pathogen attack (Chew, 1988). In addition, the glucosinolate degradation products have toxic effects on animals and humans (Fenwick et al., 1983). The last step in glucosinolate biosynthesis is catalyzed by a desulfoglucosinolate SULT (DSG-SULT). A DSG-SULT has been partially purified from

*Lepidium sativum* (Glendening and Poulton, 1990) and *Brassica juncea* (Jain et al., 1990), and its activity was detected in cell-free extracts of several crucifers, including *A. thaliana* and other *Brassica* species (Glendening and Poulton, 1990). The *L. sativum* desulfoglucosinolate SULT had apparent $K_m$ for PAPS and desulfo-benzylglucosinolate of 60 and 82 µM, respectively (Glendening and Poulton, 1990).

The presence of glucosinolates in *A. thaliana* and the availability of its genome sequence facilitated the identification of the three DSG-SULTs present in this plant (Varin and Spertini, 2003; Piotrowski et al., 2004; Klein et al., 2006). Different names have been used to describe these proteins (AtST5a-c and AtSOT16-18) and in the following section, the nomenclature described in Table 1 will be used. Two different approaches were used to identify the genes encoding DSG-SULTs in *A. thaliana*. Varin and Spertini (2003) used a systematic approach of cloning all SULT-coding genes from *A. thaliana* and assay the recombinant proteins with desulfoallyl-, desulfobenzyl- and desulfoindolylglucosinolates. Three DSG-SULTS (SULT201B1, SULT201B2, and SULT201B3) with different substrate specificities were identified using this approach. Piotrowski et al. (2004) isolated a cDNA clone induced by the phytotoxin coronatine, a structural homolog of jasmonic acid, using differential mRNA display. The cDNA was found to correspond to *SULT201B1*. Upon wounding, an immediate and transient increase in *SULT201B1* transcript was observed both locally and systemically. The coronatine-induced SULT was shown to prefer desulfoindolylglucosinolates, whereas long chain desulfoglucosinolates derived from methionine are the preferred substrates of the two close homologs, SULT201B2 and SULT201B3. The substrate preference of SULT201B1 and its induction by coronatine and methyl jasmonate correlates with a previous report demonstrating that *B. napus* responds to methyl jasmonate treatment by increasing the accumulation of indolylglucosinolates (Doughty et al., 1995). The $K_m$ values ranged from 50 to 100 µM for desulfoglucosinolates and 25 to 100 µM for PAPS (Klein et al., 2006). Green fluorescent protein fusions were used to study the sub-cellular localization of the DSG-SULTs of *A. thaliana*. The results indicated that the three SULTs have a cytosolic localization (Klein et al., 2006).

### *d. Flavonoid SULT (At3g45070)*

To characterize the biochemical function of At3g45070, a large number of substrates were tested including desulfo-derivatives of most of the known plant sulfated metabolites as well as a collection of metabolites for which no sulfated derivatives have been reported in the literature (Gidda and Varin, 2006). When expressed in *E. coli*, At3g45070 exhibited specificity for position 7 of flavones, flavonols and their monosulfate derivatives. The substrate specificity studies clearly indicate that At3g45070 prefers flavonols over flavones. Furthermore, among the flavonols tested, At3g45070 was shown to have a higher affinity for the monosulfate derivatives as compared with the corresponding aglycones, suggesting the existence in *A. thaliana* of another flavonol SULT that may sulfonate the 3-hydroxyl group. The fact that At3g45070 could sulfonate flavones (apigenin, chrysin, 7-hydroxyflavone), flavonols with different substitution patterns on ring B (kaempferol, isorhamnetin, quercetin, myricetin 3, 5, 7, 3′, 4′, 5′ hexahydroxyflavone) and an isoflavone (genistein) albeit with different catalytic efficiencies suggested that the enzyme does not exhibit strict structural requirements for defined ring B and C structures. However, the position specificity exhibited by At3g45070, and the low enzyme activity observed with 6- or 8-hydroxylated flavonols indicated that the enzyme exhibits strict structural requirements for the flavonoid A ring. In contrast, the two isoforms of the previously characterized flavonol 7-SULT from *F. bidentis* were found to exhibit strict structural requirements for the three rings of the flavonol skeleton (Varin and Ibrahim, 1991). The fact that At3g45070 exhibits a broad substrate specificity suggests that sulfonation of flavonols in *A. thaliana* does not follow a stepwise or sequential order as observed in *Flaveria* species (Varin and Ibrahim, 1989). The $K_m$ values for kaempferol 3-sulfate and PAPS were found to be 2 and 1 µM, respectively with a $V_{max}$ of 285 pkatal/mg$^{-1}$.

The substrate specificity of At3g45070 is consistent with the reported occurrence of kaempferol, quercetin and myricetin in *A. thaliana* (Shirley et al., 1995; Burbulis et al., 1996). However, the absence of any reports on the presence of flavonol sulfates in *A. thaliana* suggests that these compounds might be present in very low quantities and

possibly in specific tissues or at specific developmental stages or desulfated during isolation because of their labile nature. RT- PCR experiments indicated that *At3g45070* is expressed only at the early stage of seedling development and in siliques and inflorescence stem in mature plants. Using biochemical and visualization techniques it has been shown that flavonoid accumulation in *A. thaliana* seedlings is developmentally regulated, and in the mature plants, flavonoid accumulation is restricted to flowers, immature siliques and upper inflorescence stems (Peer et al., 2001). The pattern of expression of *At3g45070* therefore parallels flavonoid accumulation in *A. thaliana*. The discovery of a flavonol sulfotransferase from *A. thaliana* suggests that flavonol sulfates are more widely distributed in plants than once thought and this model plant could be used to study their biological significance.

The At3g45070 and At3g45080 enzymes exhibit 86% amino acid sequence identity, suggesting that they might represent isoenzymes sharing similar substrates. Attempts to characterize the biochemical function of At3g45080 were unsuccessful (Gidda and Varin, unpublished results). Therefore, it remains to be determined if this enzyme exhibits overlapping or different substrate and position preferences as compared with those of At3g45070.

## B. Membrane-Bound Plant Sulfotransferases

In mammals, 12 membrane-associated SULTs have been characterized extensively (Chapman et al., 2004). These enzymes are known to play critical functions in anticoagulation, angiogenesis, leucocyte adhesion, cartilage development, corneal transparency, neuronal function, lymphocyte binding, T-cell response, and HSV-1 entry. They catalyze the sulfonation of sugar residues in heparan, keratan, dermatan and chondroitin to mention a few, and of tyrosyl residues in peptides and proteins. Because of their biological importance and medical relevance, there is intense interest in understanding their properties and mode of action. In contrast, very little is known about membrane-associated plant SULTs. To date, only two have been partially characterized. Interestingly, both are involved in the sulfonation of regulatory or signalling molecules.

### 1. Gallic Acid Glucoside Sulfotransferase *from* Mimosa pudica

*Mimosa pudica* has the ability to close its leaves at night (nyctinasty) or in response to a mechanical stimulus (seismonasty). Following mechanical stimulation, the leaves close in 1 to 2 s, and in the absence of further stimulation, they recover their original position during a recuperation phase lasting 2 to 5 min. The movements are taking place at the motor organs (pulvini) localized at the base of the petiole and leaflets. A combination of electrical and chemical signals has been shown to be involved in this movement (Satter, 1990).

Early in the 20th century, scientists reported that an extract from *Mimosa*, or other plants exhibiting nyctinastic movement was able to induce leaf closure when applied to the cut stem of the seisomonastic plant *M. pudica*. The substance that could induce the movement was later found to be gallic acid 4-*O*-(β-D-glucopyranosyl-6′-sulfate) and was named the periodic leaf movement factor 1 (PLMF-1) (Schildknecht and Schumacher, 1981). Structure–activity relationship studies of PLMF-1 revealed that the presence of the sulfate group was required for biological activity.

A SULT that catalyzes the transfer of the sulfuryl group from PAPS to gallic acid glucoside was characterized from *M. pudica* plasma membrane protein preparations (Varin et al., 1997). The enzyme was found to exhibit strict specificity and high affinity for the substrate gallic acid glucoside ($K_m$ 3.0 μM) and cosubstrate PAPS ($K_m$ 0.5 μM). The PLMF-1 SULT activity was detected only in plasma membrane protein extracts from pulvini, suggesting that the site of synthesis of this compound is restricted to the motor organs. Although the PLMF-1 SULT has never been purified to apparent homogeneity, the availability of antibodies that reacted with this protein allowed detecting a 42 kDa band in plasma membrane fractions exhibiting PLMF-1 SULT activity. The 7 to 12 kDa difference between the molecular mass of the *Mimosa* SULT and the 30–35 kDa reported for the plant and animal cytosolic SULTs might reflect the presence of an additional transmembrane domain anchoring the protein in the plasma membrane of the pulvini. Indirect immunogold labeling of sections from primary and secondary pulvini indicated the presence of the gold particles on the plasma membrane of

the sieve tubes. The localization of the PLMF-1 SULT in the phloem cells of the motor organs coincides with the site of transport of the electrical signal that is triggered following stimulation and supports the model that has been proposed for the mode of action of PLMF-1 (Schildknecht and Meir-Augenstein, 1990).

### *2. Tyrosylprotein Sulfotransferase from Oryza sativa*

Protein tyrosine *O*-sulfonation is one of the most frequent posttranslational modifications occurring in many secretory and membrane bound proteins of eukaryotes (Niehrs et al., 1994). In mammalian cells, this sulfonation reaction is catalyzed by tyrosylprotein sulfotransferase (TPSULT), a membrane-bound enzyme localized in the *trans*-Golgi network (Lee and Huttner, 1983; Baeuerle and Huttner, 1987). Human and mouse TPSULT cDNAs have been cloned and shown to encode type II transmembrane proteins of 370 amino acids with apparent molecular masses of 54 kDa (Ouyang et al., 1998).

In plants, the first report on the characterization of a TPSULT had to await the isolation of a disulfated pentapeptide [Tyr($SO_3$H)-Ile-Tyr($SO_3$H)-Thr-Gln] named phytosulfokine-α (PSK-α) which was purified from a conditioned medium derived from asparagus cell culture (Matsubayashi and Sakagami, 1996). PSK-α triggers cell proliferation at nanomolar concentrations synergistically with other plant hormones. Structure–activity relationship studies demonstrated that the sulfonation of the tyrosine residues is essential for the mitogenic activity of PSK-α. Subsequently, PSK-α was shown to be present in monocotyledonous and dicotyledonous cell cultures, and PSK-α encoding genes from rice and *A. thaliana* have been cloned and characterized (Yang et al., 1999, 2001). Putative receptor proteins for this autocrine-type growth factor were identified by photoaffinity labelling of plasma membrane fractions derived from rice, carrot and tobacco cells suggesting the widespread occurrence of the binding proteins (Matsubayashi and Sakagami, 1999, 2000). A PSK receptor was purified and cloned from carrot microsomal fractions, belonging to the leucine-rich repeat receptor-like kinases (Matsubayashi et al., 2002).

An *in vitro* enzyme assay to detect TPSULT activity was developed using a 14 amino acid synthetic oligopeptide precursor derived from the rice PSK-α cDNA sequence (Hanai et al., 2000). The precursor contained the mature YIYTQ sequence and the acidic amino acid residues at the N-terminal of the first tyrosine residue of mature PSK-α. TPSULT activity was found in microsomal membrane preparations from rice, asparagus and carrot cells (Hanai et al., 2000). The widespread distribution of TPSULT activity in higher plants suggests that tyrosine *O*-sulfonation is a ubiquitous posttranslational process involved in the modification of proteins and peptides. The asparagus enzyme exhibited a broad pH optimum of 7.0–8.5, required manganese ions for optimal activity and appeared to be membrane-localized in the Golgi apparatus. The apparent $K_m$ of the rice enzyme for the peptide precursor was 71 μM, with a $V_{max}$ of 1.0 pmol $min^{-1}$ $mg^{-1}$. Substrate specificity studies revealed that acidic amino acid residues adjacent to the tyrosine residues of the acceptor peptide were essential for activity, indicating that, as in mammals, tyrosine sulfonation takes place prior to the proteolytic processing and maturation of the peptide hormone. Structural requirements for the peptide substrate and the properties of the rice enzyme are similar to those reported for mammalian TPSULT enzymes. So far, the cloning of a plant TPSULT has not been reported, and no candidate sequences are available. Cloning of a plant TPSULT will allow studying the regulation of its expression and its tissue distribution to better understand its function. A cloned plant TPSULT may also assist the identification of additional substrates beside the PSK precursor. An improved understanding of the biology and substrate specificity of the plant TPSULT is also relevant to the production of pharmaceutical proteins in plants by molecular farming, since a large number of therapeutic proteins and peptides are modified by tyrosine sulfonation in their native form (Niehrs et al., 1994).

## VI. Sulfotransferases in Algae and Phototrophic Bacteria

### *A. Algae*

Several genera of marine macroalgae synthesize sulfated polysaccharides which constitute the major compound of their cell wall (see Chapters

15 and 22 in this book). These polyanionic molecules chelate metallic ions, and provide a hydration shell to the organism. The commercially valuable sulfated polysaccharides are the carageenans from red algae (Rhopophyta) and sulfated fucans from brown algae (Phaeophyceae). The sulfated fucans have anticoagulant properties, and present a possible alternative to medical treatment with heparin (Mourao, 2004). Carageenans are used in various products due to their properties as hydrocolloids, including as food thickeners, but also in cosmetic and pharmaceutical products, and in various industrial applications (McHugh, 2003). Carageenans are sulfated galactans composed of alternating 1,3-α-1,4-β D-galactose units substituted with one (κ-), two (ι-) or three sulfate residues (λ-) per disaccharide monomer. Sulfated fucans are polysaccharides mainly constituted of sulfated L-fucose, with substitutions at positions 2, 3 and/or −4, depending on the species (Berteau and Mulloy, 2003).

Sulfated fucans are also present in echinoderms, such as sea urchins and sea cucumbers, while sulfated galactans are found in marine invertebrates ascidians. Aquino et al. (2005) reported that marine angiosperms, the seagrassses, contain sulfated polysaccharides in their cell walls, whereas they are absent from land plants. The structure of the polysaccharide from *Ruppia maritima* was determined to be a sulfated galactan containing regular tetrasaccharide units. This discovery suggested that biosynthesis of sulfated polysaccharides may be a result of physiological adaptation to marine environments.

In mammals, numerous genes have been cloned coding for Golgi-localized, membrane-bound SULTs involved in the biosynthesis of sulfated oligosaccharides and glycosaminoglycans (Fukuda et al., 2001). Carbohydrate SULTs have also been characterized from symbiotic rhizobacteria, and participate in the biosynthesis of nodulation (Nod) factors acting as signals eliciting developmental programs leading to nodule formation in the plant root. NodH SULT from *Sinorhizobium meliloti* catalyzes the sulfonation of the 6-reducing end of *N*-acetylglucosamine in an oligosaccharide (Ehrhardt et al., 1995), while NoeE from *Rhizobium* sp. NGR234 sulfates a 2-*O*-methyl fucose residue substituting the reducing end of the *N*-acetylglucosamine oligomer (Hanin et al., 1997). The mammalian and rhizobacterial carbohydrate SULTs share consensus sequences which are the hallmark of all SULTs characterized to date (Marsolais and Varin, 1995; Kakuta et al., 1998). Yet, at present, no algal candidate genes are available for the SULTs involved in the biosynthesis of sulfated polysaccharides. This may be due to the scarcity of genomic information available from algae, although a few expressed sequence tag sequencing projects have been initiated (Grossman, 2005). Alternatively, algal carbohydrate SULT genes may have diverged to the extent that they cannot be readily identified using bioinformatic search tools. Algal carbohydrate SULT genes may constitute a highly attractive target for future research, to produce existing or new sulfated carbohydrate polymers, through enzymatic synthesis or metabolic engineering. However, carrageenans and sulfated fucans can act as elicitors of plant defence responses (Mercier et al., 2001; Klarzynski et al., 2003), and therefore, their production in plants may have deleterious effects. This is indeed a yet unexplored area.

### *B. Phototrophic Bacteria*

Recent sequencing of bacterial genomes has uncovered a multitude of SULT genes of unknown function. They are particularly well represented in cyanobacteria, but also in other phototrophic bacteria. A search of the Entrez conserved protein domain database (http://www.ncbi.nlm.nih.gov) revealed that cyanobacterial sequences are represented in proteins defined by five different Pfam SULT conserved domains. As compared with plant and mammalian SULTs, the unique feature of bacterial sequences is that the SULT domain may be associated with other functional domains in a single protein. The SULTs present in phototrophic bacteria can be classified into four structural types, based on their domain architecture (visualized in the Pfam database, http://www.sanger.ac.uk/Software/Pfam/) (Table 3, Fig. 5). For each type illustrated in Fig. 5, a cyanobacterial SULT was arbitrarily chosen as a model. Information about possible function of the SULTs may sometimes be inferred from their domain structure and genomic context. In the first structural type, the SULT domain is located C-terminal to a series of tetratricopeptide (TPR) repeats (Table 3, Fig. 5). The number, subtype and arrangement of the TPR repeats can be variable. Among pro-

*Table 3*. Sulfotransferase sequences from phototrophic bacteria. Sequences are listed with their NCBI protein database annotation, species of origin, accession no. and length. aa: amino acid.

| Sulfotransferase structural type | Protein model | Other selected proteins from cyanobacteria | Presence in other phototrophic bacteria | Taxonomic distribution |
|---|---|---|---|---|
| TPR: sulfotransferase | Hypothetical protein, *Gloeobacter violaceus* (NP_924330.1) (631 aa) | Sulfotransferase TPR repeat, *Prochlorococcus marinus* (NP_893237.1) (584 aa); Probable TPR domain protein, *Synechococcus* sp. (ZP_0108118) (606 aa) | Purple non-sulfur bacteria, for example sulfotransferase, *Rhodospirillum rubrum* (YP_427300.1) (656 aa) | Present in other proteobacteria; sulfotransferase, *Chromohalobacter salexigens* (ABE59338) (1415 aa) has fused TPR, *O*-GlcNAc- and sulfo- transferase domains |
| Sulfotransferase | Hypothetical protein, *Synechocystis* sp. (NP_942202.1) (316 aa) | Sulfotransferase, *Crocosphera watsonii* (ZP_00516056.1) (273 aa); unknown, *Prochloron didemni* (AAT37520.1) (277 aa); hypothetical protein, *G. violaceus* (NP_924845) (320 aa) | Purple non-sulfur bacteria, for example sulfotransferase, *Rhodospirillum rubrum* (ABC21658) (289 aa); green sulfur bacteria, for example sulfotransferase, *Chlorobium phaeobacteroides* (ZP_00533423) (343 aa) | Present in other proteobacteria and Gram-positive bacteria |
| Polyketide synthase: sulfotransferase | CurM, *Lyngbya majuscula* (AAT70108) (2147 aa) | None | Absent | Unique; Polyketide synthase, *Pseudomonas entomophila* (YP_610919.1) (1217 aa) has a related sulfotransferase domain |
| GAF: sulfotransferase | GAF, Synechococcus sp. (ABB34328) (529 aa) | None | Absent | Unique |

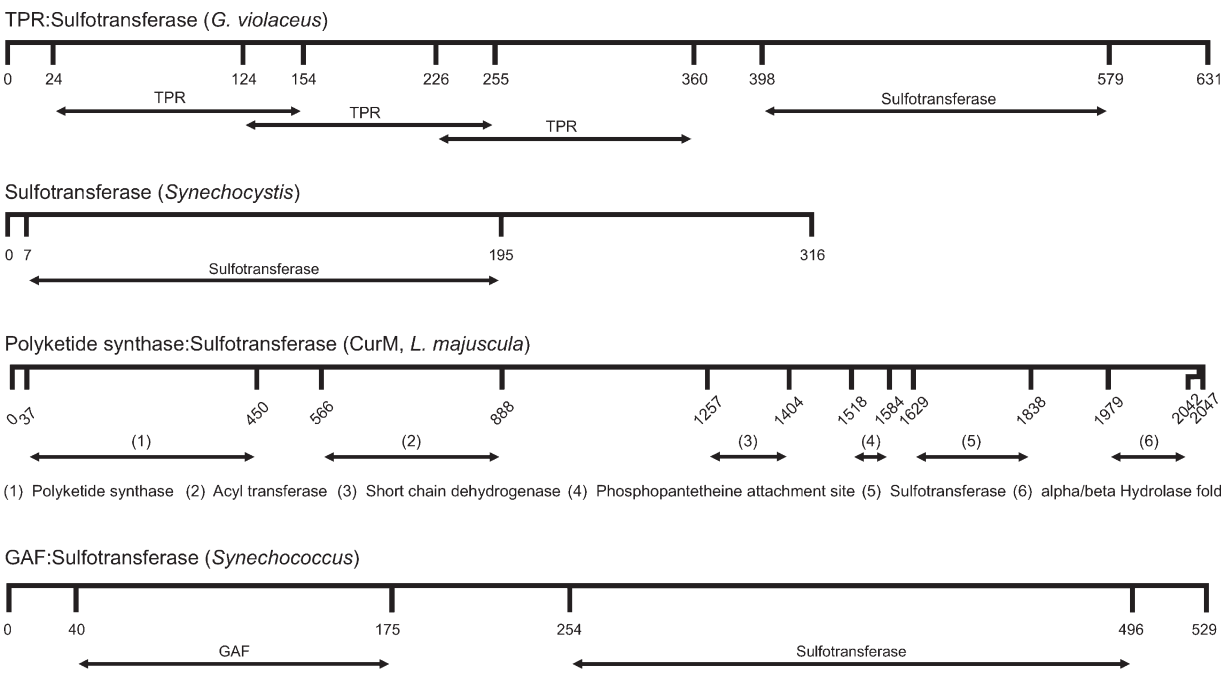



*Fig. 5*. Protein domain architecture of SULTs from phototrophic bacteria. Schematic representation of the four structural types of SULT present in phototrophic bacteria, based on their domain structure. The protein models represented are listed in Table 1.

teins of known function, the TPR repeat domains present in the *Gloeobacter* and *Rhodospirillum* SULTs listed in Table 3 are most similar to that of eukaryotic *O*-linked *N*-acetylglucosamine (*O*-GlcNAc) transferases (data not shown), which catalyze posttranslational modification of proteins on serine and threonine residues. In *O*-GlcNAc transferases, the TPR repeat domain is involved in protein-protein interaction, and assists substrate recognition (Iyer and Hart, 2003). Interestingly, a related hybrid protein from *Chromohalobacter salexigens* has a fusion of the TPR, *O*-GlcNAc transferase and SULT domains (Table 3). Proteins from the second structural type have no additional domains beside the SULT domain (Table 3, Fig. 5). The *Synechocystis* SULT used as a protein model is encoded on a plasmid, pSYSM, and its gene located within a cluster comprised of eight glycosyl transferases and two homologues of polysaccharide transporters, which may participate in the biosynthesis and secretion of unknown exopolysaccharides (Kaneko et al., 2003). Within the same structural type, the SULT gene from the cyanobacterial symbiont *Prochloron didemni* (Table 3) is located in a cluster with a nonribosomal peptide synthetase, which may be involved in the biosynthesis of bioactive nonribosomal cyclic peptides (Schmidt et al., 2004). The third structural type is characterized by a fusion of polyketide synthase and SULT domains, represented by CurM from *Lyngbya majuscula* (Table 3, Fig. 5). The *curM* locus is part of large gene cluster involved in the biosynthesis of curacin A, comprised of a nonribosomal peptide synthetase and multiple polyketide synthases (Chang et al., 2004). CurM is the only example of its type within the phototrophic bacteria. However, a protein harbouring a polyketide synthase:sulfotransferase fusion has also been reported from *Pseudomonas entomophila* (Table 3). In the fourth structural type, represented by the GAF protein from *Synechococcus* sp. (Table 3), the SULT domain is fused with a GAF domain, involved in ligand binding in phytochromes and cGMP-dependent 3′,5′-cyclic phosphodiesterase (Fischer et al., 2005; Gross-Langenhoff et al., 2006). The multitude of SULT sequences and multiplicity of protein domain architectures suggest the presence of a large variety of sulfated macromolecules and secondary metabolites in cyanobacteria and phototrophic bacteria, which are yet completely uncharacterized.

## VII. Future Prospects in Sulfotransferase Research

Structural and functional genomic approaches have demonstrated great potential for identifying enzymes catalyzing the sulfonation reaction. While only four plant SULT sequences were known 10 years ago, there are now more than 200. The acquisition of genomic information allowed developing powerful tools to study the biological function of the SULT-coding genes. For example, a centralized database of *A. thaliana* microarray data provides researchers with valuable information to understand how SULT-coding genes are regulated during development, in adaptation to stress, as well as in various mutant backgrounds. In addition, international efforts to produce collections of T-DNA insertion mutants may allow an assessment of *in vivo* gene function. Despite the availability of functional genomic tools, and although many plant SULTs have been expressed heterologously in the past few years, the catalytic activity of only a few has been demonstrated *in vivo*. The low abundance of functional studies conducted with SULTs from phototrophic bacteria is even more dramatic. Studies conducted with *A. thaliana* so far illustrate the problems encountered when trying to assess the biochemical function of plant SULTs. In the case of DSG-SULTs, substrates could be predicted based on prior knowledge of glucosinolates accumulating in *A. thaliana*. However, no other sulfated metabolites were known from this plant, making difficult the selection of potential substrates that could be assayed with the remaining 14 uncharacterized SULTs. Although the approach of testing a collection of potential substrates proved to be successful for the characterization of the hydroxyjasmonate SULT, it did not allow demonstrating unambiguously the function of other SULTs present in *A. thaliana*. In future, unbiased metabolite profiling experiments will need to be conducted to resolve this problem. Specific protocols allowing the isolation and characterization of labile sulfated compounds by mass spectrometry will need to be developed, and systematically applied to have a more advanced understanding of the sulfonated compounds accumulating in plants. This knowledge will greatly facilitate future studies to elucidate the biochemical function of the SULTs, and the biological significance of the accumulation of their sulfonated enzymatic products.

## Acknowledgements

We thank the Natural Sciences and Engineering Research Council of Canada for its support. We thank Alex Molnar for expert graphical assistance in the preparation of a figure.

## References

Aquino RS, Landeira-Fernandez AM, Valente AP, Andrade LR, Mourao PA (2005) Occurrence of sulfated galactans in marine angiosperms: evolutionary implications. Glycobiology 15: 11–20

Atha DH, Lormeau JC, Petitou M, Rosenberg RD, Choay J (1985) Contribution of monosaccharide residues in heparin binding to antithrombin III. Biochemistry 24: 6723–6729

Baeuerle PA, Huttner WB (1987) Tyrosine sulfation is a trans-Golgi-specific protein modification. J Cell Biol 105: 2655–2664

Berteau O, Mulloy B (2003) Sulfated fucans, fresh perspectives: structures, functions, and biological properties of sulfated fucans and an overview of enzymes active toward this class of polysaccharide. Glycobiology 13: 29R–40R

Blanchard RL, Freimuth RR, Buck J, Weinshilboum RM, Coughtrie MW (2004) A proposed nomenclature system for the cytosolic sulfotransferase (SULT) superfamily. Pharmacogenetics 14: 199–211

Bones AMR, Rossiter JT (1996) The myrosinase-glucosinolate system, its organisation and biochemistry. Physiol Plant 97: 194–208

Burbulis IE, Iacobucci M, Shirley BW (1996) A null mutation in the first enzyme of flavonoid biosynthesis does not affect male fertility in Arabidopsis. Plant Cell 8: 1013–1025

Chang Z, Sitachitta N, Rossi JV, Roberts MA, Flatt PM, Jia J, Sherman DH, Gerwick WH (2004) Biosynthetic pathway and gene cluster analysis of curacin A, an antitubulin natural product from the tropical marine cyanobacterium *Lyngbya majuscula*. J Nat Prod 67: 1356–1367

Chapman E, Best MD, Hanson SR, Wong CH (2004) Sulfotransferases: structure, mechanism, biological activity, inhibition, and synthetic utility. Angew Chem Int Ed Engl 43: 3526–3548

Chew FS (1988) Biologically Active Natural Products. American Chemical Society, Washington, D.C.

Clarke C, Thorburn P, McDonald D, Adams JB (1982) Enzymic synthesis of steroid sulphates. XV. Structural domains of oestrogen sulphotransferase. Biochim Biophys Acta 707: 28–37

Cronquist A (1988) Evolution and Classification of Flowering Plants. New York Botanical Garden, New York

Doughty KJ, Kiddle GA, Pye BJ, Wallsgrove RM, Pickett JA (1995) Selective induction of glucosinolates in oilseed rape leaves by methyl jasmonate. Phytochemistry 38: 347–350

Driscoll WJ, Komatsu K, Strott CA (1995) Proposed active site domain in estrogen sulfotransferase as determined by mutational analysis. Proc Natl Acad Sci USA 92: 12328–12332

Ehrhardt DW, Atkinson EM, Faull KF, Freedberg DI, Sutherlin DP, Armstrong R, Long SR (1995) In vitro sulfotransferase activity of NodH, a nodulation protein of Rhizobium meliloti required for host-specific nodulation. J Bacteriol 177: 6237–6245

Farzan M, Mirzabekov T, Kolchinsky P, Wyatt R, Cayabyab M, Gerard NP, Gerard C, Sodroski J, Choe H (1999) Tyrosine sulfation of the amino terminus of CCR5 facilitates HIV-1 entry. Cell 96: 667–676

Faulkner IJ, Rubery PH (1992) Flavonoids and flavonol sulfates as probes of auxin transport regulation in *Cucurbita pepo* hypocotyl segments and vesicles. Planta 186: 618–625

Fenwick GR, Heaney RK, Mullin WJ (1983) Glucosinolates and their breakdown products in food and food plants. Crit Rev Food Sci Nutr 18: 123–201

Fischer AJ, Rockwell NC, Jang AY, Ernst LA, Waggoner AS, Duan Y, Lei H, Lagarias JC (2005) Multiple roles of a conserved GAF domain tyrosine residue in cyanobacterial and plant phytochromes. Biochemistry 44: 15203–15215

Fukuda M, Hiraoka N, Akama TO, Fukuda MN (2001) Carbohydrate-modifying sulfotransferases: structure, function, and pathophysiology. J Biol Chem 276: 47747–47750

Gachon CM, Langlois-Meurinne M, Saindrenan P (2005) Plant secondary metabolism glycosyltransferases: the emerging functional analysis. Trends Plant Sci 10: 542–549

Gidda SK, Miersch O, Levitin A, Schmidt J, Wasternack C, Varin L (2003) Biochemical and molecular characterization of a hydroxyjasmonate sulfotransferase from *Arabidopsis thaliana*. J Biol Chem 278: 17895–17900

Gidda SK, Varin L (2006) Biochemical and molecular characterization of flavonoid 7-sulfotransferase from Arabidopsis thaliana. Plant Physiol Biochem 44: 628–636

Glendening TM, Poulton JE (1990) Partial purification and characterization of a 3′-phosphoadenosine 5′-phosphosulfate: desulfoglucosinolate sulfotransferase from cress (*Lepidium sativum*). Plant Physiol 94: 811–818

Gross-Langenhoff M, Hofbauer K, Weber J, Schultz A, Schultz JE (2006) cAMP is a ligand for the tandem GAF domain of human phosphodiesterase 10 and cGMP for the tandem GAF domain of phosphodiesterase 11. J Biol Chem 281: 2841–2846

Grossman AR (2005) Paths toward algal genomics. Plant Physiol 137: 410–427

Hanai H, Nakayama D, Yang H, Matsubayashi Y, Hirota Y, Sakagami Y (2000) Existence of a plant tyrosylprotein sulfotransferase: novel plant enzyme catalyzing tyrosine O-sulfation of preprophytosulfokine variants in vitro. FEBS Lett 470: 97–101

Hanin M, Jabbouri S, Quesada-Vincens D, Freiberg C, Perret X, Prome JC, Broughton WJ, Fellay R (1997) Sulphation of Rhizobium sp. NGR234 Nod factors is dependent

on noeE, a new host-specificity gene. Mol Microbiol 24: 1119–1129

Hanson AD, Rathinasabapathi B, Chamberlin B, Gage DA (1991) Comparative physiological evidence that beta-alanine betaine and choline-O-sulfate act as compatible osmolytes in halophytic Limonium species. Plant Physiol 97: 1199–1205

Hanson AD, Rathinasabapathi B, Rivoal J, Burnet M, Dillon MO, Gage DA (1994) Osmoprotective compounds in the Plumbaginaceae: a natural experiment in metabolic engineering of stress tolerance. Proc Natl Acad Sci USA 91: 306–310

Iyer SP, Hart GW (2003) Roles of the tetratricopeptide repeat domain in O-GlcNAc transferase targeting and protein substrate specificity. J Biol Chem 278: 24608–24616

Jacobs M, Rubery PH (1988) Naturally occurring auxin transport regulators. Science 241: 346–349

Jain JC, GrootWassink JWD, Kolenovsky AD, Underhill EW (1990) Purification and properties of 3′-phosphoadenosine 5′-phosphosulfate: desulphoglucosinolate sulphotransferase from *Brassica juncea* cell cultures. Phytochemistry 29: 1425–1428

Kakuta Y, Pedersen LG, Pedersen LC, Negishi M (1998) Conserved structural motifs in the sulfotransferase family. Trends Biochem Sci 23: 129–130

Kaneko T, Nakamura Y, Sasamoto S, Watanabe A, Kohara M, Matsumoto M, Shimpo S, Yamada M, Tabata S (2003) Structural analysis of four large plasmids harboring in a unicellular cyanobacterium, Synechocystis sp. PCC 6803. DNA Res 10: 221–228

Klarzynski O, Descamps V, Plesse B, Yvin JC, Kloareg B, Fritig B (2003) Sulfated fucan oligosaccharides elicit defense responses in tobacco and local and systemic resistance against tobacco mosaic virus. Mol Plant Microbe Interact 16: 115–122

Klein M, Reichelt M, Gershenzon J, Papenbrock J (2006) The three desulfoglucosinolate sulfotransferase proteins in Arabidopsis have different substrate specificities and are differentially expressed. FEBS J 273: 122–136

Lacomme C, 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

Lee RW, Huttner WB (1983) Tyrosine-*O*-sulfated proteins of PC12 pheochromocytoma cells and their sulfation by a tyrosylprotein sulfotransferase. J Biol Chem 258: 11326–1334

Lerouge P, Roche P, Faucher C, Maillet F, Truchet G, Prome JC, Denarie J (1990) Symbiotic host-specificity of Rhizobium meliloti is determined by a sulphated and acylated glucosamine oligosaccharide signal. Nature 344: 781–784

Marsolais F, Boyd J, Paredes Y, Schinas AM, Garcia M, Elzein S, Varin L (2007) Molecular and biochemical characterization of two brassinosteroid sulfotransferases from *Arabidopsis*, AtST4a (At2g14920) and AtST1 (At2g03760). Planta 225: 1233–1244

Marsolais F, Gidda SK, Boyd J, Varin L (2000) Plant soluble sulfotransferases: Structural and functional similarity with mammalian enzymes. Rec Adv Phytochem 34: 433–456

Marsolais F, Hernández Sebastià CH, Rousseau A, Varin L (2004) Molecular and biochemical characterization of BNST4, an ethanol-inducible steroid sulfotransferase from *Brassica napus*, and regulation of BNST genes by chemical stress and during development. Plant Sci 166: 1359–1370

Marsolais F, 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, Varin L (1997) Mutational analysis of domain II of flavonol 3-sulfotransferase. Eur J Biochem 247: 1056–1062

Marsolais F, Varin L (1998) Recent developments in the study of the structure–function relationship of flavonol sulfotransferases. Chem Biol Interact 109: 117–122

Matsubayashi Y, Ogawa M, Morita A, Sakagami Y (2002) An LRR receptor kinase involved in perception of a peptide plant hormone, phytosulfokine. Science 296: 1470–1472

Matsubayashi Y, Sakagami Y (1996) Phytosulfokine, sulfated peptides that induce the proliferation of single mesophyll cells of *Asparagus officinalis* L. Proc Natl Acad Sci USA 93: 7623–7627

Matsubayashi Y, Sakagami Y (1999) Characterization of specific binding sites for a mitogenic sulfated peptide, phytosulfokine-alpha, in the plasma-embrane fraction derived from Oryza sativa L. Eur J Biochem 262: 661–671

Matsubayashi Y, Sakagami Y (2000) 120- and 160-kDa receptors for endogenous mitogenic peptide, phytosulfokine-alpha, in rice plasma membranes. J Biol Chem 275: 15520–15525

McHugh D (2003) A guide to the seaweed industry. FAO fisheries technical paper no 441. FAO, Rome

Mercier L, Lafitte C, Borderies G, Briand X, Esquerré-Tugayé M-T, Fournier J (2001) The algal polysaccharide carrageenans can act as an elicitor of plant defence. New Phytol 149: 43–52

Mourao PA (2004) Use of sulfated fucans as anticoagulant and antithrombotic agents: future perspectives. Curr Pharm Des 10: 967–981

Niehrs C, Beisswanger R, Huttner WB (1994) Protein tyrosine sulfation, 1993 – an update. Chem Biol Interact 92: 257–271

Ouyang Y, Lane WS, Moore KL (1998) Tyrosylprotein sulfotransferase: purification and molecular cloning of an enzyme that catalyzes tyrosine *O*-sulfation, a common posttranslational modification of eukaryotic proteins. Proc Natl Acad Sci USA 95: 2896–2901

Peer WA, Brown DE, Tague BW, Muday GK, Taiz L, Murphy AS (2001) Flavonoid accumulation patterns of transparent testa mutants of Arabidopsis. Plant Physiol 126: 536–548

Piotrowski M, Schemenewitz A, Lopukhina A, Muller A, Janowitz T, Weiler EW, Oecking C (2004) Desulfoglucosinolate sulfotransferases from *Arabidopsis thaliana* catalyze the final step in the biosynthesis of the glucosinolate core structure. J Biol Chem 279: 50717–50725

Rivoal J, Hanson AD (1994) Choline-O-sulfate biosynthesis in plants: identification and partial characterization of a salinity-inducible choline sulfotransferase from species of *Limonium* (Plumbaginaceae). Plant Physiol 106: 1187–1193

Roche P, Debelle F, Maillet F, Lerouge P, Faucher C, Truchet G, Denarie J, Prome JC (1991) Molecular basis of symbiotic host specificity in Rhizobium meliloti: nodH and nodPQ genes encode the sulfation of lipo-oligosaccharide signals. Cell 67: 1131–1143

Rouleau M, Marsolais F, Richard M, Nicolle L, Voigt B, Adam G, Varin L (1999) Inactivation of brassinosteroid biological activity by a salicylate-inducible steroid sulfotransferase from *Brassica napus*. J Biol Chem 274: 20925–20930

Sakakibara Y, Takami Y, Nakayama T, Suiko M, Liu MC (1998) Localization and functional analysis of the substrate specificity/catalytic domains of human M-form and P-form phenol sulfotransferases. J Biol Chem 273: 6242–6247

Satter RH (1990) Leaf movement: an overview of the field. In: Satter RL, Gorton, HL and Vogelman TC (eds) The Pulvinus: Motor Organ For Leaf Movement, pp 1–9. American Society of Plant Physiologists, Rockville, MD

Schildknecht H, Schumacher K (1981) Ein och wirksamer leaf movement factor aus *Acacia karroo*. Chem-Ztg 105: 287–290

Schildknecht H, Meier-Augenstein W (1990) Role of turgorins in leaf movement. In: Satter RL, Gorton HL and Vogelman TC (eds) The Pulvinus: Motor Organ For Leaf Movement, pp 205–213. American Society of Plant Physiologists, Rockville, MD

Schmidt EW, Sudek S, Haygood MG (2004) Genetic evidence supports secondary metabolic diversity in *Prochloron* spp., the cyanobacterial symbiont of a tropical ascidian. J Nat Prod 67: 1341–1345

Shirley BW, Kubasek WL, Storz G, Bruggemann E, Koornneef M, Ausubel FM, Goodman HM (1995) Analysis of Arabidopsis mutants deficient in flavonoid biosynthesis. Plant J 8: 659–671

Smith DW, Johnson KA, Bingman CA, Aceti DJ, Blommel PG, Wrobel RL, Frederick RO, Zhao Q, Sreenath H, Fox BG, Volkman BF, Jeon WB, Newman CS, Ulrich EL, Hegeman AD, Kimball T, Thao S, Sussman MR, Markley JL, Phillips GN, Jr. (2004) Crystal structure of At2g03760, a putative steroid sulfotransferase from *Arabidopsis thaliana*. Proteins 57: 854–857

Varin L, Chamberland H, Lafontaine JG, Richard M (1997) The enzyme involved in sulfation of the turgorin, gallic acid 4-*O*-(β-D-glucopyranosyl-6′-sulfate) is pulvinilocalized in *Mimosa pudica*. Plant J 12: 831–837

Varin L, DeLuca V, Ibrahim RK, Brisson N (1992) Molecular characterization of two plant flavonol sulfotransferases. Proc Natl Acad Sci USA 89: 1286–1290

Varin L, Ibrahim RK (1989) Partial purification and characterization of three flavonol-specific sulfotransferases from *Flaveria chloraefolia*. Plant Physiol 90: 977–981

Varin L, Ibrahim RK (1991) Partial purification and some properties of flavonol 7-sulfotransferase from *Flaveria bidentis*. Plant Physiol 95:1254–1258

Varin L, Ibrahim RK (1992) Novel flavonol 3-sulfotransferase. purification, kinetic properties, and partial amino acid sequence. J Biol Chem 267:1858–1863

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

Varin L, Spertini D (2003) Desulfoglucosinolate sulfotransferases, sequences coding the same and uses thereof for modulating glucosinolate biosynthesis in plants. Patent WO 03/010318-A1 Concordia University, Canada

Yang H, Matsubayashi Y, Nakamura K, Sakagami Y (1999) Oryza sativa PSK gene encodes a precursor of phytosulfokine-alpha, a sulfated peptide growth factor found in plants. Proc Natl Acad Sci USA 96: 13560–13565

Yang H, Matsubayashi Y, Nakamura K, Sakagami Y (2001) Diversity of Arabidopsis genes encoding precursors for phytosulfokine, a peptide growth factor. Plant Physiol 127: 842–851

Yoshihara T, Omer E-SA, Koshino H, Sakamura S, Kikuta Y, Koda Y (1989) Structure of tuber inducing stimulus from potato leaves. Agric Biol Chem: 2835–2837

Yoshinari K, Petrotchenko EV, Pedersen LC, Negishi M (2001) Crystal structure-based studies of cytosolic sulfotransferase. J Biochem Mol Toxicol 15: 67–75