Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Molecular Graphics and Modelling 23 (2004) 77–87
Journal of Molecular Graphics and Modelling
www.elsevier.com/locate/JMGM

The three-dimensional structure of *Arabidopsis thaliana* *O*-methyltransferase predicted by homology-based modelling

Heejung Yang [a], Joong-Hoon Ahn [a], Ragai K. Ibrahim [b], Sangsan Lee [c], Yoongho Lim [a,*]

[a] *Bio/Molecular Informatics Center, Konkuk University, Seoul 143-701, South Korea*
[b] *Plant Biochemistry Laboratory, Centre for Structural-Functional Genomics, Concordia University, Montreal, Que., Canada H3G 1M8*
[c] *Supercomputing Center, Korea Institute of Science and Technology Information, Daejon 305-350, South Korea*

Received 2 January 2003; accepted 11 February 2004
Available online 20 March 2004

Abstract

O-Methylation of flavonoid compounds is an important enzymatic reaction since it not only reduces the chemical reactivity of their phenolic hydroxyl groups but also increases their lipophilicity and, hence, their intracellular compartmentation. Several genes encoding flavonoid *O*-methyltransferases (OMTs) have been isolated and characterized both at the molecular and biochemical levels. In contrast with mammalian enzymes, plant OMTs exhibit narrow substrate specificities as well as position-specific activities, so that the homology comparison, derived using programs such as BLAST can not provide sufficient information on the enzyme function or its substrate preference. In order to study these characteristics, therefore, another approach, homology-based modelling is being carried out. We report here the determination of the 3-D structure of *Arabidopsis thaliana O*-methyltransferase, AtOMT1 as well as its dynamics when complexed with its substrate. The predicted structure obtained by homology-based modelling is conserved during molecular dynamics simulations. AtOMT1 exhibits a structure similar to that of caffeic acid *O*-methyltransferase, COMT when the latter was used as a template. Whereas COMT includes 20 α-helices and nine β-sheets, AtOMT1 has 16 and 9, respectively. Although the homology between both proteins is higher than 77% and all amino acids surrounding the active sites, except one residue, are similar in their primary sequences, the two proteins exhibit different substrate preferences. The differences in substrate specificity may be explained on the basis of the predicted structures of the protein and its complex with the substrate. In addition, docking the substrate into the active site of the protein allowed the study of the structural change of the active site on the dihedral angle distribution of the residues surrounding the active site.

Keywords: Modelling; Molecular dynamics; Homology; Methyltransferase; AtOMT1; COMT

1. Introduction

Plants elaborate over 100,000 secondary metabolites which are not considered essential for the basic metabolic processes of the plant [1,2], most of which belong to the isoprenoids, phenylpropanoids, alkaloids, and polyketides. Among these the flavonoid compounds, which are synthesized via phenylpropanoid pathway, exhibit extensive diversity, and play an important role in plant growth and development as well as the interactions of plants with their environment [3]. The structural diversity of flavonoids is the result of a number of substitution reactions, such as hydroxylation, glycosylation, methylation, acylation and/or prenylation. These are catalyzed by substrate-specific, position-oriented enzymes. Among these modification reactions, enzymatic methylation is mediated by a family of methyltransferases which involve the transfer of the methyl group of *S*-adenosyl-L-methionine (AdoMet) to the oxygen, nitrogen, or carbon atoms of various acceptor molecules with the formation of the methylated derivative and *S*-adenosyl-L-homocysteine (AdoHcy) [4,5]. *O*-methylation of flavonoids not only reduces the chemical reactivity of their phenolic hydroxyl groups but also increases their lipophilicity and, hence, their intracellular compartmentation.

* Corresponding author.
E-mail addresses: ragibra@vax2.concordia.ca (R.K. Ibrahim), yoongho@konkuk.ac.kr (Y. Lim).

1093-3263/$ – see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2004.02.001

Many genes encoding flavonoid OMTs have been isolated and characterized both at the molecular and biochemical levels with the aim to study their function and their substrate specificity [6–13], as well as their role in plant metabolism [14–18]. In contrast with mammalian enzymes, however, plant OMTs show narrow substrate specificities as well as position-specific activities, so that the homology comparison, derived using programs such as BLAST may not provide sufficient information on the biological function of individual enzymes as to its preferred substrate and product [19]. As a result, in order to predict the possible biological functions of OMTs, a new tool is necessary. Recently, the 3-D structures of three OMTs have been determined using X-ray crystallography, which are the chalcone *O*-methyltransferase (ChOMT), isoflavone *O*-methyltransferase (IOMT), and caffeic acid *O*-methyltransferase (COMT) [19–22]. The crystal structures of these OMTs revealed the structures of the complex of enzymes and their substrates as well as the structures of their binding sites. The 3-D structures of these proteins can help us to predict the function of the other OMTs which belong to the phenylprophanoid pathways [20].

As genome sequences are analyzed, biological functions of the gene products become the center of interest. In general, the function of a protein is related to its primary sequence, although we may be faced with the dilemma that their described functions are putative. Therefore, the 3-D structures of the expressed gene products should be studied for the determination of their functions. Unfortunately, however, while the complete genomic sequences increase geometrically, the determination of the 3-D structures increases arithmetically. To overcome this bottleneck, structure prediction by molecular modelling can serve as a valuable tool for these studies. Since the structural modelling is based on the homology, a few criteria are required. That is, the X-ray crystal or NMR solution structure showing high sequence homology should be known for the template, which should be the structure obtained without its substrate.

The primary sequence of *Arabidopsis thaliana* OMT (AtOMT1) shows 77.4% homology with COMT [13]. The enzyme activity of COMT for the substrate, caffeic acid is 57.5 nmol/(min mg) protein [18]. COMT recognizes caffeic acid as well as caffeoyl alcohol, caffeoyl aldehyde, 5-hydroxyferulic acid, 5-hydroxyconiferyl alcohol, and 5-hydroxyconiferaldehyde [22]. Because of the high homology, the substrate of AtOMT1 was expected to be one of caffeic acid derivatives. However, the enzymatic study showed that AtOMT1 transfers the methyl group of AdoMet to 3′-hydroxyl group of the pentahydroxyflavone, quercetin and results in the formation of its 3′-*O*-methyl derivative, isorhamnetin. Therefore, in order to know why the two proteins with the high homology exhibit different substrate preferences, we described here attempts to determine the 3-D structure of AtOMT1. Since there are five hydroxyl groups in quercetin, it is intriguing to know how AtOMT1 interacts with quercetin and specifically methylates the 3′-hydroxyl group. In addition, understanding the 3-D structure and substrate binding mechanism of AtOMT1 provides valuable information as to how other OMTs specify their substrates and which amino acid residues are important for conferring such a specificity. Although AtOMT1 was expressed and its preferred substrate was determined, its 3-D structure and active site anatomy have not been studied yet. This prompted us to study the 3-D structure of AtOMT1 using molecular modelling and determine its dynamics corresponding to the complex of the protein and its substrate.

2. Methodology

2.1. Target and template proteins

The amino acid sequence of the target protein, AtOMT1 [13] was obtained from GenBank (Accession No. U70424) and is composed of 363 residues. The template protein used was a caffeic acid/5-hydroxyferulic acid OMT, COMT deposited in Protein Data Bank (PDB code: 1KYZ.pdb) [21]. 1KYZ includes three polymer chains: A, C, and E. While A and E chains form a dimmer, the polymer chain C is separated from the others. Because AtOMT1 functions as a monomer [13], the chain C was chosen as a template. The 3-D structure of AdoHcy was adapted from 1KYZ.pdb. Since the 3-D crystallographic structure of the enzymatic product, isorhamnetin could not be found it was obtained by molecular modelling calculation based on the structure of quercetin included in Anthocyanidin Synthase (PDB code: 1GP6.pdb) [23].

2.2. Molecular modelling

The molecular modelling calculations were carried out on an O2 R12,000 Silicon Graphics workstation. The modelling software was InsightII (Accelrys, San Diego, USA). The forcefield used for molecular dynamics (MD) and energy minimization was cvff provided by Accelrys. Homology module included in InsightII was used for homology modelling. The sequence of the template protein (COMT) was extracted and aligned with the target protein (AtOMT1). The pair-wise sequence alignment of the two proteins exhibits 77.4% homology and is shown in Fig. 1. The structurally conserved regions (SCRs) were determined automatically using the method provided by InsightII/Homology module. Based on SCRs, the sequences of the template and target proteins were aligned. The loops and the variable regions not to be included in SCRs were built. The conformations of side chains were determined by conformational search of rotamers. The protein was embedded in a 5 Å shell of 2,079 water molecules to imitate aqueous solvent conditions. The assembled molecules were subjected to energy minimization by Discover module included in InsightII. Steepest descents were carried out until maximum derivative of 1.0 kcal/(mol Å), and conjugate gradients were followed until maximum derivative of

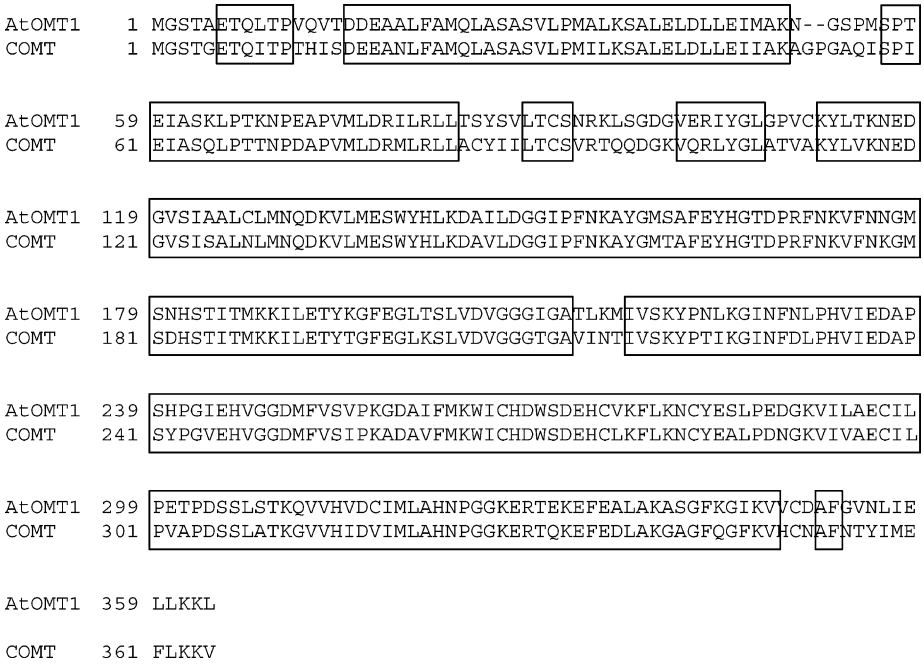

Fig. 1. The pair-wise sequence alignment of the target protein (AtOMT1) and the template protein (COMT). Boxes denote structure conserved regions (SCRs) used for homology modelling in InsightII/Homology module.

0.1 kcal/(ol Å) After energy minimization, MD was performed at 300 K, 1 atm for 500 ps with 1 fs each step. The output conformers were collected at every 4 ps and 125 conformers were saved in the history file. The energy profile was analyzed using the Analysis module included in InsightII. Among 125 conformers, 20 conformers showing the low total energy were selected and their 3-D structures were superimposed. Their root mean square deviation (RMSD) value was 1.1 Å. Of these, the conformer with the lowest energy was chosen and PROCHECK was applied for a statistical evaluation.

2.3. Docking of substrate and co-factor into the protein

The protein used for the docking was the conformer chosen by the analysis of MD results, and it was soaked in a 5 Å shell of 2,079 water molecules to imitate aqueous solvent conditions. The 3-D structure of the substrate, quercetin, was adopted from Anthocyanidin Synthase (PDB code: 1GP6.pdb). The 3-D structure of methylated quercetin, isorhamnetin was obtained by molecular modelling calculation because its 3-D structure was not known. To find the binding sites between the protein and the substrate, InsightII/Binding_site module was used. Among several binding sites suggested by InsightII/Binding_site module, the most similar site to the binding site of COMT determined from the X-ray crystal structure was chosen. Compared with residues neighboring the binding site of the COMT substrate, the residues of AtOMT1 participating in the docking were selected as follows: Met128, Asn129, Phe174, Met178, His267, Glu295, Val314, Ile317, Met318, Asn322, and Glu327. After determining the binding site, the co-factor was docked into the protein. Since the crystal structure of COMT provided the 3-D structure of AdoHcy contained in the protein, in this experiment, the same structure and the same position of AdoHcy were used instead of AdoMet. The docking of the co-factor into the protein was carried out using InsightII/Docking module on a Silicon Graphics O2 R12,000 workstation. The product, isorhamnetin was docked into an AdoHcy and protein assembly by the same method used for the docking of AdoHcy into the protein. As mentioned above, the substrate was tried to be docked into the most similar site to the binding site of COMT. After a substrate, AdoHcy, and protein assembly was created, the assembly was embedded in a 5 Å shell of 1,960 water molecules to imitate aqueous solvent conditions. The assembled molecules were subjected to energy minimization by the Discover module included in InsightII. Steepest descents were carried out until maximum derivative of 1.0 kcal/(mol Å), and conjugate gradients were followed until maximum derivative of 0.1 kcal/(mol Å). After energy minimization, MD was performed at the same condition and by the same method used for the MD calculation of the protein for 500 ps with 1 fs each step. The output conformers were collected at every 4 ps and 125 conformers were saved in the history file. The energy profile was then analyzed using the Analysis module included in InsightII.

3. Results and discussion

3.1. Structure prediction of AtOMT1

As shown in Fig. 2, AtOMT1 gives similar structure to COMT. In order to compare the secondary structures

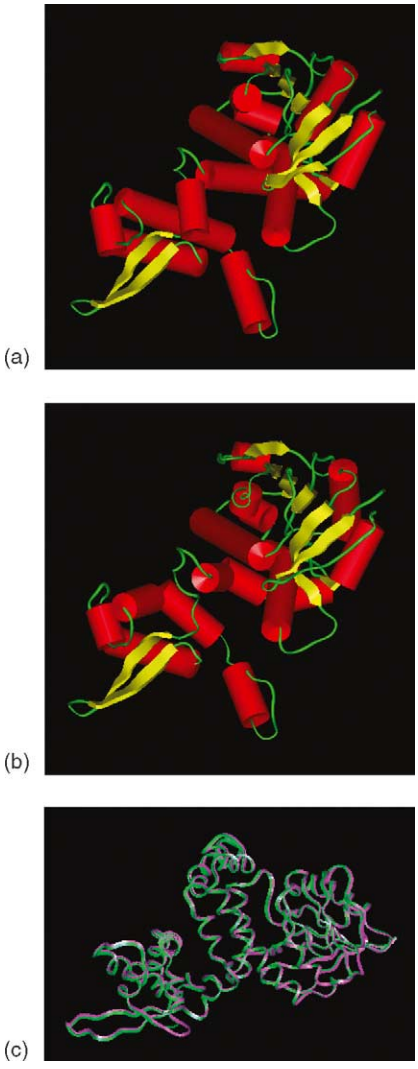

Fig. 2. (a) The three dimensional structure of COMT obtained by X-ray crystallography (1KYZc.pdb), (b) the predicted structure of AtOMT1 based on homology-based modelling, and (c) superimposition of two proteins. Green ribbon is COMT and purple one is AtOMT1.

of two proteins, Kabsch and Sander method provided by InsightII was applied because the secondary structure of COMT can be obtained from PDB but that of AtOMT1 cannot be. Even though the residues showing α-helices and β-sheets of COMT determined by Kabsch and Sander method are different with those listed in the header of PDB file, the residues determined by Kabsch and Sander method were used because the same method should be applied on two proteins. Whereas COMT includes 20 α-helices and nine β-sheets, AtOMT1 has 16 and 9, respectively. However, since the superimposition of backbones of the two proteins gives the RMSD value of 5.3 Å, it can be considered that they have a high similarity. As shown in Table 1, while in COMT the last two sheets are followed by helix 20, in AtOMT1 helix 16 is placed between the last two sheets. Four helices of COMT, missed in AtOMT1, are helices 6, 9, 15, and 17. As to the nine β-sheets, both AtOMT1 and COMT do not exhibit any difference except for the last two criteria mentioned above. That is, seven strands are mixed and two strands are anti-parallel.

Table 1
A comparison of α-helices and β-sheets included in AtOMT1 with those included in COMT

	AtOMT1			COMT	
	Start	End		Start	End
Helix 1	ALA19	ALA27	Helix 1	GLU18	SER28
Helix 2	VAL31	GLU41	Helix 2	SER30	LEU42
Helix 3	LEU44	ASN51	Helix 3	ASP43	ALA51
Helix 4	PRO57	SER62	Helix 4	SER58	SER64
Helix 5	ALA71	SER84	Helix 5	ASP72	TYR87
Sheet 1	LEU88	LYS94	Sheet 1	LEU90	THR96
Sheet 2	VAL100	LEU106	Sheet 2	VAL102	LEU108
			Helix 6	THR110	VAL116
Helix 6	ALA123	ASN129	Helix 7	ILE124	GLN132
Helix 7	LYS132	TRP138	Helix 8	ASP133	GLU138
			Helix 9	SER139	TYR141
Helix 8	LEU141	ASP147	Helix 10	HIS142	GLY150
Helix 9	PRO151	ALA155	Helix 11	ILE152	TYR158
Helix 10	ALA160	TYR163	Helix 12	THR161	HIS166
Helix 11	PRO168	THR192	Helix 13	ASP169	TYR195
Sheet 3	LEU202	VAL205	Sheet 3	LEU204	VAL207
Helix 12	ALA211	LYS214	Helix 14	GLY212	THR217
			Helix 15	ILE218	TYR222
Sheet 4	GLY225	ASN229	Sheet 4	GLY227	ASP231
Helix 13	PRO231	GLU235	Helix 16	LEU232	ASP238
Sheet 5	ILE243	GLY247	Sheet 5	VAL245	GLY249
Sheet 6	ALA259	PHE261	Sheet 6	VAL262	PHE263
			Helix 17	ILE267	TRP271
Helix 14	ASP271	SER284	Helix 18	SER272	LEU287
Sheet 7	LYS290	LEU298	Sheet 7	LYS292	LEU300
Helix 15	LEU306	HIS321	Helix 19	SER307	ASN324
Sheet 8	ARG328	THR329			
Helix 16	GLU330	ALA339	Helix 20	GLN332	ALA342
			Sheet 8	VAL349	ALA353
Sheet 9	VAL354	LEU360	Sheet 9	THR356	LEU362

The structure of AtOMT1 was evaluated using PROCHECK. Among the 305 residues two amino acids, Ser86 and Lys132 were found in the disallowed regions of Ramachandran plot (Fig. 3). When their phi and psi angles were fixed manually in the conformer obtained from MD calculation they were observed in generously allowed regions. The statistical analysis of Ramachandran plot showed that 70.0% are in the most favored regions, 25.9% in additional allowed regions, and 3.0% in generously allowed regions.

3.2. Complex structure of AtOMT1 and substrate

OMTs have a common co-factor, AdoMet. AtOMT1 receives methyl group from AdoMet and catalyzes the methylation of quercetin. The methylation results in the formation of AdoHcy and isorahmnetin, respectively. As mentioned in Section 2.3, an AdoHcy, isorhamnetin, and protein assembly was created by docking experiments. In the case of AdoHcy, the residues surrounding AdoHcy in

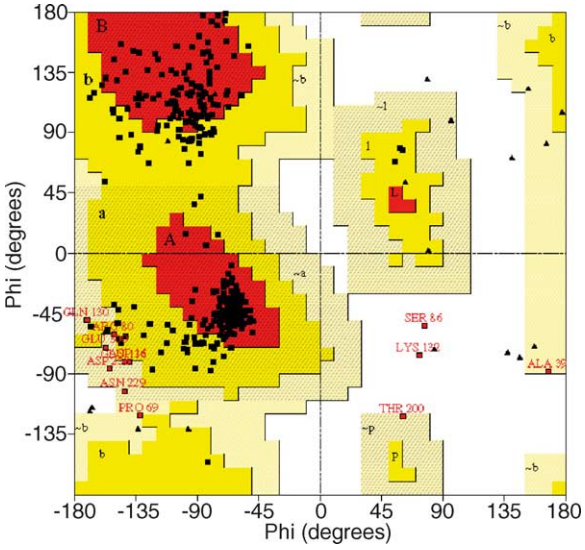

Fig. 3. Ramachandran plot of AtOMT1 obtained by PROCHECK.

COMT are matched exactly with those in AtOMT1 as listed in Table 2, so that it does not matter that the same position of AdoHcy obtained from the crystal structure of COMT was used for the docking procedure of AtOMT1 and AdoHcy. However, in the case of isorhamnetin, one residue of residues surrounding the binding site of isorhamnetin in AtOMT1 is different with that in COMT as listed in Table 2. In order to confirm whether the binding site determination for isorhamnetin is correct, the residues surrounding the substrate binding site of COMT as well as those of ChOMT and IOMT were compared in Table 2. Based on these comparisons, it can be concluded that the binding site of isorhamnetin in AtOMT1 is reliable.

Even though AdoHcy and isorahmnetin are docked into the protein, the 3-D structure of the protein was not changed much and a comparison of the backbone of the complex protein with that of the protein only gave less RMSD value than 1 Å. Fig. 4a and b show the close-up view of the binding sites of the product, isorhamnetin and AdoHcy in AtOMT1, and that of the binding sites of the product, ferulic acid and AdoHcy in COMT, respectively. A transportation of the methyl group attached in sulfur of AdoMet methylates 3′-hydroxyl group of the substrate, quercetin. While the distance between sulfur of AdoHcy and 3′-methoxy group of isorhamnetin in AtOMT1 was measured at 5.97 Å, the distance between sulfur of AdoHcy and 3-hydroxyl group of ferulic acid in COMT was 11.41 Å. Ile316 participating in the substrate binding site of COMT is switched with Val314 in AtOMT1. As shown in Fig. 4a and b, these two residues are positioned at the opposite site to the phenyl ring containing the hydroxyl group as a methyl group acceptor. Therefore, the substitution of Ile316 with Val314 may change the shape inside the substrate binding site and cause an enzyme activity.

3.3. Secondary structure of AtOMT1

A plot of the total energy versus time of MD is shown in Fig. 5. The total energy reaches equilibrium after approximately 60 ps. Fourteen ribbon structures obtained every 30 ps after 30 ps are superimposed (Fig. 6). Their RMSD is 1.0 Å which indicates that the 3-D structure is conserved during MD. The change of the numbers of helices and sheets during MD was investigated and the results are shown in Fig. 7a and b, respectively. The averages of the numbers of α-helical residues and helices are 105 and 16, respectively. In case of β-sheets, the averages are 42 and 9, respectively. The number of helices and that of sheets do not show drastic changes during MD. These results indicate that the secondary structure of AtOMT1 is conserved during MD.

COMT includes 20 α-helices and nine β-sheets whereas AtOMT1 has 16 and 9, respectively. While in the case of α-helices, the change of the number is between 13 and 21, in β-sheets, the change is between 8 and 11. This indicates that during MD a fluctuation of α-helix is major but a fluctuation of β-sheet is minor. A comparison between COMT and AtOMT1 (Table 1) shows that four helices making a difference are positioned at helices 6, 9, 15, and 17. Among these, helix 17 includes one of residues neighboring the AdoMet binding site, Trp271. However, in the case of AtOMT1 the corresponding residue is Trp269 which is not positioned in the helix. As a result, Trp269 in AtOMT1 might be flexible. In order to investigate its flexibility, its phi and psi dihedral angles are plotted against MD time. As shown in Fig. 8a, variations of phi and psi dihedral angles are

Table 2
A comparison of residues neighboring binding sites

Enzyme	Residues neighboring the binding site of AdoHcy							
COMT	Asp206	Gly208	Asp231	Leu232	Asp251	Met252	Lys265	Trp271
ChOMT	Asp215	Gly217	Asp240	Leu241	Asp260	Met261	Lys274	Trp280
IOMT	Asp194	Gly196	Asp219	Arg220	Asp239	Met240	Lys253	Trp259
AtOMT1	Asp204	Gly206	Asn229	Leu230	Asp249	Met250	Lys263	Trp269
	Residues neighboring the binding site of substrate							
COMT	Met130	Asn131	Phe176	Met180	Ile316	Ile319	Met320	Asn324
ChOMT	Phe138	Leu139	Phe185	Met189	Leu325	Leu328	Met329	Thr332
IOMT	Cys117	Val118	Phe164	Met168	Met307	Asn310	Met311	Leu314
AtOMT1	Met128	Asn129	Phe174	Met178	Val314	Ile317	Met318	Asn322

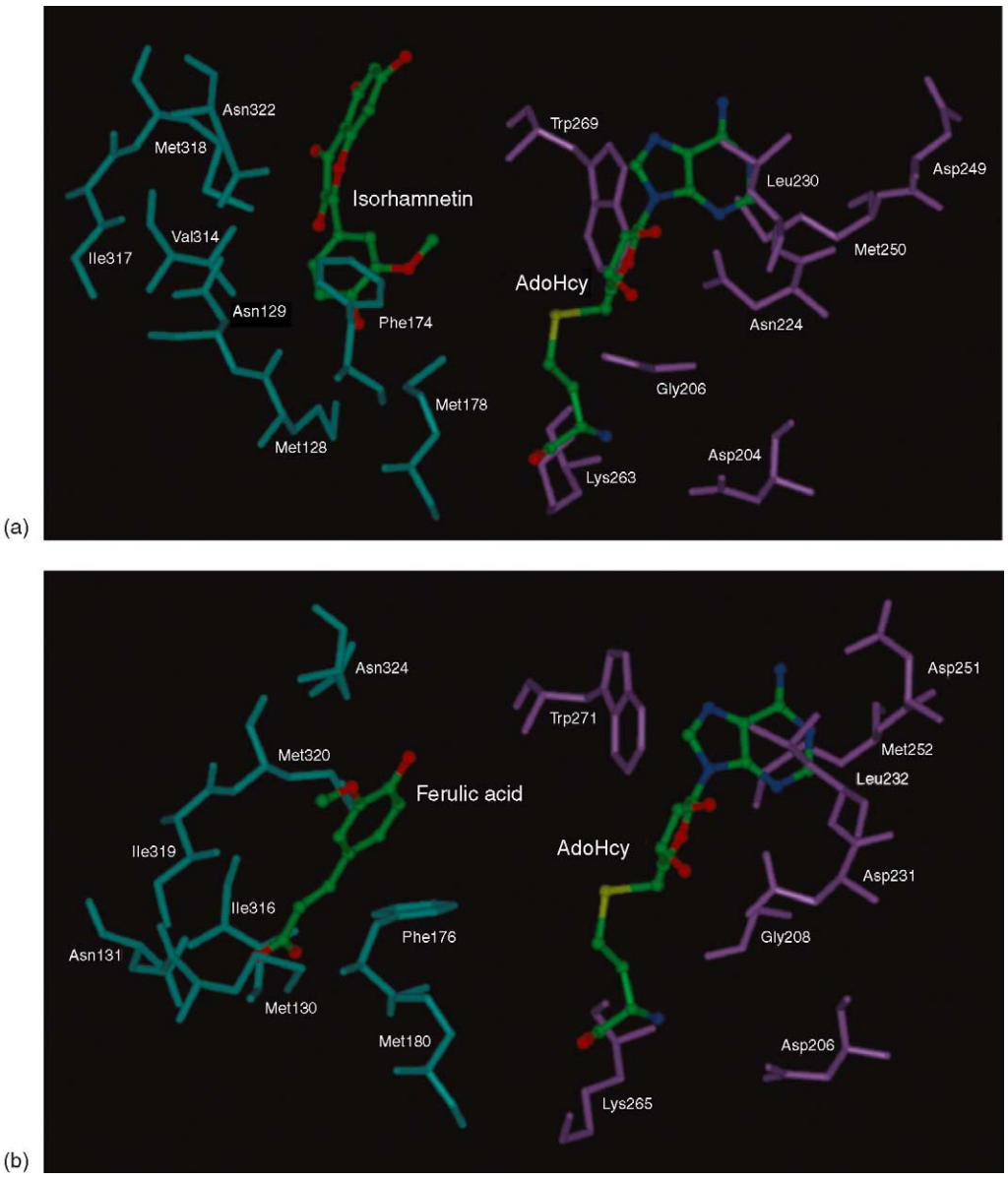

Fig. 4. A comparison of (a) the binding sites of the product, isorhamnetin and AdoHcy in AtOMT1 and (b) the binding sites of the product, ferulic acid and AdoHcy in COMT.

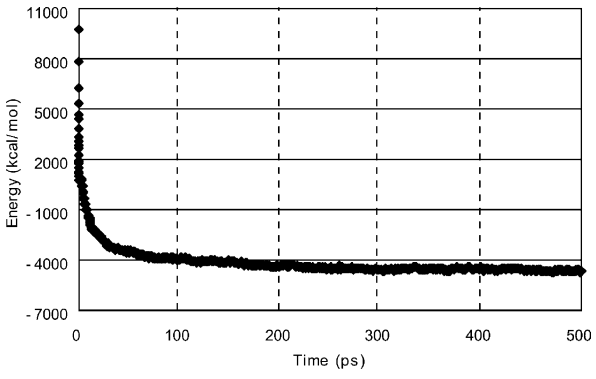

Fig. 5. A plot of total energy vs. time of molecular dynamics.

Fig. 6. Superimposition of fourteen ribbon structures obtained every 30 ps after 60 ps.

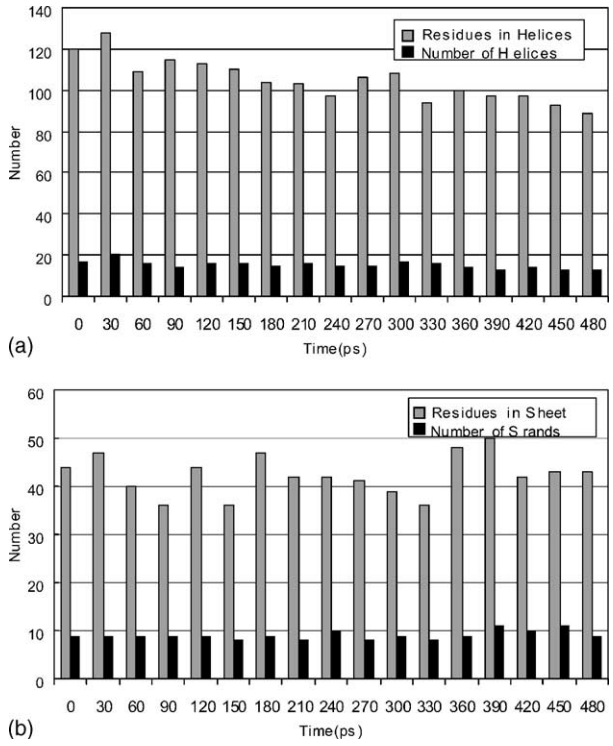

Fig. 7. The change of the numbers of (a) helices and (b) sheets during MD.

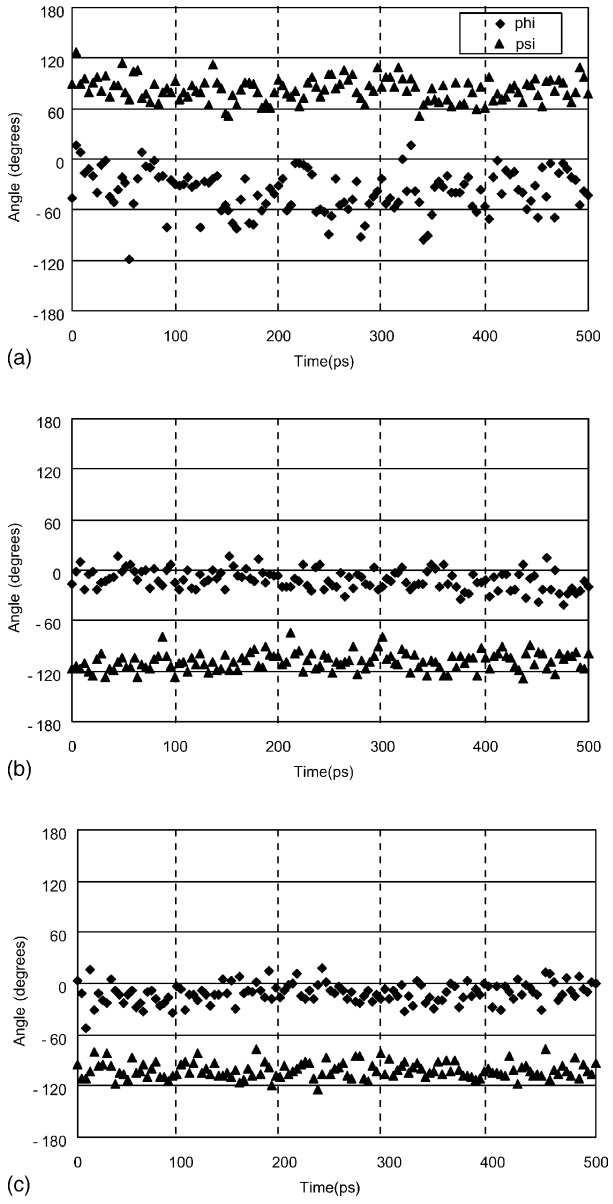

Fig. 8. Phi (◆) and psi (▲) dihedral angles of (a) Trp269, (b) Cys274 in AtOMT1, and (c) Trp271 in COMT plotted against molecular dynamics time.

placed within 120°. However, phi and psi dihedral angles of Cys274 positioned in helix14 do not show variation, which are placed within 50° (Fig. 8b). Likewise, in the case of Trp271 included in the helix in COMT, the dihedral angles are placed within 50° (Fig. 8c). As a result, the difference between the residue Trp271 in COMT included in helix17 and Trp269 in AtOMT1 not to be included in a helix may make a difference in the stability of AdoMet, which can result in different binding affinities in COMT and AtOMT1.

3.4. Substrate preference of AtOMT1

As mentioned before, COMT recognizes caffeic acid as well as caffeoyl alcohol, caffeoyl aldehyde, 5-hydroxyferulic acid, 5-hydroxyconiferyl alcohol, and 5-hydroxyconiferaldehyde [22]. Likewise, AtOMT1 has different substrate preferences over quercetin (3′,4′-dihydroxyflavonol), myricetin (3′,4′,5′-trihydroxyflavonol), and luteolin (3′,4′-dihydroxyflavone) with relative activities of 100, 70, and 15%, respectively [13]. We tried to explain the substrate preference of AtOMT1 based on its 3-D structure and docking experiment. Luteolin and myricetin were docked into the protein. The 3-D structure of myricetin was adapted from 1E7U.pdb [23], and that of luteolin was obtained from a modification of quercetin using molecular modelling.

Investigating the binding sites of AtOMT1, it was found that there is H-bond between the 3′-hydroxyl group of quercetin and the NH group of Asn129 in AtOMT1. Of course, in a protein-substrate complex structure, the substrate is isorhamnetin, so that 3′-hydroxyl group of quercetin is considered 3′-methoxy group of isorhamnetin. However, in an actual condition, H-bond should be observed between hydrogen of 3′-hydroxyl group of quercetin and nitrogen of the NH group of Asn129. The distance between proton and nitrogen constituting H-bond amounted to 2.2 Å (Fig. 9). In order to clarify the existence of H-bond, the case of COMT was studied, where H-bond was observed also between the carboxylic groups of caffeic acid and Asn131. However, since there is no 3′-hydroxyl group in luteolin, H-bond cannot be observed. It results in the formation of less stable complex of the enzyme and substrate. As a result, the substrate binding affinity of AtOMT1 for luteolin should be lower than that for quercetin. On the other hand, in the protein-myricetin complex the 5′-hydroxyl group of the

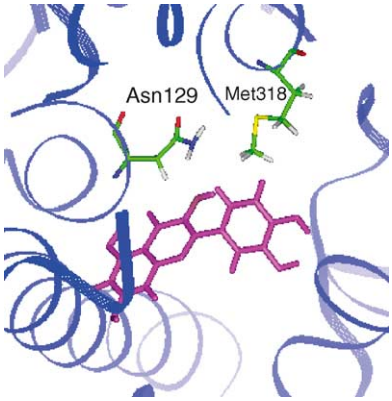

Fig. 9. The residues showing H-bond and van der Waals interaction with the substrate inside the enzyme binding site.

substrate faces the residue, Met318. The van der Waals clouds of Met318 and myricetin can conflict with each other. Therefore, the substrate binding affinity of AtOMT1 for myricetin should be lower than that for quercetin. An overlay of the docking results for quercetin, luteolin, and myricetin with AtOMT1 is shown in Fig. 10.

3.5. Substrate specificity of OMTS

It is predictable that the AdoMet binding sites are very much conserved in all of the four OMTs investigated, except for Arg220 in IOMT (Table 2) since AdoMet is the common co-substrate. In contrast, the residues neighboring the binding sites of the substrates for AtOMT1, COMT, ChOMT, and IOMT are as variable as their substrates, although caffeic acid, 3-(3,4-dihydroxyphenyl)-2-propenoic acid, forms ring B and the 3-C side chain of quercetin, 2-(3,4-dihydroxyphenyl)-3,5,7-trihydroxy-4H-1-benzopyran-4-one. However, in the cases of AtOMT1 and COMT, residues neighboring the binding site of the substrate of AtOMT1 show only one different residue, Val314 instead of Ile316 in COMT. This raises the question as to how such a change may determine substrate specificity.

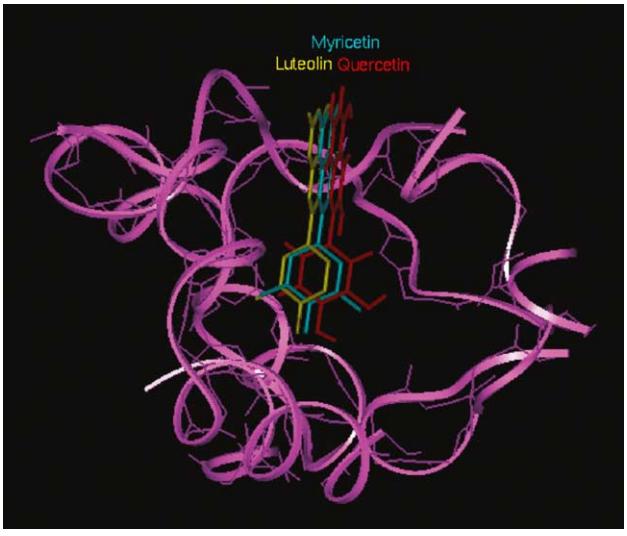

Fig. 10. An overlay of the docking results for quercetin (red), luteolin (yellow), and myricetin (cyan) with AtOMT1.

Table 3
A comparison of the residues participated in dimerization of AtOMT1 and COMT with those of ChOMT and IOMT

Enzyme	Residues		
ChOMT	Met29	Thr32	Thr33
IOMT	Tyr25	Ile28	Asp29
COMT	Met24	Ala27	Ser28
AtOMT1	Met24	Ala27	Ser28

Several possibilities to explain such strict substrate specificity can be considered. The first possible explanation would be that one amino acid difference in the binding site may be sufficient to determine substrate specificity. This is supported by the fact that chavicol, 4-(2-propenyl)-phenol, and eugenol, 2-methoxy-4-(2-propenyl)-phenol, OMTs differ in their substrate binding sites by one amino acid [24], and that switching of this single residue in both OMTs resulted in the opposite substrate preference. In the case of AtOMT1, switching of a single residue causes a change of the volume within the binding site. Ile316 of COMT has one more methylene group than Val314 of AtOMT1. The difference between their molar volumes is 17.6 cm^3, which may justify this first inference. The second possibility is caused by the dimerization interface [19] i.e., dimerization causes insertions of some amino acid residues into the catalytic domain and modifies the overall structure of the active site. The crystallographic study of ChOMT and IOMT, reveals that three residues at the N-terminal participate in dimerization. A comparison of the corresponding residues of AtOMT1 and COMT with ChOMT and IOMT is listed in Table 3. In the cases of ChOMT and IOMT, three residues show drastic difference so that it might contribute the substrate specificity of these two enzymes. In the cases of AtOMT1 and COMT, however, three residues, which are predicted to be inserted into the catalytic site when they form a dimer, are exactly similar. Therefore, this possibility may be ruled out.

On the other hand, residues consisting of a gate for the entrance of the substrate may be important. As shown in Fig. 4, substrate binding site is placed inside a cleft where a substrate docked into the enzyme must pass through. Therefore, residues positioned inside the cleft are expected not to be flexible. Likewise, the residues holding the methylated benzene ring should not be flexible either. However, Asp268 corresponding to one side of the gate must be flexible because the substrate may enter into the cleft through the gate. In order to test this hypothesis, phi and psi dihedral angles of several residues are plotted against MD time. As shown in Fig. 11a, the fluctuation of phi and psi dihedral angles of the residue placed inside a cleft, Met128 is placed within 65°. Likewise, variation of

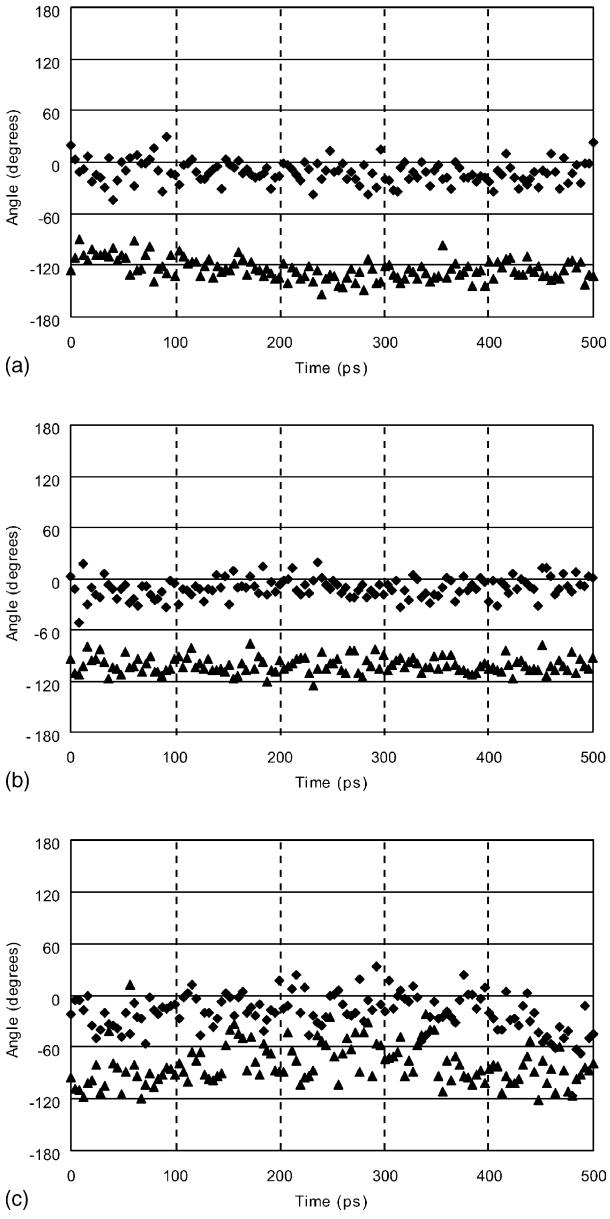

Fig. 11. Variations of phi (◆) and psi (▲) dihedral angles of (a) the residue placed inside a cleft, Met128, (b) the residue holding benzene ring methylated in the substrate, Met178, and (c) Asp268 corresponding to the gate.

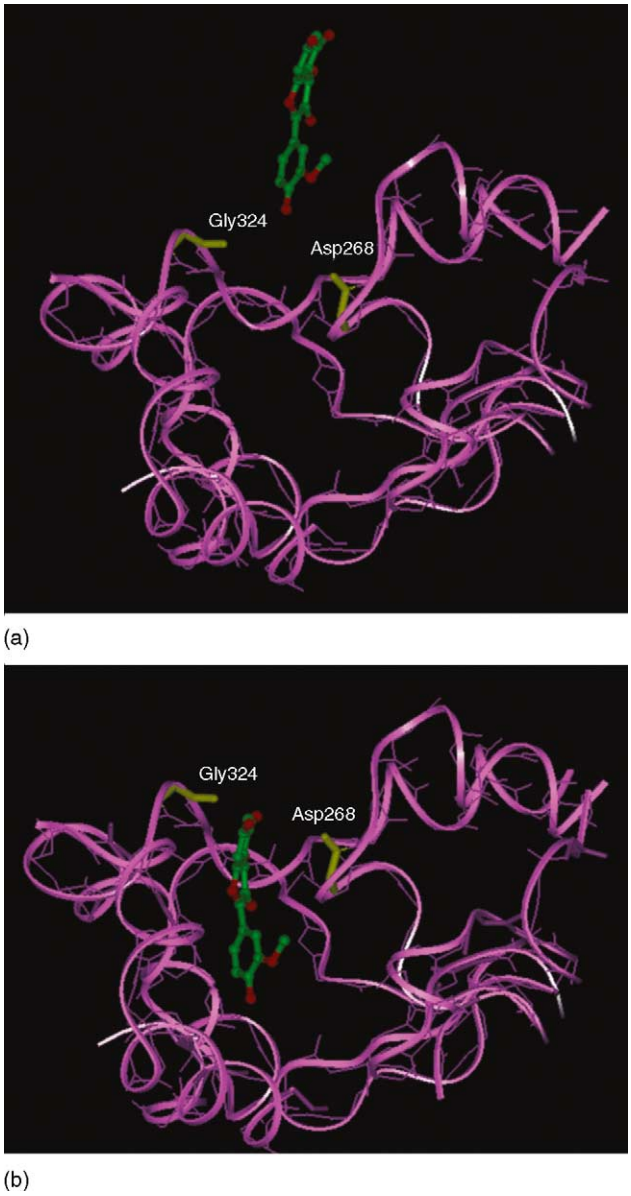

Fig. 12. The conformers (a) before and (b) after a substrate enters in a cleft surrounding the binding site.

the residue holding the methylated benzene ring, Met178 is placed within 60° (Fig. 11b). However, variation of Asp268 located at the gate is placed within 140° (Fig. 11c). The other side of the gate is composed of Gly324. In case of Gly, Phi and psi dihedral angles in glycine residues cannot be provided automatically by the software used in this experiment, and therefore, a fluctuation of the residue corresponding to the other side of the gate could not be investigated. Studies on the dynamics of residues affecting docking of the substrate into the protein reveal that one side of the gate is composed of Asp268. As mentioned in the experimental section, the docking of isorhamnetin and the protein was carried out using InsightII/Docking module. During the docking, MD calculations were performed, and the results were analyzed using InsightII/Analysis module. In this experiment, 20 conformers showing the docking progress were collected. Fig. 12a and b show the conformers before and after a substrate enters in a cleft surrounding the binding site, respectively. Likewise, two residues, Gly326 and Asp270, corresponding to the gate of the cleft are observed in COMT. While the distance between two residues of COMT is 5.0 Å, that of AtOMT1 is 7.5 Å. As a result, the gate of COMT may favor smaller a substrate than that of AtOMT1.

Therefore, we conclude that the substrate specificity of the two proteins, COMT and AtOMT1 may originate from the substrate binding site difference and the substrate entrance size.

4. Conclusions

The 3-D structure of AtOMT1 obtained by homology-based modelling shows a high similarity to COMT that was used a template. While COMT includes 20 α-helices and nine β-sheets, AtOMT1 has 16 and 9, respectively. Superimposition of the backbones of the two proteins gives RMSD value of 5.3 Å. Four helices of COMT, missed in AtOMT1, are helices 6, 9, 15, and 17. In addition, the last two sheets are followed by helix 20 in COMT, whereas helix 16 is situated between the last two sheets in AtOMT1. In the case of nine β-sheets, both AtOMT1 and COMT do not show any difference except the last two mentioned above, i.e. seven strands are mixed and two strands are anti-parallel. A plot of the total energy versus MD time shows that the total energy reaches equilibrium after approximately 60 ps. A superimposition of fourteen ribbon structures obtained every 30 ps after 30 ps gives RMSD of 1.0 Å. Therefore, the 3-D structure is conserved during MD.

Even though there is only a single amino acid difference in the residues neighboring the substrate binding site of COMT and AtOMT1, Ile316 instead of Val314, respectively, both enzymes exhibit distinct substrate specificities. In order to explain this phenomenone, several possibilities were investigated, and it was concluded that only one factor cannot give enough information about the substrate specificity. We suggest, therefore, that the substrate specificities of COMT and AtOMT1 can be explained on the basis of substrate binding site difference and the size of the gate of the cleft surrounding the binding site. As a result, the reason why two proteins exhibiting high sequence homology have different substrates should be explained based on their 3-D structural investigation. The 3-D structure of AtOMT1 obtained by homology modelling in this study was deposited in the Protein Data Bank (PDB ID: NII1.pdb).

Acknowledgements

This work was supported by Frontier 21 Crop Functional Genomics Center, Korea Ministry of Science and Technology (CG2214).

References

[1] P.M. Miller (Ed.), Phytochemistry: Organic Metabolies, vols. I and II, Van Nostrand Reinhold, New York, 1973.
[2] R.A. Dixon, Natural products and plant disease resistance, Nature 411 (2001) 843–847.
[3] B.A. Bohm, Introduction to Flavonoids. Chemistry and Biochemistry of Natural Products Series, vol. 2, Harwood Academic Publishers, Amsterdam, 1998.
[4] J.E. Poulton, Transmethylation and demethylation reactions in the metabolism of secondary plant products, in: P.K. Stumpf, E.E. Conn (Eds.), The Biochemistry of Plants: A Comprehensive Treatise, vol. 7, Academic Press, London, 1981, pp. 667–723.
[5] R.K. Ibrahim, I. Muzac, The methyltransferase gene superfamily, a tree with multiple branches, Rec. Adv. Phytochem. 34 (2000) 349–384.
[6] A.B. Christenssen, P.L. Gregersen, C.E. Olsen, D.B. Collinge, Flavonoid 7-*O*-methyltransferase is expressed in barley leaves in response to pathogen attack, Plant Mol. Biol. 36 (1998) 219–227.
[7] C.A. Maxwell, M.J. Harrison, R.A. Dixon, Molecular characterization and expression of alfalfa isoliquiritigenin 2′-*O*-methyltransferase, an enzyme specifically involved in the biosynthesis of an inducer of *Rhizobium meliloti* nodulation genes, Plant J. 4 (1993) 971–981.
[8] M. Haga, T. Akashi, T. Aoki, S. Ayabe, A cDNA for *S*-adenosyl-L-methionine:isoliquiritigenin 2′-*O*-methyltransferase from cultured licorice (*Glycyrrhiza echunata*) cells, Plant Physiol. 113 (1997) 663–665.
[9] X.-E. He, J.T. Reddy, R.A. Dixon, Stress responses in alfalfa. XXII. CDNA cloning and characterization of an elicitor-inducible isoflavone 7-O-methyltransferase, Plant Mol. Biol. 36 (1998) 219–227.
[10] A. Gauthier, P.J. Gulick, R.K. Ibrahim, cDNA cloning and characterization of 3′,4′-*O*-methyltransferase for partially methylated flavonols from *Chrysosplenium americanum*, Plant Mol. Biol. 32 (1996) 1163–1169.
[11] A. Gauthier, P.J. Gulick, R.K. Ibrahim, Characterization of two cDNA clones which encode *O*-methylatransferases for the methylation of both flavonoid and phenylpropanoid compounds, Arch. Biochem. Biophy. 351 (1998) 243–249.
[12] H. Chiron, A. Drouet, A.C. Claudot, C. Eckerskorn, M. Trost, W. Heller, D. Ernst, H. Sandermann, Molecular cloning and functional expression of a stress-induced multifunctional *O*-methyltransferase with pinosylvin methyltransferase activity from Scots pine (*Pinus sylvestris* L.), Plant Mol. Biol. 44 (2000) 733–745.
[13] I. Muzac, J. Wang, D. Anzellotti, H. Zhang, R.K. Ibrahim, Functional expression of an *Arabidopsis* cDNA clone encoding a flavonol 3′-*O*-methyltransferase and characterization of the gene product, Arch. Biochem. Biophy. 375 (1999) 385–388.
[14] X.-Z. He, R.A. Dixon, Genetic manipulation of isoflavone 7-*O*-methyltransferase enhances biosynthesis of 4′-*O*-methylated isoflavonoid phytoalexins and disease resistance in alfalfa, Plant Cell. 12 (2000) 1689–1702.
[15] K. Inoue, K. Parvathi, R.A. Dixon, Substrate preferences of caffeic acid/5-hydroxyferulic acid 3/5-*O*-methyltransferases in developing stems of alfalfa (*Medicago sativa* L.), Arch. Biochem. Biophys. 375 (2000) 175–182.
[16] K. Parvathi, F. Chen, D. Guo, J.W. Blount, R.A. Dixon, Substrate preferences of *O*-methyltransferases in alfalfa suggest new pathways for 3-*O*-methylation of monolignols, Plant J. 31 (2001) 193–202.
[17] J. Wang, E. Pichersky, Identification of specific residues involved in substrate discrimination in two plant *O*-methyltransferases, Arch. Biochem. Biophy. 368 (1999) 172–180.
[18] G. Schroder, E. Wehinger, J. Schroder, Predicting the substrates of cloned plant *O*-methyltransferases, Phytochemistry 59 (2002) 1–8.
[19] C. Zubieta, X.Z. He, R.A. Dixon, J.P. Noel, Structures of two natural product methyltransferases reveal the basis for substrate specificity in plant *O*-methyltransferases, Nat. Struct. Biol. 8 (2001) 271–279.
[20] N.A. Eckardt, Probing the mysteries of lignin biosynthesis: the crystal structure of caffeic acid/5-hydroxyferulic acid 3/5-*O*-methyltransferase provides new insights, Plant Cell. 14 (2002) 1185–1189.
[21] R. Wilmouth, J. Turnbull, R. Welford, I. Clifton, A. Prescott, C. Schofield, Structure and mechanism of anthocyanidin synthase from *Arabidopsis thaliana*. The Protein Data Bank, Structure (Lond.) 10 (2002) 93–103.

[22] C. Zubieta, P. Kota, J.L. Ferrer, R.A. Dixon, J.P. Noel, Structural basis for the modulation of lignin monomer methylation by caffeic acid/5-hydroxyferulic acid 3/5-*O*-methyltransferase, Plant Cell 32 (2002) 1265–1277.

[23] E.H. Walker, M.E. Pacold, O. Perisic, L. Stephens, P.T. Hawkins, M.P. Whymann, R.L. Williams, Structural determinations of phosphoinositide 3-kinase inhibition by Wortmannin, Ly294002, quercetin, myricetin and staurosporine. The Protein Data Bank, Mol. Cell. Biol. 6 (2000) 909–919.

[24] D.R. Gang, N. Lavid, C. Zubieta, F. Chen, T. Beuerle, E. Lewinsohn, J.P. Noel, E. Pichersky, Characterization of phenylpropene *O*-methyltransferases from sweet basil: facile change of substrate specificity and convergent evolution within a plant *O*-methyltransferase family, Plant Cell 14 (2002) 505–519.