

Substrate and Positional Specificity in Cation Dependent *O*-methyltransferases

Dissertation

For obtaining the academic degree of
doctor rerum naturalium
(Dr. rer. nat.)

submitted to
Naturwissenschaftlichen Fakultät I - Biowissenschaften
der Martin-Luther-Universität Halle-Wittenberg

Jakub Kopycki

Table of contents

List of abbreviations.....	iii
Index of tables.....	v
Index of figures.....	vi
1. Summary.....	1
2. Summary in German (Zusammenfassung).....	4
3. Theoretical introduction.....	8
3.1. Principle of the methylation reaction.....	8
3.2. Classification of methylating enzymes.....	8
3.3. S-adenosyl-L-methionine (AdoMet) regeneration cycle.....	8
3.4. The relevance of methylation in nature.....	9
3.5. <i>O</i> -Methyltransferases.....	11
3.5.1. Animal <i>O</i> -Methyltransferases.....	12
3.5.2. Plant <i>O</i> -Methyltransferases.....	12
3.5.3. Microbial <i>O</i> -methyltransferases.....	15
3.6. The three dimensional structure of proteins as a tool for studying protein functions.....	16
3.6.1. General structure elucidation techniques.....	16
3.6.2. X-ray crystallography.....	18
3.7. Structural characterization of Mg ²⁺ -dependent OMTs.....	21
3.8. Mechanism of the methylation reaction catalyzed by Mg ²⁺ -dependent OMTs.....	25
3.9. Rationale of the thesis.....	26
3.9.1. Substrate specificity of Mg ²⁺ -dependent <i>O</i> -methyltransferases.....	26
3.9.2. Position specificity of Mg ²⁺ -dependent <i>O</i> -methyltransferases.....	28
4. Materials and methods.....	30
4.1. Molecular biological procedures.....	30
4.1.1. Preparation of gene inserts.....	30
4.1.2. Preparation of gene fragments for hybrid proteins.....	30
4.1.3. Ligation.....	31
4.1.4. DNA purification by agarose gel electrophoresis.....	31
4.1.5. Preparation of plasmid DNA.....	32
4.1.6. Sequencing.....	32
4.1.7. Site directed mutagenesis.....	33
4.2. Protein expression.....	35
4.2.1. Characterization of expression strains.....	35
4.2.2. Storage strains.....	35
4.2.3. Preparation and transformation of competent <i>E. coli</i> cells.....	35
4.2.4. Expression tests.....	35
4.2.5. Expression of PFOMT and N-terminal Hybrids.....	36
4.2.6. Expression of loop Hybrids.....	36
4.2.7. Expression of SynOMT.....	36
4.2.8. Expression of <i>M. sativa</i> CCoAOMT.....	37
4.3. Protein purification.....	37
4.3.1. Preparation of soluble protein fraction.....	37
4.3.2. Immobilized metal affinity chromatography.....	37
4.3.3. Gel filtration.....	38
4.3.4. Hydrophobic interaction chromatography.....	38
4.3.5. Concentration of protein solutions.....	39
4.3.6. Determination of protein concentration.....	39
4.4. Protein analysis.....	40

4.4.1. SDS PAGE	40
4.4.2. Western blot	40
4.5. Protein modification	41
4.5.1. His-tag cleavage	41
4.6. Activity assays.....	42
4.7. Isolation of the enzymatic reaction products of SynOMT	43
4.8. Crystallographic procedures.....	43
4.8.1. Crystallization of proteins	43
4.8.2. Data collection and processing.....	44
4.8.3. Structure solution	44
4.8.4. Model building, refinement and visualization.....	44
4.8.5. <i>In silico</i> substrate docking.....	45
5. Results	46
5.1. Structural characterization of PFOMT	46
5.2. <i>In silico</i> docking experiments	49
5.3. Structural and amino acid sequence similarity of plant OMTs	51
5.4. Hybrid proteins.....	53
5.4.1. N-terminus of PFOMT controls substrate specificity	53
5.4.2. Loop Hybrid	53
5.5. Characterization of hybrid proteins.....	54
5.5.1. Production and purification of hybrid proteins	54
5.5.2. His tag removal	55
5.5.3. Activity assays.....	56
5.6. The structure and characterization of SynOMT	60
5.6.1. Purification of SynOMT.....	60
5.6.2. Identification of the para methylation carried out by SynOMT.....	61
5.6.3. Kinetic characterization of SynOMT	65
5.6.4. Crystallization and structure determination of SynOMT	66
5.6.5. Structural details.....	67
5.6.6. Topology of the SynOMT	68
5.6.7. Binding sites	70
5.6.7.1. Ion binding site of SynOMT	70
5.6.7.2. Substrate binding site	71
5.6.7.3. AdoMet binding site.....	75
5.6.8. N-terminal mutations.....	76
5.6.9. Expression and purification of the SynOMT K3A mutant	78
5.6.10. The kinetic parameters of the K3A SynOMT mutant	78
6. Discussion	80
6.1. Substrate specificity of cation dependent OMTs	80
6.2. Position specificity of cation dependent OMTs	84
6.3. Summary and outlook	88
7. List of References.....	92

List of abbreviations

5-CH3-THF	N5-methyltetrahydrofolate
AdoHcy	S-adenosyl-L-homocysteine
AdoMet	S-adenosyl-L-methionine
APS	ammonium persulfate
ATP	adenosine triphosphate
BCIP	5-bromo-4-chloro-3-indolyl phosphate
BSA	bovine serum albumin
CCoA	coenzyme A
CCoAOMT	coenzyme A <i>O</i> -methyltransferase
CFeSP	cobamide-dependent methyl carrier protein
COMT	catechol <i>O</i> -methyltransferase
COMTD1	catechol- <i>O</i> -methyltransferase domain containing 1
DMF	dimethylformamide
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
EDTA	ethylenediaminetetraacetic acid
ESI-MS	electrospray ionization mass spectrometry
GC/TOFMS	gas chromatography coupled with time-of-flight mass spectrometry
HMBC	heteronuclear multiple bond correlation
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HIC	hydrophobic interactions chromatography
HPLC	high-performance liquid chromatography
HSQC	heteronuclear single quantum coherence
IMAC	immobilized metal affinity chromatography
IPTG	isopropyl-beta-D-thiogalactopyranoside
$K_{m\ app}$	apparent K_m
KPi	potassium phosphate buffer
LB	Lysogeny broth (LB), also known as Luria broth or Luria-Bertani broth
LiOMT	OMT from <i>Leptospira interrogans</i>
MAD	multiple wavelength anomalous diffraction
MES	2-(N-morpholino)ethanesulfonic acid
Met	methionine
NBT	nitro blue tetrazolium

NMR	nuclear magnetic resonance
OMT	<i>O</i> -Methyltransferase
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PFOMT	phenylpropanoid and flavonoid <i>O</i> -Methyltransferase
RCF	relative centrifugal force
RMSD	root mean square deviation
rOMT	catechol OMT from rat (<i>Rattus norvegicus</i>)
SAD	single-wavelength anomalous diffraction
SDS	sodium dodecyl sulfate
SDS-PAGE	polyacrylamide gel electrophoresis
SeMet	seleno-methionine
SGT	sinapic acid glucose transferase
SOMT-2	flavone OMT from Soybean (<i>Glycine max</i>)
SynOMT	<i>O</i> -Methyltransferase from <i>Synechocystis</i> sp. PCC 6803
THF	tetrahydrofolate
TBS buffer	Tris buffered saline
TEMED	N,N,N',N'-Tetramethylethylenediamine
TMS	trimethylsilyl derivative
Tris	trishydroxymethylaminomethane
UDP-glucose	uridine diphosphate glucose
w/v	weight by volume

Index of Tables

Table 1. The Table of Primers used.....	34
Table 2. Crystallographic data, phasing and refinement statistics for PFOMT	47
Table 3. Results of in silico docking of quercetin, caffeic acid, and 5-hydroxyferuloyl CoA into the active sites of PFOMT and CCoAOMT	50
Table 4. Approximate yields of the protein production.....	55
Table 5. The Km app values for investigated proteins. The values in parentheses represent the error of the curve fitting result.	57
Table 6. The kcat/Km app values for investigated proteins	59
Table 7. The NMR data of 3,5-dihydroxy-4-methoxycinnamic acid.	63
Table 8. GC/TOFMS data of derivatized hydroxycinnamic acid standards and products of 3,4,5-trihydroxycinnamic acid methylation catalyzed by SynOMT.	64
Table 9. The kinetic parameters of SynOMT	65
Table 10. Crystallographic data, phasing and refinement statistics for SynOMT (PDB code 3CBG).....	67
Table 11. Cation interacting residues	71
Table 12. Residues comprising the binding cavity for AdoHcy molecule	75
Table 13. The percentage of specific activity of SynOMT K3A mutant in comparison to the wild type SynOMT (100%).....	79

Index of figures

Figure 1. AdoMet regeneration cycle.....	9
Figure 2. Schematic diagram showing the biochemical pathways leading to the production of lignin precursors (Humphreys et al., 2002).....	13
Figure 3. Commonly used setups for protein crystallization.....	18
Figure 4. General topology of Mg ²⁺ -dependent OMTs according to (Martin et al., 2002; Schluckebier et al., 1995).....	22
Figure 5. Structural comparison of OMTs involved in phenylpropanoid metabolism.	22
Figure 6. Representation of structural differences between plant and animal Mg ²⁺ dependent OMTs.	23
Figure 7. Dimer representation of PFOMT	24
Figure 8. The dimerization of COMT, a cation independent plant OMT	24
Figure 9. Mechanism of the reaction catalyzed by Mg ²⁺ dependent OMTs.	25
Figure 10. Three dimensional representation of the catalytic center of rOMT from rat (Vidgren et al., 1994).	26
Figure 11. The amino acid sequence alignment of substrate specific CCoAOMTs and promiscuous OMTs (<i>M. crystallinum</i> PFOMT and <i>S. longipes</i> OMT).....	27
Figure 12. Differences in positional specificity between plant and microbial OMTs.	28
Figure 13. Dimer representation of PFOMT.....	46
Figure 14. Representation of the AdoMet/AdoHcy binding site of PFOMT.....	48
Figure 15. Manual Docking of the substrate quercetagenin in the active site of PFOMT.....	49
Figure 16. Amino acid sequence alignment of substrate specific and promiscuous cation dependent OMTs.....	51
Figure 17. The overlay of structures of PFOMT yellow and <i>M. sativa</i> CCoAOMT green.	52
Figure 18. Schematic representation of cloning strategy which was used to produce the hybrid proteins.....	54
Figure 19. Coomassie stained SDS PAGE gel of purified hybrid proteins.....	55
Figure 20. Protease Factor Xa His-tag cleavage reaction.	56
Figure 21. Sample chromatograms of the enzymatic reaction of NO.....	57
Figure 22. The charts show the results of the curve fitting for NO and two substrates A caffeoylglucose and B caffeic acid.....	58
Figure 23. The IMAC purification of SynOMT.....	60
Figure 24. The chromatograms and UV spectra of SynOMT and PFOMT reaction mixtures with trihydroxy cinnamic acid.....	61
Figure 25. Reaction scheme contrasting the position specificity of methylation carried by A, SynOMT from <i>Synechocystis</i> sp. strain PCC6803 and B, PFOMT from <i>M. crystallinum</i>	62
Figure 26. The crystal of SynOMT for which a data set was collected.	66
Figure 27. Dimeric structure of SynOMT.	68
Figure 28. Topology of SynOMT in spatial orientation.....	69
Figure 29. Superposition of structures of plant OMTs and the microbial OMT.....	69
Figure 30. Representation of SynOMT cation binding site.....	70
Figure 31. Binding of two reaction products the active site of SynOMT.	72
Figure 32. Interactions of substrate bound the active site of SynOMT.....	73
Figure 33. Electrostatic surface representation of SynOMT monomer with substrates docked in different orientations.	74
Figure 34. Binding of a AdoHcy cofactor on the example of SynOMT.	76
Figure 35. IMAC purification of the SynOMT K3A mutant.	78
Figure 36. Dimer representation of A LiOMT (monomers colored in blue and cyan) and B PFOMT (monomers colored in pale yellow and brown)	81

Figure 37. Catalytic efficiency profile for the investigated proteins towards different substrates.	83
Figure 38. Structure based amino acid sequence alignment of Mg ²⁺ dependent OMTs from different organisms.....	85
Figure 39. The scaffold of Cryptophycin-1.....	88
Figure 40. Bootstrapped cladogram of selected CCoAOMTs and CCoAOMT-like proteins from pro- and eukaryotes (Kopycki et al., 2008b).	90

1. Summary

In plants, *O*-methylation is a common step in the biosynthetic pathways of many naturally occurring compounds. The enzymes catalyzing methylation reactions of phenolics with a vicinal dihydroxy systems in plants can be divided into two major classes, low molecular weight Mg^{2+} dependent and the higher molecular weight enzymes that do not require Mg^{2+} for enzymatic activity. Magnesium dependent *O*-methyltransferases (OMTs) can be further divided into two groups according to their substrate specificity. On the one hand a group of substrate specific Caffeoyl Coenzyme A *O*-methyltransferases (CCoAOMTs), which were determined to be one of the key elements in lignin biosynthesis, can be distinguished. On the other there is a an array of promiscuous OMTs, that are able *in vitro* to accept as substrates a wide variety of chemical compounds with vicinal dihydroxy systems of phenolics. The representative of the second group is PFOMT – phenylpropanoid and flaovonoid *O*-methyltransferase. The amino acid sequence analysis of the members of both groups indicates a vary high level of similarity. Such sticking similarity can also be observed at the structural level. Although very similar, the members of both groups of enzymes have a different substrate preference. The first aim of this work was to elucidate the mechanisms governing the substrate specificity of plant Mg^{2+} dependent OMTs.

The methylation reactions carried out by plant Mg^{2+} dependent OMTs have very strict positional preference. The methyl group is transferred only to the hydroxyl group located in *meta* position of the phenyl ring moiety. Such situation is not observed for microbial and animal enzymes. The animal and microbial counterparts of plant Mg^{2+} dependent OMTs are able to methylate the hydroxyl groups in both *meta* and *para* positions of the phenyl ring. The discovery of an algal cation dependent OMT capable of methylation in both positions gave an opportunity to attempt an explanation of the mechanisms responsible for positional specificity in cation dependent OMTs. This was the second theme of the thesis.

To address the question of substrate specificity in plant cation dependent OMTs a comparison of the representatives from each group was carried out. The amino acid sequence alignments of the Mg^{2+} dependent OMTs belonging to both groups indicate that there are two regions where the least similarity is found: the N-terminus and the flexible loop located close to the C-terminus. The structural data further support the hypothesis that those parts of the protein may be responsible for determination of the substrate specificity. Artificial proteins were designed to investigate the influence of the mentioned parts on substrate specificity. In this way the N-terminal hybrid was created combining the N-terminus of the substrate specific CCoAOMT from *M. sativa* with the scaffold of the promiscuous PFOMT. To test the

influence of the flexible loop region a “loop” hybrid was created where the amino acid sequence coding for the flexible loop of CCoAOMT region was introduced into the promiscuous enzyme. The “double” hybrid is the combination of both previous modifications. The hybrid proteins were expressed as N-terminal his tag fusions and because the N-terminus is a focus of the experiments a protease cleavage site was introduced between the his tag and the protein sequence. This enabled production of enzymes with no tags and determination whether the his tag has an influence on the enzymatic activity. Because of the problems with expression this procedure was only possible for the N-terminal hybrid and PFOMT.

The artificial proteins were expressed, purified and used to determine the kinetic parameters towards the compounds commonly accepted as substrates for these enzymes. The k_{cat} and the $K_{m\ app}$ values were recorded for each of the investigated protein towards five substrates to produce a substrate specificity profile. The values obtained from the hybrid proteins were compared with PFOMT and *M. sativa* CCoAOMT. The inspection of the obtained data revealed the major influence of the loop part on the substrate specificity. The Hybrids containing the loop part exhibited higher catalytic efficiency towards the CCoA than caffeic acid and caffeoylglucose. This fact makes them similar in catalytic behavior to substrate specific CCoAOMT, for which the CCoA is the preferred substrate and caffeoylglucose together with caffeic acid are very poorly methylated.

The results obtained for the N-terminal hybrid are less conclusive. Those proteins showed largely reduced affinity towards all substrates tested. In the obtained data the influence of the N-terminal his tag can be observed. The proteins where this affinity tag was cleaved off show k_{cat} reduction towards the caffeic acid and CCoA. For flavonoid substrates the decrease in catalytic efficiency was not so pronounced. Taking into consideration the presented results it can be speculated that the N-terminal residues of PFOMT are responsible for the recognition of the flavonoid substrates while the mentioned loop part plays an essential role in the positioning and methylation of CCoA.

The question concerning the positional specificity in cation dependent OMTs could be answered by the comparison of structures of those enzymes coming from plants and cyanobacteria. The SynOMT from cyanobacterium *Synechocystis* sp. strain PCC 6803 was expressed in *E. coli* cells. While the natural substrate of this enzyme is currently unknown, it was shown that SynOMT is capable of methylation in both *meta* and *para* positions of the vicinal dihydroxyl systems. This protein was affinity purified and characterized. The *para* methylation carried out by SynOMT was discovered and proven by the analysis of the methylation products of an artificial substrate, 3,4,5-Trihydroxycinnamic acid. In order to

obtain the structure, SynOMT was successfully co-crystallized with caffeic acid and the methyl group donor (AdoMet) as well as the bivalent cation (Mg^{2+}). The crystal structure of SynOMT was determined for the 2Å resolution data set. Because the crystallization setup included AdoMet, the electron density reveals caffeic acid monomethylation products (ferulic acid and isoferulic acid) bound in the active site of the enzyme. The presence of the two reaction products methylated in different positions of the phenolic moiety further confirm the unusual positional specificity exhibited by SynOMT.

It was not possible to unequivocally determine the precise structural features of the SynOMT that contribute to this unexpected positional methylation specificity. One possibility is provided by the presence of side chains of His174 and Lys176 in the neighborhood of the propenoic acid moiety, equivalent to Asn and Ser in the plant CCoAOMTs, respectively. The positively charged side chains of those residues may be responsible for forming additional interactions with substrate molecule and immobilizing it in the position where methylation of the hydroxyl group in *para* position can take place.

2. Summary in German (Zusammenfassung)

Ein häufiger Modifizierungsschritt, den Metaboliten in pflanzlichen Stoffwechselwegen durchlaufen, ist die *O*-Methylierung. Die Enzyme die diesen Schritt in Pflanzen katalysieren, lassen sich dabei in der Regel einer von zwei Gruppen zuordnen. Während Vertreter der ersten Gruppe ein Molekulargewicht von 23 – 27 kDa aufweisen und für ihre katalytische Aktivität die Anwesenheit Mg^{2+} -Ionen voraussetzen, katalysieren die etwa 38 – 42 kDa großen Vertreter der zweiten Gruppen den Methylgruppentransfer über einen Mg^{2+} -unabhängigen Mechanismus. In Abhängigkeit von ihrer Substratspezifität werden die Magnesium-abhängigen *O*-Methyltransferasen (OMTs) dabei nochmals in zwei Untergruppen unterteilt. Dabei handelt es sich zum einen um die Koffoyl-Coenzym A-*O*-Methyltransferasen (CCoAOMT), die sich durch eine strenge Substratspezifität auszeichnen und unter anderem eine zentrale Rolle in der Ligninbiosynthese einnehmen. Daneben gibt es aber auch noch eine weitere Gruppe von Kationen-abhängigen OMTs, wie die PFOMT (Phenylpropanoid- und Flavonoid *O*-Methyltransferase) aus *Mesembryanthemum crystallinum*, für die durch *in vitro*-Versuche aufgezeigt werden konnte, dass sie ein wesentlich breiteres Spektrum an phenolischen Substanzen mit vicinalen Hydroxygruppen als Substrate akzeptieren. Sequenzvergleiche ergaben, dass Vertreter dieser beiden Gruppen auf Aminosäureebene viele Gemeinsamkeiten aufweisen. Auch auf der Ebene der Proteinstrukturen gleichen sich die beiden Enzymgruppen. Trotz dieser strukturellen Gemeinsamkeiten unterscheiden sich die Vertreter dieser beiden Gruppen jedoch bezüglich ihrer Substratspezifität.

Ein Ziel dieser Dissertation war es daher, aufzuklären, welche Mechanismen diese Unterschiede in der Substratspezifität der Mg^{2+} -abhängigen OMTs bedingen.

Der zweite Teil der Arbeit fokussierte die Positionsspezifität der von den Kationen-abhängigen OMTs katalysierten Methylierung. So konnte für alle bisher untersuchten pflanzlichen Vertreter dieser Enzymgruppe eine strenge Positionsspezifität aufgezeigt werden, wobei sie ausschließlich den Methylgruppentransfer auf die *meta*-Hydroxygruppe des Phenyl-Ring-Systems ihrer Substrate katalysieren. Tierischen und mikrobiellen Mg^{2+} -abhängigen OMTs fehlt diese strikte Positionsspezifität. Zwar weisen auch die tierischen Catechol-OMTs (COMTs) eine starke Präferenz für die Methylierung der *meta*-Hydroxygruppe auf, doch können sie in geringerem Maße auch den Transfer in *para*-Stellung katalysieren. Bei einigen der prokaryotischen Vertreter konnte, für einen Teil der untersuchten Substrate, sogar eine Umkehr der Präferenz hin zur Bevorzugung der Methylierung in *para*-Position beobachtet werden. Einer dieser prokaryotischen Vertreter ist die SynOMT aus dem Cyanobakterium *Synechocystis* sp. PCC 6803. Im Rahmen dieser Dissertation wurde die Kristallstruktur dieses

Enzyms mit dem Ziel aufklärt, durch einen Vergleich mit den Strukturen pflanzlicher Kationen-abhängiger OMTs, neue Einblicke in die strukturellen Hintergründe für die Positionsspezifität dieser Enzyme zu gewinnen.

Um der Frage nach der Substratspezifität der pflanzlichen Kationen-abhängigen OMTs nachzugehen, wurden die Sequenzen und Strukturen von Vertretern beider Gruppen miteinander verglichen. Der Abgleich der Aminosäuresequenzen offenbarte, dass die größten Unterschiede zwischen den beiden Enzymgruppen im N-terminalen Abschnitt und einer flexiblen *loop*-Region am C-terminalen Ende auftreten. Die Vermutung, dass diese beiden Proteinabschnitte möglicherweise für die beobachtete Substratspezifität verantwortlich sind, wurde durch die Betrachtung der Strukturinformationen weiter gestützt.

Zur Untersuchung des Einflusses dieser Regionen auf die Substratspezifität, wurden verschiedene artifizielle Hybridproteine entworfen. Bei einem dieser Hybriden wurde dabei die Grundstruktur von einem Vertreter aus der Untergruppe der Kationen-abhängigen OMTs mit der geringeren Substratspezifität (PFOMT) und die N-terminale Region eines der Kaffeoyl-CoA-spezifischen CCoAOMT-Proteine (*M. sativa*-CCoAOMT) miteinander kombiniert. Für die Überprüfung des Einflusses der flexiblen *loop*-Region, wurde im Weiteren eine Kombination aus der Aminosäuresequenz dieses Bereichs ausgehend von der CCoAOMT mit der Sequenz der PFOMT erstellt. Bei der zuletzt bereitgestellten Doppel-Hybrid-Variante, handelte es sich um ein Enzym, das die beiden zuvor genannten Modifikationen in sich vereinte.

Zur Abtrennung des, für die Aufreinigung der rekombinant hergestellten Proteine benötigten, N-terminalen His-Tags, wurde zwischen dem His-Tag und der eigentlichen Protein-Sequenz eine zusätzliche Protease-Schnittstelle eingeführt. Auf diese Weise konnten mittels einer, der Aufreinigung nachgeschalteten, Protease-Behandlung Enzyme hergestellt werden, die keinen His-Tag mehr enthielten. Auf Grund von Expressionsproblemen, konnte diese Methode jedoch nur für die PFOMT und das N-Terminus-Hybridprotein angewendet werden.

Zur Quantifizierung der Einflüsse der verschiedenen Proteinregionen, wurden die rekombinant hergestellten und über Metallaffinitätschromatographie aufgereinigten Hybridproteine zur Bestimmung der kinetischen Parameter verwendet und im Anschluss mit den Parametern der beiden Ausgangsenzyme verglichen. Als Testsubstrate wurden dabei eine Reihe von Substanzen eingesetzt, die von Vertretern der beiden Enzymklassen für gewöhnlich als Substrate akzeptiert werden.

Zur Untersuchung des Einflusses des N-terminalen Abschnittes bzw. der C-terminalen *loop*-Region Mg^{2+} -abhängiger OMTs auf deren Substratspezifität, wurden Hybridproteine erzeugt, in denen die entsprechenden Abschnitte von zwei Vertretern der beiden Untergruppen auf unterschiedliche Weise miteinander kombiniert wurden. Zur Erstellung spezifischer Substratspezifitätsprofile, wurden die Hybridproteine kinetisch untersucht und die k_{cat} - und $K_{m\ app}$ - Werte gegenüber 5 verschiedenen Substraten bestimmt. Durch den Vergleich der dabei erhaltenen Parameter mit denen der beiden Ausgangsenzyme (PFOMT, *M. sativa*-CCoAOMT) ergab sich, dass die Substratspezifität hauptsächlich von der C-terminalen *loop*-Region bestimmt wird. So wiesen Hybridproteine, in denen die PFOMT-Grundstruktur mit der C-terminalen CCoAOMT-*loop*-Region kombiniert war, eine höhere katalytische Aktivität gegenüber Kaffeoyl-CoA (CCoA) als gegenüber Kaffeensäure oder Kaffeoyl-Glucose auf. In dieser Eigenschaft glichen sie eher den CCoAOMTs, die eine deutliche Präferenz gegenüber dem Kaffeoyl-CoA-Ester aufweisen und die korrespondierende Säure bzw. das Kaffeoyl-Glucose nur schwach umsetzen. Die Ergebnisse der Hybriden, in denen der N-terminale Abschnitt der CCoAOMT mit der PFOMT kombiniert wurde, waren hingegen weniger eindeutig. Für diese Proteine konnte eine grundlegende Absenkung der Affinität gegenüber allen getesteten Substraten beobachtet werden. Ausgehend von den erhaltenen Daten ist ein signifikanter Anteil an dieser Affinitätsabsenkung jedoch auf den N-terminal angefügten His-Tag zurückzuführen. Für Proteine, in denen dieser His-Tag entfernt worden war, konnte eine signifikante Absenkung des k_{cat} -Wertes gegenüber Kaffeensäure und Kaffeoyl-CoA beobachtet werden. Für die eingesetzten Flavonoide war scheinbar der Abfall in der katalytischen Effizienz hingegen deutlich geringer zu sein.

Zusammenfassend scheint der N-terminale Abschnitt der PFOMT demnach maßgeblich an der Erkennung der Flavonoide beteiligt zu sein, während der C-terminalen *loop*-Region eine entscheidende Rolle bei der Positionierung und Methylierung von Kaffeoyl-CoA zukommt.

Um der Frage nach den strukturellen Hintergründen für die Positionsspezifität der Kationen-abhängigen OMTs nachzugehen, wurden die Kristallstrukturen einiger pflanzlicher Vertreter und der cyanobakteriellen SynOMT aus *Synechocystis* sp. PCC 6803, die eine davon abweichende Positionsspezifität aufweist, miteinander verglichen. Zum Erhalt der SynOMT-Röntgenkristallstruktur, wurde rekombinant in *E. coli* hergestelltes und über Metall-Affinitätschromatographie aufgereinigtes Enzym zusammen mit bekannten Substraten, dem Methylgruppendonator *S*-Adenosyl-Methionin und Mg^{2+} als bivalentem Kation kokristallisiert. Zur Überprüfung der Positionsspezifität der SynOMT, wurde ein Teil der rekombinant hergestellten SynOMT außerdem auch zur kinetischen Charakterisierung verwendet.

Die Fähigkeit der SynOMT, Methylierungen in *para*-Position vornehmen zu können, wurde durch den Einsatz des artefiziellen Substrates 3,4,5-Trihydroxymethylsäure und der anschließenden Analyse der Produkte entdeckt.

SynOMT konnte zusammen mit Kaffeesäure, S-Adenosyl-Methionin und MgCl₂ erfolgreich kokristallisiert und die Kristallstruktur des Enzyms, mit einer Auflösung von 2 Å, aufgeklärt werden. Die Zugabe von Substrat, Methylgruppendonor und Mg²⁺ führte dazu, dass in der resultierenden Elektronendichte auch die, im aktiven Zentrum der SynOMT assoziierten, Methylierungsprodukte der Kaffeesäure (Ferula- und Isoferulasäure) beobachtet werden konnten. Das parallele Auftreten dieser beiden, an unterschiedlichen Positionen des aromatischen Ringes methylierten Produkte, ist ein weiterer Hinweis auf die ungewöhnliche Positionsspezifität der SynOMT.

Die Frage, welche strukturellen Eigenschaften der SynOMT zu der beschriebenen Positionsspezifität führen, konnte im Rahmen dieser Dissertation nicht vollständig geklärt werden. Eine Möglichkeit könnte der Austausch eines konservierten Asn- bzw. Ser-Restes aus pflanzlichen CCoAOMTs durch His174 bzw. Lys176 in SynOMT sein, deren Seitengruppen sich in unmittelbarer Nähe zum Propensäure-Rest des Substrates befinden. Die positiv geladene Seitenketten dieser Aminosäurereste können für zusätzliche Interaktionen mit dem Substrat verantwortlich sein, die ihn in einer Position immobilisieren, wo die Methylierung in *para*-Position möglich ist.

3. Theoretical introduction

3.1. Principle of the methylation reaction

Methylation in which a methyl moiety is introduced into a chemical compound is one of the most common chemical reactions that take place in living organisms. In organic chemistry, methylation is referred to as an alkylation reaction where only one CH₃ group is delivered (Morrison et al., 1992). The most commonly used compounds that can deliver a methyl group are methyl halogens, dimethyl sulfate and dimethyl carbonate. This reaction is neither regio- nor stereoselective. In living organisms, methylation is carried out by specialized enzymes that are capable of transfer of the methyl group from a donor molecule, often S-adenosyl-L-methionine (AdoMet), to an acceptor molecule, the methyltransferases.

3.2. Classification of methylating enzymes

Within the group of enzymes that catalyze transfer reactions a subclass of enzymes that transfer one carbon (E.C.2.1.-.-) is distinguished. A prominent member of this subclass are the methyltransferases (E.C.2.1.1.-). These enzymes catalyze the transfer of a methyl group from a donor molecule (usually AdoMet or N⁵-methyltetrahydrofolate) to an acceptor molecule. The members of this family are differentiated according to the atom that is methylated and the compound that they methylate. There are four major types of methyltransferases, *N*-, *S*-, *O*- and *C*-methyltransferases. *N*-methyltransferases add a methyl group to the nitrogen atom of an acceptor molecule, which may be an amino acid amine moiety or nucleotide hetero nitrogen. Examples of those *N*-methyltransferases may be histone arginine N-MT, lysine N-MT (Guo et al., 2007), adenine m⁶A MT and cytosine m⁵C MT (Timinskas et al., 1995). *S*-methyltransferases usually add methyl function to sulfur atom of cysteine, homocysteine as well as methionine like betaine-homocysteine *S*-methyltransferase (Rao et al., 1998). *O*-methyltransferase methylate the hydroxy group oxygens. Finally the *C*-methyltransferases like Cytosine m⁴C MT (Timinskas et al., 1995) add the methyl group to the carbon atoms of various chemical compounds.

3.3. S-adenosyl-L-methionine (AdoMet) regeneration cycle

The most wide spread methyl group donor molecule used in nature is AdoMet. It is produced in living organisms during the activated methyl cycle or S-methionine cycle (Figure 1) by the action of methionine adenosyltransferase (Hanson et al., 2001; Lu, 2000). The methyl group

donor in the AdoMet regeneration cycle is N5-methyltetrahydrofolate (5-CH₃-THF). Once AdoMet donates the methyl group in a methylation reaction it is converted into the demethylated form S-adenosyl-L-homocysteine (AdoHcy) (Giovanelli, 1987). The AdoHcy molecule is then degraded to L-homocysteine (Hcy) by the action of S-adenosylhomocysteine hydrolase [EC 3.3.1.1], which in turn can accept the methyl group from N5-methyltetrahydrofolate (5-CH₃-THF) with the help of methionine synthase [EC 2.1.1.14] to produce methionine (Met) and the tetrahydrofolate (THF). AdoMet is regenerated by the action of methionine adenosyltransferase [EC 2.5.1.6]. In plants there is no evidence that this process takes part anywhere else but in the cytosol, however, regeneration of methionine can also take part in the chloroplasts (Ravanel et al., 2004).

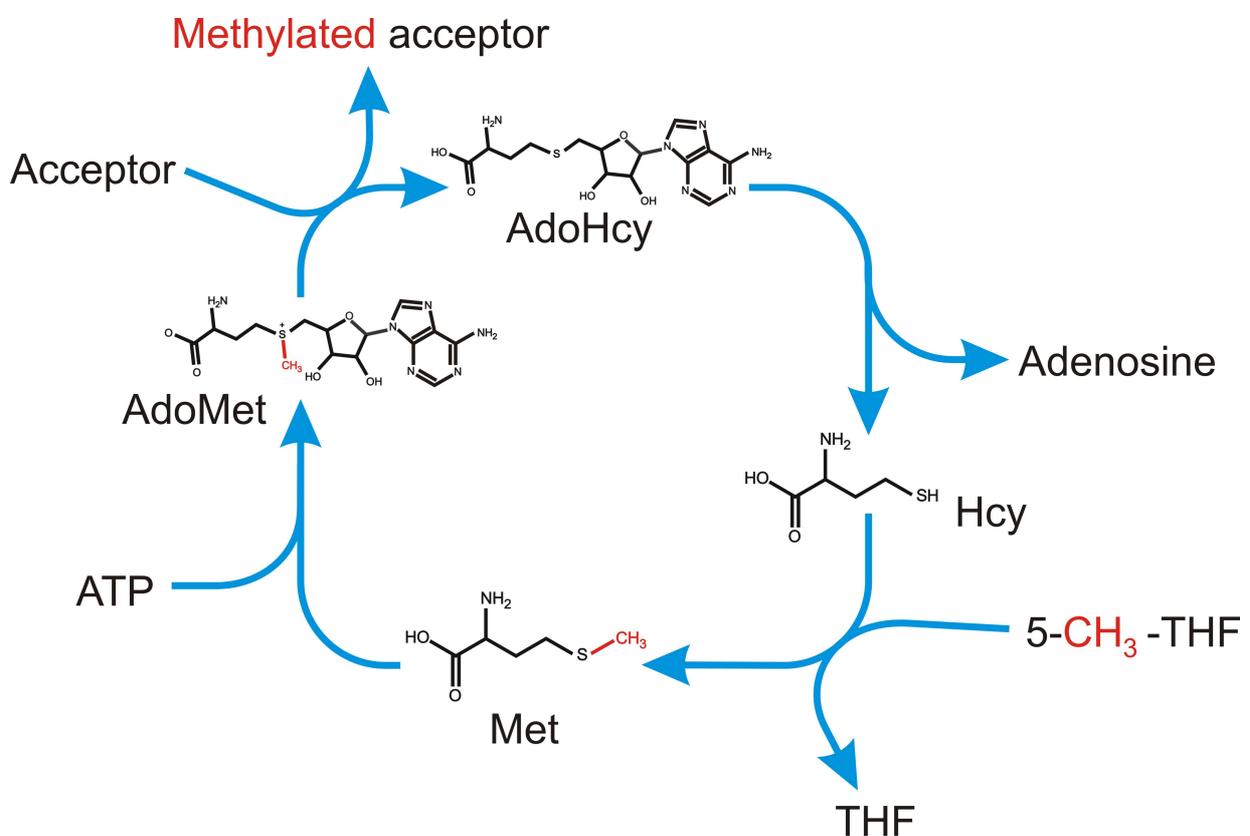


Figure 1. AdoMet regeneration cycle. The methyl group that is transferred during the regeneration of AdoMet is colored red. Compounds that take part in this cycle are Met methionine, ATP adenosine triphosphate, AdoMet S-adenosyl-L-methionine, AdoHcy S-adenosyl-L-homocysteine, Hcy L-homocysteine, Adenosine, 5-CH₃-THF N5-methyltetrahydrofolate, THF tetrahydrofolate.

3.4. The relevance of methylation in nature

Methylation is a chemical reaction widely spread in nature. Enzymatically catalyzed addition of a methyl group can be found in many biochemical pathways and processes. A few

examples may illustrate their relevance in DNA and protein modification as well as decoration of smaller molecules.

Methylation of DNA is usually performed on adenosine or cytosine nucleotides. DNA methylating enzymes use AdoMet as a methyl group donor and can be divided into different groups accordingly to the reaction they catalyze. In this way, there are adenine-n6-DNA-methyltransferase (EC 2.1.1.72) that generate N6-methyladenine, the cytosine-n4-DNA-methyltransferase (EC 2.1.1.113) generating N4-methylcytosine and cytosine 5-methyltransferase (EC 2.1.1.37) that produce C5-methylcytosine.

In prokaryotes, DNA methylation plays an important role in the restriction modification system, which is used for protection against foreign DNA (Gromova et al., 2003). A methylation pattern specific for a given organism is created. This specific methylation of DNA molecules protects it from digestion by endonucleases. Any foreign nucleic acids whose methylation pattern is different will be readily degraded (Noyer-Weidner et al., 1993).

In plants as well as in animals, methylation of DNA plays a role in the regulation of gene expression by modification of cytosine residues (Dodge et al., 2002; Haines et al., 2001; Pradhan et al., 2003). Cytosine methylation in mammals also plays a role in regulation of development. The mutation in DNA methyltransferase gene or in the proteins interacting with methylated DNA may lead to sicknesses such as ICF and Rett syndrome (Bestor, 2000). The hypomethylation of DNA can lead to chromosome instability and increased mutation rates (Jones et al., 2002).

Upon translation, proteins often undergo chemical modifications which include addition of chemical moieties, linkage to other protein molecules via disulphide bonds as well as chemical modification of amino acid residues (deamidation of glutamine and asparagines). Methylation of proteins is also one of their several post-translational modifications (Krause et al., 2007) and is usually carried out on lysine or arginine residues. Methylation is one of the modifications of histone proteins (Bauer et al., 2002; Strahl et al., 2000) that alter the chromatin structure and thus influence many of genomic processes, such as gene expression (Zhang et al., 2001).

Methylation reactions take part in a way that living organisms utilize CO₂ to synthesize complex chemical compounds. Under anaerobic conditions microorganisms utilize the reductive acetyl CoA pathway known as Wood-Ljungdahl pathway (Menon et al., 1999) to fix CO₂.

Methyltransferases are also involved in amino acid metabolism. Betaine-homocysteine methyltransferase is involved in the production of methionine and dimethylglycine from

homocysteine and betaine (Szegedi et al., 2008). There are only two known enzymes in mammals that are capable of methylating homocysteine (Garrow, 1996). It is speculated that the defective betaine-homocysteine methyltransferase may among other factors contribute to genetic disorders associated with faulty metabolism of homocysteine. Glycine N-methyltransferase is another example of a methyltransferase involved in amino acid metabolism (Garrow, 1996; Luka et al., 2002). This enzyme catalyzes the addition of a methyl group to a glycine. As for majority of methyltransferases this AdoMet is the methyl group donor. Glycine N-methyltransferase is important element in regulation of the amounts of methionine and AdoMet/AdoHcy ratios in the bodies of animals.

The biological activity of many chemical compounds is also controlled by methylation. Bioactive molecules such as hormones or signal carrying molecules (neurotransmitters) are inactivated by the addition of a methyl group. Methylation is also able to regulate the biological activity of exogenous compounds (Zhu et al., 1994). The transfer of a methyl group in this reaction is proposed to be the means of shuffling between the active CoI and inactive CoII state of cobamide-dependent methyl carrier protein CFeSP.

Enzymatic addition of a methyl group can be observed during the synthesis of many natural compounds e.g., creatine (Walker, 1960) or phosphatidylcholine (Pessi et al., 2004; Vance et al., 1988). The process of an addition of a methyl group takes also part in biosynthetic pathways leading to production of many chemical compounds which include important cofactor molecules. Among the most important ones are ubiquinone (Avelange-Macherel et al., 1998; Poon et al., 1999), antibiotics in microorganisms, and secondary metabolites found in plants, such as subunits of lignin monomers and compounds responsible for pigmentation and scent of flowers (Lavid et al., 2002). Methyltransferase activities in plants have also been identified to be connected with adaptive processes in plants such as cold acclimation (Ndong et al., 2003) and pathogen response (Zubieta et al., 2003).

3.4.1. O-Methyltransferases

As mentioned earlier the enzymes that transfer a methyl group to an oxygen atom are called *O*-methyltransferases (OMTs). This group of enzymes is present in all organisms. Modification by an addition of a methyl group to a oxygen atom is a common reaction in the secondary metabolism of plants and microbes. Enzymatic methylation catalyzed by OMTs can be carried out on hydroxyl and carboxyl group oxygen atoms.

3.4.2. Animal *O*-Methyltransferases

Regulatory functions performed by action of various compounds in animals may be controlled by methylation. The prominent example of such case is the action of mammalian catechol OMT (EC 2.1.1.6). This enzyme catalyzes the methylation of vicinal dihydroxy systems like those found in catechol estrogens, dihydroxyindolic intermediates of melanin, catecholamine neurotransmitters such as dopamine or norepinephrine and epinephrine. It was first described in 1958 by J. Axelrod (Axelrod et al., 1958), who later in 1970 was awarded the Nobel prize for his work on neurotransmitter metabolism. Methylation of neurotransmitters serves as a means of regulation of their biological activity. The catechol OMT from rat (rOMT) was the first Mg²⁺ dependent OMT from animals to be characterized structurally (Vidgren et al., 1994). Deficiency in the action of this enzyme may lead to many neurological diseases. The polymorphism in the gene encoding catechol OMT is implicated in schizophrenia (Herken et al., 2001). Together with its regulatory function this enzyme is also responsible for inactivation of foreign substances in the liver (Zhu et al., 1994).

Another example of an important OMT found in variety of organisms is L-isoaspartyl-*O*-Methyltransferase OMT. This OMT is responsible for the repair of proteins containing abnormal (L-isoaspartyl) residues which are results of spontaneous aging. These protein repairing enzymes are found in wide range of organisms including animals, plants as well as microbes (Kagan et al., 1997; Mudgett et al., 1997; Thapar et al., 2000). The study carried out on *Caenorhabditis elegans* larvae deficient in L-isoaspartyl OMT (Gomez et al., 2007) showed that the larvae lacking this OMT have lower survival and recovery rate after incubation in M9 medium without nutrients than wild type ones. The OMT deficient larvae were also observed to show decreased longevity. In tissue preparations from mice with downregulated levels of this enzyme the accumulation of the substrates for L-isoaspartyl OMT (protein molecules containing L-isoaspartyl residues) was found. Those animals also showed retarded growth and succumbed to fatal seizures (Kim et al., 1997).

3.4.3. Plant *O*-Methyltransferases

In plants, *O*-methylation is a common step in the biosynthetic pathways of many naturally occurring compounds. Methylation reactions are important steps in a wide network of enzymatic reactions referred to as phenylpropanoid metabolism (Hahlbrock et al., 1989). Among the many products of this set of metabolic pathways are flavonoids, alkaloids and polymers like lignin. OMTs that play a very important part in lignin biosynthesis. It is, after cellulose, the second most abundant biopolymer in nature. Its composition and amounts is

very important for the paper industry as well as farming. Attempts have been made to produce transgenic plants with reduced lignin content by downregulation of enzymes taking part in early steps of the biosynthetic pathway (Anterola et al., 2002). Lignin biosynthesis is thought to occur through a metabolic grid (Figure 2) (Humphreys et al., 2002) where, among other reactions, methylation of caffeic and 5-hydroxyferulic acids and their derivatives takes place. The most important OMT involved in this biochemical pathway is caffeic acid OMT (COMT), which is capable of methylation of caffeic and its corresponding aldehydes and alcohols. This enzyme which does not require a bivalent cation for activity and, in principle, is able to perform *in vitro* all methylation reactions during lignin biosynthesis (Edwards et al., 1991).

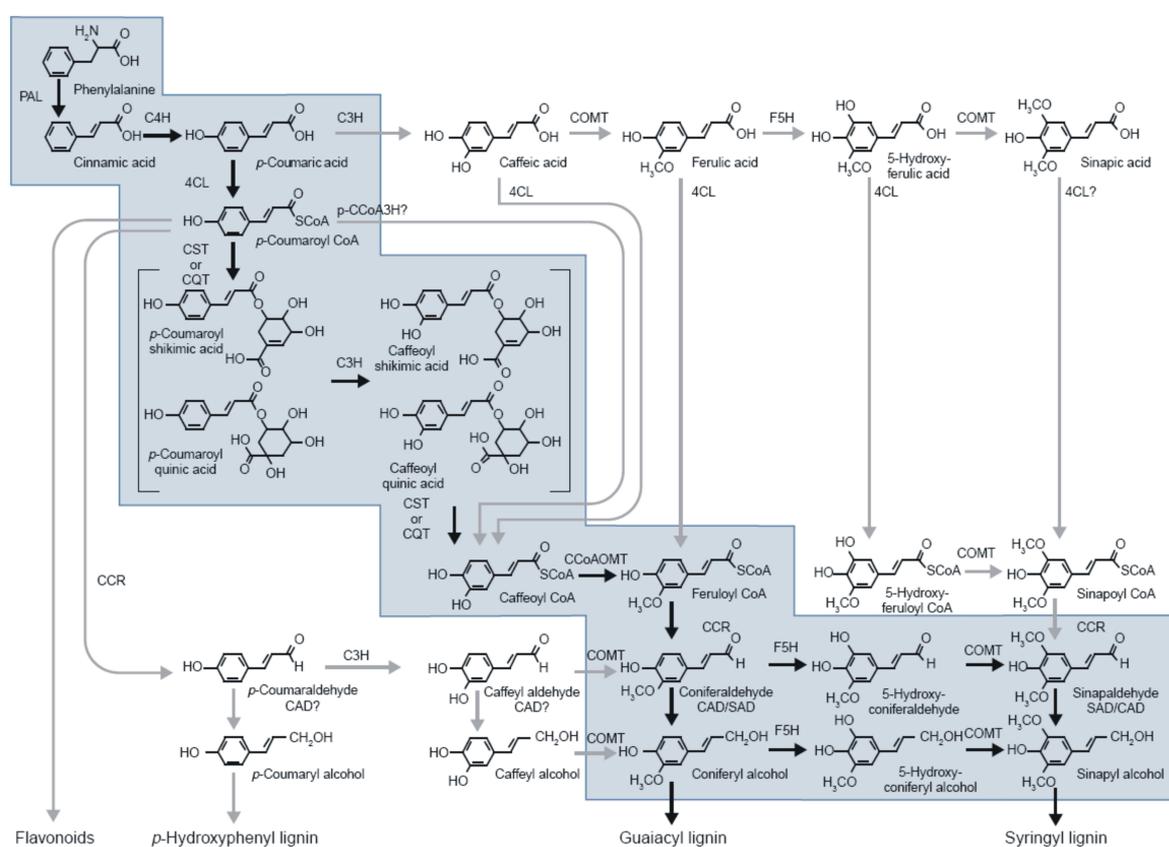


Figure 2. Schematic diagram showing the biochemical pathways leading to the production of lignin precursors (Humphreys et al., 2002). Reactions thought to be key in lignin biosynthesis are indicated with black arrows. The most prominent pathway leading to the production of lignin is shaded in blue. Enzyme abbreviations used: 4CL, 4-(hydroxy)cinnamoyl CoA ligase; C3H, *p*-coumarate 3-hydroxylase; C4H, cinnamate 4-hydroxylase; CAD, cinnamyl alcohol dehydrogenase; CCoAOMT, caffeoyl CoA *O*-methyltransferase; CCR, cinnamoyl CoA reductase; COMT, caffeic acid/5-hydroxyferulic acid *O*-methyltransferase; CQT, hydroxycinnamoyl CoA: quinate hydroxycinnamoyltransferase; CST, hydroxycinnamoyl CoA: shikimate hydroxycinnamoyltransferase; F5H, ferulate 5-hydroxylase; PAL, phenylalanine ammonia-lyase; *p*CCoA3H, *p*-coumaroyl CoA 3-hydroxylase; SAD, sinapyl alcohol dehydrogenase.

However, downregulation experiments of COMT showed that it is not the case *in vivo* (Guo et al., 2001). Several investigations point to an additional magnesium dependent CCoA *O*-methyltransferase (CCoAOMT) as an important enzyme involved in this pathway (Parvathi et al., 2001; Ye et al., 2001; Zhong et al., 1998; Zhong et al., 2000). This enzyme is at present thought to play a central role in this pathway.

In addition to lignin biosynthesis, *O*-methylation is also an important step in production of a variety of metabolites essential to plant growth and development. The addition of a methyl group to different chemical compounds renders them more volatile. An interesting example of a biosynthetic pathway where OMTs activity is crucial is the formation of volatile compounds in plant leaves and flowers. Methylations take place during the synthesis of 3,5-dimethoxytoluene in the rose flowers. This compound makes up almost 60% of the total volatiles in some rose varieties. Biosynthesis of 3,5-dimethoxytoluene involves two subsequent methylation reactions catalyzed by two distinct OMTs (Lavid et al., 2002). In this way orcinol is converted to its methylated derivative, 3,5-dimethoxytoluene. Both reactions are catalyzed by related enzymes with different substrate specificities. Those two enzymes share 96.5% identity at the amino acid level. It was determined that a single amino acid mutation is sufficient for the change of substrate specificity between those OMTs (Scalliet et al., 2008).

The function of OMTs in production of plant flavors and scents is not restricted to roses. The products of the phenylpropanoid pathway are also key flavoring elements of many herbs. An elegant example may be the reaction of eugenol OMT and chavicol OMT in sweet basil (Gang et al., 2002). The products of the methylation reaction, methyleugenol and methylchavicol are used as food flavorings and are responsible for spicy flavor and “clove-like” aroma of basil and also other herbs.

Flavonoids are another very important class of compounds in plants. These products of the phenylpropanoid metabolism besides their antioxidant activities act as antimicrobial defense compounds, allelopathic agents and regulators of plant growth and seed germination (Aliotta et al., 1993). There is a great abundance of methylated flavonoid derivatives in plants (Ibrahim et al., 1998). Methylation of the reactive hydroxyl groups is thought to be responsible for reduction of the toxicity of those compounds. Enzymes that catalyze methylation of hydroxyl groups of flavonoids have been known for over two decades (De Luca et al., 1985). Methylation of flavonoids modifies their solubility and therefore, controls their intercellular compartmentalization.

The action of flavonoids is usually accredited to their antioxidant activity. The production of flavonoids is often a part of a general stress response. Flavonoids accumulate under stress conditions that include pathogen attack, low temperature and high UV radiation (Gregersen et al., 1994; Winkel-Shirley, 2002). The UV-B absorbing qualities of those compounds are thought to be responsible for accumulation in response to elevated UV-B radiation levels. Plants protect themselves by accumulation of flavonoids together with their glycosylated and methylated conjugates in the epidermal tissues. Knock out of chalcone synthase and chalcone isomerase, enzymes that play a crucial role in the biosynthesis of flavonoids in *Arabidopsis*, produce UV-hypersensitive phenotypes (Li et al., 1993). Similarly a UV-B resistant *Arabidopsis* mutant found to accumulate high levels of flavonoids and other phenolics, was known to exhibit an upregulation of the chalcone synthase gene (Bieza et al., 2001). There is an indisputable involvement of flavonoids in the UV protection of plants, however it is not known whether a specific structure of flavonoid is responsible for the protective effects or an entire array of compounds is involved.

OMTs in plants can be divided into two major classes, the higher molecular weight enzymes that do not require Mg^{2+} for enzymatic activity and low molecular weight Mg^{2+} dependent ones. As described above the cation independent enzymes are known to have many functions. They methylate flavonoids, coumaric and caffeic acid, as well as alkaloids. The COMT known to take part in the lignin biosynthesis is a very important member of this group. On the other hand there is a class of low molecular weight Mg^{2+} dependent OMTs. Cation dependent OMTs methylate substrates sharing a phenolic moiety with vicinal dihydroxy system, like those found in quercetin or caffeic acid. This group includes substrate specific CCoAOMTs, responsible for methylation of CoA esters during lignin biosynthesis. Among Mg^{2+} dependent OMTs a promiscuous OMT from *Mesembryanthemum crystallinum*, can be found (Ibdah et al., 2003). This plant commonly known as the ice plant, responds by accumulation of betacyanins and elevated levels of methylated and glycosylated flavonol conjugates to elevated light radiation. This CCoAOMT-like protein which was purified from UV irradiated leaves of this plant was found to exhibit a broad substrate specificity, accepting caffeic acid, caffeoyl CoA, flavonols and their glucosylated derivatives. Such broad spectrum of accepted substrates was found to be a novelty among cation dependent OMTs in plants.

3.4.4. Microbial *O*-methyltransferases

In various bacteria and fungi, the *O*-methylation is involved in antibiotic (Bauer et al., 1988) and aflatoxin (Keller et al., 1993; Yabe et al., 1989) synthesis. Other secondary metabolites

such as isobutyraldoxime methyl ether, may also be substrates for enzymatic methylation (Harper et al., 1985). Catechol OMT-like protein was also found in the fungus *Streptomyces griseus* (Dhar et al., 2000). This enzyme belongs to the Mg^{2+} -dependent group and it accepts various catechol substrates as well as coumarins and caffeic acid. Mg^{2+} -dependent OMTs are also involved in the biosynthesis of tylosin, a macrolide antibiotic (Kreuzman et al., 1988). The recent structure of a microbial OMT (Hou et al., 2007) shows high similarity to the plant Mg^{2+} -dependent OMTs.

3.5. The three dimensional structure of proteins as a tool for studying protein functions

3.5.1. General structure elucidation techniques

There are many of techniques that can be used for structure elucidation of chemical compounds. One of the techniques employed is Nuclear Magnetic Resonance (NMR). This technique is based on the magnetic properties and quantum mechanics of atomic nuclei. The nuclei of atoms that have an odd number of protons or neutrons possess an intrinsic property of spin (Morrison et al., 1992). For those atoms, when placed in an external magnetic field, the resonance absorption will occur at the frequency, which matches the energy requirement for the nucleus to change between the two spin states. This absorption is recorded during the NMR experiment. The neighboring atoms influence the frequency at which resonance absorption takes place. By knowing the influence of the atomic neighborhood of the absorbing atom it is possible to obtain the structure of the chemical compound. This technique was used to determine the structure of small molecules. Later when more sophisticated equipment was developed and methods of obtaining isotopically labeled proteins it was employed to determine the three dimensional structures of proteins (Wuthrich, 1990). The present limitation of this technique is the size of the investigated protein molecule.

Cryo-Electron microscopy is another technique which allows the determination of three dimensional structures of proteins and their complexes (Frank, 2002). The basis for this technique is imaging of biological samples by electron microscopy under the temperature of liquid nitrogen. The low temperature ensures the adequate stability of the investigated sample. The molecule of interest is present in the sample in multiple copies which assume different spatial orientations. The recorded images of the protein molecules show them in different orientations. Such state can be described as series of rigid body movement of a single object. The collection of images can then be assembled to yield a complete three dimensional picture of molecule's surface. This simplified explanation does not mention all the complicated

mathematical procedures and difficulties involved in use of cryo-electron microscopy to elucidate the structures of single biomolecules. Cryo-electron microscopy is usually employed to study large complexes of proteins which would be too difficult to investigate using different methods. This technique is often combined with the other structure elucidation methods (Topf et al., 2005). The models of proteins obtained by other methods may be assembled together to form a structure that was determined with electron microscopy.

Homology modeling, also referred to as comparative modeling, should rather be called structure prediction rather than elucidation technique. This method is based on already known structures and assumes that proteins of similar amino acid sequences will also have similar fold (Chung et al., 1996). The amino acid sequence of the target protein is compared with the sequences of related protein of known structure. The similarities detected at the amino acid level allows the assumption of similarity in 3D structure (Marti-Renom et al., 2000). The amino acid sequence of interest is compared with many sequences present in databases. Out of those a template for the protein fold is selected. This step includes preparation of an alignment between the template and the sequence of interest. It is advantageous to use multiple templates when assigning the fold of the protein. In such case the structural alignment of the templates is carried out first. Once the template is assigned to the target sequence the building of the model can start. The positions of side chains are predicted using the information from related structures as well as the geometrical and energetic restraints.

When the amino acid sequence of the protein of interest has a very low similarity to sequences of known structures the modeling methods based on sequence comparison may fail to predict the correct fold of the target sequence. In such case fold recognition or threading may be employed. This method is based on a theory which states that there is a limited number of protein folds. In order to carry out threading a library of unique representative structures is searched for structure analogs to the target sequence.

Modeling methods also include *ab initio*- or *de novo*- protein modeling (Bonneau et al., 2001; Bonneau et al., 2002). This method bases on energy function rather than similarity to known structures to obtain a three dimensional model of the protein. The computational approaches try to simulate the process of protein folding with the use of molecular dynamics simulation. A physically reasonable energy function is optimized during this process. Those methods are very demanding in terms of computing power and are now used to predict the structures of small proteins.

3.5.2. X-ray crystallography

Among the available structure determination methods, crystallography is the most widely used for biological macromolecules. The structure determination workflow begins with crystallization of the purified proteins. Crystallization can be seen as a special case of precipitation. Crystals, whether protein, small organic compounds or inorganic salts are a form of precipitate with highly ordered structure. In order for the crystals to grow a solution of protein is mixed with a solution of precipitating agent at a concentration just below the one sufficient for precipitation of protein. A precipitating agent is a chemical substance which decreases the solubility of the protein such as ammonium sulfate or polyethylene glycol. Subsequently the conditions are slowly changed to bring the concentration of protein into a super saturated stage at which the crystal growth may occur. The most widely used crystallization method for proteins is vapor diffusion (Figure 3) . Drops of protein and precipitant solution are mixed together and placed above a large reservoir of precipitant solution in a sealed vessel. The diffusion of water vapor between the drop and the reservoir is the means of approaching the supersaturation conditions. The driving force of this process is the difference between the precipitant concentration in the protein drop and in the reservoir solution. Other approaches to achieve supersaturated solutions of proteins include dialysis and free interface diffusion. The protein – precipitant solution is dialyzed against the solution of precipitant of higher concentration. During dialysis the membrane plays a similar role as diffusion of water vapor. The free interface diffusion method is based on the diffusion of precipitant into the protein solution on an interface between two solutions with high differences of viscosity. In this manner, a gradient of precipitant is formed and suitable conditions for crystal growth are achieved. This setup is usually performed in capillaries.

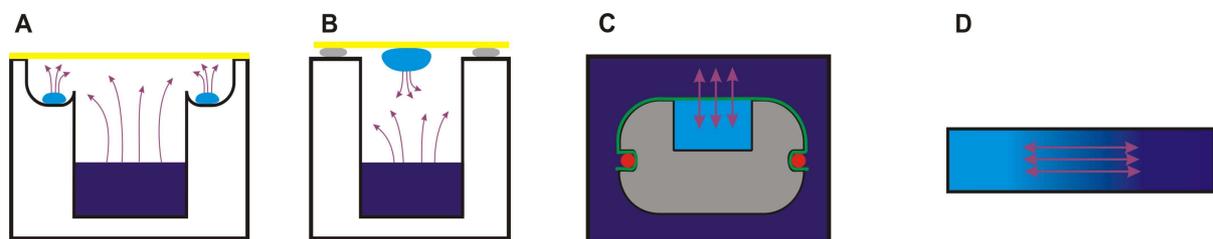


Figure 3. Commonly used setups for protein crystallization. A – sitting drop vapor diffusion, B – hanging drop vapor diffusion, C – microdialysis, D – free interface diffusion. The crystallization solution is colored violet while the protein solution is blue. The purple arrows represent the diffusion of water during the crystallization.

The second step in the process of structure determination is collection of the diffraction data. The X-ray diffraction experiment is performed and the diffraction images are recorded. The crystal is usually mounted in the stream of cooled nitrogen at 100K. Freezing of the crystal protects it from radiation damage and allows for longer data collection times. Once the diffraction experiment is performed the diffraction images are merged and scaled.

The major problem encountered when attempting to determine the structure of a protein by crystallography is the inability to measure the phases of the diffracted beams. In order to calculate the electron density, into which an atomic model of the protein structure is built two parameters are needed: the amplitude of the diffracted beam and its phase. During the diffraction experiment only intensities are measured. The missing phase information has to be obtained from other sources. There are two major approaches to solve the “phase problem” the direct methods and methods basing on a Patterson function. The firstly mentioned direct methods employ statistical relationships to directly solve the phase problem by the use of phase relationships based on the observed intensities. The assumption for this method is that the crystal is made of similarly shaped atoms which have positive electron density around them. Hence there is a statistical relationship between the sets of structure factors. Latter methods, as mentioned, employ the Patterson function as a basis for solving the phase problem. The Patterson function is a Fourier transform of structure factors calculated only with intensities. The intensities, which are the squared amplitudes, can be measured experimentally. The result is so called Patterson map, which is a map of vectors between atoms.

There are three major methods that employ the Patterson function, heavy atom methods, MAD (Multiple wavelength Anomalous Diffraction) and molecular replacement. Heavy atom methods are based on the presence of strongly diffracting atoms in the protein crystals. The position of those atoms can be identified and serve as a reference for the rest of the protein structure. There are two ways of introducing strongly diffracting atoms, such as Hg, Pt or Au, into the crystals of proteins. The first one is soaking of the crystals in the solutions of ions of heavy atoms. During soaking the ions or ionic complexes of heavy atoms enter the solvent channels in the crystal and interact with the protein. An example of such method can be Isomorphous replacement. The other way is co-crystallization of proteins with compounds containing heavy atoms. If the protein is known to bind any ions or ligands it is advantageous to include them in the crystallization solution. Such additives not only stabilize protein but also may help in structure determination. Prosthetic groups binding Fe or Zn may be a prominent example.

MAD is a method of solving the phase problem that also employs use of heavy atoms. The principle underlying this method is the fact that atoms present in the structure of the investigated protein can absorb X-rays of a certain very specific wavelength (Walsh et al., 1999). This phenomena happens when the wavelength of the X-ray photon is close to that is close to that of an electronic transition in a bound atomic orbital. electronic transition. It perturbs the amplitude and phase of the diffracted X-rays and gives rise to anomalous scattering. The effect of anomalous scattering for small atoms is negligible and can not be easily measured. However, for the heavy atoms anomalous scattering can be recorded and measured. For non-centrosymmetric structures, the anomalous scattering causes the differences in the intensities of the symmetry related reflections. Friedel's law, $[I(hkl) = I(-h-k-l)]$ no longer holds and the resultant intensity differences can be used to obtain phase information. One of the most commonly used methods for obtaining a heavy atom protein derivative suitable for MAD is production of selenomethionine crystals where the protein contains selenium atom instead of sulfur in the methionine residues (Doublié, 1997).

If a molecular model of a similar protein is available it may be used to provide the initial estimates for phases for a new protein. Such technique of obtaining phase information is called molecular replacement. During this procedure the known protein, called phasing model, is placed inside the unit cell of a new protein. The experimentally measured intensities and the phases from the model are used to calculate the electron density. For the phasing model a protein should share as much structural (sequence) homology as possible. It is considered that if the model is fairly complete and shares at least 40% sequence identity with the unknown structure, the molecular replacement will be fairly straightforward. It becomes progressively more difficult as the model becomes less complete or shares less sequence identity.

This approach is especially useful if the object of interest is the conformational changes when a known protein binds to small ligands. This understanding is based on the assumption that introduction of the ligand to the protein produces a complex that is similar in structure to the free protein. The free molecule is used as a phasing model for the protein-ligand complex. In such case, where the similarity of the phasing model is very high, Fourier methods will suffice.

The major drawback of crystallography as a method of structure determination is the requirement of crystals. The crystallization of the proteins is a major bottle neck of the entire process. The common obstacles encountered are inability of finding proper crystallization conditions or the inability to produce crystals of sufficient quality. Growing of crystals of

membrane proteins presents a big challenge. For those, quite insoluble proteins, it is necessary to reconstitute membrane-like surroundings with the help of additives, like detergents. Soaking of the crystals may be detrimental to their structure. Radiation damage is the next factor one must consider when collecting diffraction images. The X-rays illuminating the crystal also damage the protein out of which it is composed. Data collection at low temperatures and ability to control the intensity of the X-ray beam are tools that help to overcome this problem.

When interpreting crystal structures one must keep in mind that what one sees is a snapshot of the reality. Crystal structure captures just one state of a sometimes very dynamic object. The crystal environment may influence the conformation of the protein and the observed situation may deviate from the actual state. Despite many limitations and pitfalls, crystallography remains the method of choice for structural studies on proteins.

3.6. Structural characterization of Mg²⁺ dependent OMTs

A number of OMTs taking part in the phenylpropanoid metabolism have been structurally characterized. (Ferrer et al., 2005; Ferrer et al., 2008; Zubieta et al., 2002). While the structures of cation dependent and independent OMTs show distinct differences, the structures enzymes within each group are very similar (Ferrer et al., 2005; Zubieta et al., 2001). This structural similarity extends to catechol OMT from rat (rOMT) (Vidgren et al., 1994), which can be perceived as an animal counterpart of plant magnesium dependent OMTs. From the group of plant Mg²⁺dependent *O*-methyltransferases only two have been characterized structurally (Ferrer et al., 2005; Kopycki et al., 2008a). Both of the proteins have core α/β Rossmann fold which provides a scaffold for binding the AdoMet cofactor (Figure 4) (Martin et al., 2002; Schluckebier et al., 1995). It consists of a β -sheet structure surrounded by eight α -helices. This structural motive is highly conserved among all methyltransferases that use AdoMet as a methyl group donor. (Figure 5)

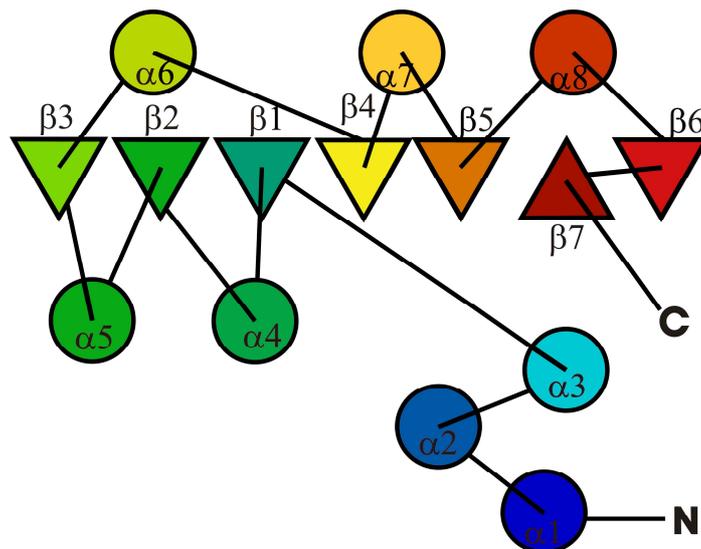


Figure 4. The general topology of Mg^{2+} -dependent OMTs according to (Martin et al., 2002; Schluckebier et al., 1995). The centrally placed β -sheet is surrounded by eight α -helices. This structure constitutes the AdoMet binding domain, common for all AdoMet dependent methyltransferases.

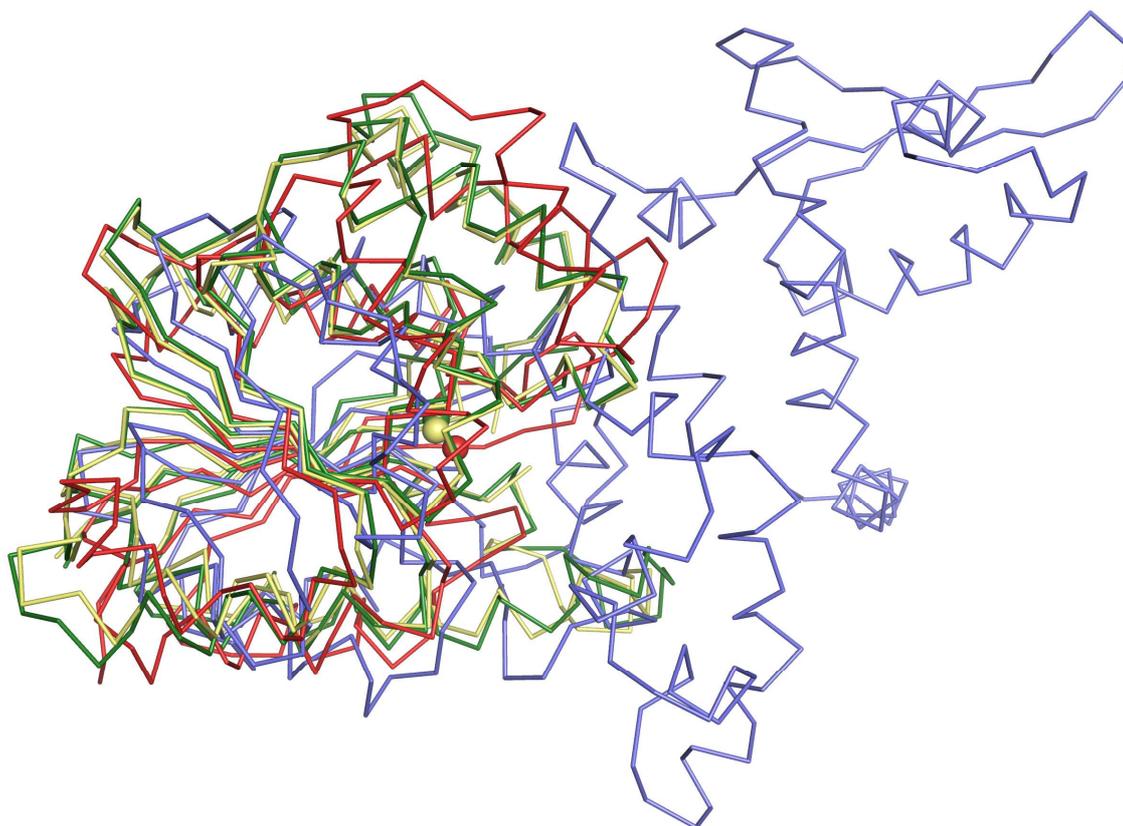


Figure 5. The structural comparison of OMTs involved in phenylpropanoid metabolism. The alignment of the structures of plant OMTs. Blue COMT (PDB code 1KYZ) (Zubieta et al., 2002), green CCoAOMT (PDB code 1SUI) (Ferrer et al., 2005), pale yellow PFOMT (PDB code 3C3Y) (Kopycki et al., 2008a) (only monomers are shown) and red for comparison rOMT (PDB code 1VID) (Vidgren et al., 1994). The colored spheres are the bivalent cations found in the structures of the corresponding proteins. Please note the high degree of similarity for the AdoMet binding domain for all of the structures.

The mode coordination of the metal ion which is the prerequisite for the enzymatic activity is highly conserved. The general scaffold of the active site and the catalytic machinery is virtually the same for all Mg^{2+} dependent OMTs. Despite the similarities, however, there are few structural differences between plant and animal Mg^{2+} dependent OMTs (Figure 6). In rOMT the first two helices show a distinctly different topological arrangement. The position of helix $\alpha 1$ in the plant enzymes is in part accomplished by a deletion between strands $\beta 6$ and $\beta 7$ with respect to the mammalian enzyme and contributes significantly to dimer formation. This could be to one reason that the plant enzymes form dimers while the animal ones do not. The second notable difference is the presence of the extended Loop region in the plant enzymes that can support the substrates binding to the active site of the enzyme.

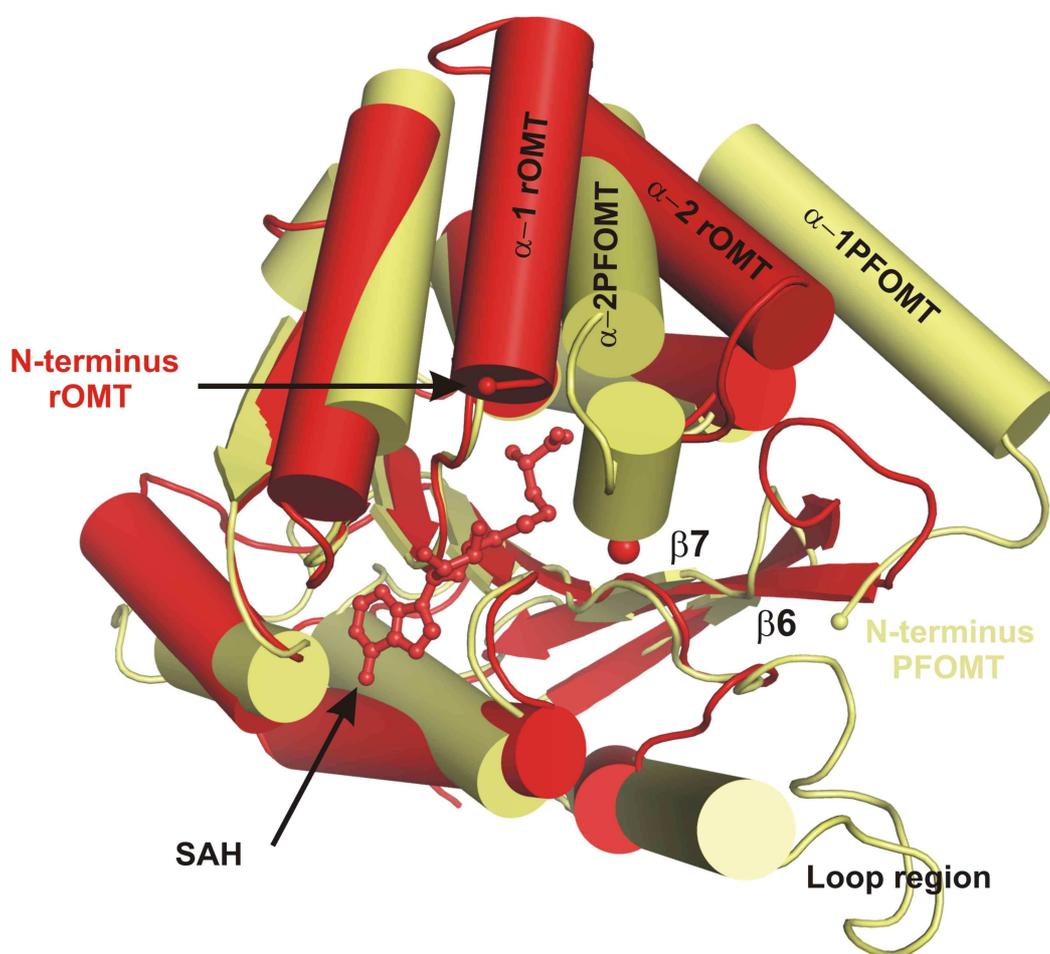


Figure 6. The representation of structural differences between plant and animal Mg^{2+} dependent OMTs. Pale yellow PFOMT (PDB code 3C3Y) (Kopycki et al., 2008a) (only monomer is shown) and red rOMT (PDB code 1VID) (Vidgren et al., 1994). The spheres represent the ion bound to the structure. The red ball and stick structure is AdoHcy bound to the structure of rOMT. The different positioning of the fist two N-terminal helices for both proteins can be seen.

Plant OMTs are functional dimers. The active sites of cation dependent enzymes can act independently from one another (Figure 7). This is in contrast to the cation independent enzymes, where the active site is composed of the residues belonging to the both subunits forming a dimer (Figure 8). On the other hand the high flexibility of the N-terminus makes it plausible that the N-termini belonging to the different monomers communicate with the active site of corresponding molecule.

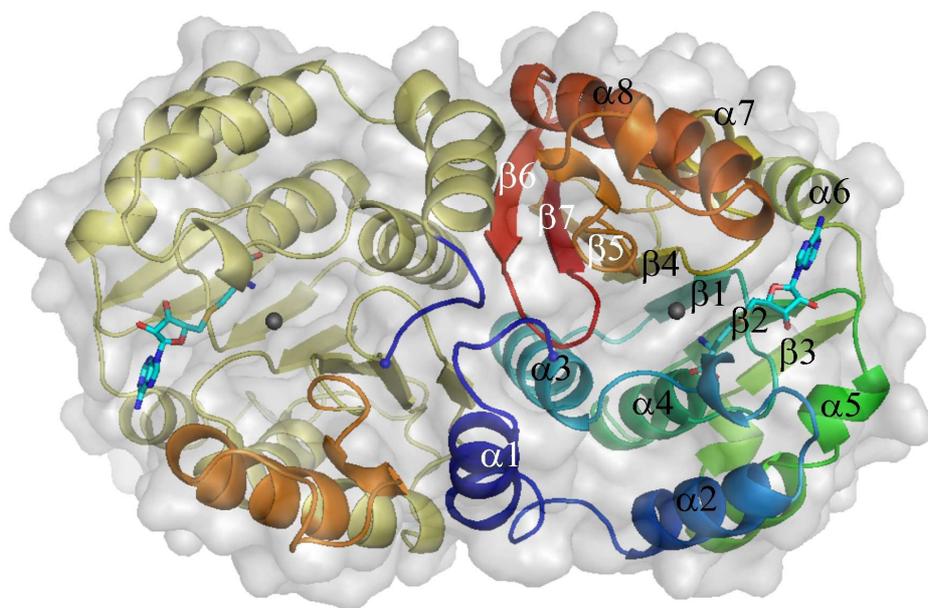


Figure 7. Dimer representation of PFOMT (PDB code 3C3Y). The gray spheres represent the Mg²⁺ binding site and bound AdoMet is shown as cyan sticks (Kopycki et al., 2008a).

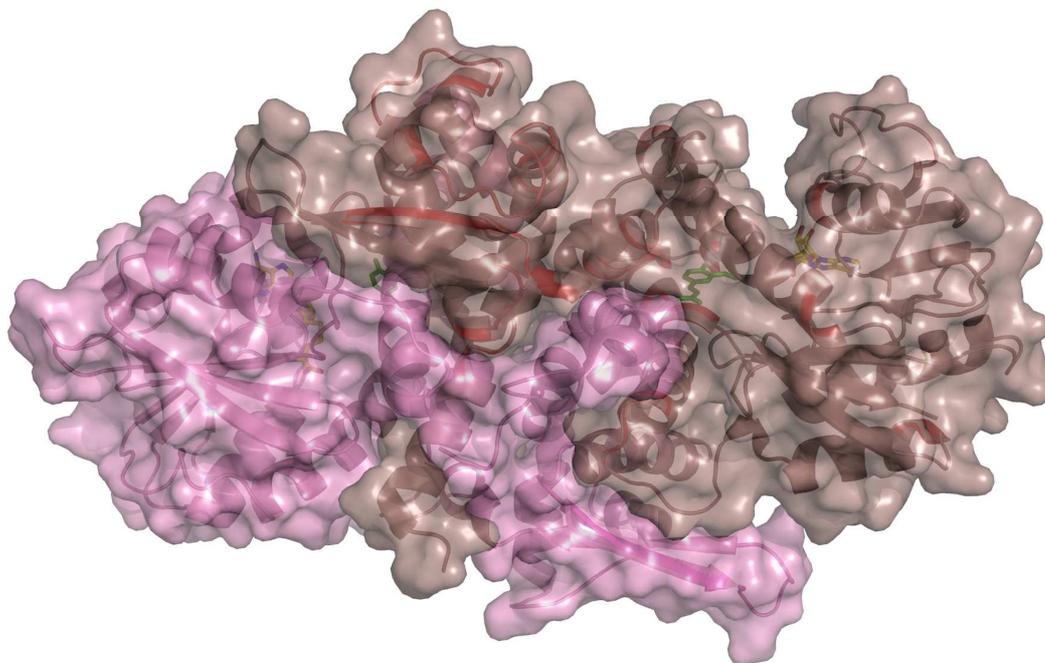


Figure 8. The dimerization of COMT, a cation independent plant OMT. (PDB code 1KYZ) The monomers are colored purple and red. The Yellow sticks represent the AdoHcy while green ones ferulic acid.

3.7. Mechanism of the methylation reaction catalyzed by Mg^{2+} dependant OMTs

The reaction mechanism found in magnesium dependent *O*-methyltransferases is based on the general acid-base reaction. This catalytic mechanism was first described in 1994 (Vidgren et al., 1994) for rOMT from rat. The transfer of the methyl itself proceeds via an S_N2 (Substitution Nucleophilic Bimolecular) like transition state (Woodard et al., 1980). The reaction involves replacement of the leaving group with a nucleophilic one. The lone electron pair from a nucleophile attacks an electron deficient electrophile. In the case of the enzymatic reaction the nucleophile is the hydroxyl group of the substrate and the positively charged sulfur atom on the AdoMet molecule is the electron deficient electrophilic center (Figure 9).

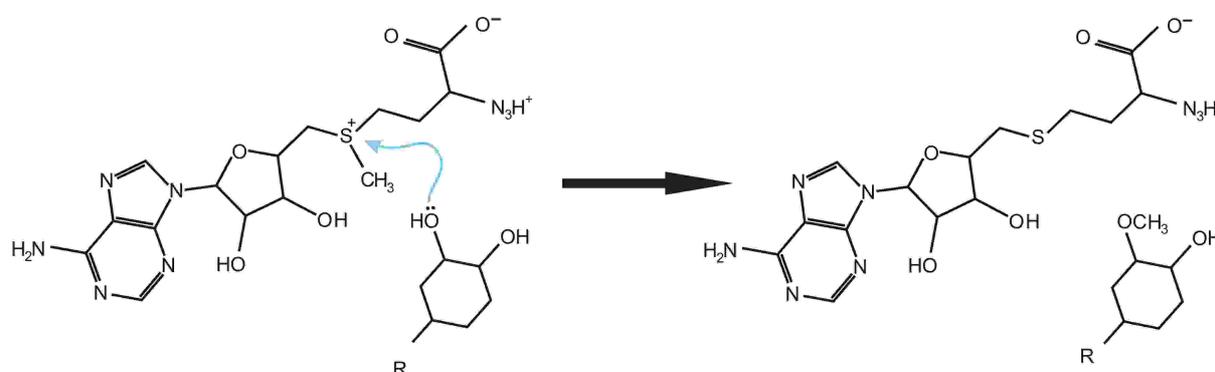


Figure 9. Mechanism of the reaction catalyzed by Mg^{2+} dependent OMTs. The blue arrow represents the nucleophilic attack

The presence of a bivalent cation is the prerequisite for the methylation. The binding of Mg^{2+} is accomplished by interactions of residues Asp141, Asp169, Asn170, (Figure 10) it improves the ionization of the two hydroxyl groups of catechol. Lys144 accepts the proton of one of the hydroxyl groups, it acts as a catalytic base for the nucleophilic methyl transfer reaction. The hydrophobic residues Trp38, Trp143 and Pro174 are thought to form a hydrophobic environment that facilitates the binding of more lipophilic substrates in the active site. These residues may also determine the positioning of the substrate in the active site and thus determine the *meta/para* ratio of the methylation with different substrates.

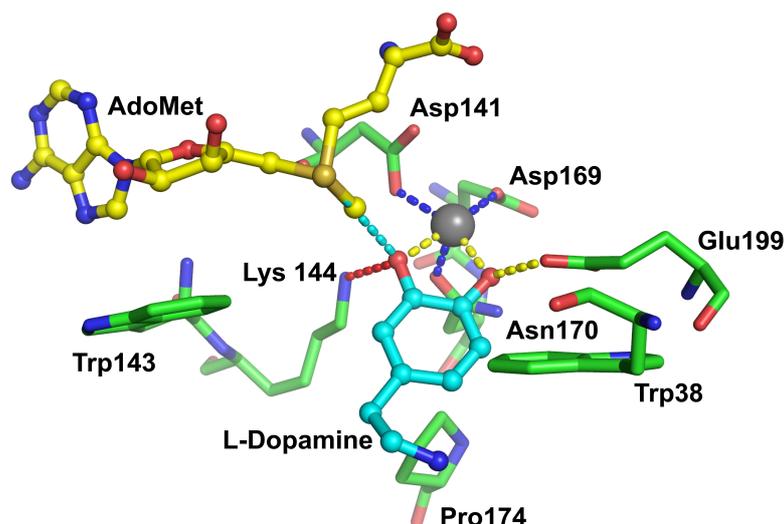


Figure 10. Three dimensional representation of the catalytic center of rOMT from rat (Vidgren et al., 1994). The blue dashed lines represent the coordination of the magnesium ion (gray sphere) by the protein, the yellow ones stand for the interactions with the substrate. The red dashed line represents the interaction of the catalytic lysine residue while the cyan points to the carbon atom of the methyl group which is transferred. The inhibitor dinitro-catechol present in the structure was replaced with the molecule of L-Dopamine, which is to simulate the real case reaction scenario.

The AdoMet cofactor was determined to bind first to the enzyme, then a magnesium ion binds to the complex and finally the methyl group acceptor molecule binds as the last step. The methylation is achieved by a direct nucleophilic attack by one of the hydroxyl groups of the substrate at the methyl carbon of AdoMet (Zhu, 2002). A similar mechanism of reaction and cation binding has been postulated in the case of CCoAOMT (Ferrer et al., 2005).

3.8. Rationale of the thesis

3.8.1. Substrate specificity of Mg^{2+} dependent *O*-methyltransferases

Among plant Mg^{2+} dependent OMTs two major groups of enzymes can be distinguished: the substrate specific ones that almost exclusively methylate the CoA thioesters of caffeic acid (the CCoAOMTs) and the promiscuous ones that methylate caffeic acid itself as well as its derivatives with similar efficiency. The amino acid sequences of the members of those two groups share a large degree of similarity (Figure 11).

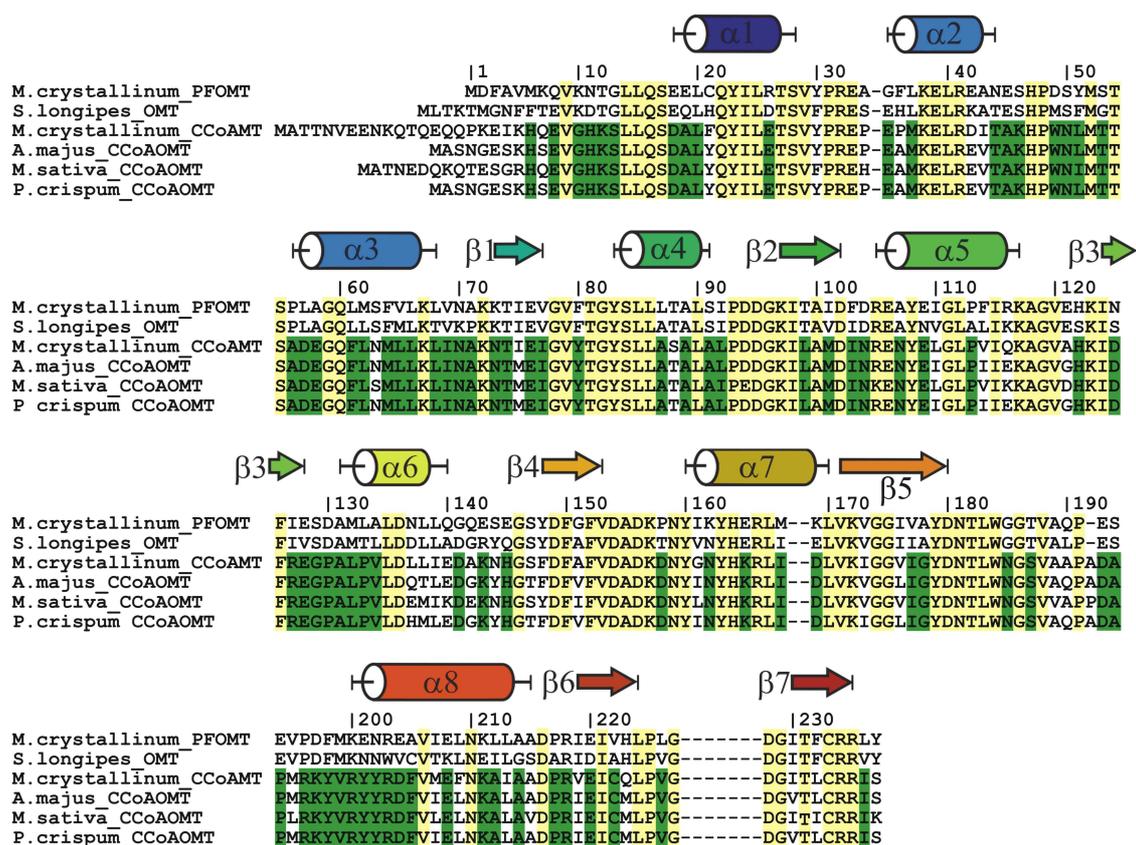


Figure 11. The amino acid sequence alignment of substrate specific CCoAOMTs and promiscuous OMTs (*M. crystallinum* PFOMT and *S. longipes* OMT). The residues shaded in yellow are common for all of the sequences, while the ones shaded green are conserved only among CCoAOMTs. The secondary structure elements are shown for PFOMT.

The high degree of similarity between those two groups suggests the existence of a very subtle mechanism, which governs the substrate specificity. Small changes in otherwise conserved regions may be sufficient to accomplish this shift in substrate specificity. The structure based amino acid sequence alignment of Mg²⁺ dependent OMTs from different plants shows just how similar those proteins really are. Despite the high degree of sequence identity the substrate specificity for those OMTs is quite different. The observed phenomena concerning these two groups of OMTs give rise to many questions. What is the mechanism that determines the substrate specificity? What structural features determine the substrate specificity? It seems plausible that for both groups of plant OMTs exists a mechanism which involves a change in small regions of proteins. This may be sufficient to determine the substrate specificity and position of methylation. If this is the case, which regions of the protein are responsible for this phenomenon? May such small variations be a result of different pathways through which both groups of enzymes evolved? If those mechanisms are

elucidated it may be possible to predict and engineer the specificity of OMTs by exchange of individual amino acid residues or larger fragments of the protein structures.

A comparison of an enzymatic specialist *M. sativa* CCoAOMT and a promiscuous enzyme, PFOMT enables the detailed analysis of the mechanisms governing the substrate specificity. A model system consisting of those two enzymes has many advantages. Both proteins are well studied (Ferrer et al., 2005; Ibdah et al., 2003) and no major problems were encountered during their expression and purification. Moreover, the availability of the structures of both OMTs makes it possible to attempt a rational redesign the substrate specificity.

3.8.2. Position specificity of Mg²⁺ dependent *O*-methyltransferases

Plant Mg²⁺ dependent OMTs carry out methylation of the vicinal dihydroxyl systems of their substrates exclusively in the *meta* position of the phenyl ring. Substrate specific CCoAOMTs as well as promiscuous OMTs methylate their substrates exclusively in this position (Figure 12). On the other hand, for cation independent OMTs the methylation in the *para* position is not uncommon (De Luca et al., 1985; Kim et al., 2005). In case of animal enzymes the position of hydroxyl group which is methylated may vary. The methylation in position *para* of the phenyl ring is observed for rOMT (Zhu et al., 1994). This positional promiscuity is observed in varying degree between soluble and membrane bound isoforms of this enzyme. Methylation is less regiospecific with OMT in prokaryotes.

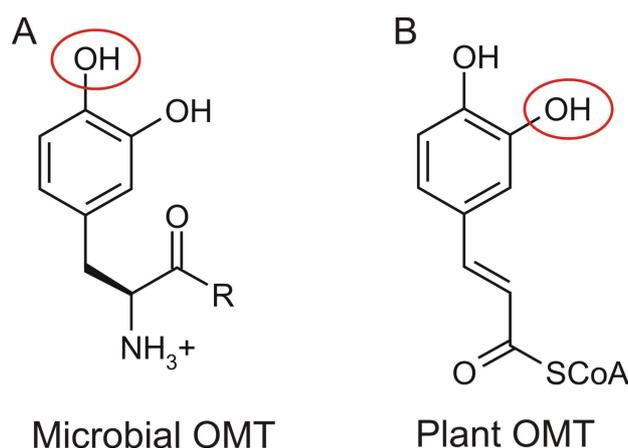


Figure 12. Differences in positional specificity between plant and microbial OMTs. The methylation carried out by characterized microbial enzymes A, is preferentially directed to *para* position (marked red) of the phenyl ring. In the case of plant OMTs B, the methylation is restricted to *meta* position (also marked red)

The recent research indicates an existence of catechol OMT like protein of fungal origin, thought to take part in the synthesis of saframycin in *Myxococcus xanthus* (Nelson et al., 2007). This OMT methylates its substrates (L-dihydroxyphenylalanine and other catechols, such as caffeic acid) preferentially in position *para* of the phenyl ring (Figure 12). The high stringency in the position to which a methyl group is added is observed only for plant enzymes, regardless of the preferred substrate. OMTs coming from microbes or animals seem to show more relaxed behavior. This raises a question: What are the key factors influencing the positional specificity in cation dependent OMTs? The availability of the cyanobacterial cation dependent OMT provides a unique opportunity to study this phenomenon, since the preliminary tests have indicated that this enzyme may have the *para*-methylation activity towards some phenylpropanoids. A comparison of the three dimensional structures of the plant cation dependent OMT and microbial one may provide information needed to address the question stated earlier.

4. Materials and methods

4.1. Molecular biological procedures

4.1.1. Preparation of gene inserts

The inserts coding for the proteins of interest were prepared by PCR reaction carried out according to the protocols provided by the manufacturers of the DNA polymerases used. The proofreading enzymes Pfu DNA polymerase (Promega, Mannheim, Germany) or Isis Proofreading DNA Polymerase (Q-Biogene, Heidelberg, Germany) were used. The cDNA coding for the PFOMT was already available and prepared as described in (Ibdah et al., 2003). The cDNA coding for the SynOMT was also available, cloned into the expression vector as described (Kopycki et al., 2008b). The restriction sites for the ligation to the vectors were introduced by proper primer sequences (Table 1). The BamHI restriction site was introduced at the 5' end of the sequence with the forward primer and HindIII at the 3' end with the reverse primer. The additional stop codons were added to the reverse sequence to ensure termination of transcription.

4.1.2. Preparation of gene fragments for hybrid proteins

The DNA coding for the PFOMT – CCoAOMT hybrid protein containing the loop fragment from *M. sativa* (Ferrer et al., 2005) was initially prepared from two pieces. The NheI restriction site contained in the sequence of original PFOMT was critical in this procedure. The N-terminal piece ending with the NheI site was amplified from original full length PFOMT gene with the use proper primers 9, 10, (Table 1). The DNA sequence encoding the replaced loop was chemically synthesized (Geneart, Regensburg, Germany). The sequence synthesized by Geneart was provided already cloned into pPCR-SCRIPT vector. The fragment of interest was amplified with the use of primers 11, 12 (Table 1). The NheI site was present in the sequence, the reverse primer introduced the stop codons and the HindIII restriction site. Both amplified fragments were then digested with NheI. The restriction digestion reaction products were purified by Agarose gel electrophoresis. Fragments were ligated using the Quick Ligation Kit (Promega, Mannheim, Germany). The ligation mixture served as a template for subsequent amplification with primers specific for PFOMT (Table 1 primers 5 and 6). The products of the PCR reaction that were of the appropriate size were then digested with BamHI and HindIII in double digestion reaction and ligated to the expression vector.

The N-terminal hybrid was obtained by DNA amplification with a long primer (Table 1). A proofreading DNA polymerase mix was used for this reaction, Platinum Supermix HiFi (Invitrogen, Karlsruhe, Germany). The 92bp primer introduced the desired N-terminal sequence into the PFOMT gene. The sequence of the BamHI site preceding the sequence of the protein was present in the primer. The reaction products were purified by gel electrophoresis.

4.1.3. Ligation

The DNA sequences were ligated into pQE vectors for expression. Vectors with N terminal his tag pQE30, pQE9 and those providing C terminal his tag pQE60 and pQE70 (Qiagen, Hilden, Germany) were used for expression of the proteins. This vector system is based on the T5 promoter transcription-translation system. Vectors of the pQE family feature ampicillin resistance. The expression of the proteins is controlled by phage T5 promoter, two strong transcriptional terminators: t_0 from phage lambda and T1 from the *rrnB* operon of *E. coli* and two lac operator sequences (The QIAexpressionist - Qiagen, Hilden, Germany). The vector DNA was digested with the restriction endonucleases BamHI and HindIII. The reaction products were purified by agarose electrophoresis. The ligation reaction was carried out with the use of Rapid Ligation Kit (Promega, Mannheim, Germany) according to the protocol supplied by the manufacturer with the exception that 10 minutes incubation time at room temperature was applied. This reaction mixture was used to transform M15 pREP4 (Qiagen, Hilden, Germany) chemically competent cells (see Section 4.2.3 Preparation and transformation of competent *E. coli* cells). PFOMT protein without a his tag was obtained with the use of pQE60 and pQE70 vectors. The proteins cloned into those vectors were expressed without C terminal, vector derived his tag due to stop codons introduced into the DNA sequences with PCR primers.

4.1.4. DNA purification by agarose gel electrophoresis

The DNA amplification and cleavage reactions were analyzed by agarose gel electrophoresis on 1,2% agarose gel. To stain the DNA bands ethidium bromide (Roth, Karlsruhe, Germany) was applied to the gel in a concentration of 0.4 mg/ml. The gels were run in TBE buffer and visualized under UV light (BioDocAnalyze system, Biometra, Goettingen, Germany). The purified DNA was eluted from a gel using QIAEX II gel extraction kit (Qiagen, Hilden, Germany) according to the protocol supplied by the manufacturer. This kit relies on binding

of DNA to glass beads under high salt conditions and then subsequent elution with low salt buffer.

TBE buffer x 10 concentrate:

108 g Tris base

55 g Boric acid

9.3 g Na₄EDTA

Deionised water is added to 1 liter.

4.1.5. Preparation of plasmid DNA

The plasmids were prepared from the expression strain using QIAprep Spin Miniprep Kit (Qiagen, Hilden, Germany). The plasmids were isolated from 10 ml of overnight cell culture. The standard protocol supplied by the manufacturer was used with minor modifications. Two lysis reactions were applied to a single column in order to obtain more DNA. The principle of isolation of plasmid DNA using this kit is the selective binding of nucleic acid molecules to a silica membrane in the presence of high salt conditions. The DNA was eluted with 10 mM Tris/HCl buffer pH 8.5. The quantity of obtained plasmid was estimated by the measurement of absorbance at 260 nm. The quality of the DNA preparation was also assessed by the absorbance ratio at 260 nm to 280 nm. The measurements were carried out on Beckman DU 640 (Beckman Coulter, Krefeld, Germany) spectrophotometer.

4.1.6. Sequencing

The BigDye Terminator v1.0 (ABI Applied Biosystems, Foster City, U.S.A) was used for sequencing. The sequencing reactions were carried out with the optimized protocol using one sixteenth of the reaction volume suggested by the manufacturer. Primers 1 - 4 (Table 1) were used.

BigDye premix:

0,5 µl BigDye mix;

1,75 µl Reaction Buffer;

1,75 µl Sterile water

The sequencing reaction composition:

4 µl BigDye premix

1 µl Sequencing primer (3-10 pmol)

1-5 µl DNA (600ng plasmid DNA)

0-5 µl Sterile water

Thermocycler program:

Preheat the thermocycler to 96 °C 4min

25 cycles of:

96 °C - 10 sec

50 °C - 5sec

60 °C – 4min

4 °C –

The sequencing reaction products were purified by gel filtration. Dry sephadex G-50 Superfine (20-80 mm) 45 μ l was distributed to multiscreen 96 well plate. Each well was filled with 300 μ l sterile water. The plate was then covered parafilm and incubated 3 h in 4 °C. The water was removed by centrifugation at 910 g for 5min. Subsequently 10 μ l sterile water were added to each 10 μ l of sequencing reaction and the entire volume (20 μ l) filled to the middle of each of the Sephadex columns. The multiscreen plate was placed on top of the sequencing 96 well plate and centrifuged at 910 g for 5min. Samples purified in such way were then ready for analysis. The reaction products were analyzed by capillary electrophoresis. The ABI sequencer was used to analyze the reactions. The resulting sequences were compared with the theoretical ones. Each plasmid was sequenced in forward and reverse direction with two different primers (Table 1). The first was the primer used for amplification of the insert and the remaining one corresponded to the sequences flanking the multicloning site of the vectors used.

4.1.7. Site directed mutagenesis

Site directed mutagenesis was carried out according to the protocol from Quick change Kit (Stratagene, La Jolla, U.S.A). For the mutagenesis PCR AccuPrime Pfx Supermix (Invitrogen, Karlsruhe, Germany) was used and the set of primers 5 and 6 (Table1). The following PCR reaction was carried out to generate plasmids with altered sequences:

The PCR reaction composition:

1 μ l Plasmid DNA (15 ng) original PFOMT sequence in pQE30 vector
1 μ l Forward primer (100 pmol/ μ l)
1 μ l Reverse primer (100 pmol/ μ l)
1 μ l AccuPrime Pfx Supermix (Invitrogen, Karlsruhe, Germany)

Thermocycler program:

95 °C 5 min - initial deanturation and DNA polymerase activation
18 cycles of
 63 °C 30 sec - annealing
 68 °C 7 min - extension
 95 °C 15 sec - melting
4 °C

All PCR reactions were digested with 15U of DpnI restriction endonuclease for 1 hour in 37 °C in order to remove any remaining template DNA. Following the digestion, the reactions were used to transform chemically competent XL1Blue *E. coli* cells. The transformed cells were plated out on LB agar supplied with ampicillin/carbencillin at the concentration of 50 μ g/ml overnight and then re-grown overnight in 37 °C in liquid LB medium. On the

following day the colonies were regrown in 3 ml volume and plasmids sequenced. The plasmids that carried the desired sequence were used to transform the *E. coli* expression strain M15 pREP4. The transformed M15 pREP4 cells were again reselected on LB agar supplied with 50 µg/ml ampicillin/carbencillin.

Table 1. The Table of Primers used

Primer	Name	Direction	Sequence
1	pQE30/9 sequencing	Forward	5' GGA TCG CAT CAC CAT CAC 3'
2	pQE30/9 sequencing	Reverse	5' CCA AGC TCA GCT AAT TAA 3'
3	pQE60/70 sequencing	Forward	5' GAA TTC ATT AAA GAG GAG 3'
4	pQE60/70 sequencing	Reverse	5' CAG GAC TCC AGA CTC AG 3'
5	PFOMT	Forward	5' CGG GAT CCA TGG ATT TTG C 3'
6	PFOMT	Reverse	5' ATT AAG CTT TCA TCA ATA AAG ACG 3'
7	SynOMT K3A	Forward	5' CCA TCA CGG ATC CAT GGG TGC GGG CAT CAC CGG TTT TGA TCC 3'
8	SynOMT K3A	Reverse	5' GGA TCA AAA CCG GTG ATG CCC GCA CCC ATG GAT CCG TGA TGG 3'
9	N-terminal Fragment of Loop Hybrid	Forward	5' CGG GAT CCA TGG ATT TTG CTG TG 3'
10	N-terminal Fragment of Loop Hybrid	Reverse	5' CAA GAG CTA GCA TAG CAT C 3'
11	C-terminal Fragment of Loop Hybrid	Forward	5' GAA TTG GGT ACC TCG GAT G 3'
12	C-terminal Fragment of Loop Hybrid	Reverse	5' GCT GGA GCT CCA AGC TTT C 3'
13	PFOMT factor Xa site	Forward	5' CGG GAT CCA TCG AGG GAA GGA TGG ATT TTG CTG TGA TGA 3'
14	M. sativa factor Xa site	Forward	5' CGG GAT CCA TCG AGG GAA GGA TGG CAA CCA ACG AAG ATC A 3'

4.2. Protein expression

4.2.1. Characterization of expression strains

E. coli strain M15[pREP4] was used for the expression of the recombinant proteins. It is derived from *E. coli* K12 and has the phenotype NaIS, StrS, RifS, Thi⁻, Lac⁻, Ara⁺, Gal⁺, Mtl⁻, F⁻, RecA⁺, Uvr⁺, Lon⁺. Protein expression in this strain is regulated by the presence of the lac repressor plasmid pREP4. *E. coli* strain M15 does not harbor a chromosomal copy of the lacI_q mutation, so pREP4 plasmid must be maintained by selection on kanamycin resistance.

4.2.2. Storage strains

XL1 Blue *E. coli* strain was used for all cloning procedures. It harbors the lacI_q mutation, and produces enough lac repressor to efficiently block transcription. It is recommended for storing and propagating pQE plasmids..

4.2.3. Preparation and transformation of competent *E. coli* cells

Competent cells used for transformation were prepared according to the manufacturer (Qiagen, Hilden, Germany).

The competent cells were transformed as described by the manufacturer (Qiagen, Hilden, Germany) protocol with small modifications. To 50 ml freshly thawed competent cells 3 ml of the reaction mixture were added. The cells were allowed to rest on ice for 10 minutes. Subsequently a heat shock was performed, 50 sec in 43 °C, after the heat shock the tubes containing the cell suspension was allowed to cool down on ice for additional 15 min. After incubation on ice 500 µl of SOC medium was added to each of the tubes and the cells were shaken for 1h in 37 °C. The final step was plating of the cells to LB agar plate containing the proper antibiotic selection marker, either ampicillin or carbenicillin 50 mg/ml was used. The plates were left at 37 °C for overnight incubation. The plates were inspected for the growth of colonies. Usually, 6 colonies were picked from a plate and allowed to grow in liquid LB medium under the selection pressure of the antibiotic. Colonies developed in this manner were tested for expression.

4.2.4. Expression tests

To test for the expression of the protein colonies that were grown on the LB agar plate were picked and used to inoculate 3 ml of liquid LB medium supplied with 50 mg/ml ampicillin/carbenicillin antibiotic. The inoculated culture tubes were allowed to grow

overnight. After the overnight incubation 50-100 ml of the culture was used to inoculate fresh 3 ml of liquid LB medium. The fresh culture was incubated for 3.5 h at 37 °C and then the protein overexpression was induced with 1mM IPTG for 3 h. 50 µl aliquots of the culture were mixed with 50 µl SDS sample buffer. Those probes were then thermally denatured for 10 min in 80 °C and applied directly to SDS-PAGE to check for protein expression. The remainder of the cultures was centrifuged to collect the cells. The cell pellets were resuspended in 50 mM Tris/HCl pH 7,5 buffer containing 10% glycerol and 0,02% sodium azide. Subsequently the cells were lysed by sonication and the crude protein preparation was clarified by centrifugation. The total soluble protein preparation was then tested for activity by the chromatographic analysis of the enzymatic reactions. The colonies producing the highest amount of active protein were chosen for subsequent work.

4.2.5. Expression of PFOMT and N-terminal Hybrids

The proteins were cloned into pQE30 or pQE9 vector and transformed to M15 pREP4 *E. coli* strain. The production of the proteins was carried out in shaken flasks in 400ml volume. The cells were allowed to grow in liquid LB medium for 3 h, after that time the OD₆₀₀ reached the value around 0.6. At this time the cultures were induced with 1mM IPTG. The cultures were incubated for additional 3,5 h and then the cells were harvested by centrifugation.

4.2.6. Expression of loop Hybrids

The procedure for expression of loop hybrids was carried out similarly to the PFOMT proteins. The proteins containing the loop fragment were expressed in large part as insoluble proteins. Two different *E. coli* expression systems were tested. The pET vector system (Invitrogen, Karlsruhe, Germany) and pQE (Qiagen, Hilden, Germany) were used. The expression was carried out in appropriate *E. coli* hosts strains, BL21 DE3 for pET 28a vectors and M15 pREP4 for pQE30 vectors. The cells were grown in liquid LB medium with the addition of 50 mg/ml carbenicillin as a selection pressure. The growth time of the culture was extended to 5h in 37 °C. After that time the cultures were induced with 1 mM IPTG and allowed to incubate at 4 °C over weekend.

4.2.7. Expression of SynOMT

SynOMT protein was expressed analogously to the PFOMT proteins. However, in this case only the freshly transformed cells were used as inoculum for expression cultures.

4.2.8. Expression of *M. sativa* CCoAOMT

The reference OMT from *M. sativa* cDNA was a kind gift from Joseph P. Noel (The Salk Institute for Biological Studies, La Jolla, USA) and Jean-Luc Ferrer (Institut de Biologie Structurale, Grenoble, France). The DNA sequence cloned into pET15b was transformed into *E. coli* DE3 strain for expression. This protein was also expressed analogously to PFOMT.

4.3. Protein purification

4.3.1. Preparation of soluble protein fraction

The cultures which were used for protein overexpression were harvested by centrifugation. The culture medium was centrifuged for 5 min at RCF 16274 g in 4 °C. The supernatant was discarded and the cell pellet was collected for further processing. Subsequently cell pellets were resuspended in the buffer containing 50 mM Tris/HCl pH 7,5 buffer containing 10% glycerol and 0,05% sodium azide. Following, the resuspension lysozyme was added to the working concentration of 10 µg/ml then the tube containing the cell suspension was placed on ice for 15 min. Cell lysis was completed by sonication using the Sonoplus sonicator (Bandelin, Berlin, Germany) at 60% intensity. The sonication was performed in six 30 s bursts allowing the sample to cool down for additional 30 s in between the sonication cycles. The cell lysate was clarified by centrifugation for 10min 4 °C at RCF 16274 g. The clarified lysate containing all soluble cell materials was furthermore treated with 0,02% protamine sulphate solution in order to precipitate the majority of DNA contaminations. The protamine sulphate solution preheated to 45 °C was added drop wise to the stirred protein solution to a working concentration of 0,02%. The precipitated contaminations were then centrifuged for 10 min at 4 °C at RCF 1284 g and the crude soluble protein solution was used for further purification.

4.3.2. Immobilized metal affinity chromatography

All overexpressed proteins, that were expressed with vector derived his tag, were purified by Immobilized Metal Affinity Chromatography (IMAC) as a first step of purification. During the purification of methyltransferases Talon affinity matrix (BD Biosciences) was used with Äkta Explorer chromatography system (GE Healthcare, Munich, Germany). This metal affinity resin uses cobalt ions to complex proteins that have a his tag. Elution of proteins from the matrix can be achieved by a shift to low pH (pH 4-6) or by application of imidazole which binds competitively to the column matrix. Increasing concentrations of imidazole were used

to prevent the pH dependent denaturation of protein. The soluble protein preparation was applied to the column and washed with Buffer A. The loaded column was then washed once with the mixture of buffers A and B with the resulting concentration of imidazole equal to 30 mM. This washing step ensured the removal of unspecifically bound proteins. The elution of the protein of interest was accomplished by applying a stepwise gradient of increasing the imidazole concentration to 240 mM. At this point pure recombinant his tagged protein was eluted. The column regenerated with the aqueous buffer containing 30 mM MES pH 5,8 and 0.02% NaN₃ was then ready for another run or storage at 4 °C after flooding with 20% ethanol solution.

Buffer compositions:

Buffer A	Buffer B
50 mM KPi pH 7,5	50 mM KPi pH 7,5
200 mM NaCl	200 mM NaCl
10 glycerol	300 M imidazole
	10% glycerol

4.3.3. Gel filtration

Gel filtration (Rosenberg, 2005) was the final step of purification of protein samples used for crystallization. This chromatographic method separates the protein on the basis of their size. The concentrated protein sample from previous purification step was applied in the amount of 2 ml to a HiLoad 16/60 Superdex75 column (Pharmacia). The separation was carried out with the use of degassed KPi buffer ran at 1.5 ml/min at the pressure of 0.3 MPa. To accurately determine the molecular mass, the elution times of the samples was compared to the protein standards, BSA (67 kDa), ovoalbumin (43 kDa), Chymotripsin (17 kDa) (Serva, hidelberg Germany).

Buffer composition:

50 mM KPi pH 7,5
150 mM NaCl
10% glycerol

4.3.4. Hydrophobic interaction chromatography

To purify proteins that were expressed without a his tag, the hydrophobic interaction chromatography (HIC) with a phenylsepharose matrix used was (GE Healthcare) was used.

The purification of the proteins was achieved in two steps. The first one was batch hydrophobic interactions purification on a Buechner funnel filled with phenyl sepharose packing material (GE Healthcare). The crude soluble protein preparation was prepared with 50 mM KPi pH 7.5, 150 mM NaCl, 10% glycerol buffer and adjusted to 0.4 M ammonium sulphate concentration. Before loading the column was also equilibrated with this buffer containing 0.4 M ammonium sulphate. After loading the column the elution was carried out in steps with the buffers of decreasing ammonium sulphate concentration, 0.2 M, 0 M and 10% methanol solution in water. After that procedure the collected solutions were tested for enzymatic activity. The active fractions were pooled and applied to high performance HIC column. The second purification step was carried out on Phenylsepharose 6 Fast Flow (high sub) 16/10 (GE Healthcare). The column was adjusted to 0.5 M ammonium sulphate with the same buffer as used previously. The elution was performed at 1.5 ml/min in two gradient steps of decreasing ammonium sulphate concentration 0.5 M to 0.25 M in 20min and 0.25 M to 0 M in the following 50 min. The collected fractions were tested for purity by SDS-PAGE.

4.3.5. Concentration of protein solutions

The proteins of interest were concentrated by ultra filtration in Amicon centrifugal filter units (Millipore, Schwalbach/Ts, Germany) using 10 kDa cutoff membranes. The purified protein from IMAC purification was concentrated to a volume of 1 ml, then diluted with appropriate storage buffer (50 mM Tris/HCl pH 7.5 buffer containing 10% glycerol and 0.05% sodium azide), and reconcentrated to a desired protein concentration. After this step, PFOMT proteins were ready to be frozen for storage, with the addition of 10% glycerol to the concentration buffer as a cryoprotectant.

4.3.6. Determination of protein concentration

The amounts of protein were estimated using the absorption at 280 nm and the calculated extinction coefficient based on an amino acid sequence calculated by the program PROTEAN, a part of DNA star (Madison, USA) software package.

Protein concentration was also determined by the Bradford assay (Bradford, 1976; Stoscheck, 1990) based on the shift of absorbance for a Coomassie brilliant blue G-250 dye at 595 nm. When the dye binds to arginine and hydrophobic residues present in the protein its color changes to blue (absorbing at 595 nm) from red color of unbound (anionic form) which has an absorbance maximum at 470 nm. The absorbance of different samples was measured on a Beckman spectrophotometer.

4.4. Protein analysis

4.4.1. SDS PAGE

The purified protein was analyzed for purity by SDS-PAGE (Laemmli, 1970). This electrophoresis analysis method separates the proteins basing on their electrophoretic mobility, which in turn is dependent on their size. The gel for the analysis was assembled on a Protean II (Biorad, Hercules, U.S.A). The protein bands are visualized by staining with Serva Blue R (Serva, Heildelberg, Germany) stain.

Buffer compositions:

The sample buffer 4x

40% (w/v) Glycerol

50 mM DDT

2% (w/v) SDS

0.0625 M Tris/HCl pH 8.8

0.0004% (w/v) Bromophenol blue

The running buffer x10 stock 1 l

30 g Tris

144 g Glycin

10 g SDS sodium salt

Deionised water to 1 l

Stacking gel 4.7%

3 ml Water

1.25 ml 0.5M tris pH6.8

0.75 ml 30% Acrylamide/bisacrylamide (30:1) Solution

25 µl Ammonium persulphate (APS)

5 µl TEMED

(N,N,N',N' tetramethylenediamine)

Running gel 14%

2.8 ml Water

2,5 ml 1M Tris pH 8.8

4, 7ml 30% Acrylamide/bisacrylamide (30:1) solution

50 µl APS

5 µl TEMED

4.4.2. Western blot

The principle of this method is immuno-detection of proteins which are immobilized on a membrane (usually nitrocellulose). The protein preparations are firstly separated by size with SDS PAGE, and then the proteins are transferred to a nitrocellulose membrane (Invitrogen) by electroblotting. Subsequently the membrane is probed with the antibodies specific to the protein of interest. The protein specific antibodies against PFOMT prepared in rabbit were purchased from Eurogentec. The last step is visualization, which employs subsequent washings with the solution of an enzyme linked antibody specific to the antibody used to probe for the protein. The enzymatic reaction is used for visualization.

For visualization the antibody against the rabbit coupled to alkaline phosphatase was used. The detected protein bands were visualized by the color reaction with BCIP (5-bromo-4-

chloro-3-indolyl phosphate) and NBT (nitro blue tetrazolium). The nitrocellulose membrane was treated with 15 ml of blocking solution for 2 h, and after that time a protein specific antibody was added to the solution. The membrane was incubated with protein specific antibody overnight. The membrane was washed 3 times with TBS pH 7.5 buffer. The antibody against the rabbit coupled to the reporter enzyme was also dissolved in 15 ml of the blocking solution and incubated for 2 h with the membrane. The membrane was then washed again 3 times in TBS pH 7.5 and transferred to TBS pH 9.5. 50 ml of BCIP solution and 25 ml of NBT solution were added. The membrane was incubated until signals were detected. The reaction was stopped with 0.5 M EDTA. The membrane was dried and photographed.

Buffer compositions:

Electroblotting buffer	Blocking solution
Tris buffered saline TBS pH 7.5 and pH 9.5	5% Powdered milk in TBS buffer 1% BCIP in 100% DMF (dimethylformamide) 1.5% NBT in 70% DMF

4.5. Protein modification

4.5.1. His-tag cleavage

In order to remove the his tag from the N-terminus of the protein a Factor Xa protease cleavage site was introduced. The amino acid sequence IEGR, the cleavage site recognized by this site specific protease, was placed in front of the methionine where the protein sequence starts. This was accomplished by the use of the PCR primers which encoded the sequence containing Factor Xa cleavage site (Table1) The purified proteins containing the Factor Xa cleavage site were re-buffered into protease Factor Xa reaction buffer. The cleavage reactions were conducted in a volume of 400 μ l. The amount of protein was adjusted to 0.25 mg/ml. Two units of Factor Xa protease (Qiagen, Hilden, Germany) were used per reaction (400 μ l). The reactions were incubated at 37 °C for 24 h to achieve complete cleavage. After the digestion reaction the protease was removed using Factor Xa Removal resin (Qiagen, Hilden, Germany). This resin is linked to benzamidine which is a trypsin-like serine protease inhibitor and binds covalently to the active site of Factor Xa. The reactions were purified by Talon-affinity resin. The cleaved protein eluted with the flow through. It was concentrated and rebuffered to 50 mM Tris/HCl pH 7.5, 10% glycerol and subsequently checked for enzymatic activity. The cleavage and the purity were confirmed by SDS-PAGE. In order to verify the

absence of any His-tag, Western Blot analysis was also performed with His-tag specific antibodies (Qiagen, Hilden, Germany).

Protease Factor Xa reaction buffer:

20 mM Tris/HCl pH 6.5

50 mM NaCl

1 mM CaCl₂

4.6. Activity assays

Enzyme assays were performed in a buffer containing 100 mM KPi (pH 7.5) and 10% glycerol with 10 μ M substrate (dissolved in 30% DMSO), 0.5–2 μ g (and up to 10 μ g in case of SynOMT K3A mutant) of total protein, and 400 μ M AdoMet in a total volume of 50 μ l. The assays were incubated at 30 °C for 60 to 3600 seconds (dependent on the protein and substrate tested). The reactions were stopped by the addition of 20 μ l 7% trichloroacetic acid in 50% acetonitrile/water. The reaction products were analyzed as described (Ibdah et al., 2003). Caffeoylglucose was prepared as described previously from caffeic acid and UDP-glucose with the purified recombinant sinapic acid glucosyltransferase (SGT) from *Brassica napus* (Milkowski et al., 2000). CCoA, was prepared based on published methods (Strack et al., 1987). Quercetagenin was obtained from Extrasynthese (Genay, France). Caffeic acid and quercetin were obtained from Serva (Heidelberg, Germany) and Roth (Karlsruhe, Germany) respectively. The reaction products were analyzed by reversed phase liquid chromatography on a Nucleosil 5- μ m C18 column (50 mm length x 4 mm inner diameter; Macherey & Nagel, Düren, Germany), as described previously (Stockigt et al., 1975; Vogt et al., 1999). Compounds were analyzed with linear gradients from 10% B (100% acetonitrile) in A (1.5% aqueous phosphoric acid) to 70% B in A (for phenolics), from 5% B to 50% B in A (for free acids and CoA esters), from 5% B to 30% B (for glucose esters), and from 20% B in A to 80% B in A in 4 min (for flavonoids) at a flow rate of 1 ml min⁻¹. Detection of flavonoids, catechol, coumarins, and hydroxycinnamic acid esters was performed between 260 and 400 nm. Identification and quantification was achieved with reference compounds from our Institute collection or from external sources. For $K_{m\ app}$ determination of methyl group acceptors, acceptor concentrations were chosen between 2 and 20 μ M, while AdoMet was kept constant at 1.5 μ M. The apparent $K_{m\ app}$ and V_{max} values were calculated by nonlinear curve fitting assuming steady state Michaelis-Menten kinetics and from Lineweaver-Burk plots. All enzyme assays were recorded in triplicates.

4.7. Isolation of the enzymatic reaction products of SynOMT

The enzymatic reaction products were prepared in similar way as the enzymatic reactions used for activity tests. The amounts of reagents were upscaled and incubated in 37°C for 10 minutes, several enzymatic reactions were pooled together. Subsequently the reaction mixtures were separated using semi preparative HPLC. The same solvent system as described for the activity assays was used to achieve the separation of the reaction products. The gradient times were adjusted to accommodate for the longer column. The fraction containing the compound of interest was collected. When all of the reaction mixture was separated, the collected fraction was applied to a Solid Phase Extraction mini-cartridge (Waters, Eschborn, Germany) pre-equilibrated with water. The cartridge was washed with water and 10% methanol then the bound compounds eluted with 100% methanol. The eluate was concentrated and dried under vacuum.

When the separation of the different reaction products was not required (GC-TOF/MS), the enzymatic reactions were extracted with ethylacetate. The organic phase was separated and then the solvent evaporated under vacuum. The dried mixture of products and substrates was subsequently redissolved in 100% methanol and dried under vacuum. Samples prepared in this way were given to NMR analysis to verify the structures of the products of SynOMT reaction. The ¹H and ¹³C NMR analyzes were performed by Dr. Andrea Porzel. Additional GC-MS analyzes were performed by Dr. Juergen Schmidt and Dr. Willibald Schliemann.

4.8. Crystallographic procedures

4.8.1. Crystallization of proteins

Purified, concentrated (4-8 mg/ml) and rebuffed PFOMT protein preparations were used for crystallization. The proteins were crystallized by vapor diffusion method. To find out initial crystallization conditions, a screening procedure was used. Commercially available crystallization screens Hampton Research (Aliso Viejo U.S.A), Sigma-Aldrich (Hamburg, Germany), Jena Bioscience (Jena, Germany) were used. For the screening of conditions a Cartesian pipetting robot was used to prepare the crystallization setups. The 96 well format sitting drop plates were used, with 200 nl drop size. The drops set up by the robot were scanned by Veeco LC3 imaging robot (Veeco Instruments GmbH, Mannheim, German) The subsequent refinement of the crystallization conditions was carried out by hand on 24 well

format in 4 μ l drop volume. The proteins were always co crystallized with their ligands, magnesium ion and DMSO.

PFOMT was crystallized under the following conditions 20% PEG 4000, 0.2 M CaCl₂ in 100 mM HEPES/NaOH pH 7.0 using a concentration of 3 mg/ml PFOMT, 250 μ M MgCl₂, 250 μ M AdoMet, 25 μ M quercetin, and 2.5% DMSO. The sitting drop method was used.

The SynOMT was also crystallized by vapor diffusion method from a sitting drop setup. The crystallization conditions screens were carried out identically as in the case of PFOMT. The protein solution contained 250 μ M MgCl₂, 250 μ M AdoMet, 25 μ M caffeic acid and 2.5% DMSO. The crystals were obtained from 0.2 M MgCl₂ x 6H₂O, 0.1 M Tris hydrochloride pH 8.5, 30%w/v polyethylene glycol 4000.

4.8.2. Data collection and processing

The diffraction data were collected at 100 K on MSC Rigaku diffractometer equipped with Raxis IV++ imageplate detector, with copper rotating anode X-ray source and Osmic optics. (Rigaku MSC, Sevenoaks, England). The synchrotron data were collected at BW 6 DESY (Hamburg). The scaling and processing of data was achieved with the use of HKL2000 software package (Otwinowski et al., 1997) and XDS (Kabsch, 1993)

4.8.3. Structure solution

The crystal structure of PFOMT was solved by SAD experiment using seleno-methionine (SeMet) protein derivative. CNS (Brunger et al., 1998) was used for the structure determination. To produce this protein derivative the bacterial cells were grown in the presence of SeMet (50 μ g/liter). The SeMet derivative crystallized in the same conditions as the native protein. The PFOMT structure (PDB code 3C3Y) was solved by Dr. Daniel Rauh (Chemical Genomics Centre of the Max Planck Society, Dortmund). Two datasets at different wavelengths were collected for SeMet derivative and a data set for native protein crystal at DESY synchrotron facility (Hamburg, Germany) (BW6). (Table 2)

The structure of SynOMT (PDB code 3CBG) was solved by molecular replacement using PFOMT structure as a search model, which was performed with the use of program Phaser (Mccoy et al., 2007), part of the CCP4 (Bailey, 1994) crystallographic software package.

4.8.4. Model building, refinement and visualization

Model building was performed using the program O (Jones et al., 1991) and Coot (Emsley et al., 2004), and the structures refined using CNS and REFMAC 5 (Murshudov et al., 1997).

The coordinates for the AdoHcy cofactor and other ligands were obtained from HicUp server (Kleywegt et al., 1998). The twinning of the SynOMT crystals was identified by “identify_twin” input from CNS and then twinned refinement was carried out using previously identified twinning law and fraction. Visualization of the structures was performed by Pymol (Delano, 2002).

4.8.5. *In silico* substrate docking

In silico substrate docking experiments were carried out using AutoDock4 (Morris et al., 1998) with AutoDockTools GUI [<http://mgltools.scripps.edu/>]. The starting position and orientation of the substrate was modeled in by hand. The Lamarckian Genetic Algorithm was used to search the active site of the protein for alternative substrate conformations. 250 runs were conducted with population size of 150 and 250000 energy evaluations. 27000 generations were analyzed with only one conformation surviving to the next generation. The searches carried out with the use of Lamarckian Genetic Algorithm were complemented with local search scheme. 300 iterations of local search were carried out for each run. The results of docking were manually analyzed. The productive docking of a substrate in the active site with the smallest docking energy was considered probable.

5. Results

5.1. Structural characterization of PFOMT

PFOMT is a 237 amino acid protein with the predicted molecular mass of 26.6 kDa. The recombinant protein was expressed as a his tag fusion. The vector derived N-terminal his tag adds additional 12 amino acids to the N-terminus of the protein. PFOMT exists as a homodimer in the solution (Figure 13) as confirmed by previous research. (Ibdah et al., 2003). The initial coordinates coming from the SAD experiment were obtained by Dr. Daniel Rauh. Table 2 summarizes the data collection and refinement statistics for the PFOMT structure (PDB code 3C3Y).

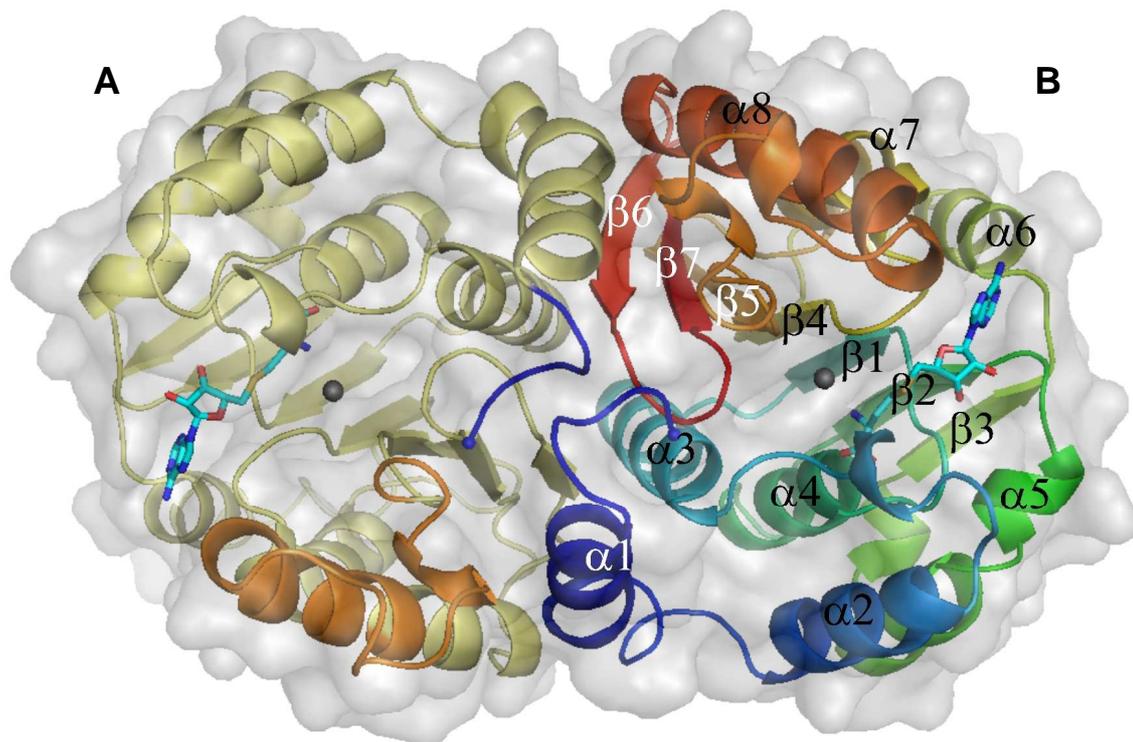


Figure 13. Dimer representation of PFOMT. The gray spheres represent the Mg^{2+} in the binding site and bound AdoMet is shown as cyan sticks. The secondary structure elements are labeled for the monomer on the right (B). The conserved fold of Mg^{2+} dependant OMTs can be recognized, with centrally located β -sheet surrounded by α -helices. For the monomer on the left the (A) N-terminus is colored blue and the catalytically relevant loop region in orange.

Table 2. Crystallographic data, phasing and refinement statistics for PFOMT**Data Collection**

Dataset	native	peak	high remote	
Wavelength (Å)	1.05	0.97905	0.95	
Resolution (Å)	1.37	1.93	1.93	
Total reflections	436990	247123	259057	
Unique reflections	93389	65998	72580	
Completeness ¹ (%)	97.6 (92.9)	99.2 (87.1)	99.2 (88.8)	
$I/\sigma(I)$ ¹	27.88 (2.33)	32.14 (10.46)	28.95 (3.71)	
R_{sym} ²	0.069	0.071	0.067	
Redundancy	4.67	3.74	3.60	
Space group	P2 ₁ 2 ₁ 2 ₁			
Cell dimensions (Å)	a=	48.89	49.47	49.47
	b=	71.83	71.78	71.78
	c=	128.12	128.15	128.15

Refinement

Refinement reflections	89982
R^3	0.18850
R_{free} ⁴	0.22427
Protein atoms	3577
Ligand atoms	52
Solvent molecules	480
Ion atoms	4
Rmsd bond lengths ⁵ Å	0.008
Rmsd bond angles ⁵ °	1.245
<B-factor> protein (Å ²) chain A	17.395
<B-factor> protein (Å ²) chain B	17.207
<B-factor> ligand (Å ²)	14.238
<B-factor> solvent (Å ²)	29.079
<B-factor> ion (Å ²)	14.450

Ramachandran plot:

most favoured regions	92.1%
additional allowed regions	6.9%
generously allowed regions	1.0%
disallowed regions	0%

¹ The values in the parentheses represent the value for the highest resolution shell² $R_{\text{sym}} = \sum |I_h - \langle I_h \rangle| / \sum I_h$, where $\langle I_h \rangle$ is the average intensity over symmetry equivalent reflections³ $R = \sum ||F_{\text{obs}}| - |F_{\text{calc}}|| / \sum |F_{\text{obs}}|$.⁴ R_{free} is calculated the same way as R using 4.3% of the data (4041 reflections) that is excluded from the refinement⁵ The rmsd for bonds and angles are the root mean squared deviation from the ideal values.

The protein crystallizes in $p2_12_12_1$ space group where two monomers forming a dimer are visible in the asymmetric unit. The protein exhibits α/β Rossmann fold. The first 12 amino acids of the PFOMT protein with additional 12 residues derived from the vector are not visible in the electron density. The absence of interpretable electronic density for the N-terminus of the protein might be explained by the high flexibility of this region. Comparison of PFOMT structure with the structure of *M. sativa* CCoAOMT (Ferrer et al., 2005) shows almost identical fold of this protein.

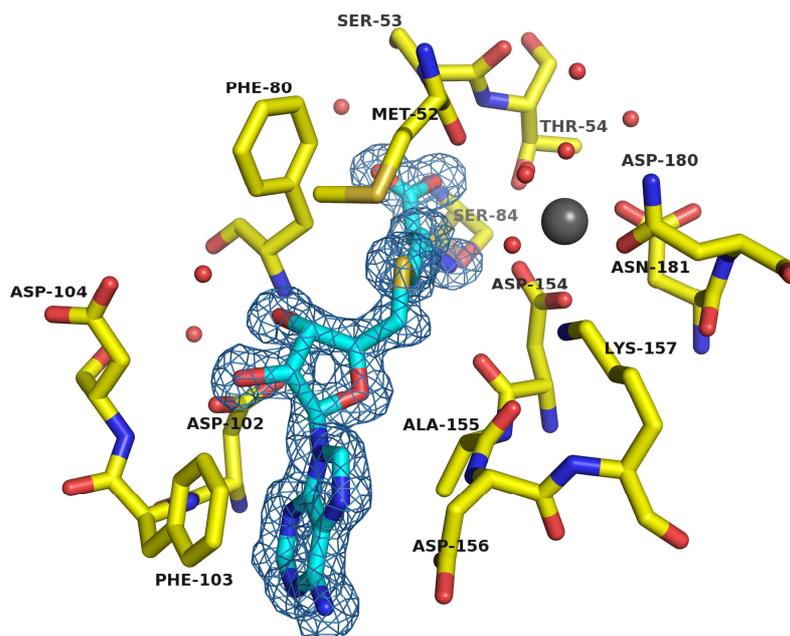


Figure 14. Representation of the AdoMet/AdoHcy binding site of PFOMT. The protein carbon atoms are shown in yellow while carbon atoms belonging to the ligand are colored cyan. The nitrogen atoms are colored blue, and oxygen, red. The gray sphere represents the Ca^{2+} ion. The electron density at $\sigma=1$ is shown only for the ligand molecule. The red spheres represent the oxygen atoms belonging to water molecules.

The binding site for AdoMet/AdoHcy is very conserved in this group of proteins (Figure 14). The amino acid identified to take part in AdoMet binding in the case of CCoAOMT structure are as well conserved in the PFOMT protein. Moreover the superposition of the two structures results in very similar orientation of the AdoMet cofactor. In the active site of PFOMT a clear indication of a bound Ca^{2+} ion is found. The presence of calcium ion is the result of high concentration of CaCl_2 in the crystallization solution. The mode of ion binding is almost identical to those of CCoAOMT (Ferrer et al., 2005). In case of PFOMT also the acidic residues take part in octahedral coordination of the metal ion. Those residues

include, Asp 154, Asp 180 and Asn 181, as well as Thr 54. Unfortunately no electron density is visible for the methylated substrate in the active center of the protein.

5.2. *In silico* docking experiments

In silico docking of substrates into the active centers of PFOMT and CCoAOMT was carried out to obtain semiquantitative data about the relative substrate specificities of those two plant enzymes. Table 3 summarizes the results of the docking experiments. The molecules of quercetin, caffeic acid and 5-hydroxyferuloyl CoA were chosen. The preliminary results and the previous experiments (Ferrer et al., 2005) show that an enzymatic specialist CCoAOMT would prefer binding of CoA esters and not caffeic acid. PFOMT as an enzyme with broad substrate specificity would show high activity towards flavonoid substrates as well as caffeic acid and its CoA esters. The results of the docking experiments were in agreement with previous observations, despite the fact that the structural data for catalytically relevant N-terminus was missing. Quercetin, a representative of flavonoids could be docked in a substrate-like (Figure 15) orientation into the active site of PFOMT with reasonable binding energies, while no enzymatically reasonable binding mode could be found for CCoAOMT.

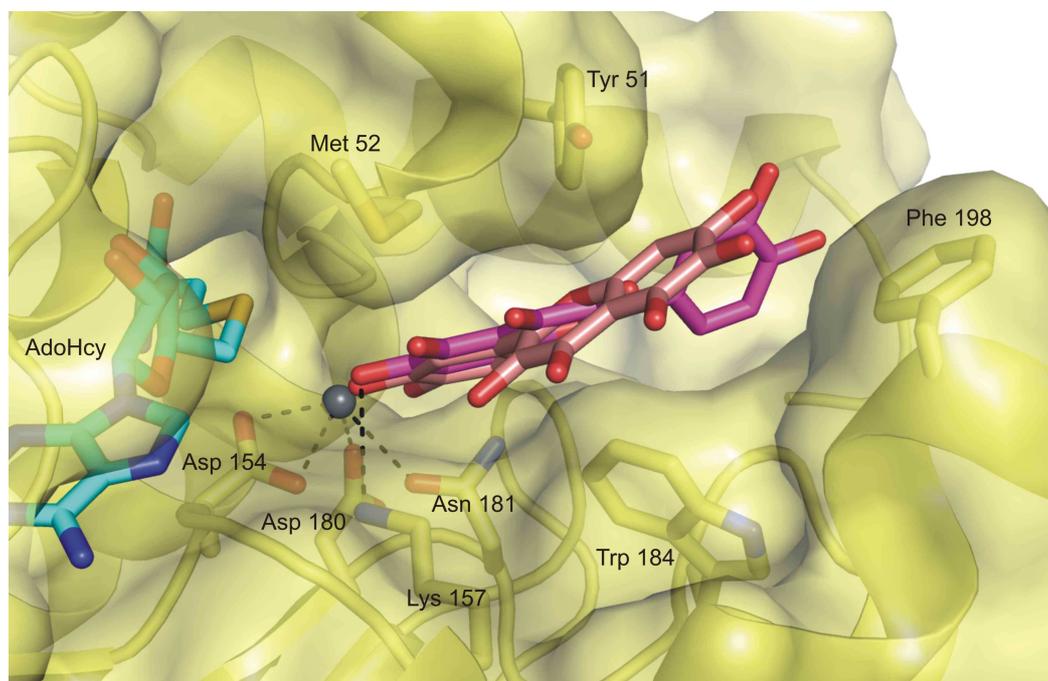


Figure 15. Manual Docking of the substrate quercetagenin in the active site of PFOMT. Quercetagenin can be methylated by PFOMT in two positions (Ibdah et al., 2003) requiring opposing orientations of the substrate (pink and purple). In both cases Both orientations aromatic stacking interactions are observed between the substrate and the side chains of Tyr 51, Trp 184, and Phe 198. The Mg^{2+} ion is represented by gray sphere, dark lines represent the binding of the cation by residues Asp 154, Asp 180 and Asn 181 as well as interaction between the oxygen of the hydroxyl group and the catalytically relevant Lys 157. AdoHcy is shown in cyan.

For both structures caffeic acid and 5-hydroxyferuloyl CoA could be docked into the active centers. The binding energies of those dockings clearly distinguish between those two enzymes. Caffeic acid is preferred by PFOMT in terms of binding energy, cluster ranking, and cluster size, while the results for CCoAOMT are less conclusive. For the CoA ester the situation is reversed. In this case the docking to the CCoAOMT is preferred. The results for PFOMT show that only a fraction of the docking attempts has favorable (negative) binding energies.

Table 3. Results of *in silico* docking of quercetin, caffeic acid, and 5-hydroxyferuloyl CoA into the active sites of PFOMT and CCoAOMT

PFOMT				CCoAOMT			
Cluster rank	Lowest binding energy	Mean binding energy	Number in cluster	Cluster rank	Lowest binding energy	Mean binding energy	Number in cluster
Quercetin							
1	-7.93	-6.81	53	1	-8.18	-7.44	67
2	-7.51	-6.72	41	2	-7.94	-7.47	63
3	-7.00	-6.56	23	3	-7.73	-7.31	16
4	-6.93	-6.32	15	4	-7.73	-7.03	32
5	-6.84	-6.49	13	5	-7.48	-7.10	25
5.1	-6.84			6	-7.21	-6.92	3
6	-6.83	-6.45	8	7	-7.17	-6.82	15
7	-6.75	-6.31	8	8	-7.16	-7.16	1
8	-6.68	-6.37	14
9	-6.61	-6.37	5
10	-6.54	-6.08	10
11	-6.52	-6.25	44
12	-6.49	-6.18	7	18	-6.40	-6.36	2
13	-6.38	-6.32	3	19	-6.37	-6.37	1
14	-6.34	-6.34	1	20	-6.35	-6.35	1
15	-6.17	-5.80	3	21	-6.02	-6.02	1
16	-5.82	-5.78	2	n.d.			
Caffeic acid							
1	-5.44	-4.92	195	1	-4.53	-4.23	16
1.1	-5.44			2	-4.50	-4.20	13
2	-4.66	-4.21	32	3	-4.47	-4.22	11
3	-4.41	-4.31	5	4	-4.45	-4.31	3
4	-4.33	-4.21	3	5	-4.43	-4.09	83
5	-4.18	-4.07	2	5.1	-4.43		
6	-4.18	-4.18	1
7	-4.17	-4.17	1
8	-4.09	-3.98	5
9	-3.94	-3.77	6	19	-3.63	-3.63	1
5OH-feruloyl CoA							
1	-0.35	+0.86	1	1	-2.85	-2.17	144
1.1	-0.35			1.1	-2.85		
2	+0.42	+1.22	20	2	-2.64	-2.16	3
.
.
.
55	+2.65	+2.69	2	11	-1.90	-1.68	5
56	+2.92	+2.92	1	12	-1.90	-1.90	1
57	+2.96	+2.96	1	13	-1.25	-1.19	2
58	+3.08	+3.08	1	14	-1.23	-1.23	1
59	+3.25	+3.25	1	15	+0.66	+0.66	1
The highest ranking substrate-like binding modes are indicated in boldface, together with the corresponding cluster and binding energy (kcal mol ⁻¹). n.d., not detected.							

5.3. Structural and amino acid sequence similarity of plant OMTs

The alignment of amino acid sequences for several CCoAOMTs show that sequences of CCoAOMTs from different plants are very similar (Figure 16). The amino acid sequence of PFOMT and promiscuous CCoAOMT-like enzyme from *S. longipes* show some differences. There are only two regions where PFOMT differs considerably from CCoAOMTs. The first one is the variable N-terminal region shaded in blue. The N-terminus varies for all protein even those coming from CCoAOMT family. It is worth noticing that the length of N-terminal domain is much shorter in PFOMT than the other enzymes. The second region where sequence differences are found is a loop region, shaded orange, spanning from amino acid residue 185 to 207 (the orange box Figure 16) of PFOMT located between strand $\beta 5$ and helix $\alpha 8$. It could be noticed by the amino acid sequence analysis that this region demonstrates the lowest sequence homology between CCoA-specific and promiscuous enzymes (30% identity for residues 185–207).

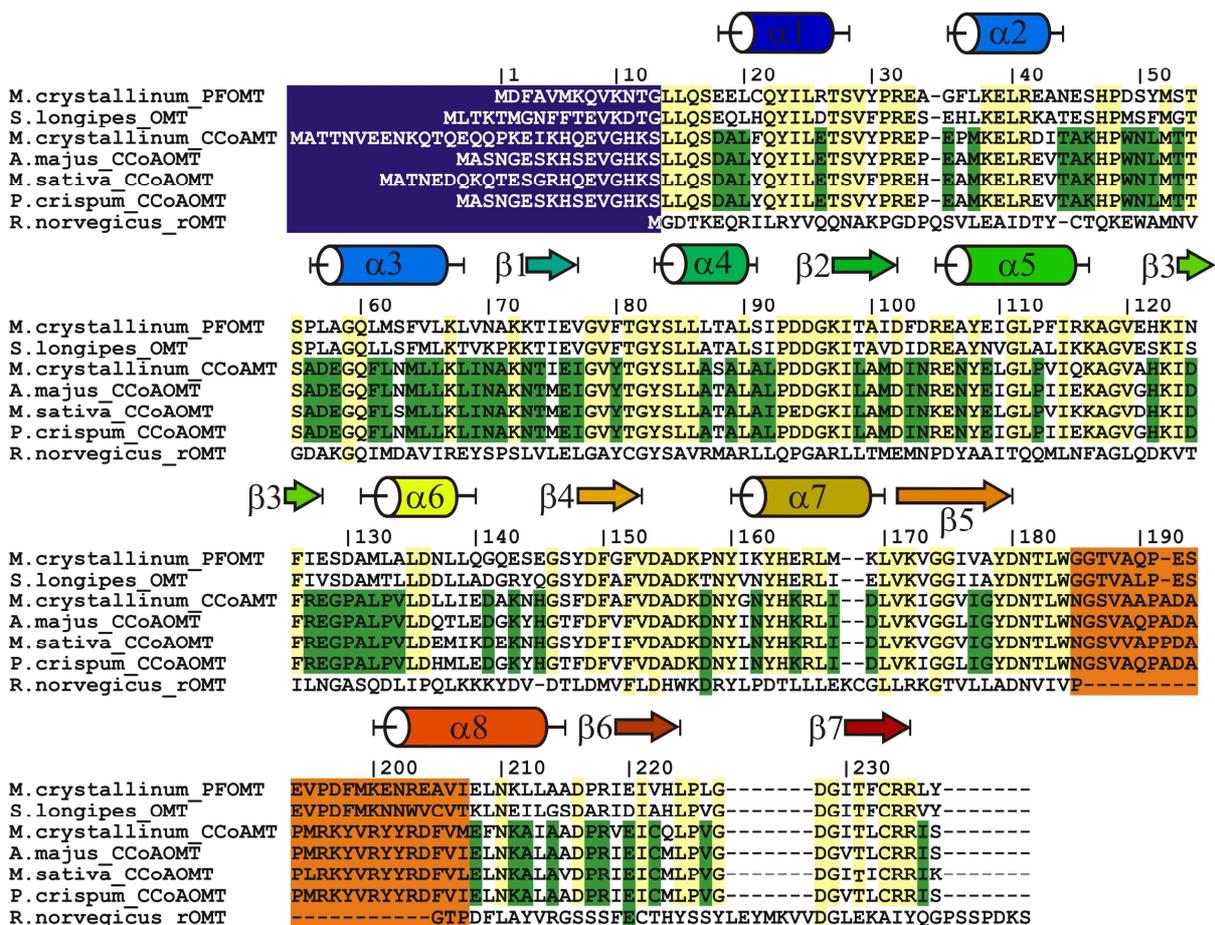


Figure 16. Amino acid sequence alignment of substrate specific and promiscuous cation dependent OMTs. The sequence of rOMT is introduced for reference. Residues shaded in yellow are common for all OMTs, in green are common only for CCoAOMTs. The N-terminal region is shaded blue and the variable loop region in orange. The numbering and the secondary structure elements are depicted for PFOMT.

The superposition of the PFOMT structure with CCoAOMT (Figure 17), shows the structural similarity of these two proteins. The RMSD between the positions of all atoms in both structures is 0.82Å. CCoAOMT from *M. sativa* was chosen as a representative of CCoAOMT family because of the availability of the structural information. However, there are regions of proteins where structural differences can be observed. The conformation of the loop region in both structures is quite different, as is the sequence of amino acid residues comprising this region (Figure 16 orange box). In addition, the modeling of the substrate in the structure of CCoAOMT shows that this region of protein may form some interactions upon substrate binding.

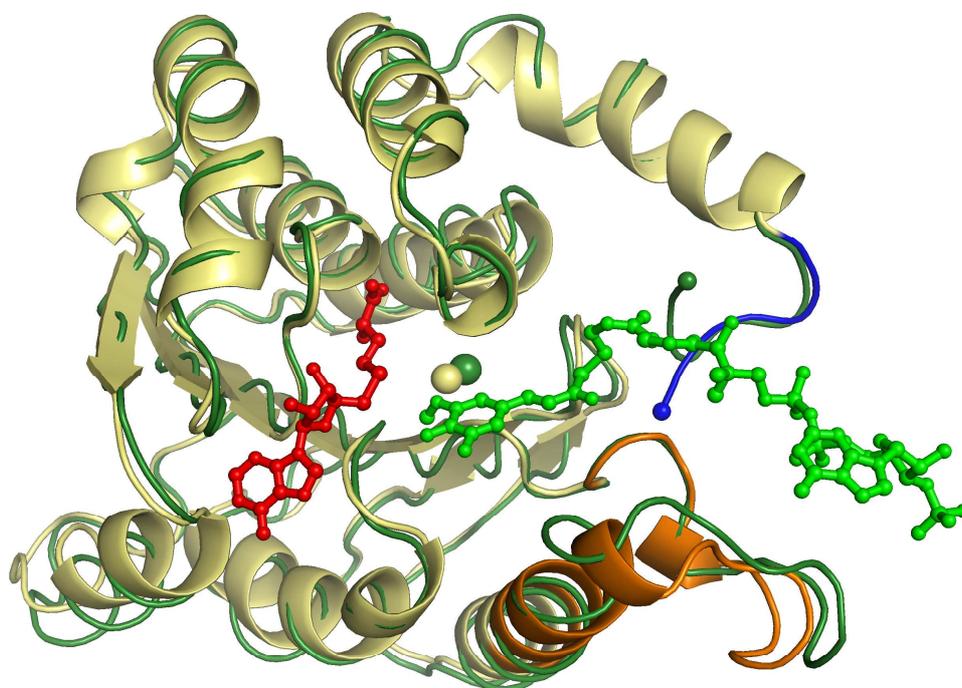


Figure 17. The overlay of structures of PFOMT yellow and *M. sativa* CCoAOMT green. The Loop spanning from amino acids 185 to 207 of PFOMT is shown in orange and the N-terminus in blue. AdoHcy molecule is represented in red. Feruloyl CoA is shown in green. The Spheres represent the cations bound to the structures.

The N-terminus of both structures also adopts a different conformation. It should be noted that the first amino acid residue for which the electron density is observed is Gly 13 in case of the PFOMT structure and Lys 21 for the structure of CCoAOMT. Due to this fact, the exact information how N-termini of those proteins interacts with the substrate is missing. It seems

plausible that differences in this loop and in the N-termini of the two enzyme classes (which also exhibit low sequence conservation and are observed neither in the present structure nor in that of CCoAOMT) could provide a clue to the promiscuous nature of PFOMT.

5.4. Hybrid proteins

5.4.1. N-terminus of PFOMT controls substrate specificity

The previous work carried out on PFOMT shows that the recombinant protein differs from the native protein isolated from the plant. Sequencing shows that the N-terminal domain of native PFOMT is 11 amino acid residues shorter when compared to recombinant PFOMT. The recombinant enzyme, in addition to a longer N-terminus, has also a N-terminal his tag. A further investigation of this phenomena shows that the length of the N-terminal domain is responsible for differences in regiospecificity, substrate affinity and methylation efficiency (Vogt, 2004). In order to investigate the effect of this N-terminal domain on the substrate specificity of the protein the following Hybrid protein was prepared. The first 13 amino acids of PFOMT were replaced with the first 22 amino acids from *M. sativa* CCoAOMT sequence. This resulted in a hybrid protein carrying the CCoAOMT N-terminal domain and PFOMT core (Figure 18). The his tag which was used to purify the protein is also located at the N-terminus. A sequence containing the Factor Xa cleavage site was designed in order to produce the protein where it was possible to completely remove the his tag. The addition of a factor Xa recognition sequence just before the Met1 residue allowed to cleave off the his tag and produce pure full length protein.

5.4.2. Loop Hybrid

In order to investigate the influence of the variable loop region a hybrid protein containing the sequence coding for the loop 194-217 of the *M. sativa* CCoAOMT was produced. This protein contains the sequence of PFOMT and only the amino acids 185-207 are replaced with the corresponding sequence (amino acid residues 194-217) from CCoAOMT (Figure 18). Due to very low yield of soluble protein, the Factor Xa cleavage site was not introduced and the cleavage reaction could not be performed. The conjunction of the two constructs resulted in a double hybrid protein. This protein contains the N-terminal sequence and the loop sequence of CCoAOMT combined with the core sequence of PFOMT. In theory, this protein should be closest in kinetic properties to CCoAOMT, since it contains both regions that are believed to be crucial for protein activity and substrate specificity.

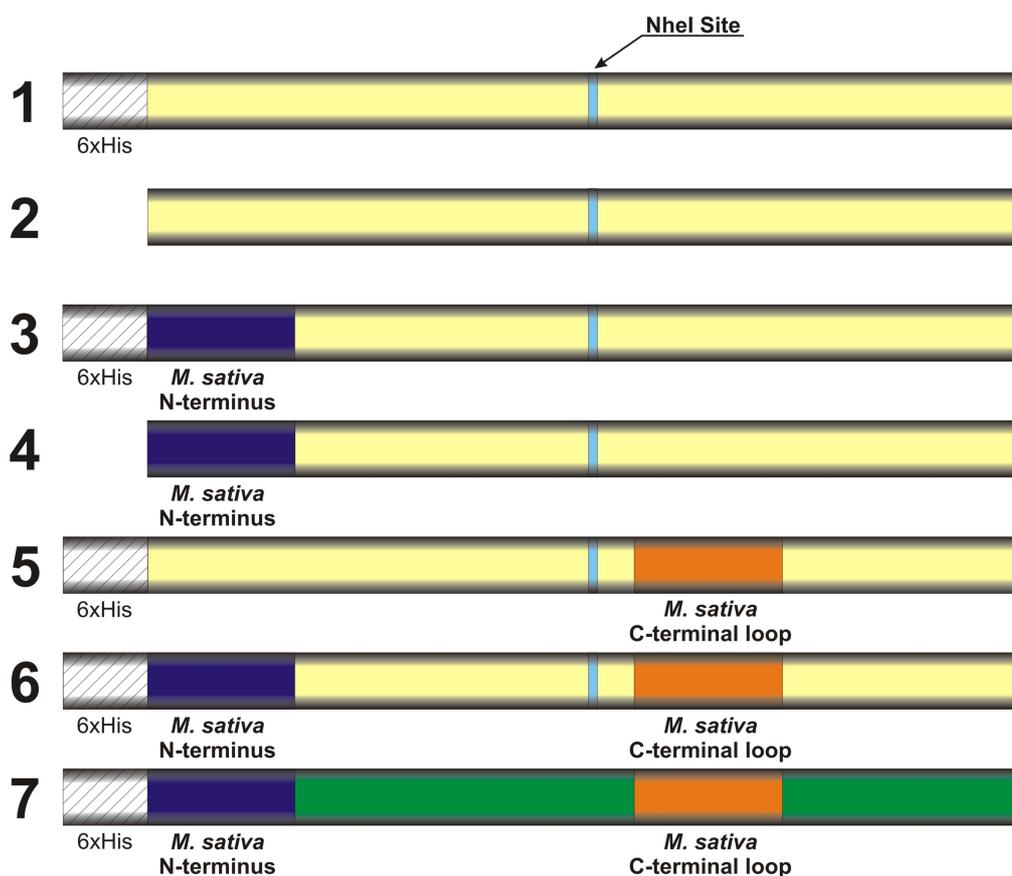


Figure 18. Schematic representation of cloning strategy which was used to produce the hybrid proteins. The NheI restriction site was used for ligation of DNA fragments (See Materials and Methods section). 1 PFOMT N0, 2 PFOMT N0Xa1 his tag cleaved, 3 N22 hybrid, 4 N22Xa1 hybrid his tag cleaved, 5 Loop Hybrid, 6 Double Hybrid, 7 CCoAOMT.

5.5. Characterization of hybrid proteins

5.5.1. Production and purification of hybrid proteins

The differences in the structures of PFOMT and *M. sativa* CCoAOMT show that certain parts of otherwise conserved protein structure adopt a different conformation. Those regions were thought to be responsible for differences in substrate specificity among Mg^{2+} dependent OMTs. The artificial, chimera proteins should help to reveal the structural differences responsible for differences in activities and substrate specificity. Those proteins consisted of PFOMT amino acid sequence containing parts of the CCoAOMT sequence. The following naming scheme (see Figure 18) was proposed: full length PFOMT, PFOMT N0; full length PFOMT with factor Xa cleavage site, N0Xa1; N-terminal hybrid, N22; N-terminal hybrid with factor Xa cleavage site, N22Xa1; hybrid containing the loop region from CCoAOMT, Loop Hybrid; hybrid containing the N-terminus and the loop region from CCoAOMT, Double Hybrid.

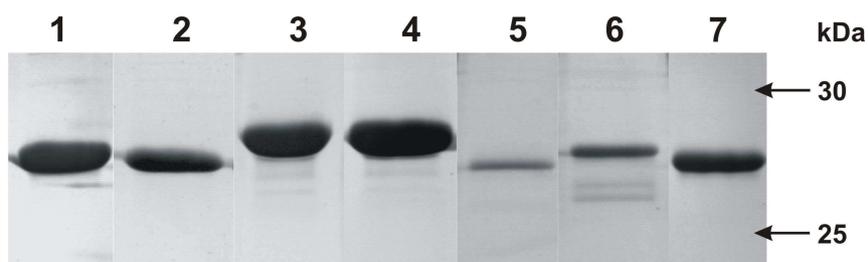


Figure 19. Coomassie stained SDS PAGE gel of purified hybrid proteins. The numbered lanes are: 1 PFOMT N0, 2 PFOMT N0Xa1 his tag cleaved, 3 N22, 4 N22Xa1 his tag cleaved, 5 Loop Hybrid, 6 Double Hybrid, 7 CCoAOMT.

The overexpression of the proteins of interest was confirmed by SDS PAGE. The plasmids from the expression strains were isolated and sequenced to confirm the correctness of the construct. All of the investigated proteins were purified, and their purity was assessed by SDS PAGE (Figure 19). Table 4 summarizes the yields of obtained proteins.

Table 4. Approximate yields of the protein production

Protein	Concentration mg/ml	Amount produced mg	Yield mg/l
1 N0	5.39	8.0	20.0
2 N0Xa1	4.75	8.3	20.8
3 N22	4.38	5.3	13.1
4 N22Xa1	15.1	7.6	18.9
5 Loop Hybrid	1.6	0.8	2
6 Double Hybrid	4	2	5
7 CCoAOMT	1.37	4.3	10.8

5.5.2. His tag removal

Because the N-terminus is the focus of the investigation the need to create proteins with a free N-terminus was seen. The removal of the N-terminal affinity tag would ensure that the observed kinetic parameters of the N-terminal hybrid proteins could be attributed to either the varying amino acid sequence or the presence of the his tag itself. Two approaches to overcome this problem were attempted. The first was shifting to another vector. The N0 and N22 were cloned into pQE60 and pQE70 vectors which provide the C-terminal his tag. The introduction of the stop codons at the end of the protein sequence would ensure that the vector

derived his tag was not expressed. The expression in these vectors was very poor, and the resultant proteins were insoluble. Expression in the pET vector yielded insoluble protein as well. Therefore, to obtain the proteins with free N-terminus the removal of the his tag by cleavage with site specific protease was used. The cleavage site for Factor Xa protease was introduced in front of the Met1 of the investigated proteins. After the successful cleavage with Factor Xa the proteins were analyzed by western blot and SDS PAGE (Figure 20) for the absence of the his tag by the use of his tag specific antibody.

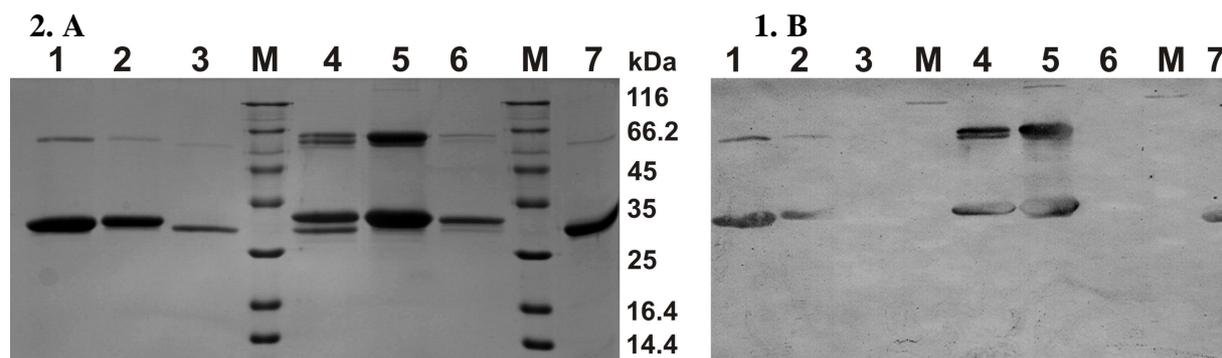


Figure 20. Protease Factor Xa His-tag cleavage reaction. **A** SDS PAGE gel, **B** western blot probed with anti his tag antibodies. The numbered lanes are: 1 PFOMT N0, 2. PFOMT NOXa1 not cleaved, 3 PFOMT NOXa1 his tag cleaved, 4 N22, 5 N22Xa1 not cleaved, 6 N22Xa1 his tag cleaved, 7. PFOMT N0 reference

Additional bands around 60 kDa correspond to the dimerized PFOMT. Their appearance is an artifact caused by which was not fully reduced by DTT present in the gel loading buffer and formed dimeric aggregates. Subsequent addition of β -mercaptoethanol as the reducing agent lowered the amount of aggregates present in the gel (data not shown). There are clear signals of anti his antibodies with the protein bands which contain a his tag while the lane with the cleaved protein shows no reaction (Figure 20).

It was observed that the lane where reference N22 protein was run there are double bands (lanes 4 and 5 Figure 20). This may be explained by the protein degradation or more likely, impurities present in the protein preparation. However, when the protein was prepared and purified by IMAC (See Materials and Methods section 1.9.2) a single homogenous band on SDS Page gel was present which corroborated the pure protein (Figure 19).

5.5.3. Activity assays

All obtained OMTs were tested for activity. The detection of the reaction products was achieved by RP-HPLC (Figure 21). The apparent K_m ($K_{m\text{ app}}$) and k_{cat} values were determined.

Values were compared with the ones obtained for full length recombinant PFOMT and *M. sativa* CCoAOMT.

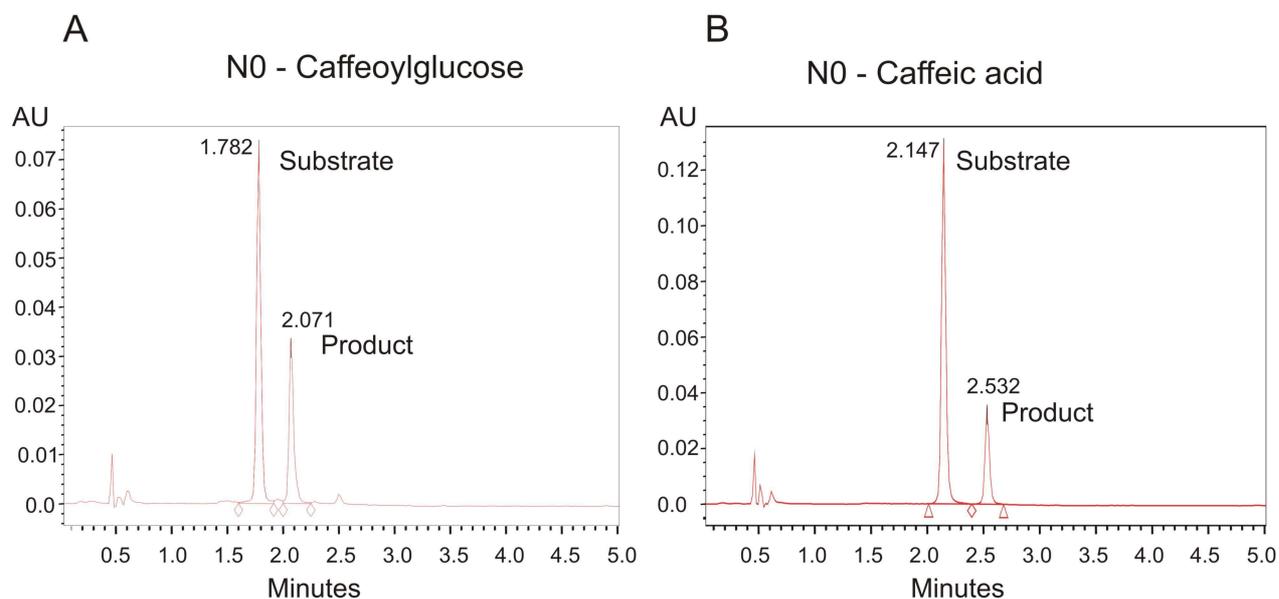


Figure 21. Sample chromatograms of the enzymatic reaction of N0. **A** reaction with caffeoylglucose forming the product feruloylglucose and **B** caffeic acid forming ferulic acid.

Table 5 shows the $K_{m\ app}$ values obtained for the hybrid proteins. The experimental value of the reaction velocity was determined for different substrate concentrations. For each data point three independent reactions were carried out.

Table 5. The $K_{m\ app}$ values for investigated proteins with the error of the curve fitting result.

Protein	Substrate $K_{m\ app}$ [μ M]				
	Quercetin	Quercetagenin	Caffeic acid	CCoA	Caffeoylglucose
N0	0.78 \pm 0.26	0.51 \pm 0.01	0.99 \pm 0.07	2.86 \pm 0.54	2.02 \pm 0.15
N0Xa1 CL	1.00 \pm 0.07	1.76 \pm 0.22	8.61 \pm 0.41	6.06 \pm 1.97	1.72 \pm 0.13
N22	3.93 \pm 0.42	1.57 \pm 0.39	112 \pm 10	103 \pm 27	48,2 \pm 8.8
N22Xa1 CL	4.74 \pm 0.59	1.56 \pm 0.62	>1000	>1000	49.1 \pm 1.6
Loop	5.07 \pm 0.50	0.96 \pm 0.25	2.81 \pm 0.12	0.35 \pm 0.10	2.36 \pm 0.424
Double22	6.97 \pm 0.49	2.02 \pm 0.24	3.82 \pm 0.93	0.62 \pm 0.15	3.07 \pm 0.394
MedCCoAOMT	5.91 \pm 0.71	3.63 \pm 0.55	56.2 \pm 13.1	0.31 \pm 0.09	76.7 \pm 11.2

The obtained turnover values were averaged over there measurements and then used to determine the amount of product formed. Taking into account the time of the reaction the velocity was calculated. The $K_{m\ app}$ and V_{max} (maximum velocity of the reaction) values were obtained by nonlinear curve fitting assuming Michaelis-Menten steady state kinetics (Figure 22).

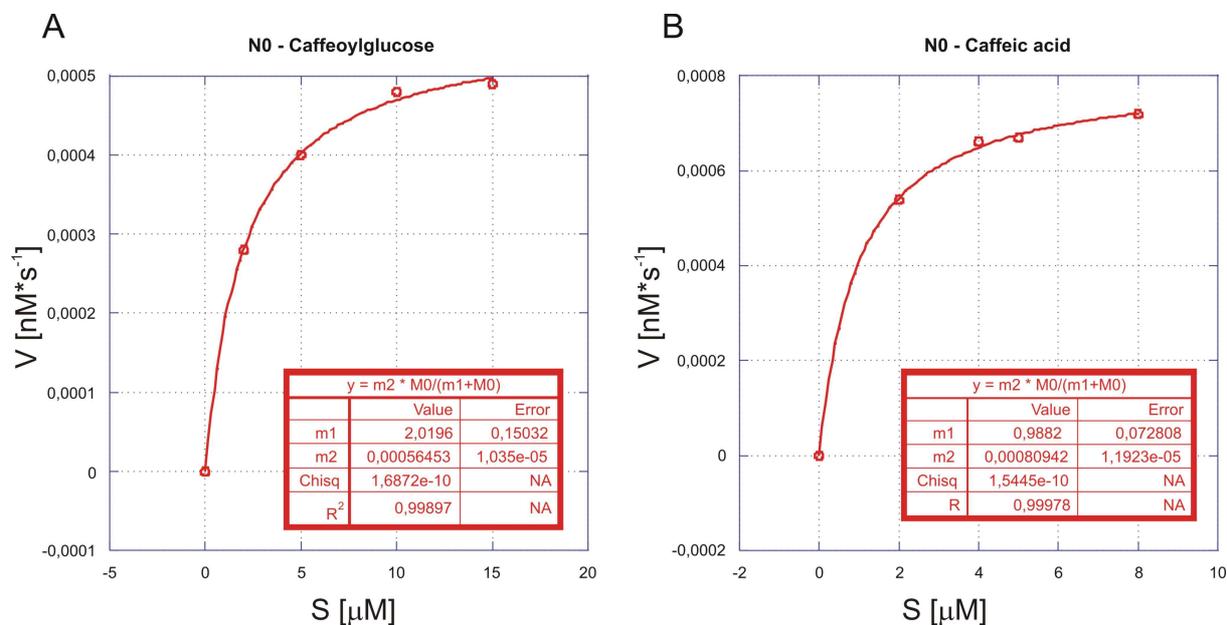


Figure 22. The charts show the results of the curve fitting for N0 and two substrates **A** caffeoylglucose and **B** caffeic acid. Nonlinear curve fitting was used to determine the kinetic parameters of the investigated proteins.

The $K_{m\ app}$ values for PFOMT N0 protein towards all substrates are quite low. This protein shows high affinity towards all substrates with the CCoA and caffeoylglucose having the lowest affinity. PFOMT N0Xa1 with the removed N-terminal his tag shows comparable results. The $K_{m\ app}$ value for caffeic acid and CCoA are the ones where large discrepancies can be observed. Removal of the his tag lowers the affinity of the protein towards almost all of the substrates with the exception of caffeoylglucose. This effect is most pronounced for caffeic acid and CCoA. Replacement of the first 22 N-terminal amino acid residues does not affect the affinity of protein towards quercetin and quercetagenin to a very large extent. This is observed regardless of the his tag. The $K_{m\ app}$ values for caffeic acid as well as its CoA ester and glucose derivatives are very high in comparison to the original PFOMT protein which except for the decreased affinity for CCoA makes them similar to *M. sativa* CCoAOMT. The replacement of the N-terminus lowered the affinity for caffeic acid and caffeoylglucose which makes those proteins similar to CCoAOMT. In case of N22 with the N-terminal his tag removed those values could not be determined. The curves for those calculations are almost

straight lines. The presence of the his tag causes a different kinetic behavior of the proteins. The Loop and Double hybrids show large increase in affinity towards caffeic acid, which makes them similar to the *M. sativa* CCoAOMT. However, unlike *M. sativa* CCoAOMT those enzymes still show quite high affinity towards caffeic acid and caffeoylglucose. The $K_{m\text{ app}}$ values for the hybrid proteins towards quercetin and quercetagenin do not show high variations and are somewhat higher than from PFOMT and PFOMT where the his-tag was cleaved off. However, in this case the values for mutations are within the same order of magnitude.

Table 6. The $k_{\text{cat}}/K_{m\text{ app}}$ values for investigated proteins

Substrate $k_{\text{cat}}/K_{m\text{ app}}$ [$\text{s}^{-1}\text{M}^{-1}$]					
Protein:	Quercetin	Quercetagenin	Caffeic acid	CCoA	Caffeoylglucose
N0	23,600 ± 7900	22,000 ± 400	22,900 ± 1700	12,400 ± 2400	15,700 ± 1200
N0Xa1 CL	21,400 ± 1700	17,200 ± 2300	2310 ± 120	2130 ± 740	16,400 ± 1300
N22	2880 ± 320	8900 ± 2300	93.2 ± 11.5	105 ± 36	233 ± 54
N22Xa1 CL	1920 ± 250	14,800 ± 6000	n.d.	n.d.	846 ± 34
Loop	727 ± 77	1510 ± 400	67.9 ± 3.1	27,900 ± 7800	154 ± 26
Double22	328 ± 25	584 ± 71	20.4 ± 5.2	8390 ± 2020	48.6 ± 6.5
MedCCoAOMT	10,700 ± 1400	8680 ± 1420	8.22 ± 2.44	200,000 $\pm 60,800$	289 ± 55

The $K_{\text{cat}}/K_{m\text{ app}}$ values represent the catalytic efficiency of the proteins (Table 6). The his tagged and non-tagged PFOMT show the highest values for quercetin and quercetagenin. The influence of the his tag is most pronounced for caffeic acid and CCoA. Replacement of the N-terminal residues lowers the catalytic efficiency towards all substrates by at least one order of magnitude. The loop hybrid despite low turnovers shows the highest catalytic efficiency towards CCoA. This fact makes it similar in catalytic properties to *M. sativa* CCoAOMT. The low efficiency of this enzyme towards caffeic acid is still higher than that of the substrate specific *M. sativa* CCoAOMT but two orders of magnitude lower than that of PFOMT. The replacement of both N-terminus and loop sequence does not produce such profound differences. However, the differences in the catalytic efficiency between the different substrates are of similar levels compared to the *M. sativa* CCoAOMT. The overall efficiency of the Double22 hybrid is usually one tenth and less than that of the *M. sativa* CCoAOMT.

5.6. The structure and characterization of SynOMT

5.6.1. Purification of SynOMT

SynOMT is a Mg^{2+} -dependent OMT from the cyanobacterium *Synechocystis* sp. strain PCC 6803 and it is coded by the *slr0095* gene. The cDNA coding for SynOMT was amplified by PCR reaction with the primers that introduced the restriction sites used for cloning (see details in (Kopycki et al., 2008b)). The gene was cloned into pQE30 vector which provides N-terminal his tag that aids in the purification and M15 pREP4 cells were used for protein expression. One step of IMAC over a Talon resin was sufficient to obtain the protein of sufficient purity. The yield of the soluble protein was approximately 1 mg from a 1 l of bacterial culture. The cultures were always inoculated with freshly transformed cells, as deterioration of protein expression was observed upon repeated re-growing of frozen bacterial stocks. This instability of the protein expression is thought to be caused by a mild toxicity effect of specific methylation carried out by SynOMT in *E. coli* cells. When a prokaryotic enzyme is expressed as a soluble protein in *E. coli*, undesired methylation reaction may take place within the bacterial cells.

After purification the protein was concentrated to 50 mM Tris 10% glycerol buffer pH 7.5. Figure 23 shows the course of recombinant purification of SynOMT. This protein preparation was used for crystallization and activity assays.

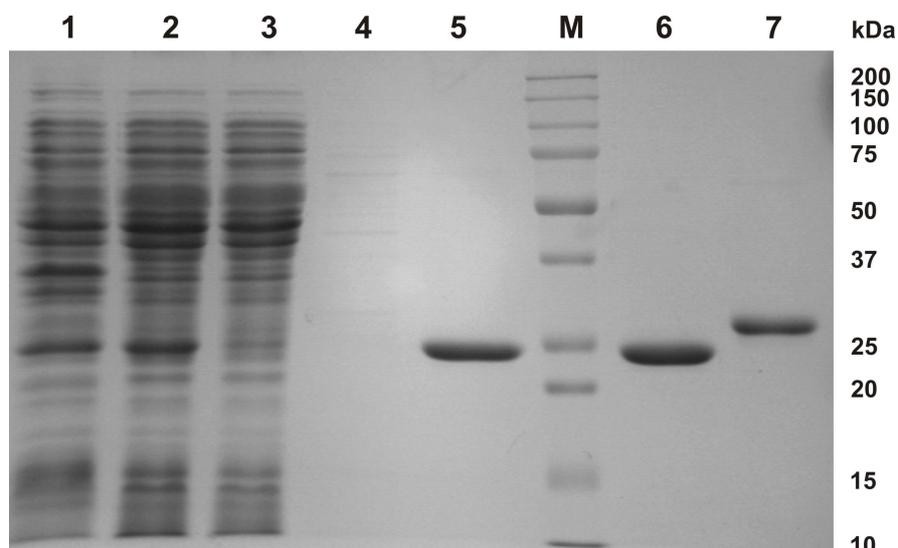


Figure 23. The IMAC purification of SynOMT. The numbered lanes are: 1 Cell Lysate, 2 soluble protein, 3 flow through, 4 30 mM imidazole wash, 5 240 mM imidazole elution, 6 Rebuffered and concentrated preparation, 7 PFOMT NO reference. The marker lane is annotated M.

5.6.2. Identification of the *para* methylation carried out by SynOMT

Like the plant enzymes SynOMT is capable of methylating the substrates with vicinal dihydroxy aromatic structures. This is consistent with the substrate specificity profile for the related plant and animal enzymes. According to sequence homology this enzyme belongs to Mg^{2+} dependent OMTs, so the presence of a magnesium ion as a prerequisite for enzymatic activity is expected. This dependence was tested by removing any bivalent cations from the reaction environment by the addition of EDTA. Several different bivalent cations were also tested to check for influence in the activity of the enzyme.

The reaction products of SynOMT with trihydroxy cinnamic acid were identified by co-chromatography with authentic standards and comparison of the UV spectra generated by diode-array UV-detection system (Figure 24).

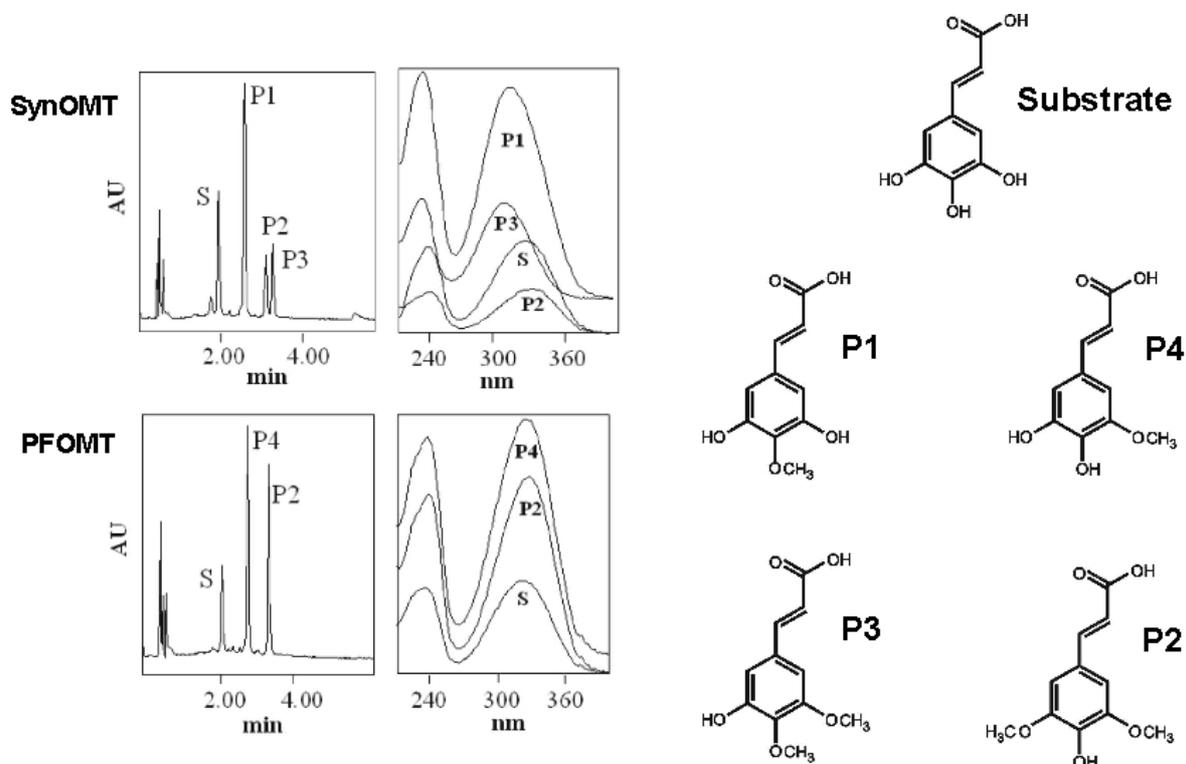


Figure 24. The chromatograms and UV spectra of SynOMT and PFOMT reaction mixtures with trihydroxy cinnamic acid. S, substrate 3,4,5-trihydroxycinnamic acid; P1, 3,5-dihydroxy-4-methoxycinnamic acid; P2, sinapic acid; P3, 5-hydroxy-3,4-dimethoxycinnamic acid; P4, 5-hydroxyferulic acid.

There were two reaction products obtained in the reaction with 5-hydroxyferulic acid. The first one showed spectral similarities to sinapic acid while the spectral properties of the second one bore no resemblance to known methylation products of plant enzymes. The reaction with 3,4,5-trihydroxycinnamic acid produced three products, one of which could be identified as sinapic acid (P2) and the other two, including the predominant one (P1). Neither

the ferulic acid nor sinapic acid, due to the lack of neighboring hydroxyl group, can not be a substrate for further methylation. This fact led to a conclusion, that the methylation could be directed into a different hydroxyl group than in the case of plant OMTs. Figure 25 depicts the methylation reactions of trihydroxy cinnamic acid carried out by SynOMT contrasted to that of PFOMT.

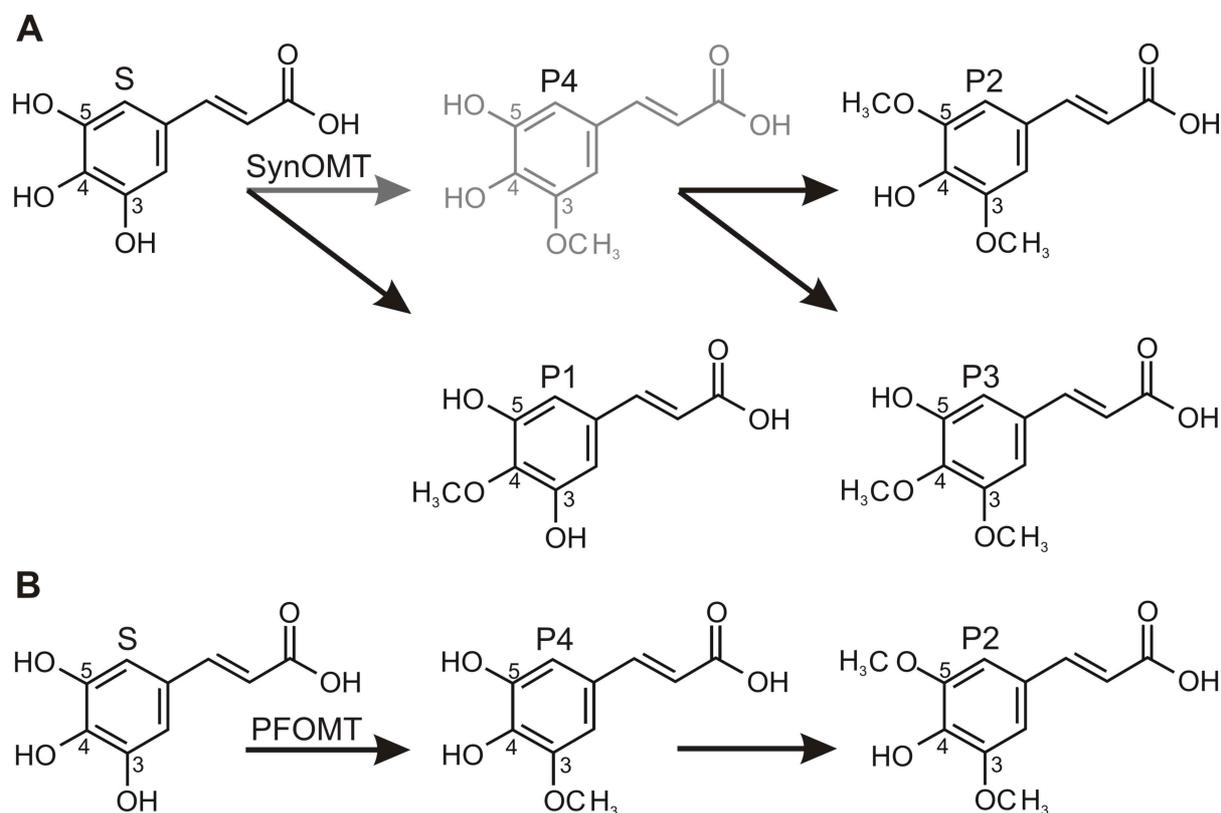


Figure 25. Reaction scheme contrasting the position specificity of methylation carried by **A**, SynOMT from *Synechocystis* sp. strain PCC6803 and **B**, PFOMT from *M. crystallinum*. S, substrate 3,4,5-trihydroxycinnamic acid; P1, 3,5-dihydroxy-4-methoxycinnamic acid; P2, sinapic acid; P3, 5-hydroxy-3,4-dimethoxycinnamic acid; P4, 5-hydroxyferulic acid. In the case of SynOMT, P4, the precursor for P2 and P3, could not be detected by HPLC and is shown in gray.

The reaction prominent products of SynOMT with 3,4,5-trihydroxycinnamic acid were isolated, purified by HPLC and analysed. by ^1H and ^{13}C NMR (Table 7) and ESI-MS as well as GC/TOFMS (Table 8). The NMR analysis and interpretation of the results were performed by Dr. A. Porzel and the mass spectrometric analyses by Dr. W. Schliemann and Dr. J. Schmidt. One- and two-dimensional NMR experiments resulted in an unambiguous assignment of all ^1H and ^{13}C NMR signals of 3,5-dihydroxy-4-methoxycinnamic acid (with the exception of the ^{13}C signal of $-\text{COOH}$, for sensitivity reasons). Only a 4-*O*-methylation is in accordance with the isochronous NMR signals of positions 2/6 and 3/5. Furthermore, the

para position of the *O*-methyl group is shown by the high-field shift of C-4 ($\delta^{13}\text{C}$ 139.0 ppm), caused by the two electron releasing hydroxyl substituents in *ortho*-position to C-4. In addition, no NOE enhancement is observed between the signal of the methoxy group and the aromatic protons.

These data are consistent with an unusual preference of this enzyme for the *para*-position, leading to the formation of 3,5-dihydroxy-4-methoxycinnamic acid (P1 in Figure 24). The simultaneous appearance of two dimethylated compounds in addition to 3,5-dihydroxy-4-methoxycinnamic acid in the HPLC analysis was confirmed by GC-MS data (Table 8; Figure 24). Those analyzes concluded unequivocally that the predominant reaction product of SynOMT and trihydroxy cinnamic acid was a 4-*O*-methyl derivative. The 3,5-dihydroxy-4-methoxycinnamic acid is a predominant reaction product which suggests the preference for methylation in this position. One of the two remaining reaction products had the same spectral properties and retention times in HPLC as well as GC- 4-hydroxy-3,5-dimethoxycinnamic acid (sinapic acid). The second product (P3) with very similar spectral properties to the *para*-methylated 3,4,5 trihydroxycinnamic acid showed different LC and GC retention times when compared to sinapic acid and, corroborated by GC-MS data was thus identified as 5-hydroxy-3,4-dimethoxycinnamic acid. Those findings were consistent with the observation of reaction products of the reaction with 5-hydroxyferulic acid. Consequently, incubation of SynOMT with 5-hydroxyferulic acid, resulted in the formation of 5-hydroxy 3,4 dimethoxycinnamic acid as well as sinapic acid as expected.

Table 7. The NMR data of 3,5-dihydroxy-4-methoxycinnamic acid.

Pos.	$\sigma^1\text{H}$ [ppm], multiplicity (J [Hz]), relative integral intensity	$\sigma^{13}\text{C}^{\text{a}}$ [ppm]
1	---	132.4
2/6	6.588, <i>s</i> , 2H	108.6
3/5	---	152.1
4	---	139.0
4-OCH ₃	3.824, <i>s</i> , 3H	60.8
7	7.448, <i>d</i> (15.9), 1H	146.5
8	6.246, <i>d</i> (15.9), 1H	118.0
9	---	n.d.
^a chemical shifts of HSQC and HMBC correlation peaks n.d. not determined		

Table 8. GC/TOFMS data of derivatized hydroxycinnamic acid standards and products of 3,4,5-trihydroxycinnamic acid methylation catalyzed by SynOMT.

Identifier	R _t (min)		m/z found	m/z calculated for	Compound
	<i>cis</i>	<i>trans</i>			
1117	30.67	32.94	396.1589	C ₁₈ H ₃₂ O ₄ Si ₃ 396.1608	Caffeic acid-3TMS
1118	29.67	32.32	338.1372	C ₁₆ H ₂₆ O ₄ Si ₂ 338.1370	Ferulic acid-2TMS
1121	33.07	34.92	484.1994	C ₂₁ H ₄₀ O ₅ Si ₄ 484.1953	3,4,5-Trihydroxycinnamic acid- 4TMS
1119	32.65	34.92	426.1749	C ₁₉ H ₃₄ O ₅ Si ₃ 426.1714	5-Hydroxyferulic acid-3TMS
1120	31.92	34.54	368.1492	C ₁₇ H ₂₈ O ₅ Si ₂ 368.1475	Sinapic acid-2TMS
1129	31.57	33.70	426.1741	C ₁₉ H ₃₄ O ₅ Si ₃ 426.1714	3,5-Dihydroxy-4-methoxy- cinnamic acid-3TMS
1130	31.20	33.57	368.1347	C ₁₇ H ₂₈ O ₅ Si ₂ 368.1475	5-Hydroxy-3,4-dimethoxy- cinnamic acid-2TMS

The appearance of the dimethylated product shows that the discrimination of the position for methylation is not strict and the enzyme is capable of methylation in position 3 of the aromatic ring as well. The absence of other monomethylated derivatives in the position 3 can be explained by the subsequent methylation reaction. When a methyl group is introduced to an hydroxyl moiety in position 3 of the phenyl ring, the two remaining hydroxyl groups are neighbouring. Such a situation allows for a subsequent methylation of one of the remaining OH groups, this phenomenon was also observed for PFOMT from *M. crystallinum*. In case of SynOMT two double methylated reaction products were found. Those findings were consistent with the observation of the products of the reaction with 5-hydroxyferulic acid. The identified sinapic acid is methylated in positions 3 and 5 and the second product was determined to be the 3,4 *O*-methyl derivative.

5.6.3. Kinetic characterization of SynOMT

Due to the expected similarity to plant enzymes, the following substrates were tested: caffeic acid, protocatechuic acid (3,4 dihydroxybenzoic acid), CCoA, quercetin, tricetin (this compound was a kind gift from Ragai Ibrahim Department of Biology, Concordia University, Montreal, Canada) and 5-hydroxyferulic acid, an artificial substrate, 3,4,5-trihydroxy cinnamic acid was also tested and kinetic parameters determined (Table 9). The low values of kinetic constants maybe the result of the fact that none of the tested substances is the natural substrate for SynOMT.

Table 9. The kinetic parameters of SynOMT

Substrates	$K_{m \text{ app}}$ [μM]	k_{cat} [1/s]	$k_{\text{cat}}/K_{m \text{ app}}$ [1/(s * μM)]
5-Hydroxyferulic acid	74.0	0.02500	0.00034
Caffeic acid	69.3	0.00551	0.00008
CCoA	32.9	0.02450	0.00074
Caffeoylglucose	106	0.00951	0.00009
3,4,5-Trihydroxycinnamic acid	20.7	0.01040	0.00050
Tricetin (3',4',5,5',7-Pentahydroxyflavone)	5.00	0.00885	0.00178
Protocatechuic acid	215	0.00802	0,00004

The apparent K_m and k_{cat} values were calculated from Lineweaver-Burk plots. The OMT from cyanobacteria shows the highest affinity to tricetin and caffeoylglucose. However the two substrates that have the fastest turnover rates are 5-hydroxyferulic acid acid and CCoA. The turnover rate of tricetin is not the highest. The k_{cat}/K_m ratio for tricetin is the highest of all of the substrates tested. It is around two to five times larger than that of the rapidly converted substrates as 5-hydroxyferulic acid and CCoA. One should however keep in mind that all of those substrates were artificial ones and the real *in vivo* substrate is not yet known. The flavone, tricetin seems to be well accepted and it could be speculated that the real *in vivo* substrate may have similar structure or at least polyhydroxylated phenyl moiety.

5.6.4. Crystallization and structure determination of SynOMT

The findings concerning the unusual position preference for methylation carried out by SynOMT lead to questions about the structural requirements. With this goal in mind efforts were undertaken to obtain suitable crystals of SynOMT. Purified recombinant SynOMT was set up for crystallisation with a standard set of crystallisation buffers. The ligands of the protein were included in the protein solution. Crystals were obtained from 0.2 M magnesium chloride hexahydrate, 0.1 M Tris hydrochloride pH 8.5, 30 %w/v polyethylene glycol 4000 with the addition of AdoMet and caffeic acid.

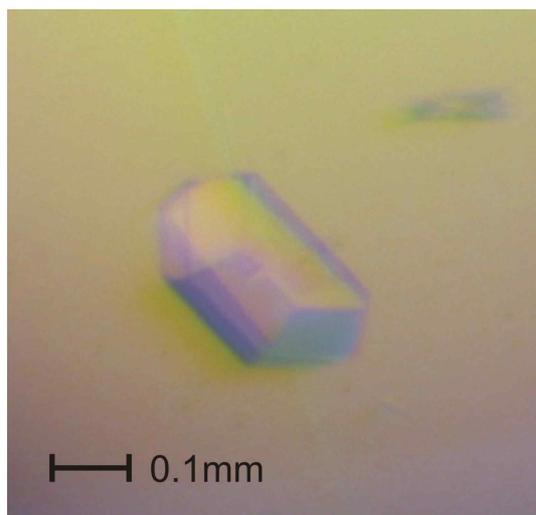


Figure 26. The crystal of SynOMT for which a data set was collected.

The best diffracting crystal (Figure 26), although quite small, yielded a data set of resolution up to 2.0 Å. This data set was collected on Raxis IV ++ with image plate detector. The P3₁21 space group with the cell constants of a=57.622 b=57.622 c=119.834 was determined. Table 10 summarizes the data collection and structure refinement parameters. The structure was solved using molecular replacement approach with the structure of PFOMT (PDB code 3C3Y) was used as search model. The refinement of the structure did not give a satisfactory results. Twinning of the crystal was detected during the re-evaluation of the diffraction data, and the twinned refinement provided by CNS program was finally used to refine the structure.

Table 10. Crystallographic data, phasing and refinement statistics for SynOMT (PDB code 3CBG)

Data Collection	
Wavelength Å (≈)	1.5418
Resolution Å (≈)	30-2.0
Total reflections	140633
Unique reflections	16207
Completeness ^a (%)	100 (99.9)
$I/\sigma(I)$	17.4 (2.7)
$R_{\text{sym}}^{\text{a,b}}$	0.124 (0.703)
Redundancy	8.7 (6.9)
Refinement	
Refinement reflections	16164
Twinned R^{c}	0.1519 (0.273)
Twinned $R_{\text{free}}^{\text{d}}$	0.2154 (0.308)
Protein atoms	1703
Ligand atoms	55
Water molecules	124
rmsd bond lengths ^e (Å)	0.006258
rmsd bond angles ^e (°)	1.23469
<B-factor> protein (Å ²)	28.93
Twinning fraction	0.284
Twinning operator	-h,-k, l
Ramachandran plot (%)	
Most favored regions	91.9
Additional allowed regions	7.0
Generously allowed regions	1.1

^a The values in parentheses represent the values for the highest resolution shell.

^b $R_{\text{sym}} = \sum |I_h - \langle I_h \rangle| / \sum I_h$, where $\langle I_h \rangle$ is the average intensity over symmetry-equivalent reflections.

^c $R = \sum ||F_{\text{obs}}| - |F_{\text{calc}}|| / \sum |F_{\text{obs}}|$

^d R_{free} is calculated as R , with 9.2 % of the data (1485 reflections) excluded from the refinement

^e The rmsd for bonds and angles are the rmsd from ideal values

5.6.5. Structural details

SynOMT exists as soluble dimer (Figure 27), but only one protein molecule was found in the asymmetric unit. The reconstruction of the functional assembly is possible by applying the crystal symmetry operations to the structure. The protein consists of 220 amino acid residues, with an additional 12 N-terminal vector derived residues containing the His-tag. The first residue visible in the electron density is Lys3. The residues comprising N-terminal his tag and first two residues of the protein are not defined, presumably due to the high flexibility of the

N-terminus. The number of the N-terminal residues for which the electron density is observed is still much higher in comparison with PFOMT or CCoAOMT structures. The structure contains the methyl group donor AdoMet in its demethylated form (AdoHcy) as well as two possible reaction products, ferulic acid and isoferulic acid. The crystallization solution included caffeic acid, which was methylated during the crystallization to form ferulic and isoferulic acid. Electron density for the Mg^{2+} ion is well defined and clearly interpretable. The electron density for the methyl derivative of caffeic acid shows that there are two possible reaction products bound to the active site, the ferulic acid and the isoferulic acid. Both possible molecules of the reaction products were placed in the same place and given 0.5 occupancy each. The spatial constraints setup during the refinement cycles were ignored for this pair of molecules. In addition 259 water molecules were found in the crystal structure.

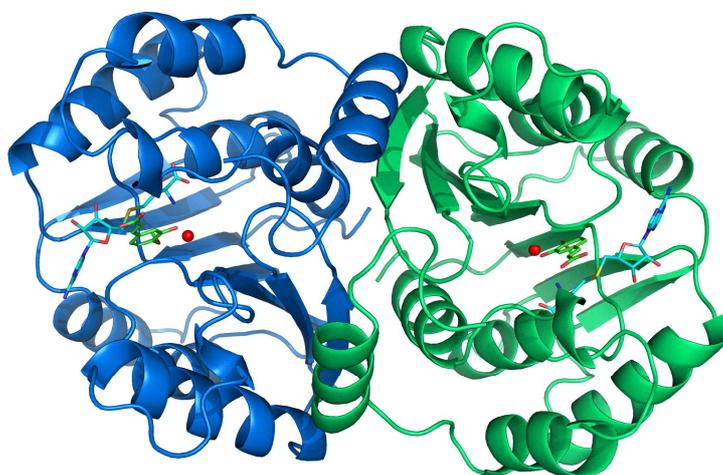


Figure 27. Dimeric structure of SynOMT. The monomers are shown in blue and green. The red spheres are the Mg^{2+} cations bound in the active sites of each of the protein molecules. Cyan sticks represent the AdoHcy and while the yellow ones one orientation of the bound substrate.

5.6.6. Topology of the SynOMT

The overall fold of the protein is almost identical to the one described for PFOMT, an α/β Rossmann fold, which is shared by all structurally characterized AdoMet dependent OMTs (Schluckebier et al., 1995) (Figure 28). The core of the protein is composed of seven β -sheet strands out of which six of them are parallel to each other and one which is in anti parallel orientation. The β -sheet structure is surrounded by eight α -helices, three on one side of the β -sheets and remaining three on the other.

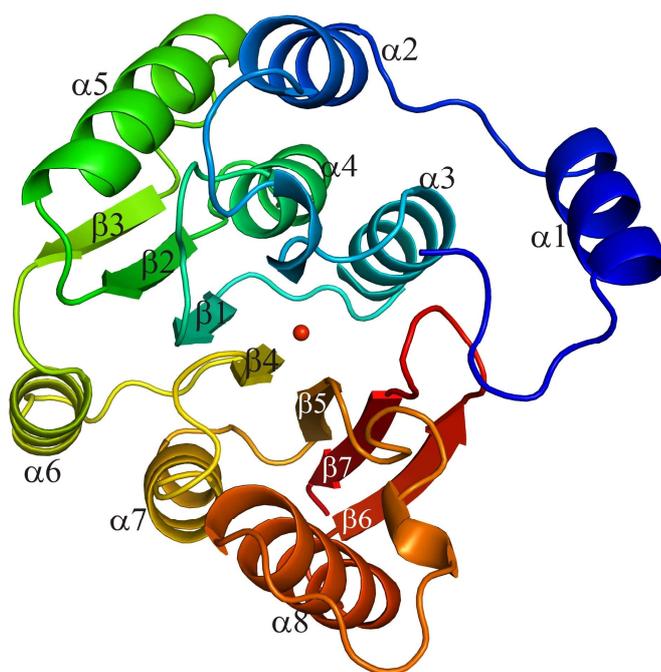


Figure 28. Topology of SynOMT in spatial orientation. The red sphere represents the Mg²⁺ ion bound in the active site of the protein.

The first view reveals the structural similarity among the plant OMTs and SynOMT. The algal enzyme shares the same fold as described for PFOMT and CCoAOMT

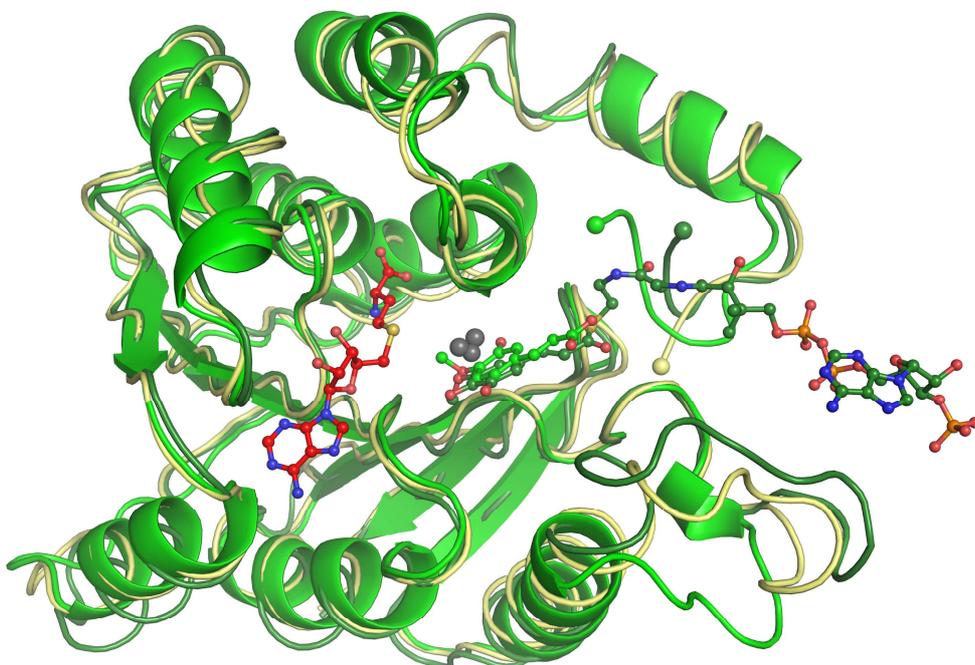


Figure 29. Superposition of structures of plant OMTs and the microbial OMT. In yellow PFOMT, dark green CCoAOMT, light green SynOMT.

5.6.7. Binding sites

5.6.7.1. Ion binding site of SynOMT

As in other cases of Mg^{2+} dependent OMTs a metal ion is present in the structure. Several polar residues take part in the binding of the metal ion. Those interactions include the hydrogen bonds formed between the metal ion and the side chains of the residues: Asp143, Asp169, Asn170 (Figure 30). An additional bonds are likely to form between the water molecule present near the active site and the backbone carbonyl oxygen of Met 42 and nitrogen atom of Ile44 (not shown). A similar water molecule is found in the PFOMT structure. In case of PFOMT Thr54 and for CCoAOMT Thr63 corresponds to Ile44, and although the side chain occupies a different position the backbone nitrogen of this amino acid is still able to take part in the binding of the mentioned water molecule. In Case of CCoAOMT the low resolution did not allow for identification of water molecules coordinated to the bivalent cation.

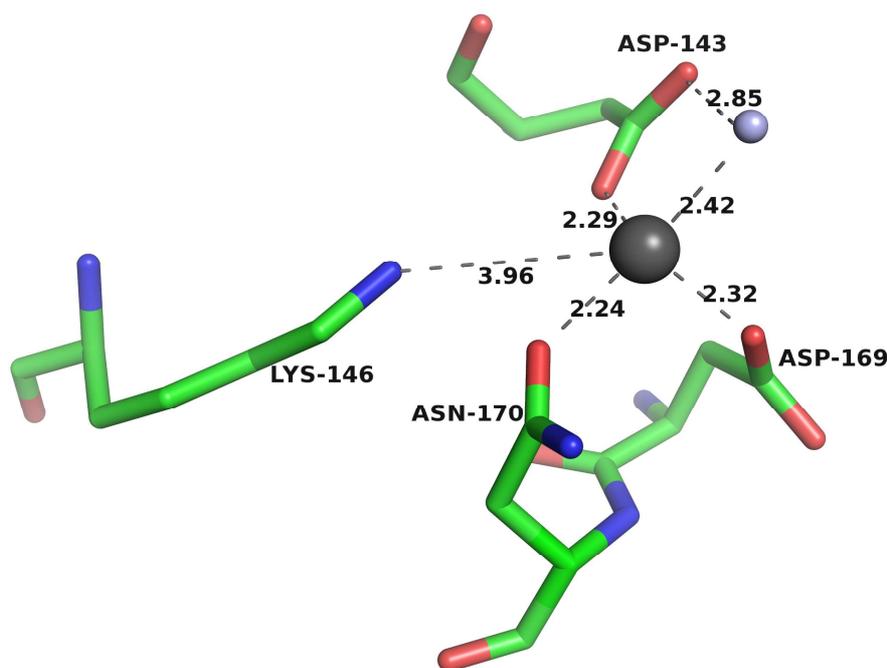


Figure 30. Representation of SynOMT cation binding site. The gray sphere represents the Mg^{2+} ion bound to the protein. The blue sphere stands for a well ordered water molecule interacting with the cation. The distances in Å from catalytically relevant amino acid residues are shown with gray dashed lines.

The superposition of the three OMT structures reveals that the metal ions are bound in very similar position in all three structures. In case of the other two OMT structures the identical residues take part in the binding of the ion. The conservation of the residues that bind the

bivalent cation results from the absolute necessity of an ion for the activity of those enzymes. Table 11 summarizes the residues interacting with the cation in three related structures.

Table 11. Cation interacting residues

CCoAOMT			PFOMT			SynOMT		
Residue	atom	distance Å	Residue	atom	distance Å	Residue	atom	distance Å
Met 61	O	4.24	Met 52	O	4.03	Lys 3	NZ	4.11
Thr 63	OG1	4.29	Thr 54	OG1	4.68	Met 42	O	3.77
Glu 67	OE1	4.80						
Asp 163	OD1	2.74	Asp 154	OD1	2.36	Asp 143	OD1	2.29
Asp 163	OD2	2.89	Asp 154	OD2	2.61	Asp 143	OD2	3.40
			Lys 157	NZ	4.08	Lys 146	NZ	3.96
Asp 189	OD1	4.34	Asp 180	OD1	4.45	Asp 169	OD1	4.44
Asp 189	OD2	2.27	Asp 180	OD2	2.36	Asp 169	OD2	2.32
Asn 190	N	4.79	Asn 181	N	4.87	Asn 170	N	4.65
Asn 190	OD1	2.61	Asn 181	OD1	2.39	Asn 170	OD1	2.24
Asn 190	ND2	3.22	Asn 181	ND2	3.76	Asn 170	ND2	3.30

Close contacts for cation interacting residues binding identified by CONTACT (part of CCP4 suite)

5.6.7.2. Substrate binding site

The structure shows a bound substrate in the active site of SynOMT (Figure 31). The hydroxyl oxygens of caffeic acid derivative form hydrogen bonds with the magnesium ion. The unmethylated hydroxyl oxygen atom in position 3 of the phenyl ring is bound by hydrogen bonds with Asn170 and Lys3. The His174 is positioned in such a way that it interacts with the carboxy function of the substrate molecule. The residue Lys146 is directly involved in the methylation reaction. This is the catalytic lysine residue that deprotonates the hydroxyl oxygen in preparation for the methyl group transfer from AdoMet donor. The action of such catalytic lysine residues was described in the reaction mechanism for other cation dependent methyltransferases (Ibrahim et al., 1998). In addition to that the hydrophobic ring of the substrate molecule forms hydrophobic interactions with following residues: Met42, Trp173, Asp143, and the methyl group donor AdoMet. The hydrophobic “gating” residues such as Met42 and Trp173, have been described for other cation dependent plant OMTs to form an important part of the substrate binding cavity .

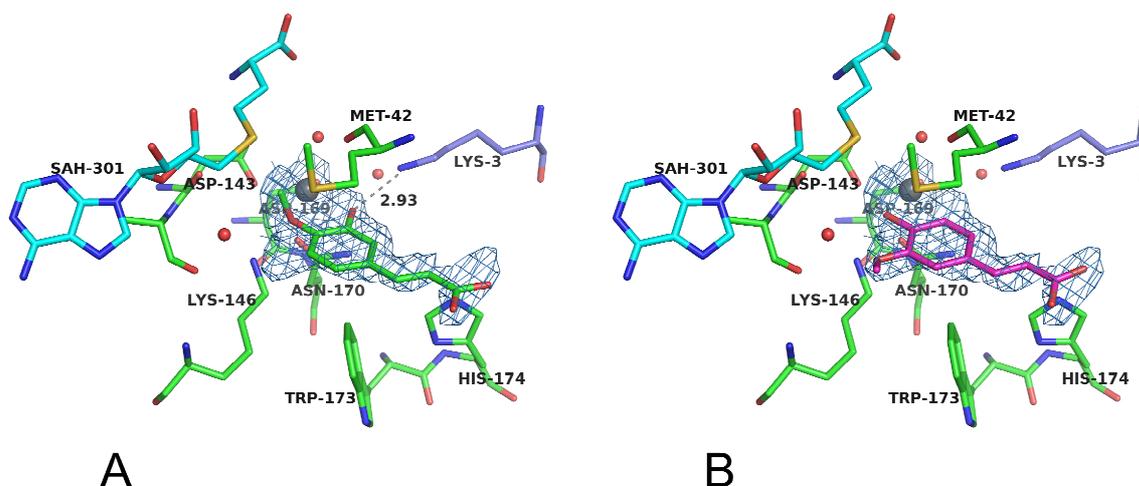


Figure 31. Binding of two reaction products the active site of SynOMT. The figure A represents binding of 4OMe-caffeic acid and B, 3OMe-caffeic acid. The catalytically important residue Lys 3 is shown in blue. For 3OMe-caffeic acid an “unproductive” binding is observed. The reaction product was found in such orientation that the methylated hydroxyl group was away from the cation.

The experimental evidence shows that methylation of caffeic acid catalyzed by SynOMT leads to formation of two reaction products 3OMe-caffeic acid (ferulic acid) and 4OMe-caffeic acid (isoferulic acid). The observed electronic density further confirms this observation. Ferulic acid can be placed into the cloud of electron density in two orientations (Figure 31). Moreover, the additional density is observed at an angle to hydroxyl group in position 4 clearly showing the presence of an additional methyl group. It should be noted, however, that the experimental evidence shows formation of only monomethylated products, no dimethylated product of caffeic acid was observed. This fact further supports the interpretation of the density for the substrate as two superimposed possible orientations. Both hydroxyl oxygens of caffeic acid form hydrogen bonds with magnesium ion. This observation is consistent with the proposed mechanism of the reaction for this group of *O*-methyltransferases. The interaction with the bivalent ion is necessary to facilitate the ionization of hydroxyl group which is to be methylated. The position of the ferulic acid is furthermore stabilized by interactions of hydroxyl oxygen atoms with Lys146 and Asn170 (Figure 32). The interactions observed for Lys146 of SynOMT are not present in the structure of *M. sativa* CCoAOMT even though the corresponding amino acid Lys166 occupies similar spatial position. In the structure of CCoAOMT this amino acid is too far away to form hydrogen bonds with the hydroxyl oxygens of 5-hydroxyferuoylCoA. The position of

substrate/product in the active site of SynOMT is additionally stabilized by interaction occurring between the carboxyl function of the substrate molecule and His 174. The closest nitrogen atom of the histidine residue is located 3.25Å away from one of the oxygen atoms of the carboxyl group. This distance and the angle at which both atoms are placed excludes the possibility of formation of hydrogen bonds. Those atoms may however interact through electrostatic integrations, that furthermore stabilize the substrate molecule in the orientation favoring the methylation in the *para* position of the benzene ring. The corresponding amino acid in the structure of PFOMT is Gly 185 and in the case of CCoAOMT is Asn 194. The absence of the charged residues in the structure of PFOMT may further explain the substrate promiscuity of this protein. There are less interaction between the substrate and the protein that are immobilizing the substrate in the proper orientation. The phenolic moiety of the substrate/product molecule fits in the hydrophobic pocket comprised of Trp 173 and Met 42. In case of CCoAOMT this hydrophobic pocket has two more components Tyr 208 and Tyr 212 besides the corresponding methionine and tryptophan. The absence of those bulky residues in the structure of SynOMT may explain the ability of the protein to produce two methylated derivatives of caffeic acid. Without those steric constraints the substrate molecule may have more freedom in the active site and be methylated in different positions of the phenyl ring.

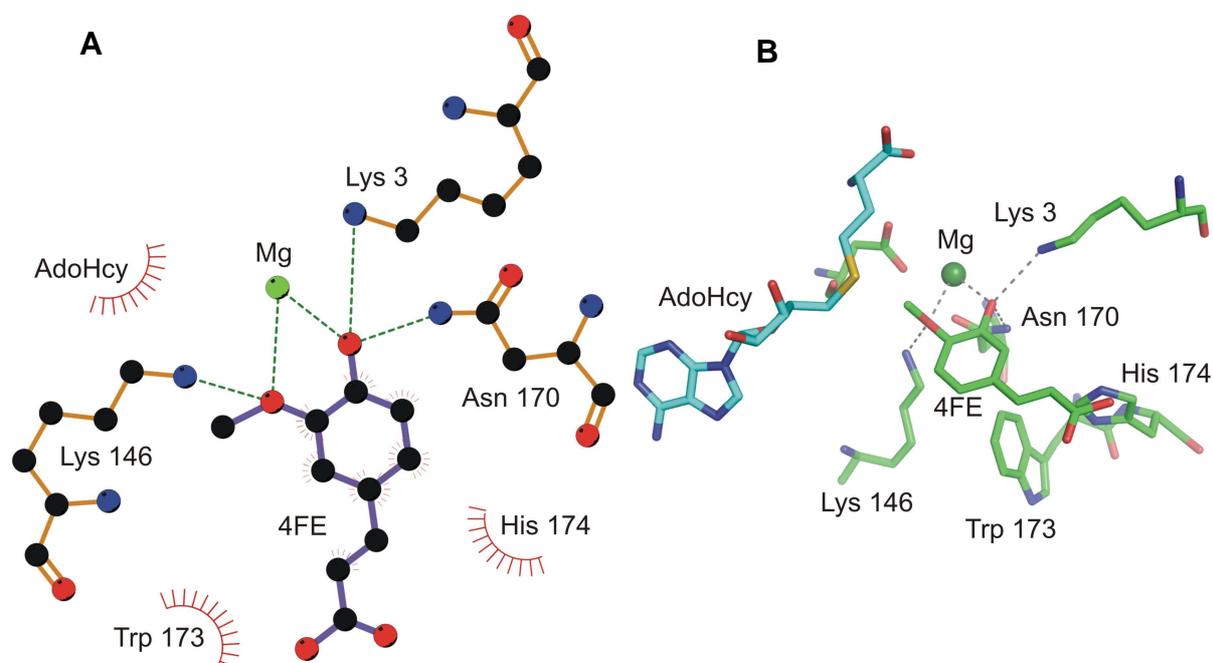


Figure 32. Interactions of substrate bound the active site of SynOMT. The schematic plot of the interactions A. The structural view B. The dashed lines represent most important electrostatic interactions. The hydrophobic interactions are shown with arches and short lines.

Ferulic acid was modeled into the active site of the of SynOMT in a position that allows for methylation in the meta position of the phenyl ring. The calculation of electrostatic potential surface shows that the substrate binding cavity is positively charged. This charge distribution shows possibility of formation of electrostatic interactions with negatively charged carboxyl groups of the substrate molecules. The substrate binding cavity allows for the positioning the caffeic acid in many orientations. The flavonoid substrates could also be docked into the active site without difficulties. It should be kept in mind that the co-crystallized substrate is only an artificial one. The natural substrate for this enzyme remains unknown. The substrate prediction approaches were carried out for cation independent OMTs of plant origin (Schroder et al., 2002). The sequence analysis and the information obtained from the structures of related enzymes enabled with structural data helped to provide a rational guess as to the *in vivo* substrate. SynOMT is the first OMT characterized from an cyanobacteria. The data presented for SynOMT suggests certain characteristics of the preferred substrates. The activity of this protein towards substrates accepted by the related plant enzymes suggests a preference for compounds having polyhydroxylated phenyl ring moieties. The overall shape of the active site allows for accommodation of a wide range of compounds. The positive overall charge of the active site furthermore facilitates the binding of negatively charged substrates (Figure 33). It should be noted, that smaller molecules such as dhydroxy benzoic acid are not readily accepted by this enzyme.

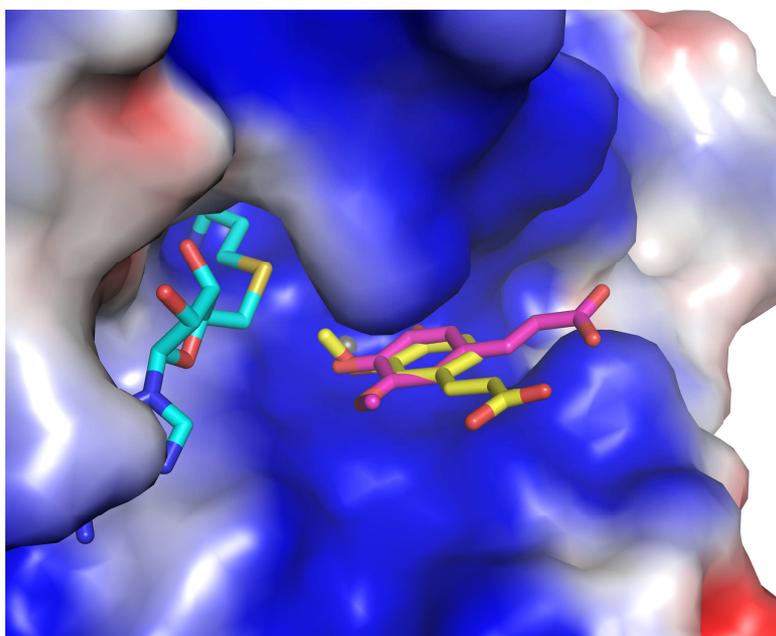


Figure 33. Electrostatic surface representation of SynOMT monomer with substrates docked in different orientations. In purple the ferulic acid bound to the structure is shown. The yellow molecule shows possible orientation for of substrate to allow for methylation in *meta* position. The blue colored surface represents the positive charge while the red color shows negative charge.

The phenylpropanoids as well as larger molecules ie. flavonoids are better substrates. This fact may be explained by an additional stabilizing interactions that a larger molecule forms with the amino acid residues forming the active site. The important structural difference between Mg²⁺dependent OMTs is the loop spanning Trp193 to Tyr211 in the structure of CCoAOMT. This region distinguishes the plant derived CCoAOMTs from structurally and functionally related animal catechol OMT. In case of SynOMT this region is also quite different. It spans from residue 173 until 185 and is 6 amino acids shorter. The previous experiments proved that this region is responsible for substrate recognition. The varying length and amino acid sequence of this part

5.6.7.3. AdoMet binding site

The binding site of AdoMet molecule is conserved among all members of AdoMet dependent OMTs. There is a number of residues involved in the formation of hydrogen bonds that bind AdoMet in place. A closer look at the AdoMet binding site of SynOMT reveals high conservation of the residues taking part in the cofactor binding (Figure 34). The spatial orientation and the types of identified hydrogen bonds are very similar to those described for CCoAOMT structure. Moreover, the placement and the orientation of cofactor itself are almost identical. In this way, Asp165 of CCoAOMT corresponds to Asp145 of SynOMT, Ala140 with Ala121, Asp111 with Asp92, Gly87 and Ser93 with Gly68 and Ser74. The Table 12 summarizes the residues forming binding cavity for AdoHcy/AdoMet molecules

Table 12. Residues comprising the binding cavity for AdoHcy molecule

CCoAOMT	PFOMT	SynOMT
Thr 63	Thr 54	Ile 44
Glu 85	Glu 76	Glu 66
Gly 87	Gly 78	Gly 68
Val 88	Val 79	Val 69
Tyr 89	Phe 80	Phe 70
Ser 93	Ser 84	Ser 74
Asp 111	Asp 102	Asp 92
Ile 112	Phe 103	Gln 93
Pro 139	Asp 130	Pro 120
Ala 140	Ala 131	Ala 121
Asp 163	Asp 154	Asp 143
Asp 165	Asp 156	Asp 145
Tyr 172	Tyr 163	Tyr 152

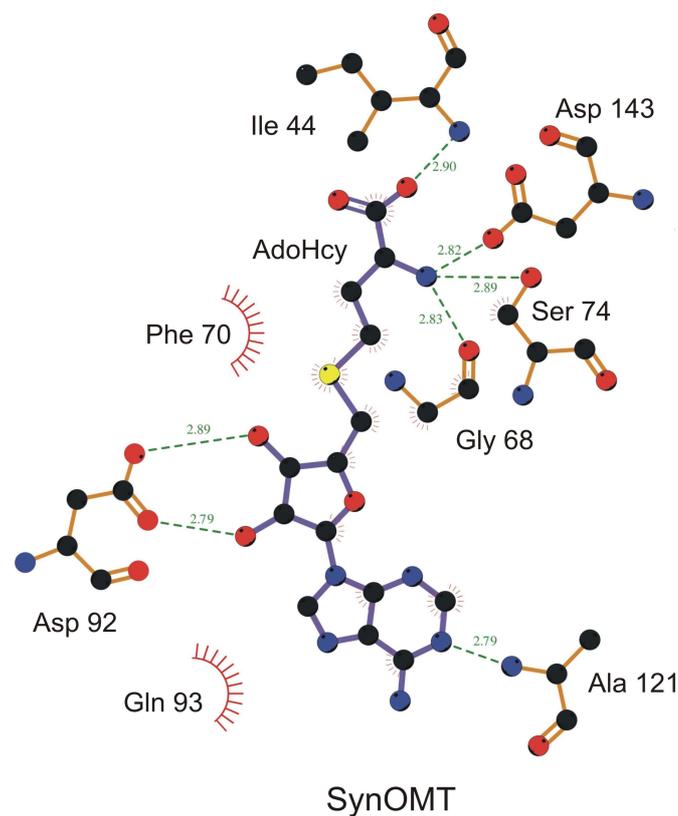


Figure 34. Binding of a AdoHcy cofactor on the example of SynOMT. Polar contacts were identified by CONTACT (part of the CCP4 suite and visualized with Ligplot) and are shown as dotted lines. The distance of each interaction is given in Å

5.6.8. N-terminal mutations

The lysine residue in the position 3 of the polypeptide chain of SynOMT occupies the position pointing inwards to the active site. It is the first structure of Mg^{2+} -dependent OMT where one of the first amino acids is visible. The electron density shows the presence of the reaction product bound in the active site of the protein. The side chain nitrogen of the Lys3 is located 2.92 Å away from the 3' oxygen of the 4OMe derivative of caffeic acid present in the structure. Such distance between those two atoms suggest that a hydrogen bond maybe formed that stabilizes the substrate in the active site exposing the *para* position to the methylation reaction. It was proposed that this residue may influence the activity of the protein as well as the methylation position specificity. Similar phenomenon was observed in case of PFOMT, since the removal of the lysine residues located on the N-terminus resulted in an altered specificity of the enzyme (Vogt, 2004). The biggest difference in enzymatic activity was observed with the PFOMT variant where 11 N-terminal residues were removed. The notable fact is that the residue number 10 in the sequence of PFOMT is also a lysine. The

structure of SynOMT suggest that the lysine residue located on the N-terminus of Mg^{2+} dependant OMTs may be an important stabilizing factor upon binding of the substrate in the active site. In order to investigate the action of this residue, a mutation where Lys3 was replaced by an alanine had been proposed.

5.6.9. Expression and purification of the SynOMT K3A mutant

The SynOMT K3A mutant was produced with the use of site directed mutagenesis. The SynOMT K3A mutant protein was purified, concentrated and rebuffered. The yield of the protein was comparable to the wild type SynOMT (1 mg/1 l culture). The purity of the protein preparations was confirmed by SDS Page (Figure 35) lane 4.

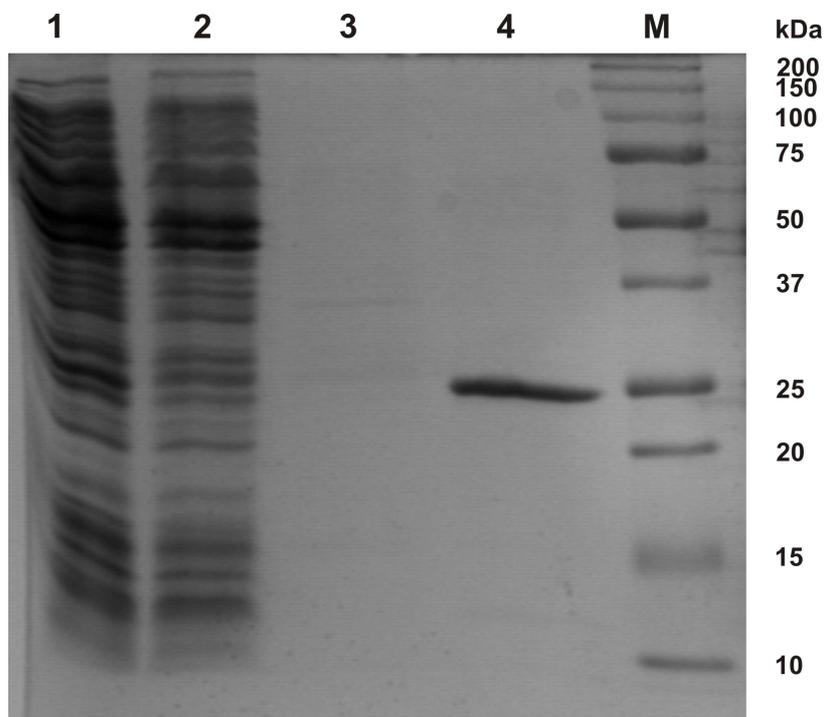


Figure 35. IMAC purification of the SynOMT K3A mutant. The numbered lanes are: 1 soluble protein, 2 flow trough, 3 30mM imidazole wash, 4 240mM imidazole elution. The molecular weight marker is annotated with M.

5.6.10. The kinetic parameters of the K3A SynOMT mutant

The specific activities of the SynOMT K3A mutant protein towards different substrates were measured and compared with the ones obtained for the purified proteins. The produced mutant shows a significant drop in the activity towards virtually all of the substrates. In two cases, values of the activity were above 0.5% of the wild type activity (Table13). The dramatic drop of activity confirms the importance of the Lys3 for the methylation reaction.

Table 13. The percentage of specific activity of SynOMT K3A mutant in comparison to the wild type SynOMT (100%)

Substrate	Percentage activity SynOMT K3A mutant
5-Hydroxyferulic acid	0.70% corresponding to 0.00146 pMol/(s* μ g)
Caffeic acid	0.61% corresponding to 0.00044 pMol/(s* μ g)
CCoA	n.d.
Caffeoylglucose	n.d.
3,4,5-Trihydroxycinnamic acid	n.d.
Tricetin (3',4',5,5',7-Pentahydroxyflavone)	n.d.

n.d. not determined

6. Discussion

6.1. Substrate specificity of cation dependent OMTs

Among the plant Mg^{2+} dependent OMTs two groups of enzymes can be distinguished. The substrate specific CCoAOMTs, which can be regarded as the enzymatic specialists, and the promiscuous enzymes that methylate a variety of substrates with a similar efficiency, the generalists. The specialization of the CCoAOMTs may be caused by the involvement in the lignin biosynthesis where CCoA is one of the most important intermediates. Current state of knowledge does not offer extensive explanation of the entire process of lignin production nor does it answer why CCoAOMT is so specific towards its substrate while presenting almost no activity towards the caffeic acid, which is also a key intermediate in this process.

Enzymes with promiscuous activities have been discovered in plants, PFOMT is an important member of this family (Ibdah et al., 2003). Other proteins of similar properties are also being identified, an example here may be a cation dependent OMT capable of methylating vicinal dihydroxyl system of flavonols found in rice (Lee et al., 2008). A recent report of the flower specific OMT from *A. thaliana* reveals its localization in tapetum and possible involvement in the methylation of phenylpropanoid spermidine conjugates detected in the flowers and mature pollen grains (Fellenberg et al., 2008). With those two groups of enzymes found at the same time it seems that there is an apparent redundancy of OMTs present in plants.

The solution of the crystal structure of PFOMT (Kopycki et al., 2008a), the enzymatic generalist allowed for comparison to CCoAOMT (Ferrer et al., 2005). The AdoMet and bivalent cation requirement are the factors influencing the remarkable similarities within the active site and overall structure of those proteins. The topology of the active site itself is also very similar because both enzymes catalyze principally the same reaction, the addition of a methyl group to a vicinal dihydroxyl system. In this way both enzymes share the same catalytic apparatus. The residues in the immediate vicinity are almost identical in both of the cation dependent enzymes. Both proteins exist as dimers where the active sites are independent of one another. It is plausible however, that that a mobile N-terminus of one monomer might communicate with the catalytic apparatus of the second monomer (Figure 36). Such situation can be observed for the recently reported structure of dimeric bacterial OMT from *Leptospira interrogans* (LiOMT) of unknown specificity (Hou et al., 2007).

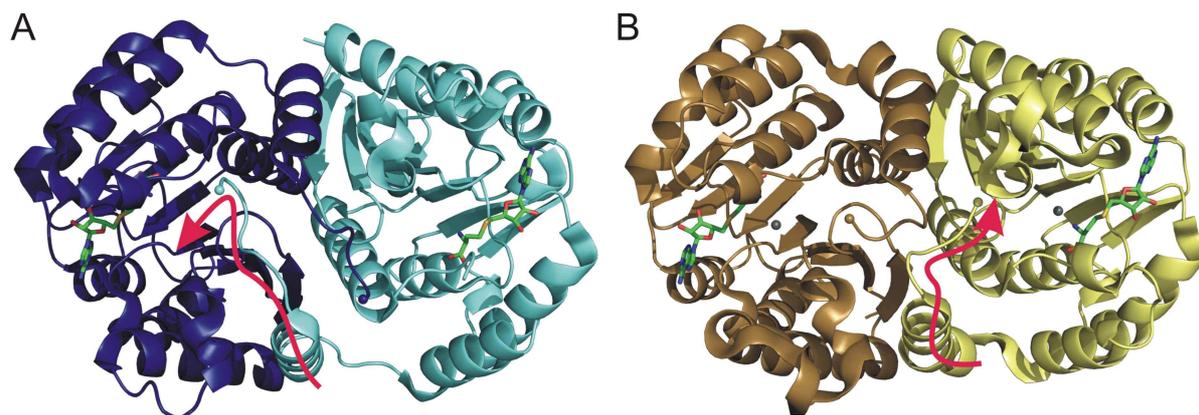


Figure 36. Dimer representation of A LiOMT (monomers colored in blue and cyan) and B PFOMT (monomers colored in pale yellow and brown). The red arrows highlight the position of the N-terminus. In case of LiOMT the N-terminus communicates with the active center of the second monomer. For PFOMT the N-terminus interacts with the monomer to which it belongs. The AdoHcy bound to both structures is shown as sticks. The gray spheres in PFOMT structure are the Ca²⁺ ions bound in the active site.

On the other hand, in case of cation independent enzymes, the active site is always composed of the residues belonging to both monomers forming a dimer. The attempts were made to introduce new activities to OMTs (Frick et al., 2001) by recombination of OMT subunits. Four of the cDNAs encoding OMTs involved in isoquinoline alkaloid biosynthesis were obtained from berberine-producing cell suspension cultures of the meadow rue *Thalictrum tuberosum*. The investigated enzymes sharing 93.2±99.7% identity were determined to have overlapping but not identical substrate specificity. The produced heterodimers in all possible combinations (10 potentially different recombinant OMTs) were tested with a library of compounds. It was found that several substrates that were accepted as substrates by heterodimers were methylated by homodimers.

The mechanisms governing the substrate specificity of cation independent OMTs were well studied. Isoeugenol OMT and COMT, which share 83% identity at the amino acid level, are a model system. Even with high amino acid sequence similarity those two OMTs have a distinct substrate preference and methylation regiospecificity. Seven amino acid residues responsible for the substrate specificity were identified. Mutation of corresponding residues on COMT with the ones from isoeugenol OMT resulted with altered substrate specificity of the mutant protein (Wang et al., 1999). A change of a single amino acid residue was reported sufficient to alter the substrate specificity of OMTs responsible for the synthesis of 3,5-dimethoxytoluene in roses. This volatile compound is responsible for a specific scent of Chinese roses and the modern hybrid varieties. 3,5-dimethoxytoluene is produced from

orcinol via two subsequent methylation reactions, catalyzed by two distinct OMTs. The substrate specificity of those two orcinol OMTs is determined by a single amino acid polymorphism (Scalliet et al., 2008).

Up to present moment no investigations were carried out that would explain how so similar in structure cation dependent OMTs have such dramatically different substrate specificities. As indicated the N-termini of Mg^{2+} dependant OMTs may play a role in the substrate specificity of those enzymes. PFOMT isolated from the natural source was determined by N-terminal sequencing to have a N-terminus that is eleven amino acid residues shorter than the protein expressed recombinantly (Ibdah et al., 2003). Further investigations carried out on recombinant PFOMT (Vogt, 2004) confirm that this part of protein is also responsible for the enzymatic activity and substrate specificity. Recombinantly expressed N-terminal truncations were shown to differ in both the catalytic efficiency as well as in substrate affinity.

In silico docking experiments were carried out to obtain semi-quantitative data concerning the substrate binding. The results of those dockings carried out on both structures (PFOMT and CCoAOMT) using compounds accepted as substrates for both enzymes are in agreement with previous observations (please see the Results section). The compounds confirmed *in vitro* to be the preferred substrates for the enzymes can be successfully docked to the structures. For quercetin, a substrate-like orientation of the flavonoid could be determined in the active center of PFOMT with reasonable binding energies, whereas no enzymatically reasonable binding mode could be found for CCoAOMT. The binding energies of docking of caffeic acid and 5-hydroxyferuloyl CoA clearly distinguish between those two enzymes. Caffeic acid is preferred by PFOMT in terms of binding energy, cluster ranking, and cluster size, while the results for CCoAOMT are less conclusive. For the CoA ester the situation is reversed. Such results were observed despite the absence of structural data for the N-terminal residues in both of the structures.

The analysis of the available structures together with structure based amino acid sequence alignments show that there are two regions where the amino acid sequence and the spatial conformation differ greatly between the promiscuous and substrate specific OMTs. Those were the N-terminus and the loop region located between strand $\beta 5$ and helix $\alpha 8$. Unfortunately in neither of the plant Mg^{2+} OMT dependent structures the N-terminal residues were seen in the electron density. This indicates that this entire region is very flexible and may easily change its conformation upon binding of the substrate. The conformation of this region, which begins with a short α -helical turn, is supported by the N-terminal helix of the opposing monomer, so that the active-site cleft appears deeper and more restricted. This

insertion loop provides a scaffold for the CoA moiety of CCoA in the CCoAOMT structure. Most interestingly, the loop region in PFOMT exhibits a conformation markedly different from that in CCoAOMT. It seems plausible that differences in this loop and in the N-termini of the two enzyme classes could provide a clue to the promiscuous nature of PFOMT.

The influence on the substrate specificity of both regions of the proteins was investigated by the production and the analysis of the hybrid proteins that combined parts of the mentioned OMTs. The investigated hybrid proteins were expressed as a N-terminal his tag fusion to aid in the purification. In order to determine the influence of the N-terminal his tag an additional construct was made with a cleavable his tag. The k_{cat} and the K_m app values were recorded for each of the investigated protein towards five substrates to produce a substrate specificity profile. The values obtained from the hybrid proteins were compared with PFOMT and CCoAOMT (Figure 37).

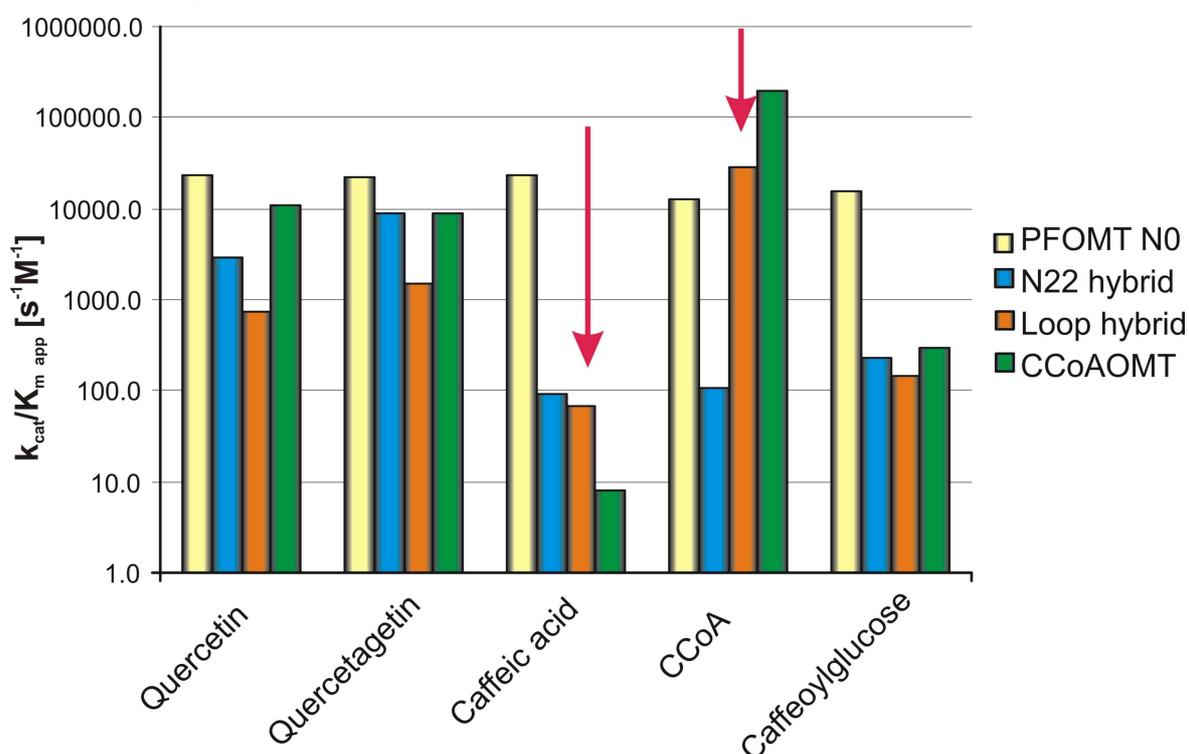


Figure 37. Catalytic efficiency profile for the investigated proteins towards different substrates. Please note the high catalytic efficiency of PFOMT N0 (yellow bars) towards all investigated substrates. The decreased (in comparison to PFOMT) values of N22 hybrid towards all substrates. The arrows indicate the similarity of the Loop hybrid to CCoAOMT. Low catalytic efficiency towards caffeic acid with a substantial increase towards CCoA (orange bar). This mimics the behavior of CCoAOMT (green bar).

The following general conclusions can be drawn. The Hybrids containing the loop part exhibited higher catalytic efficiency towards the CCoA than caffeic acid and caffeoylglucose. This observation coincided with the substrate specificity profile recorded for *M. sativa*

CCoAOMT. Where CCoA is the preferred substrate and caffeoylglucose together with caffeic acid are very poorly methylated. The N-terminal hybrids containing the N-terminal sequence of the specialized CCoAOMT showed largely reduced affinity towards all substrates tested. This effect is most pronounced for caffeic acid and its derivatives. The influence of the N-terminal his tag is also the factor to be taken into account. The proteins where this affinity tag was cleaved off show k_{cat} reduction towards the caffeic acid and CCoA. This is largely due to the increase in $K_{m\ app}$ values. For flavonoid substrates the decrease in catalytic efficiency was not so pronounced. Taking into consideration the presented results it can be speculated that the N-terminal residues of PFOMT are responsible for the recognition of the flavonoid substrates while the mentioned loop part plays an essential role in the positioning and methylation of CCoA. The conjunction of effects caused by both investigated regions may be the element that drives the specificity of those OMT. Possible explanations could also come from unpredictable changes in the scaffold conformation, changes in protein stability, and/or changes in the dynamic behavior of the hybrid enzymes. The approach for creation of new enzymatic activities through the recombination of residues or domain structures responsible for these enzymatic activities has been attempted before. The mechanism of substrate specificity of glucosyltransferase SaGT4A from *Solanum aculeatissimum* was studied by creation of hybrid proteins with sequences coming from potato (*S. tuberosum*) glucosyltransferase StSGT (Kohara et al., 2007).

6.2. Position specificity of cation dependent OMTs

SynOMT from *Synechocystis* sp. strain PCC 6803 is the first cation- and AdoMet-dependent *O*-methyltransferase to be characterized and crystallized from a prokaryote. The availability of this enzyme gives a rare opportunity to investigate its function and special characteristics in blue algae.

Preliminary tests revealed substrate specificity for vicinal dihydroxy groups of phenolics. This is comparable with the currently known eukaryotic proteins. The natural substrate of this enzyme is currently unknown. SynOMT exhibits a strikingly different position specificity when compared to eukaryotic enzymes. This OMT is capable of methylation both meta- and para-positions of the vicinal dihydroxyl systems of phenolics. This activity was discovered and proven by the analysis of the methylation products of an artificial substrate, 3,4,5-Trihydroxycinnamic acid. A comparable specificity has been described for the prokaryotic *SafC* gene product from *M. xanthus* when tested with caffeic acid, although the natural substrate for the latter enzyme appears to be L-dihydroxyphenylalanine, a biosynthetic

precursor of the isoquinoline alkaloid saframycin (Nelson et al., 2007). SynOMT, despite rather low amino acid sequence identity shares a high structural similarity to other cation dependent OMTs. This OMT, as are the other plant OMTs is a functional dimer. Similarly to PFOMT and CCoAOMT the active site of one monomer does not communicate with any residues belonging to the other. The catalytic machinery works in the same way as described for rOMT and plant enzymes. The amino acid sequence alignment (Figure 38) of OMTs from different organisms indicates that the ion binding and the catalytically relevant residues are very conserved.

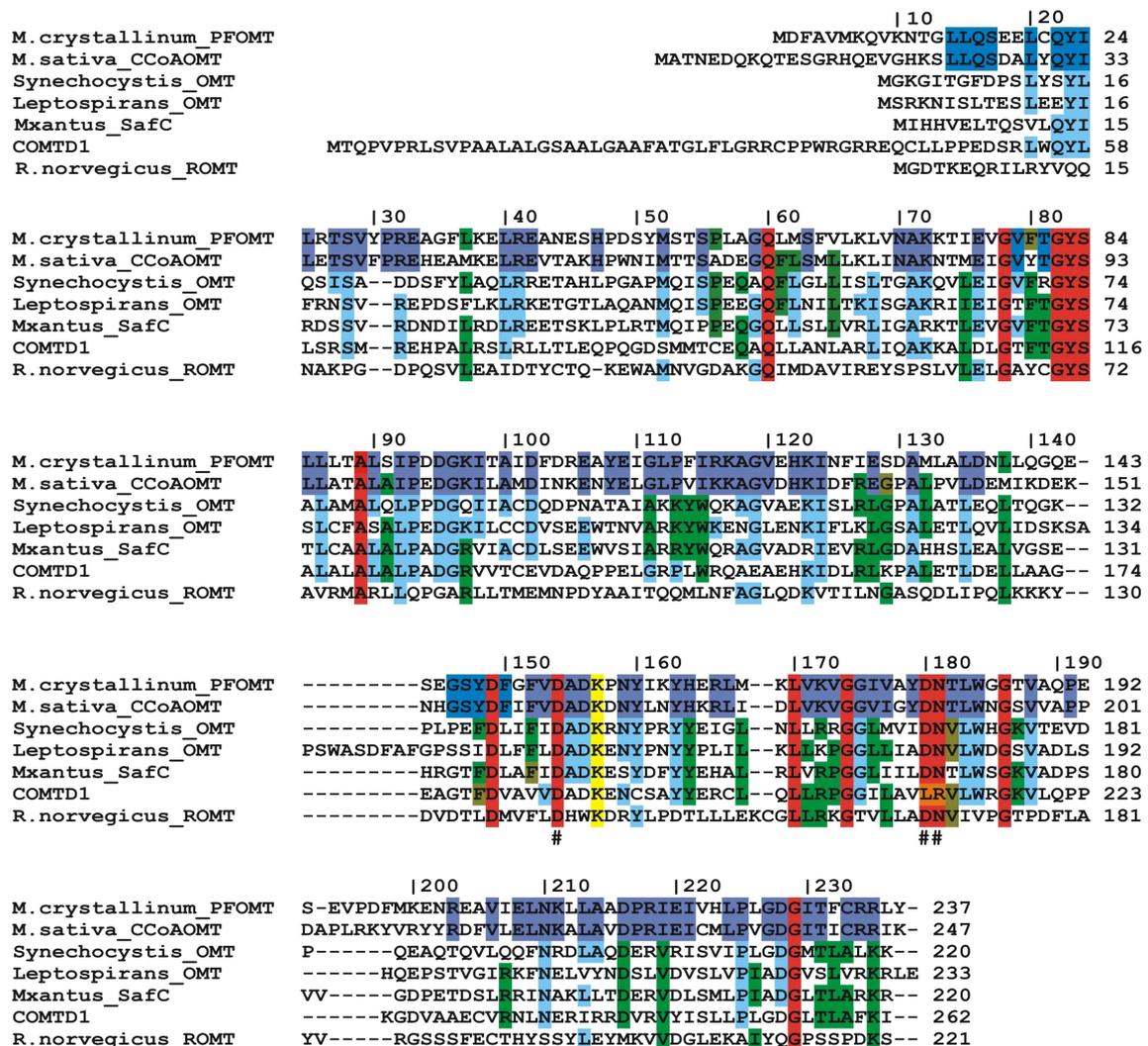


Figure 38. Structure based amino acid sequence alignment of Mg²⁺ dependent OMTs from different organisms. The residues marked in violet are the same for CoA specific and promiscuous plant OMT, in addition those marked light blue are corresponding ones from other organisms. The residues marked in red are conserved among all of the organisms. The residues common to the majority of proteins but not the same in both PFOMT and CCoAOMT are marked in green. The catalytic lysine residue is marked in yellow. The residues identified to bind the ion are marked with (#). In case of the structure of COMTD1 where ion is not present the residues corresponding to the ones involved in binding of the ion are colored orange.

This fact is not surprising since binding of the cation is essential requirement for the enzymatic activity. The only exception here is the sequence of Catechol-*O*-methyltransferase Domain Containing 1 (COMTD1), a recently released structure of a putative OMT from human, (PDB code 2avd). For this protein the residues responsible for cation binding are not conserved. Instead a side chain of an arginine residue occupies the same place as the cation in cation dependent OMTs would. Unfortunately neither the catalytic activities nor the *in vivo* substrate for this enzyme are not known. The preliminary tests of this protein with the flavonoids and phenylpropanoids commonly accepted as substrates by plant enzymes failed to show any enzymatic activity. Due to the presence of the reaction products in the active site the structure of SynOMT provides a rare opportunity to draw conclusions to enzyme - substrate interactions. While it was impossible to soak or co-crystallize the substrate into the structure of PFOMT, co-crystallization of SynOMT with caffeic acid was successful. The electron density reveals caffeic acid monomethylation products (ferulic acid and isoferulic acid) bound in the active site of the enzyme. This observation confirms the previous experiments carried out with 3,4,5-Trihydroxycinnamic acid. Two distinct binding modes for ferulic acid and isoferulic acids, were found. Whereas the positioning of the aromatic ring oxygen atoms of isoferulic acid suggest a substrate-like interaction with the active site magnesium ion and the AdoMet cofactor, ferulic acid binds with only one oxygen approaching the metal ion.

The precise structural features of the SynOMT that contribute to this unexpected positional methylation specificity are not clear. One possibility is provided by the presence of side chains of His174 and Lys176 in the neighborhood of the propenoic acid moiety, equivalent to Asn and Ser in the plant CCoAOMTs, respectively.

The proximity of the N-terminal lysine, Lys3, in the immediate vicinity of the active site and formation of molecular contacts to one of the substrate/product hydroxyl groups implicates that it is involved in the catalytic mechanism. The point mutation experiments show almost complete abolition of enzymatic activity upon mutation to alanine, which suggests a more fundamental role than determining position specificity. The importance of N-terminal residues in position specificity in the plant enzymes has already been demonstrated for PFOMT. Unfortunately the individual contribution of N-terminal lysine residues in the plant enzymes cannot be evaluated in detail due to the lack of structural data. Structure based amino acid sequence alignments of the dimeric metal-dependent OMTs reveal the presence of lysine

and histidine residues in the N-terminal regions of each of these enzymes, suggesting that minor differences in position specificity could reside there (Figure 38).

Flavonoids containing paramethylated B-rings are found throughout the plant kingdom. Methylation in *para* position in plants is sometimes performed by cation independent enzymes with strict structural requirements. For example, a flavonoid-specific OMT from *Catharanthus roseus* methylates the 4'-position only when the neighboring 3'-methoxy group of the flavonoid B-ring is present, as in case of the isorhamnetin or homoeriodictyol (Schroder et al., 2004)

A similar cation-independent OMT, cloned from rice and functionally expressed in *E. coli*, performs sequential methylation of three B-ring hydroxyl groups of the flavone tricetin *in vitro*, including the *para*-position (Zhou et al., 2006).

Another example here may be a cation-independent enzyme from *Coptis japonica*. This OMT was found to methylate the 4' hydroxyl group of (*S*)-3'-Hydroxy-*N*-methylcoclaurine to form (*S*)-Reticuline, which is an important intermediate in synthesizing isoquinoline alkaloids (Morishige et al., 2000). For the cation independent enzymes a vicinal dihydroxyl system is not always required. A flavone OMT (SOMT-2) from soybean (*Glycine max*) converts naringenin into ponciretin (Kim et al., 2005). The preferred substrate, naringenin, has only one hydroxyl group attached to the phenolic ring moiety.

When compared to plant enzymes the hydroxyl group in *para* position is an initial target of the prokaryotic enzyme, at least in the case of the phenylpropanoid esters, followed by the subsequent attack of either the 3'-OH or 5'-OH group. With trihydroxylated flavonoids, tricetin and myricetin, these subsequent methylation reactions are missing and only a single product is observed with SynOMT.

As mentioned previously the *in vivo* substrate of SynOMT remains unknown. The substrates that were tested with SynOMT do not provide any information as to the function of this enzyme in cyanobacteria. The conserved presence of at least one gene locus in most prokaryotes points to an important *in vivo* function for these organisms including *Synechocystis*.

SynOMT was shown to have the amino acid sequence similarities to the previously mentioned SafC gene product from *M. xanthus*. Another similar sequence that can be mentioned is the mdmC gene in *S. mycarofaciens*, proposed to encode the 4-*O*-methyltransferase for the 16-membered lactone ring of the macrolide antibiotic midecamycin (Hara et al., 1992). Those similarities indicate that the function of SynOMT may be implicated in the biosynthesis of complex secondary metabolites

the hybrid enzymes. A similar situation was observed for variants of trypsin designed to behave like the closely related serine proteinase Factor Xa (Rauh et al., 2004; Reyda et al., 2003). In order to fully address the question concerning the substrate specificity of plant Mg^{2+} dependent OMTs additional studies are required.

Studies concerning the catalytic behavior of enzymes may prove to be very valuable. A careful understanding of mechanisms governing the substrate specificity may provide information that can be useful in the design of artificial enzymes that would be capable of catalysis of chemical reactions not found in nature. Such enzymes could be used as recombinant protein preparations in industry or introduced into living organisms offering a possibility of engineering of metabolic pathways or introduction of new ones.

Enzyme engineering can provide industry with improved enzymes that are more stable or have higher catalytic efficiency. Carbohydrate manipulation is an important topic for industry (Hancock et al., 2006). In nature, catabolism and modification of carbohydrates is carried out by two groups of enzymes: glycosyltransferases and glycosidases. *In vitro* those enzymes are a very valuable tool that allows for manipulation of the structures of carbohydrates. Due to a large diversity in the structures of carbohydrates there are still glycosidic linkages for which a catalyst with the requisite specificity is not known. The means of accomplishing the task of production of an enzyme with an altered specificity is accomplished either by rational design or by directed evolution approaches combined with effective screening methods.

An elegant example of introduction of a biosynthetic pathway into a living organism is the story of “Golden Rice” (Al Babili et al., 2005). In an effort to battle the vitamin A deficiency in developing countries a new transgenic rice variety was produced that accumulates the β -carotene in the endosperm. The name Golden Rice comes from how the rice grains look like. Unlike wild type varieties, this one due to accumulation of β -carotene has yellow grains. An entire biosynthesis pathway of β -carotene was introduced in order to obtain this transgenic plant.

Cloning and functional expression of other members of the prokaryotic CCoAOMT-like proteins may hold further surprises as far as substrate and position specificities toward natural compounds. Those enzymes might be used as specific organic catalysts in yet difficult to achieve chemical modifications. Expression of enzymes targeting *para*-hydroxy groups in the appropriate gymnosperm or angiosperm background may allow for the production of novel types of lignin with altered properties such as changes in the degree or mode of polymerization that might effect digestibility by ruminants.

The discovery and solution of the structure of an algal Mg^{2+} dependent OMT gives an opportunity to study the properties of this class of enzymes coming from organisms other than plants. The characterization of the reaction products showed a surprising positional specificity of methylation. While plant OMTs show exclusive methylation in position *meta* of the phenyl ring, the algal OMT is not restricted to this position and is capable of methylation in both *meta* and *para* positions. Like in the case of substrate specific and promiscuous OMTs the overall structure and the catalytic mechanism utilized by the SynOMT is conserved. The subtle differences in the active site of this protein may be responsible for this novel positional specificity. Unfortunately the *in vivo* substrate methylated by this enzyme in algae remains unknown. The essential function of the slr0095 gene product might be further resolved by knock-out mutations and homologous recombination, feeding of labeled precursors, and molecular docking studies.

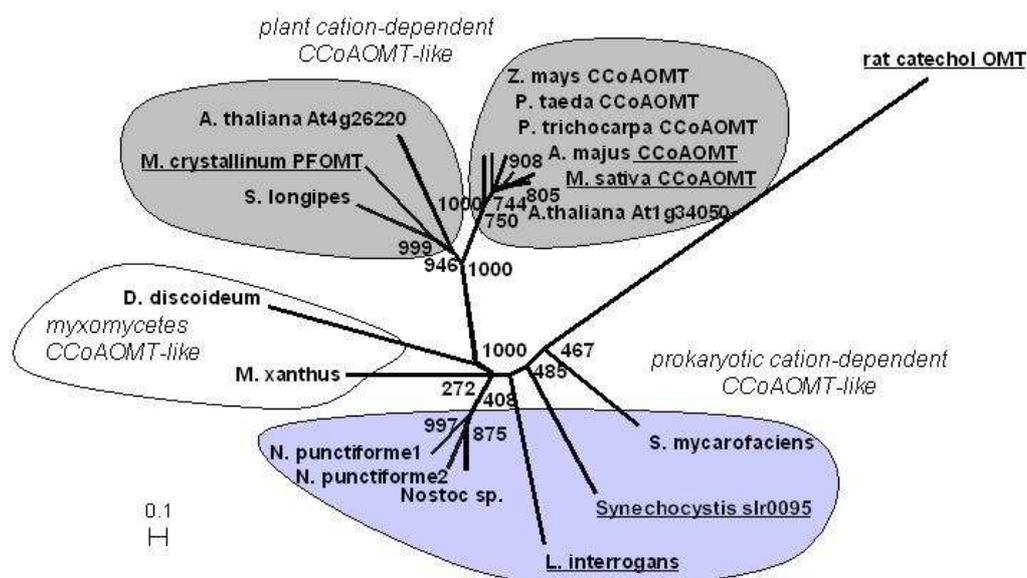


Figure 40. Bootstrapped cladogram of selected CCoAOMTs and CCoAOMT-like proteins from pro- and eukaryotes (Kopycki et al., 2008b). The neighbor joining tree was based on 1000 bootstrap trials. Notice that all bootstrap values below 500 do not allow phylogenetic relationships to be clearly defined. The cluster of six plant proteins marked as CCoAOMTs is specific for CoA methylation in lignin monomer biosynthesis. Gene products of Arabidopsis CCoAOMTs are listed based on the corresponding gene identifiers. Crystallized proteins are underlined. The mammalian catechol OMT from rat was used as an outgroup. NCBI data base accession numbers used are: BAA10567 (*Synechocystis*PCC6803 slr 0095); NP_710596 (*L. interrogans*); Q00719 (*S. mycarofaciens*); BAB76778 (*Nostoc* sp. PCC 7120); ZP_00108749 (*Nostoc punctiforme2*);ZP_00111674 (*Nostoc punctiforme1*); AAM33748 (*Dictyostelium discoideum*); AAC44130 (*M. xanthus*);AAB61680 (*Stellaria longipes*); AAN61972 (*M. crystallinum* PFOMT); AAM64800 (*Arabidopsis thaliana* At4g 26220AOMT1); AJ242980 (*Zea mays* CCoAOMT); AAD02050 (*Pinus taeda* CCoAOMT); CAA12198 (*Populus trichocarpa* CCoAOMT); AAT40111 (*Ammi majus* CCoAOMT), Q40313 (*M. sativa* CCoAOMT); AY057554, *A. thaliana* At1g34050), P22734 (rat catechol OMT).

The completion of several genome projects, shows that plants contain several members of apparently redundant, but conserved CCoAOMT-like genes. The phylogenetic tree (Figure 40) shows the close clustering (high sequence identity 80-90% among different species) of the specific CCoAOMTs. A significant, but less stringent identity (~ 40-60 %) can be observed for the promiscuous plant enzymes. Plants seem to contain closely related but nevertheless distinct proteins with overlapping substrate specificities. The results presented here suggest that other members of this class may serve functions distinct to methylation of the single lignin precursor, CCoA. The sequences of prokaryotic cation dependent OMTs are not highly clustered. Despite rather low sequence identity of prokaryotic enzymes their three dimensional structures are remarkably similar to one another as well as to plant OMTs. So far it was not possible to say why such situation takes place and in which way those groups of enzymes evolved, or what the common ancestor is. The work presented here points to the regions of proteins that are likely to have been targets for the process of evolution. It would not be unlikely that the different substrate specificities evolved by introduction of mutations in those regions while the rest of the protein remained relatively unchanged or the introduced mutations did not affect the overall structure.

7 List of References

References

- Al Babili, S. and Beyer, P.(2005) Golden Rice - five years on the road - five years to go? Trends in Plant Sci.**10** 565-573
- Aliotta, G., Cafiero, G., Fiorentino, A., and Strumia, S.(1993) Inhibition of Radish Germination and Root-Growth by Coumarin and Phenylpropanoids. J. Chem. Ecol. **19** 175-183
- Anterola, A. M. and Lewis, N. G.(2002) Trends in lignin modification: a comprehensive analysis of the effects of genetic manipulations/mutations on lignification and vascular integrity. Phytochemistry **61** 221-294
- Avelange-Macherel, M. H. and Joyard, J.(1998) Cloning and functional expression of AtCOQ3, the Arabidopsis homologue of the yeast COQ3 gene, encoding a methyltransferase from plant mitochondria involved in ubiquinone biosynthesis. Plant J. **14** 203-213
- Axelrod, J. and Tomchick, R.(1958) Enzymatic *O*-Methylation of epinephrine and other catechols. J.Biol.Chem. **233** 702-705
- Bailey, S.(1994) The Ccp4 Suite - Programs for Protein Crystallography. Acta Crystallogr., Sect. D: Biol. Crystallogr. **50** 760-763
- Bauer, N. J., Kreuzman, A. J., Dotzlaf, J. E., and Yeh, W. K.(1988) Purification, characterization, and kinetic mechanism of S-Adenosyl-L-Methionine - Macrocin *O*-Methyltransferase from *Streptomyces-Fradiae*. J.Biol.Chem. **263** 15619-15625
- Bauer, U. M., Daujat, S., Nielsen, S. J., Nightingale, K., and Kouzarides, T.(2002) Methylation at arginine 17 of histone H3 is linked to gene activation. EMBO **3** 39-44
- Bestor, T. H.(2000) The DNA methyltransferases of mammals. Hum.Mol.Genet. **9** 2395-2402
- Bieza, K. and Lois, R.(2001) An Arabidopsis mutant tolerant to lethal ultraviolet-B levels shows constitutively elevated accumulation of flavonoids and other phenolics. Plant Physiol. **126** 1105-1115
- Bonneau, R. and Baker, D.(2001) *Ab initio* protein structure prediction: Progress and prospects. Ann. Rev. Biophys. Biomol. Struct. **30** 173-189
- Bonneau, Richard, Ruczinski, Ingo, Tsai, Jerry, and Baker, David(2002) Contact order and *ab initio* protein structure prediction. Protein Sci. **11** 1937-1944
- Bradford, M. M.(1976) Rapid and Sensitive Method for Quantitation of Microgram Quantities of Protein Utilizing Principle of Protein-Dye Binding. Anal. Biochem. **72** 248-254
- Brunger, A. T., Adams, P. D., Clore, G. M., Delano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L.(1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Crystallogr., Sect. D: Biol. Crystallogr. **54** 905-921

- Chung, S. Y. and Subbiah, S.(1996) A structural explanation for the twilight zone of protein sequence homology. *Structure* **4** 1123-1127
- De Luca, Vincenzo and Ibrahim, Ragai K.(1985) Enzymatic synthesis of polymethylated flavonols in *Chrysosplenium americanum*. I. Partial purification and some properties of S-adenosyl--methionine: Flavonol 3-, 6-, 7-, and 4'-*O*-methyltransferases. *Arch. Biochem. Biophys.* **238** 596-605
- Delano, W. L.(2002) **The PyMOL Molecular Graphics System.**
- Dhar, K. and Rosazza, J. P. N.(2000) Purification and characterization of *Streptomyces griseus* catechol *O*-methyltransferase. *Appl. Environ. Microbiol.* **66** 4877-+
- Dodge, J. E., Ramsahoye, B. H., Wo, Z. G., Okano, M., and Li, E.(2002) *De novo* methylation of MMLV provirus in embryonic stem cells: CpG versus non-CpG methylation. *Gene* **289** 41-48
- Doublet, S.(1997) Preparation of selenomethionyl proteins for phase determination. *Macromolecular Crystallography, Pt A* **276** 523-530
- Edwards, R. and Dixon, R. A.(1991) Purification and Characterization of S-Adenosyl-L-Methionine - Caffeic Acid 3-*O*-Methyltransferase from Suspension-Cultures of Alfalfa (*Medicago-Sativa L*). *Arch. Biochem. Biophys.* **287** 372-379
- Eggen, M. and Georg, G. I.(2002) The cryptophycins: Their synthesis and anticancer activity. *Medicinal Research Reviews* **22** 85-101
- Emsley, P. and Cowtan, K.(2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **60** 2126-2132
- Fellenberg, C, Milkowski, C, Hause, B, Lange, P. R, Böttcher, C, and Vogt, T(2008) Tapetum-specific location of a cation-dependent *O*-methyltransferase in *Arabidopsis thaliana*. *The Plant J.* **9999**
- Ferrer, J. L., Austin, M. B., Stewart, C., Jr., and Noel, J. P.(2008) Structure and function of enzymes involved in the biosynthesis of phenylpropanoids. *Plant Physiol. Biochem.* **46** 356-370
- Ferrer, J. L., Zubieta, H., Dixon, R. A., and Noel, J. P.(2005) Crystal structures of alfalfa caffeoyl coenzyme A 3-*O*-methyltransferase. *Plant Physiol.* **137** 1009-1017
- Frank, Joachim(2002) Single-particle imaging of macromolecules by cryo-electron microscopy. *Ann. Rev. Biophys. Biomol. Struct.* **31** 303-319
- Frick, S., Ounaroon, A., and Kutchan, T. M.(2001) Combinatorial biochemistry in plants: the case of *O*-methyltransferases. *Phytochemistry* **56** 1-4
- Gang, D. R., Lavid, N., Zubieta, C., Chen, F., Beuerle, T., Lewinsohn, E., Noel, J. P., and Pichersky, E.(2002) 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** 505-519

- Garrow, T. A.(1996) Purification, kinetic properties, and cDNA cloning of mammalian betaine-homocysteine methyltransferase. *J.Biol.Chem.* **271** 22831-22838
- Giovanelli, J.(1987) Sulfur amino-acids of plants - An overview. *Methods in Enzymol.* **143** 419-426
- Gomez, Tara A., Banfield, Kelley L., Trogler, Dorothy M., and Clarke, Steven G.(2007) The l-isoaspartyl-*O*-methyltransferase in *Caenorhabditis elegans* larval longevity and autophagy. *Dev. Biol.* **303** 493-500
- Gregersen, P. L., Christensen, A. B., Sommerknudsen, J., and Collinge, D. B.(1994) A Putative *O*-Methyltransferase from Barley Is Induced by Fungal Pathogens and Uv-Light. *Plant Mol. Biol.* **26** 1797-1806
- Gromova, E. S. and Khoroshaev, A. V.(2003) Prokaryotic DNA methyltransferases: The structure and the mechanism of interaction with DNA. *Mol. Biol* **37** 260-272
- Guo, D. J., Chen, F., Inoue, K., Blount, J. W., and Dixon, R. A.(2001) Downregulation of caffeic acid 3-*O*-methyltransferase and caffeoyl CoA 3-*O*-methyltransferase in transgenic alfalfa: Impacts on lignin structure and implications for the biosynthesis of G and S lignin. *Plant Cell* **13** 73-88
- Guo, H. B. and Guo, H.(2007) Mechanism of histone methylation catalyzed by protein lysine methyltransferase SET7/9 and origin of product specificity. *Proc.Natl.Acad.Sci.U.S.A* **104** 8797-8802
- Hahlbrock, K. and Scheel, D.(1989) Physiology and molecular-biology of phenylpropanoid metabolism. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **40** 347-369
- Haines, T. R., Rodenhiser, D. I., and Ainsworth, P. J.(2001) Allele-specific non-CpG methylation of the Nf1 gene during early mouse development. *Dev. Biol.* **240** 585-598
- Hancock, Susan M., Vaughan, Mark D., and Withers, Stephen G.(2006) Engineering of glycosidases and glycosyltransferases. *Cur. Opin. Chem. Biol.* **10** 509-519
- Hanson, A. D. and Roje, S.(2001) One-carbon metabolism in higher plants. *Ann. Rev. Plant Physiol. Plant Mol. Biol.* **52** 119-137
- Hara, O. and Hutchinson, C. R.(1992) A Macrolide 3-*O*-Acyltransferase Gene from the midecamycin-producing species *Streptomyces-Mycarofaciens*. *J.of Bacteriol.* **174** 5141-5144
- Harper, D. B. and Kennedy, J. T.(1985) Purification and properties of S-adenosylmethionine: aldoxime *O*-methyltransferase from *Pseudomonas sp.* N.C.I.B. 11652. *Biochem. J.* **226** 147-153
- Herken, H. and Erdal, M. E.(2001) Catechol-*O*-methyltransferase gene polymorphism in schizophrenia: evidence for association between symptomatology and prognosis. *Psychiatr. Genet.* **11** 105-109
- Hou, X. W., Wang, Y. L., Zhou, Z. W., Bao, S. L., Lin, Y. J., and Gong, W. M.(2007) Crystal structure of SAM-dependent *O*-methyltransferase from pathogenic bacterium *Leptospira interrogans*. *J. Struct.l Biol.* **159** 523-528

- Humphreys, J. M. and Chapple, C.(2002) Rewriting the lignin roadmap. *Curr. Opin. Plant Bio.* **5** 224-229
- Ibdah, M., Zhang, X. H., Schmidt, J., and Vogt, T.(2003) A novel Mg²⁺-dependent *O*-methyltransferase in the phenylpropanoid metabolism of *Mesembryanthemum crystallinum*. *J. Biol.Chem.* **278** 43961-43972
- Ibrahim, R. K., Bruneau, A., and Bantignies, B.(1998) Plant *O*-methyltransferases: molecular analysis, common signature and classification. *Plant Mol. Biol.* **36** 1-10
- Jones, P. A. and Baylin, S. B.(2002) The fundamental role of epigenetic events in cancer. *Nat. Rev. Genet.* **3** 415-428
- Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M.(1991) Improved methods for building protein models in electron-density maps and the location of errors in these models. *Acta Crystallogr., Sect. A: Found. Crystallogr.* **47** 110-119
- Kabsch, W.(1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. App. Crystallogr.* **26** 795-800
- Kagan, R. M., McFadden, H. J., McFadden, P. N., O'Connor, C., and Clarke, S.(1997) Molecular phylogenetics of a protein repair methyltransferase - Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Comp. Biochem. Physiol.* **117** 379-385
- Keller, N. P., Dischinger, H. C., Bhatnagar, D., Cleveland, T. E., and Ullah, A. H. J.(1993) Purification of a 40-kilodalton methyltransferase active in the aflatoxin biosynthetic-pathway. *App. Environ. Microbiol.* **59** 479-484
- Kim, D. H., Kim, B. G., Lee, Y., Ryu, J. Y., Lim, Y., Hur, H. G., and Ahn, J. H.(2005) Regiospecific methylation of naringenin to ponciretin by soybean *O*-methyltransferase expressed in *Escherichia coli*. *J. of Biotechnol.* **119** 155-162
- Kim, E., Lowenson, J. D., MacLaren, D. C., Clarke, S., and Young, S. G.(1997) Deficiency of a protein-repair enzyme results in the accumulation of altered proteins, retardation of growth, and fatal seizures in mice. *Proc. Nat. Acad. Sci. U.S.A* **94** 6132-6137
- Kleywegt, G. J. and Jones, T. A.(1998) Databases in protein Crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **54** 1119-1131
- Kohara, Atsuko, Nakajima, Chiharu, Yoshida, Shigeo, and Muranaka, Toshiya(2007) Characterization and engineering of glycosyltransferases responsible for steroid saponin biosynthesis in solanaceous plants. *Phytochemistry* **68** 478-486
- Kopycki, J. G., Rauh, D., Chumanevich, A. A., Neumann, P., Vogt, T., and Stubbs, M. T.(2008a) Biochemical and structural analysis of substrate promiscuity in plant Mg²⁺-dependent *O*-methyltransferases. *J. Mol. Biol.* **378** 154-164
- Kopycki, J. G., Stubbs, M. T., Brandt, W., Hagemann, M., Porzel, A., Schmidt, J., Schliemann, W., Zenk, M. H., and Vogt, T.(2008b) Functional and structural characterization of a cation-dependent *O*-methyltransferase from the cyanobacterium *Synechocystis* SP. strain PCC 6803. *J. Biol. Chem.* **283** 20888-20896.

- Krause, C. D., Yang, Z. H., Kim, Y. S., Lee, J. H., Cook, J. R., and Pestka, S.(2007) Protein arginine methyltransferases: Evolution and assessment of their pharmacological and therapeutic potential. *Pharmacology & Therapeutics* **113** 50-87
- Kreuzman, A. J., Turner, J. R., and Yeh, W. K.(1988) Two distinctive *O*-methyltransferases catalyzing penultimate and terminal reactions of macrolide antibiotic (tylosin) biosynthesis. Substrate specificity, enzyme inhibition, and kinetic mechanism. *J.Biol.Chem.* **263** 15626-15633
- Laemmli, U. K.(1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227** 680-685
- Lavid, N., Wang, J. H., Shalit, M., Guterman, I., Bar, E., Beuerle, T., Menda, N., Shafir, S., Zamir, D., Adam, Z., Vainstein, A., Weiss, D., Pichersky, E., and Lewinsohn, E.(2002) *O*-methyltransferases involved in the biosynthesis of volatile phenolic derivatives in rose petals. *Plant Physiol.* **129** 1899-1907
- Lee, Y. J., Kim, B. G., Chong, Y., Lim, Y., and Ahn, J. H.(2008) Cation dependent *O*-methyltransferases from rice. *Planta* **227** 641-647
- Li, J. Y., Oulee, T. M., Raba, R., Amundson, R. G., and Last, R. L.(1993) Arabidopsis flavonoid mutants are hypersensitive to Uv-B irradiation. *Plant Cell* **5** 171-179
- Lu, Shelly C.(2000) S-Adenosylmethionine. *Int. J. Biochem. Cell Bio.* **32** 391-395
- Luka, Z., Cerone, R., Phillips, J. A., Mudd, S. H., and Wagner, C.(2002) Mutations in human glycine N-methyltransferase give insights into its role in methionine metabolism. *Human Genetics* **110** 68-74
- Marti-Renom, M. A., Stuart, A. C., Fiser, A., Sanchez, R., Melo, F., and Sali, A.(2000) Comparative protein structure modeling of genes and genomes. *Ann. Rev. Biophys. Biomol. Struct.* **29** 291-325
- Martin, J. L. and McMillan, F. M.(2002) SAM (dependent) I AM: the S-adenosylmethionine-dependent methyltransferase fold. *Curr. Opin. Struct. Biol.* **12** 783-793
- Mccoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J.(2007) Phaser crystallographic software. *J. Appl. Crystallog.* **40** 658-674
- Menon, S. and Ragsdale, S. W.(1999) The role of an iron-sulfur cluster in an enzymatic methylation reaction - Methylation of CO dehydrogenase/acetyl-CoA synthase by the methylated corrinoid iron-sulfur protein. *J. Biol. Chem.* **274** 11513-11518
- Milkowski, C., Baumert, A., and Strack, D.(2000) Cloning and heterologous expression of a rape cDNA encoding UDP-glucose : sinapate glucosyltransferase. *Planta* **211** 883-886
- Morishige, T., Tsujita, T., Yamada, Y., and Sato, F.(2000) Molecular characterization of the S-adenosyl-L-methionine : 3'-hydroxy-N-methylcochlorine 4'-*O*-methyltransferase involved in isoquinoline alkaloid biosynthesis in *Coptis japonica*. *J. Biol. Chem.* **275** 23398-23405

- Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K., and Olson, A. J.(1998) Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.* **19** 1639-1662
- Morrison, R. T, Boyd, R. N., and Boyd, R. K.(1992) *Organic Chemistry*, 6th edition. **6th**
- Mudgett, M. B., Lowenson, J. D., and Clarke, S.(1997) Protein repair L-isoaspartyl methyltransferase in plants (Phylogenetic distribution and the accumulation of substrate proteins in aged barley seeds). *Plant Physiol.* **115** 1481-1489
- Murshudov, G. N., Vagin, A. A., and Dodson, E. J.(1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **53** 240-255
- Ndong, C., Anzellotti, D., Ibrahim, R. K., Huner, N. P. A., and Sarhan, F.(2003) Daphnetin methylation by a novel *O*-methyltransferase is associated with cold acclimation and photosystem II excitation pressure in rye. *J. Biol. Chem.* **278** 6854-6861
- Nelson, J. T., Lee, J., Sims, J. W., and Schmidt, E. W.(2007) Characterization of SafC, a catechol 4-*O*-methyltransferase involved in saframycin biosynthesis. *App. Environ. Microbiol.* **73** 3575-3580
- Noyer-Weidner, M. and Trautner, T. A.(1993) Methylation of DNA in prokaryotes. *EXS* **64** 39-108
- Otwinowski, Z. and Minor, W.(1997) Processing of X-ray diffraction data collected in oscillation mode. *Macromol. Crystallogr., Pt A* **276** 307-326
- Parvathi, K., Chen, F., Guo, D. J., Blount, J. W., and Dixon, R. A.(2001) Substrate preferences of *O*-methyltransferases in alfalfa suggest new pathways for 3-*O*-methylation of monolignols. *Plant J.* **25** 193-202
- Pessi, G., Kociubinski, G., and Mamoun, C. B.(2004) A pathway for phosphatidylcholine biosynthesis in *Plasmodium falciparum* involving phosphoethanolamine methylation. *Proc. Natl. Acad. Sci. U.S.A* **101** 6206-6211
- Poon, W. W., Barkovich, R. J., Hsu, A. Y., Frankel, A., Lee, P. T., Shepherd, J. N., Myles, D. C., and Clarke, C. F.(1999) Yeast and rat Coq3 and *Escherichia coli* UbiG polypeptides catalyze both *O*-methyltransferase steps in coenzyme Q biosynthesis. *J.Biol.Chem.* **274** 21665-21672
- Pradhan, S. and Esteve, P. O.(2003) Mammalian DNA (cytosine-5) methyltransferases and their expression. *Clin. Immunol.* **109** 6-16
- Rao, P. V., Garrow, T. A., John, F., Garland, D., Millian, N. S., and Zigler, J. S.(1998) Betaine-homocysteine methyltransferase is a developmentally regulated enzyme crystallin in rhesus monkey lens. *J. Biol. Chem.* **273** 30669-30674
- Rauh, D., Klebe, G., and Stubbs, M. T.(2004) Understanding protein-ligand interactions: The price of protein flexibility. *J. Mol. Biol.* **335** 1325-1341

- Ravanel, S., Block, M. A., Rippert, P., Jabrin, S., Curien, G., Rebeille, F., and Douce, R.(2004) Methionine metabolism in plants - Chloroplasts are autonomous for de novo methionine synthesis and can import S-adenosylmethionine from the cytosol. *J.Biol.Chem.* **279** 22548-22557
- Reyda, S., Sohn, C., Klebe, G., Rall, K., Ullmann, D., Jakubke, H. D., and Stubbs, M. T.(2003) Reconstructing the binding site of factor Xa in trypsin reveals ligand-induced structural plasticity. *J. Mol. Biol.* **325** 963-977
- Rosenberg, I. M.(2005) Protein analysis and purification benchtop techniques second edition. **2**
- Scalliet, G., Piola, F., Douady, C. J., Rety, S., Raymond, O., Baudino, S., Bordji, K., Bendahmane, M., Dumas, C., Cock, J. M., and Huguency, P.(2008) Scent evolution in Chinese roses. *Proc. Natl. Acad. Sci. U.S.A* **105** 5927-5932
- Schluckebier, G., Ogara, M., Saenger, W., and Cheng, X. D.(1995) Universal catalytic domain-structure of adomet-dependent methyltransferases. *J. Mol. Biol.* **247** 16-20
- Schroder, G., Wehinger, E., Lukacin, R., Wellmann, F., Seefelder, W., Schwab, W., and Schroder, J.(2004) Flavonoid methylation: a novel 4'-O-methyltransferase from *Catharanthus roseus*, and evidence that partially methylated flavanones are substrates of four different flavonoid dioxygenases. *Phytochemistry* **65** 1085-1094
- Schroder, G., Wehinger, E., and Schroder, J.(2002) Predicting the substrates of cloned plant O-methyltransferases. *Phytochemistry* **59** 1-8
- Stockigt, J. and Zenk, M. H.(1975) Chemical syntheses and properties of hydroxycinnamoyl coenzyme A derivatives. *Zeitschrift fur Naturforschung C-A Journal of Biosciences* **30** 352-358
- Stoscheck, C. M.(1990) Quantitation of protein. *Methods Enzymol.* **182** 50-68
- Strack, D., Keller, H., and Weissenbock, G.(1987) Enzymatic synthesis of hydroxycinnamic acid-esters of sugar acids and hydroaromatic acids by protein preparations from rye (*Secale-Cereale*) primary leaves. *J. Plant Physiol.* **131** 61-73
- Strahl, Brian D. and Allis, C. David(2000) The language of covalent histone modifications. *Nature* **403** 41-45
- Szegedi, S. S., Castro, C. C., Koutmos, M., and Garrow, T. A.(2008) Betaine-homocysteine S-methyltransferase-2 is an S-methylmethionine-homocysteine methyltransferase. *J. Biol. Chem.* **283** 8939-8945
- Thapar, N. and Clarke, S.(2000) Expression, purification, and characterization of the protein repair L-isoadipyl methyltransferase from *Arabidopsis thaliana*. *Protein Expression and Purif.* **20** 237-251
- Timinskas, A., Butkus, V., and Janulaitis, A.(1995) Sequence motifs characteristic for DNA [Cytosine-N4] and DNA [adenine-N6] methyltransferases - classification of all DNA methyltransferases. *Gene* **157** 3-11

- Topf, M. and Sali, A.(2005) Combining electron microscopy and comparative protein structure modeling. *Curr. Opin. Struct. Biol.* **15** 578-585
- Vance, D. E. and Ridgway, N. D.(1988) The methylation of phosphatidylethanolamine. *Prog. Lipid Res.* **27** 61-79
- Vidgren, J., Svensson, L. A., and Liljas, A.(1994) Crystal-structure of catechol *O*-methyltransferase. *Nature* **368** 354-358
- Vogt, T.(2004) Regiospecificity and kinetic properties of a plant natural product *O*-methyltransferase are determined by its N-terminal domain. *Febs Letters* **561** 159-162
- Vogt, T., Ibdah, M., Schmidt, J., Wray, V., Nimtz, M., and Strack, D.(1999) Light-induced betacyanin and flavonol accumulation in bladder cells of *Mesembryanthemum crystallinum*. *Phytochemistry* **52** 583-592
- Walker, J. B.(1960) Metabolic control of creatine biosynthesis. *J. Biol. Chem.* **235** 2357-2361
- Walsh, M. A., Evans, G., Sanishvili, R., Dementieva, I., and Joachimiak, A.(1999) MAD data collection - current trends. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **55** 1726-1732
- Wang, J. H. and Pichersky, E.(1999) Identification of specific residues involved in substrate discrimination in two plant *O*-methyltransferases. *Arch. Biochem. Biophys.* **368** 172-180
- Winkel-Shirley, B.(2002) Biosynthesis of flavonoids and effects of stress. *Curr. Opin. Plant Biol.* **5** 218-223
- Woodard, R. W., Tsai, M. D., Floss, H. G., Crooks, P. A., and Coward, J. K.(1980) Stereochemical course of the transmethylation catalyzed by catechol *O*-methyltransferase. *J. Biol. Chem.* **255** 9124-9127
- Wuthrich, K.(1990) Protein structure determination in solution by NMR spectroscopy. *J. Biol. Chem.* **265** 22059-22062
- Yabe, K., Ando, Y., Hashimoto, J., and Hamasaki, T.(1989) 2 Distinct *O*-methyltransferases in aflatoxin biosynthesis. *Appl. Environ. Microbiol.* **55** 2172-2177
- Ye, Z. H., Zhong, R., Morrison, W. H., III, and Himmelsbach, D. S.(2001) Caffeoyl coenzyme A *O*-methyltransferase and lignin biosynthesis. *Phytochemistry* **57** 1177-1185
- Zhang, Yi and Reinberg, Danny(2001) Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. *Genes and Development* **15** 2343-2360
- Zhong, R. Q., Morrison, W. H., Himmelsbach, D. S., Poole, F. L., and Ye, Z. H.(2000) Essential role of caffeoyl coenzyme A *O*-methyltransferase in lignin biosynthesis in woody poplar plants. *Plant Physiol.* **124** 563-577
- Zhong, R. Q., Morrison, W. H., Negrel, J., and Ye, Z. H.(1998) Dual methylation pathways in lignin biosynthesis. *Plant Cell* **10** 2033-2045

Zhou, J. M., Gold, N. D., Martin, V. J. J., Wollenweber, E., and Ibrahim, R. K.(2006) Sequential *O*-methylation of tricetin by a single gene product in wheat. *Biochim. Biophys. Acta* **1760** 1115-1124

Zhu, B. T.(2002) Catechol-*O*-methyltransferase (COMT)-mediated methylation metabolism of endogenous bioactive catechols and modulation by endobiotics and xenobiotics: Importance in pathophysiology and pathogenesis. *Current Drug Metabolism* **3** 321-349

Zhu, B. T., Ezell, E. L., and Liehr, J. G.(1994) Catechol-*O*-Methyltransferase-catalyzed rapid *O*-methylation of mutagenic flavonoids - metabolic inactivation as a possible reason for their lack of varcinogenicity *in-vivo*. *J.Biol.Chem.* **269** 292-299

Zubieta, C., He, Xian Zhi, Dixon, Richard A., and Noel, Joseph P.(2001) Structures of two natural product methyltransferases reveal the basis for substrate specificity in plant *O*-Methyltransferases. *Nat. Struct. Mol. Biol.* **8** 271-279

Zubieta, C., Kota, P., Ferrer, J. L., Dixon, R. A., and Noel, J. P.(2002) Structural basis for the modulation of lignin monomer methylation by caffeic acid/5-hydroxyferulic acid 3/5-*O*-Methyltransferase. *Plant Cell* **14** 1265-1277

Zubieta, C., Ross, J. R., Koscheski, P., Yang, Y., Pichersky, E., and Noel, J. P.(2003) Structural basis for substrate recognition in the salicylic acid carboxyl methyltransferase family. *Plant Cell* **15** 1704-1716

Acknowledgements:

I would like to thank and acknowledge all of the people who were involved in this work

My supervisors: Prof. Dr. Milton T. Stubbs and Dr. Thomas Vogt their guidance, helpful discussions and critical comments

Prof. Dieter Strack for allowing me to carry out my lab work at Department of Secondary Metabolism, Leibniz Institute for Plant Biochemistry in Halle.

Dr. Daniel Rauh for the solution of the structure of PFOMT

Dr. Piotr Neumann for his help and useful tips during all of the crystallographic procedures

Dr. Andrea Porzel for NMR analysis confirming the products of *para* methylation carried out by SynOMT

Dr. Juergen Schmidt and Dr. Willibald Schliemann for mass analysis of SynOMT reaction products.

Jörg M. Augustin for his help in translating the summary into German

Dr. Regine Herbst-Irmer (University of Göttingen, Göttingen, Germany) for her help in solving the „tinning” problem in the structure of SynOMT

Prof. Dr. Joseph P. Noel (The Salk Institute for Biological Studies, La Jolla, USA) and Jean-Luc Ferrer (Institut de Biologie Structurale, Grenoble, France) for the cDNA coding for of *M. sativa* CCoAOMT

Prof. Ragai Ibrahim (Department of Biology, Concordia University, Montreal, Canada) for the sample of flavone with trihydroxy system- Tricetin.

All the staff, researches and fellow students at the Institute for Biochemistry and Biotechnology of Martin-Luther-University Halle-Wittenberg and Department of Secondary Metabolism, Leibniz Institute for Plantbiochemistry in Halle for the friendly atmosphere, help and encouragement during my work.

Erklärung

Ich versichere, dass ich meine Dissertation

„ Substrate and Positional Specificity in Cation Dependent *O*-methyltransferases”

selbständig ohne unerlaubte Hilfe angefertigt und mich dabei keiner anderen als der von mir ausdrücklich bezeichneten Quellen bedient habe.

Die Dissertation wurde in der jetzigen oder einer ähnlichen Form noch bei keiner anderen Hochschule eingereicht und hat noch keinen sonstigen Prüfungszwecken gedient.

I, the undersigned, declare that this dissertation entitled:

„ Substrate and Positional Specificity in Cation Dependent *O*-methyltransferases”

has been solely composed by myself, and describes my own work, unless otherwise acknowledged in the text. No part of this work has been submitted to other higher learning institutions.

Halle 6.10.2008

Jakub Kopycki

CURRICULUM VITAE

Name (f/m/l): Jakub Grzegorz Kopycki
Date & place of birth: August 25, 1980, Lodz, POLAND
Nationality: Polish

Education:

2008-2005 PhD-student at Martin-Luther-Universität Halle-Wittenberg (joint project with Leibniz-Institut für Pflanzenbiochemie)
2004 Master of Science, Specialty: technical biochemistry
Biotechnology and Food Science Faculty, Technical University of Lodz,
Studies carried out in cooperation with International Faculty of
Engineering (language of instruction – English)
2002 One semester abroad at The University Of Basque Country (Socrates-
Erasmus exchange program)
1999 Enrolled at of Biotechnology and Food Science Faculty Technical
University of Lodz
1999-1998 Public High School no. XV Lodz, Poland.
1998-1997 High School, Oxford MS USA
1997-1995 Public High School no. XV Lodz, Poland.
1995-1994 Primary School no.166 Lodz, Poland
1994-1992 Oxford Junior High School, Oxford MS USA
1992-1987 Primary School no.166 Lodz, Poland

Professional experience:

2004 Quality control laboratory assistant at PPW IFOTAM Sp. z o.o., Sienkiewicza 112,
Lodz, Poland
2004 Two months practice at UFZ-Umweltforschungszentrum Leipzig-Halle GmbH
Permoserstraße 15, Leipzig, Germany
2003 Project on Aminoacid Analysis by HPLC, Faculty of Biotechnology and Food
Sciences Technical University of Lodz
Wólczańska 171/173 Lodz Poland
- Volunteer
2002 Two weeks practice at The Center for Microbiology and Virusology of Polish
Academy of Sciences,
ul. Lodowa 106 Lodz Poland

Papers and oral presentations:

- 2008 Kopycki, J. G., Rauh, D., Chumanevich, A. A., Neumann, P., Vogt, T., and Stubbs, M.T.
Biochemical and structural analysis of substrate promiscuity in plant Mg²⁺-dependent *O*-methyltransferases. (2008) *J. Mol. Biol.* 378 154-164
- 2008 Kopycki, J. G., Stubbs, M. T., Brandt, W., Hagemann, M., Porzel, A., Schmidt, J., Schliemann, W., Zenk, M. H., and Vogt, T.
Functional and structural characterization of a cation-dependent *O*-methyltransferase from the cyanobacterium *Synechocystis* SP. strain PCC 6803. (2008) *J. Biol. Chem.* 283 20888-20896.
- 2007 Presentation at The HEC X, 10th Heart of Europe Crystallography meeting
Titled: "Promiscuous and specific methylation by plant *O*-methyltransferases"
Bedlewo (2007) Poland, Abstract Book
- 2007 Poster at Botanikertagung in Hamburg
Titled: "Biochemical and structural analysis of substrate specific and promiscuous Mg²⁺ - dependent *O*-methyltransferases"
Hamburg (2007) Poster Abstract Book
- 2004 Poster at VII Graduate Student's Poster Session
Titled : " Crystal structure refinement of cytokinin binding protein in the form of complex with bromobenzyladenine"
Lodz (2004) Abstract Book