

# Dissertation

## Biotransformations from and to methylated flavo- noids

### *Subtitle*

Benjamin Weigel  
Leibniz-Institute of Plant Biochemistry  
Department of Bioorganic Chemistry  
Weinberg 3  
06120 Halle(Saale)  
October 21, 2015

Advisor: Prof. Dr. Ludger A. Wessjohann  
[wessjohann@ipb-halle.de](mailto:wessjohann@ipb-halle.de)  
+49 (345) 5582-1301

noch nicht bekannt

---

It is what it is. Accept it and move on.

– *unknown* –

# Contents

<b>I</b>	<b>Preface</b>	1
<hr/>		
<b>1</b>	<b>Abstracts</b>	2
1.1	English Abstract . . . . .	3
1.2	Deutsche Zusammenfassung . . . . .	3
<hr/>		
<b>II</b>	<b>Thesis</b>	4
<hr/>		
<b>2</b>	<b>Introduction</b>	5
2.1	Flavonoids . . . . .	6
2.1.1	General . . . . .	6
2.1.2	The phenyl propanoid pathway . . . . .	8
2.1.3	Biological activity . . . . .	8
2.2	Methyl transferases (MTs) . . . . .	9
2.2.1	General . . . . .	10
2.2.2	S-Adenosyl-L-methionine . . . . .	13
2.2.3	Methyl transferase mechanisms . . . . .	14
2.2.4	Plant O-methyl transferases (O-MTs) . . . . .	17
2.3	Alkylations and biotransformations . . . . .	19
2.4	Motivation . . . . .	23
<b>3</b>	<b>Material And Methods</b>	24
3.1	Materials . . . . .	25
3.1.1	Chemicals . . . . .	25
3.1.2	Commonly used solutions and buffers . . . . .	25
3.1.3	Culture media used to grow bacteria . . . . .	26
3.1.4	Bacterial strains . . . . .	27
3.1.5	Plasmids . . . . .	28
3.1.6	Oligonucleotides and synthetic genes . . . . .	28
3.1.7	Instruments . . . . .	29
3.1.8	Software . . . . .	30
3.2	Molecular Biology . . . . .	30
3.2.1	Golden Gate Cloning . . . . .	31
3.2.2	Subcloning of genes . . . . .	31
3.2.3	Transformation of electrocompetent <i>Agrobacterium tumefaciens</i> cells . . . . .	32

3.3 Treatment of plant material . . . . .	32
3.3.1 Infiltration of <i>Nicotiana benthamiana</i> . . . . .	33
3.3.2 Plant material harvest . . . . .	33
3.3.3 Extraction of flavonoids from <i>N. benthamiana</i> leaves. . . . .	33
3.4 Protein biochemistry . . . . .	34
3.4.1 Determination of protein concentration . . . . .	34
3.4.2 Protein production test (expression test) . . . . .	35
3.4.3 Protein subfractionation . . . . .	36
3.4.4 Protein sample concentration by TCA precipitation . . . . .	36
3.4.5 Preparation of periplasmic protein . . . . .	37
3.4.6 Discontinuous SDS-polyacrylamide gel electrophoresis (SDS-PAGE) . . . . .	37
3.4.7 Buffer change of protein samples . . . . .	38
3.4.8 Production of recombinant protein . . . . .	38
3.4.9 Preparation of inclusion bodies (IBs). . . . .	39
3.4.10 Purification of His-tagged proteins using immobilized metal affinity chromatography (IMAC) . . . . .	40
3.4.11 Refolding of SOMT-2 on a micro scale using design of experiments (DoE) . . . . .	41
3.4.12 Enzymatic production of SAM and SAE . . . . .	43
3.5 Crystallographic Procedures . . . . .	44
3.5.1 Crystallization of proteins . . . . .	44
3.5.2 Data collection and processing . . . . .	45
3.5.3 Structure solution. . . . .	46
3.5.4 Model building, refinement and validation. . . . .	46
3.5.5 <i>In silico</i> substrate docking . . . . .	47
3.6 Analytics . . . . .	47
3.6.1 Recording of growth curves . . . . .	47
3.6.2 <i>In vitro</i> determination of glucose . . . . .	48
3.6.3 <i>In vitro</i> O-methyl transferase (O-MT) assay . . . . .	48
3.6.4 Photospectrometric assay for the methylation of catecholic moieties. . . . .	51
3.6.5 Concentration of SOMT-2 using hydrophobic interaction chromatography (HIC) . . . . .	51
3.6.6 Analytical gel filtration. . . . .	53
3.6.7 Binding experiments using Isothermal Titration Calorimetry (ITC) . . . . .	53
3.6.8 High-performance liquid chromatography (HPLC) analytics .	54
3.6.9 Liquid chromatography-tandem mass spectrometry (LC-MS/MS) measurements. . . . .	54

<b>4 Engineering of phenylpropanoid and flavonoid O-methyl transferase (PFOMT)</b>	56
4.1 Introduction . . . . .	57
4.2 Crystallization of PFOMT . . . . .	59
4.3 Substrate binding studies using ITC. . . . .	63
4.4 Study of variants for long-chain alkylations . . . . .	66
4.5 Conclusion/Discussion . . . . .	68
4.6 Contributions . . . . .	70
<b>5 Tandem mass-spectrometry studies of flavonoids</b>	71
5.1 Introduction . . . . .	72
5.2 Fragmentation of flavanones . . . . .	75
5.3 Fragmentation of flavones . . . . .	78
5.4 Fragmentation of flavonols . . . . .	83
5.5 Conclusions and Discussion . . . . .	88
5.6 Contributions . . . . .	89
<b>6 Enzymatic methylation of Non-catechols</b>	90
6.1 Introduction . . . . .	91
6.2 SOMT-2 . . . . .	92
6.2.1 <i>In vivo</i> biotransformation in <i>Nicotiana benthamiana</i> . . . . .	92
6.2.2 <i>In vivo</i> biotransformation in <i>E. coli</i> . . . . .	94
6.2.3 <i>In vitro</i> studies using recombinantly produced SOMT-2 . . . . .	98
6.3 PFOMT . . . . .	103
6.3.1 Phenolic hydroxyls . . . . .	104
6.3.2 PFOMT pH-profiles are influenced by Mg <sup>2+</sup> . . . . .	105
6.3.3 Methylation of different chemical motifs. . . . .	109
6.4 Conclusion/Discussion . . . . .	114
<b>7 Acknowledgements</b>	116
<b>III Appendix</b>	118
<hr/>	
<b>A Figures</b>	119
A.1 Engineering of PFOMT . . . . .	120
A.2 Enzymatic methylation of non-catechols . . . . .	123
<b>B Tables</b>	128
B.1 Engineering of PFOMT . . . . .	130
B.2 Enzymatic methylation of non-catechols . . . . .	132

B.3 Tandem mass-spectrometry studies of flavonoids . . . . .	134
<b>C Affidavit</b>	138
<b>Acronyms</b>	161
<b>Glossary</b>	167

# List of Figures

2.1	The central feature of the flavonoids is the chromane ring. The names of the different groups of flavonoids are derived from the substitution of this moiety. From a biosynthetic point of view, flavonoids are built up from phenylpropanoid (blue) and acetate derived moieties (orange).	7
2.2	General pathways in the biosynthesis of flavonoids. PAL – phenylalanin ammonia lyase, C4H – cinnamate-4-hydroxylase, 4CL – 4-coumarate:CoA ligase, CHS – chalcone synthase, CHI – chalcone isomerase, F3H – flavanone-3-hydroxylase, FLS – flavonol synthase, FNS – flavone synthase, IFS – isoflavone synthase.	9
2.3	Topology plots of the five major structural classes of MTs and radical SAM methyl transferases (RSMTs) (modified and extended from Schubert <i>et al.</i> [170]). Helices are depicted as circles and $\beta$ -strands as triangles. The SAM binding site is depicted as a flag. RSMTs all share a common “radical SAM”-domain, which contains the iron sulfur cluster (red checker). The individual RSMT classes are differentiated according to the other domains they contain.	11
2.4	Reactions catalyzed by methyl transferases (MTs). Different shades of grey were used to differentiate between different groups of compounds and MTs. In contrast to other MTs, the group of RSMTs also requires additional co-factors to SAM.	12
2.5	Top: Non-radical SAM dependent methyl transferases catalyze the nucleophilic transfer of a methyl group from the donor SAM to a nucleophile (Nu; e.g. O,N,C,S). A proton ( $H^+$ ) is usually abstracted through a general base (B), to achieve activation of the nucleophile. The proton is later transferred to the aqueous medium. Bottom: The $S_N2$ reaction proceeds via a single transition state, during which the methyl-carbon is $sp^2$ hybridized. After transfer of the methyl to the nucleophile the now $sp^3$ hybridized carbon’s configuration is inverted.	15
2.6	Bioenzymatic synthesis of chiral acetate, the precursor for the synthesis of SAM carrying an assymetrical methyl, as used by Woodward <i>et al.</i> [219]. Ring hydroxyls of hexoses and pentoses are omitted for easier reading. HK – hexokinase, PGI – phosphoglucose isomerase, PFK – phosphofructokinase, ALD – aldolase, TIM –triose phosphate isomerase, G3PDH – glycerinaldehyde-3-phosphate dehydrogenase, PGM – phosphoglycerate mutase, ENO – enolase, Glu – glucose, F1,6BP – fructose-1,6-bisphosphate, 3PG – 3-phosphoglycerate, PEP – phosphoenol pyruvate	16

2.7 Natural products synthesized with the help of methyl transferases. <i>O</i> -MTs were used for the synthesis of ermanine ( <b>1</b> ), 7- <i>O</i> -methyl-aromadendrin ( <b>2</b> ) and 3,5-dimethoxy-4'-fluorostilbene ( <b>3</b> ). <i>N</i> -MT catalyze the production of epinephrine ( <b>4</b> ) and <i>N</i> -methyltetrahydroisoquinoline ( <b>5</b> ), whereas 3-(benzoylamino)-8-methylumbelliferone ( <b>6</b> ) and furfuryl-methyl-sulfide ( <b>7</b> ) can be produced by <i>C</i> -MT and <i>S</i> -MT . . . . .	21
2.8 Labelling of macromolecules by using a combination of novel alkine-derivatized SAM analogues and Cu <sup>I</sup> -catalyzed azide-alkyne 1,3-dipolar cycloaddition (CuAAC). Depending on the type of label used, it can be employed for detection (e.g. through fluorophores, coupled assays) or affinity purification (e.g. biotin). This technique is also feasable for use in activity based protein profiling (ABPP) approaches. . . . .	23
3.1 Golden Gate cloning scheme for SOMT-2 . . . . .	32
3.2 Pruned <i>N. benthamiana</i> plant, with two bottom and one top leaf, ready to be infiltrated. . . . .	34
3.3 Oxidation of the reporter substrate <i>o</i> -dianisidine. Consecutive one-electron transfers lead to the fully oxidized diimine form of <i>o</i> -dianisidine. The first electron transfer is believed to produce a charge transfer complex intermediate. [31, 90] . . . . .	49
3.4 Calibration curves of different relative compositions of ferulic acid to caffeic acid, that were taken as described in 3.6.4. The total concentration was always 0.4 mM. At lower pH values around 4, the method seems to overestimate the concentration of caffeic acid. However, the slope of the curves stays the same. . . . .	52
4.1 Some crystal and pseudo-crystal shapes that were observed during the crytsallization screen. a – high (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> , b-c – CaCl <sub>2</sub> , PEG-4000, e – LiCl, PEG-6000 . . . . .	60

4.2 An overview of the features in the <i>apo</i> -PFOMT structure. <b>a</b> – The assymetric unit of <i>apo</i> -PFOMT consists of two homodimers (4 monomers). Individual monomers are rainbow colored from N- (blue) to C-terminus (red). <b>b</b> – Comparison of 3C3Y (steelblue) and <i>apo</i> -PFOMT (green). The N-terminus of <i>apo</i> -PFOMT was resolved up to the N-terminus (red) and even the His-tag (red, transparent) was partly resolved. The N-terminus fits into a cleft on the surface of the 3C3Y structure, shown as a surface model on the right. SAH (white ball-and-sticks) and Ca <sup>2+</sup> (green sphere) are featured in the published structure, whereas a sulphate ion (red/yellow spheres) was bound in the newly solved structure. . . . .	62
4.3 Positional differences between the individual residues of the solved <i>apo</i> -PFOMT and the structure with bound SAH (pdb: 3C3Y). The diffraction precision indicator [45] (DPI) of the structures was (0.137 and 0.064) Å respectively. The overall rmsd amounted to 0.9034 Å. The secondary structure of apo-PFOMT is displayed at the top. Helices are displayed as rectangles and sheets are shown as arrows. Graphical background annotations are used to display the binding sites of SAH (green) and the metal ion (plum). The orange bars indicate regions, where much movement seems to happen upon binding or release of the co-substrate. The blue bar shows the region that was annotated as "insertion loop" in previous studies [108]. . . . .	63
4.4 The binding of different SAM analogues was measured via ITC. . . . .	64
4.5 ITC measurements of PFOMT:effector binding. <b>a</b> – Binding of SAH, SAM and SAE to PFOMT. <b>b</b> – SAH is injected into a PFOMT solution, with (red) or without (black) addition of Mg <sup>2+</sup> and caffeic acid. When Mg <sup>2+</sup> and caffeic acid were already present, the binding process seems to happen quicker, but is less enthalpic. <b>c</b> – Upon addition of caffeic acid to the protein heat is produced, however no sensible binding curve could be obtained. . . . .	65
4.6 The active site of PFOMT (pdb: 3C3Y). The outline of the protein backbone is displayed, with active site residues portrayed as colored sticks (cyan – F103, red – F80, turquoise – M52, yellow – Y51, white – F198, blue – W184, orange – N202, grey – as labelled). The co-substrate SAM (ball-and-stick model) was docked into the structure. . . . .	67
4.7 Activities of different PFOMT variants towards caffeic acid methylation. Colorations correspond to the ones used in Figure 4.6. . . . .	68

4.8 Comparison of the active sites of <b>a</b> – the solved <i>apo</i> -structure (green) and <b>b</b> – the ligand-bound structure (steelblue; pdb: 3C3Y). Waters are represented as small red spheres, calcium as a green sphere (complexing bonds are dashed) and SAH is displayed as a white ball-and-stick model. A possible hydrogen bond network (blue lines) for the ligand-bound state is displayed. . . . .	69
5.1 Ion fragment nomenclature of flavonoid aglycones as proposed by Ma <i>et al.</i> , illustrated on naringenin. Ions are labelled according to the ring they contain and the positions of the C ring that were broken. Thus $^{1,3}\text{A}^+$ , contains the ring A and bonds 1 and 3 of the C ring were broken. . . . .	75
5.2 Major fragmentation pathways of flavanones. Activation using CID conditions at 45 % NCE mainly results in neutral losses of $\text{H}_2\text{O}$ and ketene ( $\text{C}_2\text{H}_2\text{O}$ ) from the molecular ion $[\text{M}+\text{H}]^+$ (bold frame). These neutral losses are scarcely observed when HCD with a NCE of 75 % or 100 % is used for activation. Here, C-ring cleavages followed by neutral losses from the cleavage fragments are dominant. . . . .	76
5.3 Proposed $\text{MS}^2$ fragmentation of $^{1,3}\text{A}^+$ after HCD activation. In high energy $\text{MS}^2$ experiments, $^{1,3}\text{A}^+$ might lose two CO followed by an unusual $\text{C}_2\text{H}_4$ . A single loss of ketene ( $\text{C}_2\text{H}_2\text{O}$ ) to afford $\text{m/z}$ 111 is also sensible. . . . .	76
5.4 Comparison of CID and HCD $\text{MS}^2$ spectra of eriodictyol ( <b>2</b> ). <b>A</b> – CID at 45 % NCE. <b>B</b> – HCD at 75 % NCE. <b>C</b> – HCD at 100 % NCE. Four different prominent peaks are annotated in each spectrum. <b>D</b> – The shift to smaller masses in HCD spectra and with increasing NCE is illustrated by the boxplot of the distribution of peaks with relative intensities above 1 % in each of the above spectra. <b>E</b> – Relationship between the activation method and the intensity of four fragments (● $^{1,3}\text{A}^+$ , ▲ $(^{1,4}\text{B}^+-2\text{H})$ , ■ $(^{1,4}\text{B}^+-2\text{H}-\text{CO})$ , + $\text{C}_7\text{H}_5^+$ ) of different flavanones. . . . .	77
5.5 Major fragmentation pathways of unmethylated and methylated flavones. Multiple neutral losses of small molecules (e.g. CO, water or ketene) and 0/4 and 1/3 C ring cleavages are predominant in the $\text{MS}^2$ spectra of unmethylated flavones. Methylated flavones loose a methyl group in CID experiments, but only in HCD experiments do other fragmentation reaction become obvious. . . . .	79



5.11 Proposed pathways of fragmentation of isorhamnetin ( <b>15</b> ). Isorhamnetin might loose methyl, methane or methanol upon activation. A similar fragmentation pathway was proposed for the analogous chrysoe-riol (compound 5.8). Some fragmentations were observed in HCD mode only (box) . . . . .	87
6.1 Semi-synthetic pathway to naringenin and 4'-O-methyl naringenin in <i>N. benthamiana</i> . Enzymes not endogenous to <i>N. benthamiana</i> are in gray. PAL - phenylalanine ammonia lyase, C4H - cinnamic acid 4-hydroxylase, 4CL - 4-coumaric acid:CoA ligase, CHS - chalcone synthase, CHI - chalcone isomerase, SOMT2 - soy O-methyl transferase 2 . . . . .	92
6.2 Scatterplot of the first two principal components from the PCA of the HPLC data obtained from leaf material extracts. The samples are colored by leaf side (left/SOMT-2: red, right/vector control: cyan) . . . . .	94
6.3 Biotransformation methods as described by Kim <i>et al.</i> ( <b>A</b> ) and developed in this work ( <b>B</b> ). OD – optiocal density at 600 nm, LB – LB-medium, AI – AI-medium . . . . .	95
6.4 The masses resulting from the fragmentation into A- and B-ring along the C-ring (dashed line, <b>b</b> ) are evidence, that the 4'-hydroxyl on the B-ring is methylated by SOMT-2 . . . . .	97
6.5 SDS-PAGE of pET28a(+) SOMT-2 expressed in <i>E. coli</i> BL21(DE3) in autoinduction medium at different temperatures (shown above). The insoluble fractions show a protein band the same height as the 40 kDa marker band, which corresponds to the SOMT-2 protein (40 425 Da). M – protein size marker, S – soluble fraction, I – insoluble fraction . . . . .	99
6.6 SDS-PAGE of the insoluble and soluble fractions of the refolding reactions. Refolding reactions 2,3,8-11 seem to mainly produce misfolded insoluble protein, while the other refolding buffers (1,4-7,12) produce soluble protein. . . . .	101

6.7 Results of <i>in vitro</i> protein refolding trials. Measured data (left) is presented alongside the ME-plots (right). <b>a</b> – Protein concentration after refolding and rebuffering into a universal buffer. The highest yield of soluble protein was achieved in buffers 5 and 7. The ME-plots ( <b>b</b> ) illustrate the connection between a factor and the measured protein concentration, illustrating that high pH and arginine concentration might be beneficial in the refolding reactions. <b>c</b> – Calculated conversion of naringenin to ponciretin by the refolded protein fractions. Protein refolded in buffers 1 and 7 seem to afford the most active protein by conversion (~volume activity). The ME-plots for the conversion ( <b>d</b> ) show that the redox state (reducing) of the refolding environment is important to achieve active protein. . . . .	102
6.8 Initial rate/pH-profiles for the conversion of three different substrates (red – caffeic acid, green – eriodictyol, blue – <i>iso</i> -ferulic acid) by PFOMT. The non-catecholic substrate <i>iso</i> -ferulic acid is converted much less quickly than the catecholic substrates. . . . .	106
6.9 pH-profiles of substrate conversion. The linear regression models (blue, dashed lines) grasp the general trend of the data reasonably well to draw inferences. 95 % prediction intervals are displayed as shaded areas. . . . .	108
6.10 Substrates from four different groups were screened for methylation by PFOMT under different conditions (pH/Mg <sup>2+</sup> ). Figure refers to Table 6.6. . . . .	111
6.11 Conversion of multiple different substrates, catecholic and non catecholic, by PFOMT. pH/Mg <sup>2+</sup> color coded from light to dark: ✕/✗, ✕/✓, ✕/✗, ✕/✓. . . . .	112
A.1 Graphical representation of the work that has been done on MTs in combination with SAM analogues. The grey areas represent individual groups of SAM analogues (aliphatic, allylic, propargylic, aromatic, SeAM analogues, nitrogen analogues and miscelaneous others). The height of the grey areas represents the number of times a member of the corresponding group has been described as tested in the MT literature. The height of the colored bars represents the times that individual substrate has been tested. The colors represent the different types of MT (red – DNA-MT, green – P-MT, lilac – small molecule MT, blue – RNA MT). The black dash across the bar shows the number of times this substrate was actually converted by either enzyme. . . . .	120

A.2 Differences in the dihedrals $\psi$ (red) and $\phi$ (black) of the solved <i>apo</i> -PFOMT and the structure with bound SAH (pdb: 3C3Y). The secondary structure is displayed at the top. Helices are displayed as rectangles and sheets are shown as arrows. Graphical background annotations are used to display the binding sites of SAH (green) and the metal ion (plum). The orange bars indicate regions, where much movement seems to happen upon binding or release of the co-substrate. The blue bar shows the region that was annotated as "insertion loop" in previous studies. . . . .	121
A.3 Differences in the regioselectivity of some PFOMT variants. The products observed in HPLC and LC/MS measurements switched from 3'-methylated (dark grey) to 4'-methylated (light grey) for the displayed variants. The height of the bars corresponds to the total conversion of substrate. . . . .	122
A.4 Additional scatterplots of the PCA of HPLC data obtained from <i>N. benthamiana</i> leaves infiltrated by <i>A. tumefaciens</i> harbouring different constructs. <b>A</b> – samples colored by leaf position (top: red; bottom: cyan), <b>B</b> – samples colored by plant (plant 1: red; plant 2: green; plant 3: blue) . . . . .	123
A.5 Growth curve of <i>E. coli</i> BL21(DE3) expressing SOMT-2 at 37 °C. Glucose is depleted about 5 hours into growth, at which point the start of SOMT-2 expression is expected. The OD <sub>600</sub> after inoculation was about 0.1. . . . .	124
A.6 SDS-PAGE gel of samples acquired during growth curve measurements. The arrow indicated the band that could correspond to the GST-tagged SOMT-2 protein. . . . .	124
A.7 Graphical representation of a SOMT-2 model obtained from the PHYRE2 web portal ( <a href="http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index">http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index</a> ) [97] . . . . .	125
A.8 Chromatogram of the gel filtration analysis of refolded SOMT-2. . . . .	126
A.9 CD-spectrum of refolded SOMT-2 (black) compared to the spectrum that was calculated by the K2D3 web service ( <a href="http://cbdm-01.zdv.uni-mainz.de/~andrade/k2d3//index.html">http://cbdm-01.zdv.uni-mainz.de/~andrade/k2d3//index.html</a> ). Secondary structure estimates from the spectrum are 12.39 % $\alpha$ -helix and 32.51 % $\beta$ -sheet. . . . .	126
A.10 . . . . .	127

## List of Tables

2.1 Defining motifs of plant <i>O</i> -MTs as described by Joshi <i>et al.</i> [91]. . . . .	18
3.1 NADES-mixtures used within this work. . . . .	26
3.3 Plasmids used in this work. . . . .	28
3.4 Primers used in this work. Recognition sites for endonucleases are underlined. Positions used for site directed mutagenesis are in lower case font. . . . .	29
3.6 Calculated extinction coefficients of proteins used in this work.. . . . .	35
3.7 Factors used in the construction of the FrFD. . . . .	41
3.8 Experimental design matrix for the FrFD. . . . .	41
4.1 Results of fitting a simple one-site binding model to the data obtained from ITC experiments. . . . .	66
6.1 Naringenin and 4'-methylated derivatives that were inquired for in the plant samples via HPLC. The core structure of the compounds is displayed on the left. . . . .	93
6.2 <i>In vivo</i> biotransformation of different flavonoids, phenylpropanoids and anthraquinones by SOMT-2 in <i>E. coli</i> . Conversion ratios were calculated for samples taken after 30 hours. Multiple substrates containing a 4'-hydroxyl were methylated. Calculation of conversion percentages are only rough estimates, because of the nature of crude medium extracts. Products were determined by LC-MS/MS. † dihydro cinnamic ketone.	96
6.3 pK <sub>a</sub> -values of phenolic hydroxyl groups exemplified by <i>p</i> -cresole derivatives. Substituent positions on the aromatic ring are arbitrary and do not reflect conventions of the International Union of Pure and Applied Chemistry (IUPAC). . . . .	104
6.4 Maximum specific activity for the conversion of three different substrates with and without addition of magnesium. The pH at which the maximal activity was reached is indicated. . . . .	106
6.5 Coefficients of the model (6.2) for activity of <i>iso</i> -ferulic acid methylation. The factor Mg is a categorical variable (addition/no addition) and can therefore only be 0 or 1. . . . .	108

6.6 Substrate grid that was tested for methylation with PFOMT. Four different groups of compounds were screened. The groups of flavones, flavanones and cinnamic acids each contained one representative of each motif (phenolic, catecholic, 3'-hydroxy-4'-methoxy (3O4M), 4'-hydroxy-3'-methoxy (4O3M). The substitution patterns refer to Figure 6.10. . . . .	110
6.8 Conversion of substrates after 16 hours incubation. Only the maximum conversion is displayed, along with the conditions it was achieved under. <sup>†</sup> – wild-type: substrate conversion was maximal for all pH/Mg <sup>2+</sup> combinations. <sup>‡</sup> – conversion of caffeic acid by the wild-type was set to 100 %. . . . .	113
A.1 Results for the ANOVA of the main effects model describing soluble protein. Significance codes: 0.05 ( * ), 0.1 ( . ) . . . . .	123
A.2 Results for the ANOVA of the main effects model describing protein activity. Significance codes: 0.05 ( * ), 0.1 ( . ) . . . . .	125
B.1 Overview over the constructs produced for the present thesis. Each step during the production of the construct is given in the workflow steps column. Primers (italic font) or restriction sites used during each step are displayed in parenthesis. . . . .	129
B.3 SAM analogues that have been used with MTs. Targets: <i>P</i> – peptide/protein, <i>D</i> – DNA, <i>R</i> – RNA, <i>S</i> – small molecule. . . . .	130
B.4 Crystallographic data, phasing and refinement statistics. . . . .	132
B.5 Coefficients of the model (6.3) for activity of catecholic methylation. The factor <b>Mg</b> is a categorical variable (addition/no addition) and can therefore only be 0 or 1. . . . .	133
B.7 Coefficients obtained for linear regression model suing the <i>iso</i> -ferulic acid subset after shrinkage using the Lasso method and 5-fold cross validation. Only non-zero coefficients (variables actually do have an effect) are retained during the Lasso. Seed was set to 1337. . . . .	133
B.9 Coefficients obtained for linear regression model using the catechols subset after shrinkage using the Lasso method and 5-fold cross validation. Only non-zero coefficients (variables actually do have an effect) are retained during the Lasso. Seed was set to 1336. . . . .	133
B.11 Key ions in the positive mode CID and HCD ESI-MS <sup>2</sup> spectra of flvanones. . . . .	135

B.12 Key ions in the positive mode CID and HCD ESI-MS <sup>2</sup> spectra of flavones.	136
B.13 Key ions in the positive mode CID and HCD ESI-MS <sup>2</sup> spectra of flavonoles. . . . .	137

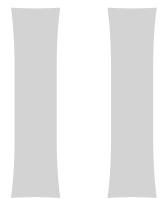


# Preface

# 1 Abstracts

## **1.1 English Abstract**

## **1.2 Deutsche Zusammenfassung**



# Thesis

## 2 Introduction

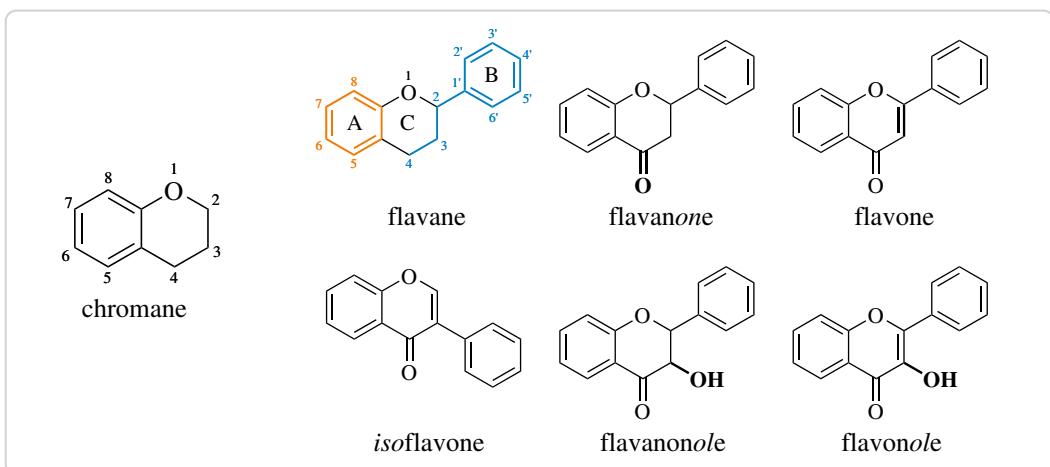
Secondary metabolites comprise a vast collection of organic compounds produced in nature, that do not directly parttake in the growth and development of an organism. Many functions of these natural products are unknown and remain to be elucidated, but it has been shown that they can be used for gene regulation, defense against biotic and abiotic stresses, (pollinator) attractant, communication and others. Natural compounds can be quite complex and show remarkable biological activities. The major classes of secondary compounds are terpenoids, alkaloids, phenylpropanoids including lignans/lignins, flavonoids and polyketides. This work is largely concerned about phenyl propanoids and flavonoids, as well as their modification, and will therefore mainly focus on these compounds.

## 2.1 Flavonoids

### 2.1.1 General

Plant phenolic compounds account for more than 40 % of the organic carbon in the biosphere and are essential for the survival of vascular plants. They are largely derived from the *phenylpropanoid* and relating pathways and take on various structural (e.g. cell walls) and non-structural roles (e.g. plant defense, flower color) [38]. The name *phenylpropanoid* describes the aromatic phenyl connected to a three-carbon chain, which biosynthetically originates from phenylalanine. Flavonoids, from the Latin *flavus* (yellow), are a diverse subclass of these phenolic compounds comprising more than 4500 different compounds described thus far and their main structural feature is the central chromane (benzodihydropyran) moiety (Figure 2.1). They consist of three rings named A, B and C. Ring A and B are of acetate and phenylpropanoid origin respectively, whereas ring C is a result of the condensation of the former.

Different types of flavonoids are named depending on the substitution pattern of the chromane ring. For example, a phenyl group at C-2 or C-3 gives flavonoids or isoflavonoids respectively, an unsubstituted C-4 means flavane, whereas a carbonyl group at C-4 indicates flavanones *et cetera*. Flavonoids are usually poly-hydroxylated, but can also carry multiple other different substitutions.



**Figure 2.1.**: The central feature of the flavonoids is the chromane ring. The names of the different groups of flavonoids are derived from the substitution of this moiety. From a biosynthetic point of view, flavonoids are built up from phenylpropanoid (blue) and acetate derived moieties (orange).

O-methylations are common at all hydroxyl positions, but flavonoids can also be C-methylated [10]. Other common derivatizations are (*O* or *C*)-prenylation, (*O* or *C*)-glycosylation, methylene-dioxy bridges (C-3'/C-4' or C-6/C-7) and various (*O* or *C*)-acylations (aliphatic and aromatic acids) [41, 76, 180, 217].

In plants flavonoids are usually produced to combat biotic or abiotic stresses. They can absorb highly energetic ultra violet (UV) light, suppress the formation of, or scavenge reactive oxygen species (ROS), once formed [2]. Furthermore, flavonoids can act as regulators during plant development [187]. A growing interest in flavonoids for the use in medicinal and nutritional applications has been spiked by their beneficial effects on health. Flavonoids possess a high antioxidant activity and also show protective effects against age-related ailments, such as cardiovascular diseases and cancers. Furthermore, they show anti-inflammatory, hepatoprotective, antimicrobial and antiviral activities [110].

A number of flavonoids are produced by the valorization of wastes and by-products of the food industry. Citrus and olive processing byproducts are especially rich in polyphenols [72, 149, 153]. However, many flavonoids are scarce in nature and/or the production from by-products is not enough to saturate the market demand, thus requiring different approaches for production. Recent developments in

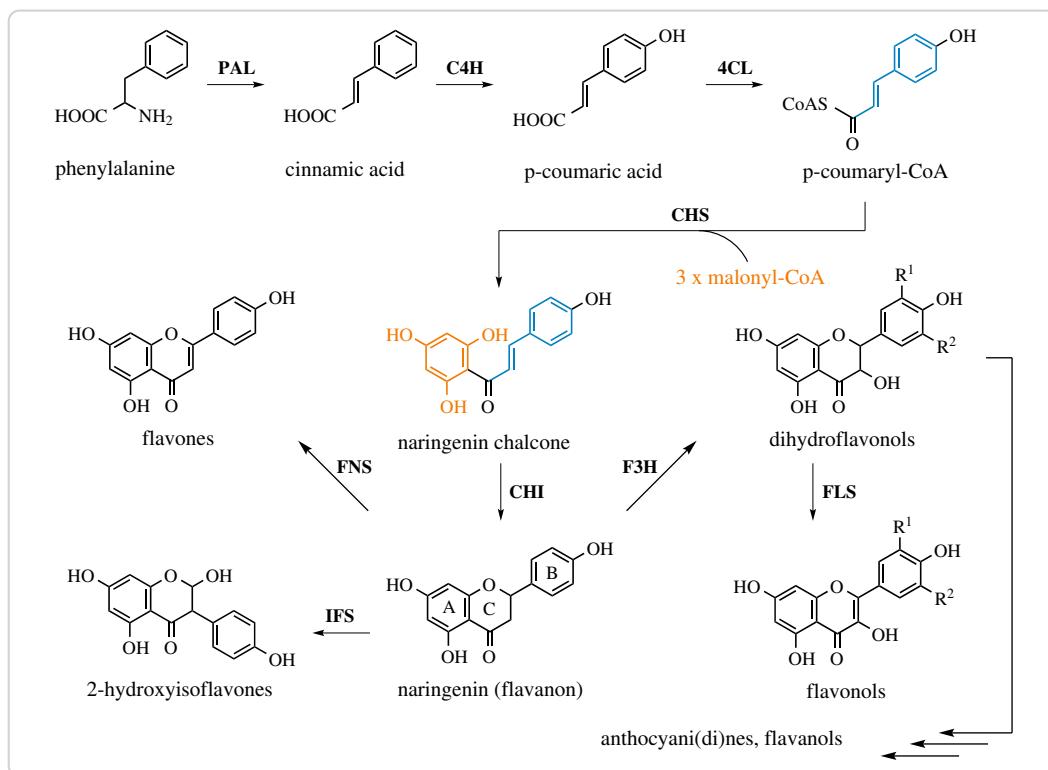
the field of metabolic engineering allowed for the high-level production of many flavonoids in microbial hosts, such as *Escherichia coli* or *Saccharomyces cerevisiae* [194]. For example, eriodictyol was produced from tyrosine in metabolically engineered *E. coli* at levels of up to 107 mg/ml [229], whereas naringenin was produced in *S. cerevisiae* from glucose at levels of 109 mg/ml [107].

### 2.1.2 The phenyl propanoid pathway

Biosynthesis of flavonoids via the phenylpropanoid pathway starts from phenylalanine, which is non-oxidatively deaminated by phenylalanine ammonia-lyase (PAL) to yield cinnamic acid (Figure 2.2) [74, 124]. Cinnamate-4-hydroxylase (C4H), a P450 monooxygenase, hydroxylates the cinnamic acid at the *para*-position and 4-coumarate:CoA ligase (4CL) converts the *p*-coumaric acid to its corresponding coenzyme A (CoA)-ester [80, 197]. Chalcone synthase (CHS) uses 3 molecules of malonyl-CoA (produced from acetyl-CoA by acetyl-CoA carboxylase) to produce naringenin chalcone from *p*-coumaryl-CoA [59]. Next, the linear chalcone can cyclize spontaneously or catalyzed by chalcone isomerase (CHI) via a MICHAEL-type addition to afford the flavanone naringenin [87]. Naringenin can serve as substrate for numerous enzymes such as flavanone-3-hydroxylase (F3H), flavone synthase (FNS) or isoflavone synthase (IFS) to afford dihydroflavonols, flavones or 2-hydroxyisoflavones respectively [69]. Dihydroflavonols are again precursors for the biosynthesis of flavonols, flavanols and anthocyanidines.

### 2.1.3 Biological activity

Flavonoids possess many properties associated with a healthy diet. They act as antioxidants and can help reduce oxidative stress. Several mechanisms that might be involved in the antioxidant activity of flavonoids are currently discussed. They can act as scavengers for free ROS or mask metal ions by chelation to suppress the production of radicals [27, 118]. The substitution pattern plays an important role in the antioxidant activity. Generally, the more free hydroxyls are present, the stronger the antioxidant activity of the flavonoid is [27]. Hydroxyls can donate an electron,



**Figure 2.2.:** General pathways in the biosynthesis of flavonoids. PAL – phenylalanine ammonia lyase, C4H – cinnamate-4-hydroxylase, 4CL – 4-coumarate:CoA ligase, CHS – chalcone synthase, CHI – chalcone isomerase, F3H – flavanone-3-hydroxylase, FLS – flavonol synthase, FNS – flavone synthase, IFS – isoflavone synthase.

or a hydrogen atom to free ROS to inactive such molecules. The resulting flavonoid radicals are stabilized by resonance, however possess prooxidant properties [142].

Numerous flavonoids possess antimicrobial activities [39]. For example, catechins from green and black teas have been shown to be effective against *Bacillus cereus* at nanomolar concentrations [64]. The mechanisms of action can be the inactivation of enzymes, binding of adhesins, membrane disruption or cell wall complexation [35].

## 2.2 Methyl transferases (MTs)

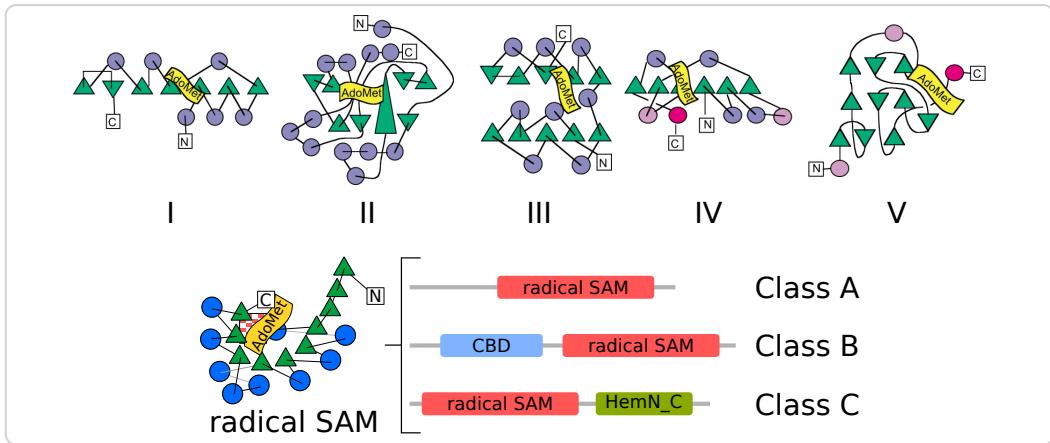
### 2.2.1 General

*S*-adenosyl-L-methionine (SAM)-dependent methyl transferases (MTs) (EC 2.1.1.x) transfer the methyl group of SAM to an activated atom of an acceptor molecule, via an S<sub>N</sub>2 displacement mechanism. SAM is converted to *S*-adenosyl-L-homocysteine (SAH), the co-product of the reaction, in the process. There are currently over 300 manually annotated MTs, each catalyzing a different reaction, included in the UniProtKB/Swiss-Prot (<http://www.uniprot.org>) database. Transfer of the methyl group to oxygen and nitrogen atoms is most common, but carbon, sulfur, selenium, arsenic atoms and even halide ions can be methylated (Figure 2.4)[167, 189]. Acceptor molecules are diverse and range from relatively small natural products (e.g. flavonoids) to bio-macromolecules such as nucleic acids or proteins. In fact, MTs are key-tailoring enzymes for many natural products of all groups (e.g. flavonoids, alkaloids or non-ribosomal peptides) [99, 108, 184, 207]. These small molecule methyl transferases (*sm*MTs) account for a significant part of the diversity present in natural products.

Other MTs, such as protein methyl transferases (P-MTs), DNA methyl transferases (DNA-MTs) and RNA methyl transferases (RNA-MTs) methylate proteins and nucleic acids respectively. In eukaryotes, DNA-MTs and P-MTs play important roles in the epigenetic regulation of gene expression and have been associated with a number of cancers and other diseases [34, 159, 160]. In bacteria, DNA-MTs are an essential part of the restriction modification system [138].

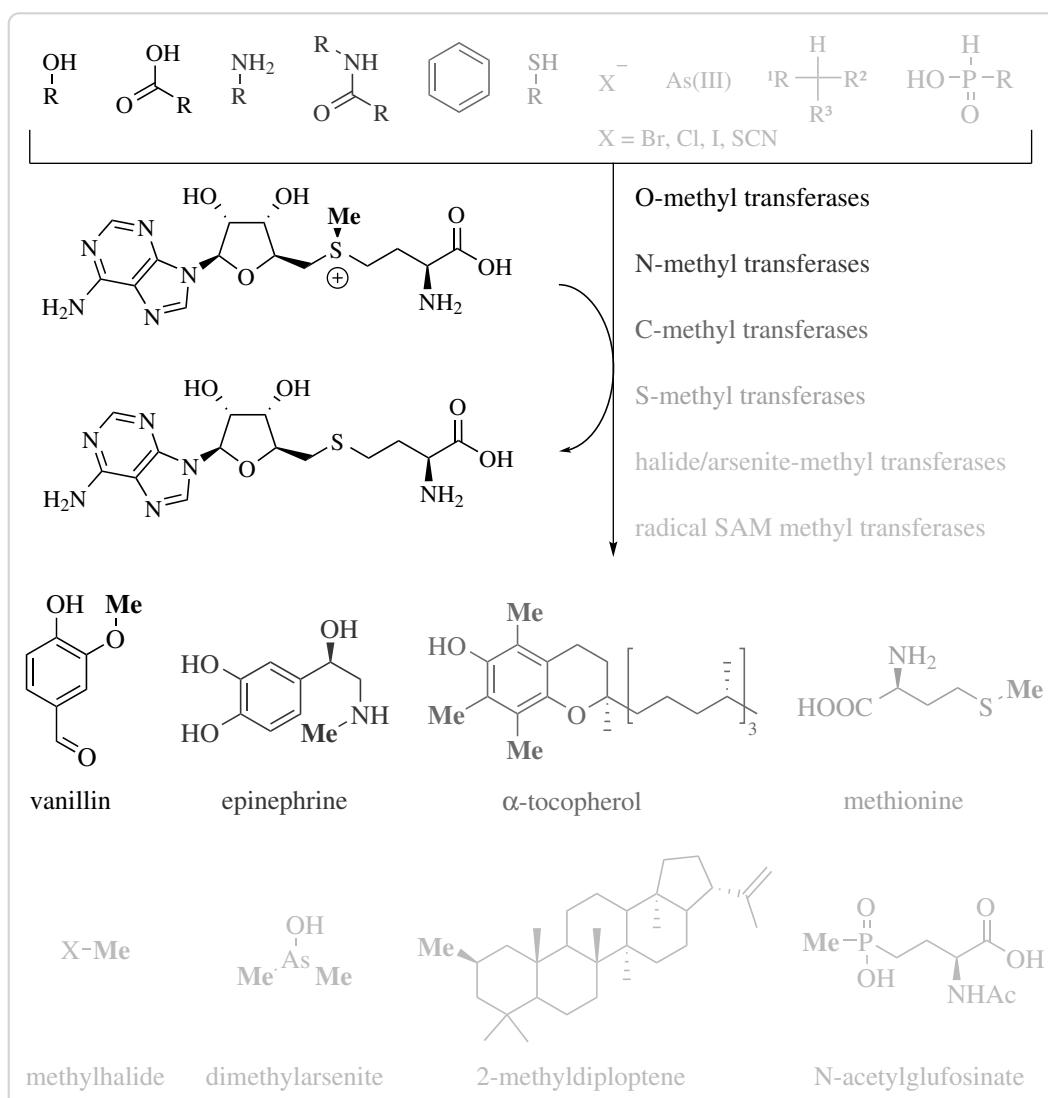
According to their structure, MTs can be classified into five main groups (I–V) (Figure 2.3)[170]. Class I MTs are the largest group of MTs and are characterized by a central Rossmann-like  $\alpha\beta\alpha$  sandwich, consisting of a seven-stranded  $\beta$ -sheet flanked by  $\alpha$ -helices. Most *sm*MTs, DNA-MTs and some P-MTs belong to class I MTs. Even though some of the enzymes belonging to class I share as little as 10 % sequence similarity, there is a pronounced structural conservation [170]. Class II MTs comprise a long anti-parallel  $\beta$ -sheet encompassed by numerous  $\alpha$ -helices [50]. In class III MTs the SAM binding site is located between two  $\alpha\beta\alpha$  domains [169]. A knot structure at the C-terminus contributes to SAM binding in the *SpoU*-TrmD (SPOUT) family of class IV RNA-MTs [143]. Protein lysine MTs make up the

largest part of P-MTs and structurally belong to class V MTs containing a suvar3-9, enhancer-of-zeste, trithorax (SET) domain [220]. Interestingly, a recent study of the



**Figure 2.3.:** Topology plots of the five major structural classes of MTs and radical SAM methyl transferases (RSMTs) (modified and extended from Schubert *et al.* [170]). Helices are depicted as circles and  $\beta$ -strands as triangles. The SAM binding site is depicted as a flag. Radical SAM methyl transferases (RSMTs) all share a common “radical SAM”-domain, which contains the iron sulfur cluster (red checker). The individual RSMT classes are differentiated according to the other domains they contain.

methyltransferome of baker’s yeast (*S. cerevisiae*) argues, that there are four more different classes of MTs [214]. Opposite to the studies by Schubert *et al.*, this work mainly relied on bioinformatical methods for structural information. 83 out of 86 MT structures in total were homology-modelled. It was shown, that two thirds of the reviewed yeast MTs belonged to class I. However, four new folding architectures, namely SSo0622-like, all- $\beta$ , all- $\alpha$  (RNA/DNA 3-helical bundle) and transmembrane, were postulated. Radical SAM methyl transferases (RSMTs) comprise another class of recently discovered MTs that contain an iron-sulfur ([4Fe-4S])-cluster coordinated by a three cysteine CxxxCxxC motif. RSMTs methylate unreactive centers through a radical mechanism [206]. Structural evidence suggests, that the mechanism is initiated by reductive cleavage of SAM into a reactive 5'-deoxyadenosyl (dAdo) radical by the [4Fe-4S] cluster [13, 24]. Three distinct classes (A, B, C) with distinct structural and mechanistic characteristics have been recognized within the RSMTs [226]. The centerpiece of RSMTs is the *radical SAM* domain, whose structure was



**Figure 2.4.: Reactions catalyzed by methyl transferases (MTs).** Different shades of grey were used to differentiate between different groups of compounds and MTs. In contrast to other MTs, the group of RSMTs also requires additional co-factors to SAM.

first described in the ribosomal ribonucleic acid (rRNA) methyl transferase RlmN of *E. coli* [13]. This domain consists of an  $\alpha_6/\beta_6$  partial barrel and contains the [4Fe-4S] cluster, as well as the SAM binding site (Figure 2.3). Class A only contains the radical SAM domain and includes mainly rRNA methyl transferases. In addition to the radical SAM domain, an N-terminal cobalamin binding domain (CBD) is proposed to be contained in RSMTs of class B. Class B RSMTs methylate numerous substrates at unreactive  $sp^3$  carbon centers, heterocycles and phosphinates. Class C RSMTs most likely contain a C-terminal domain related to the coproporphyrinogen III oxidase HemN in addition to the radical SAM domain [114]. Class C enzymes methylate aromatic heterocycles.

Radical SAM chemistry within enzymes is not confined to just methyl transfer. Instead, it has been shown that this type of chemistry is important for a number of rearrangement, cyclization, dehydrogenation, bond-formation and bond-cleavage reactions in nature [24].

### 2.2.2 **S-Adenosyl-L-methionine**

S-adenosyl-L-methionine (SAM), first described in 1953 [26], is the universal co-substrate for all SAM dependent methyl transferases. However, it is not only involved in methyl transfer, but is essential for a myriad of other reactions [24]. This makes SAM the second most ubiquitous co-substrate after adenosine triphosphate (ATP).

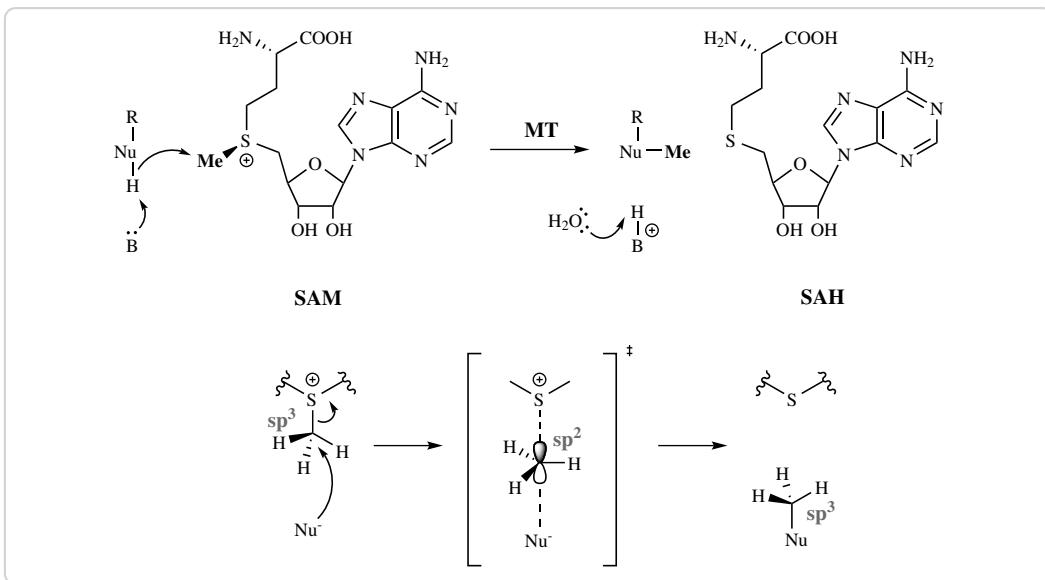
The methyl group of SAM is partially positively charged due to its position at the sulfonium center and is in consequence highly activated. The increased electrophilicity of the methyl group makes it a strong alkylation agent. Furthermore, SAH is a good leaving group. Therefore, nucleophilic transfer of the methyl group of SAM is thermodynamically highly favoured ( $\Delta G^0 \approx -70$  kJ/mol for SAM + homocysteine  $\rightarrow$  SAH + methionine) and allows the rapid and selective methylation of a range of substrates [170]. The fact that the methyl group is the least sterically hindered of all transferable carbon groups makes a methyl transfer the kinetically most favourable  $S_N2$  reaction (disregarding nucleophile and leaving group). Despite its apparent reactivity, SAM is still quite stable at physiological conditions com-

pared to other sulfonium species like the trimethylsulfonium ion, which quickly reacts with nucleophiles and is often used for derivatization prior to GC analytics [25]. Meanwhile, SAM is readily cleaved into adenine and S-ribosylmethionine under alkaline conditions [16] and other deteriorating processes such as racemization and intramolecular cleavage are to be reckoned with [77].

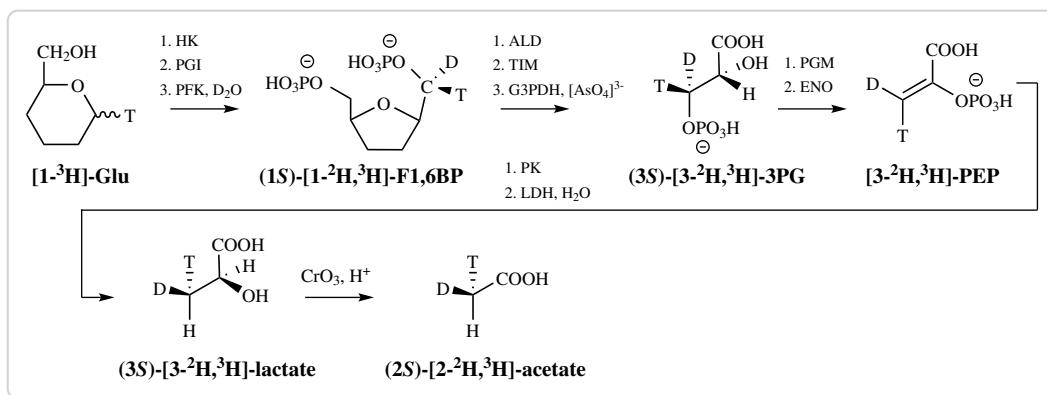
SAM is produced by the enzyme SAM synthetase (EC 2.5.1.6) from methionine and ATP in a two step reaction [186]. At first SAM is formed and the triphosphate group of ATP is cleaved off. Then, the inorganic triphosphate is hydrolyzed to monophosphate and diphosphate after which the products are released. The demethylated SAM is called *S*-adenosyl-L-homocysteine (SAH). SAH is a common side product of all SAM dependent MTs and can be further cleaved by SAH hydrolase (EC 3.3.1.1) to afford homocysteine and adenosine [126]. The cobalamin (vitamin B<sub>12</sub>) dependent methionine synthase (EC 2.1.1.13) can remethylate homocysteine to methionine using N<sup>5</sup>-methyltetrahydrofolate as a methyl donor [9]. Taken together, reactions leading from and to SAM are commonly called the activated methyl cycle.

### 2.2.3 Methyl transferase mechanisms

Non-radical SAM-dependent MTs catalyze the transfer of the methyl group of SAM to an activated nucleophile. The methylation reaction proceeds via a single displacement S<sub>N</sub>2 mechanism, through an sp<sup>2</sup> hybridized transition state and results in the inversion of configuration (Figure 2.5). The S<sub>N</sub>2 mechanism was proposed early on [73, 144], but only with the development of the chiral methyl group methodology extended mechanistic studies were made possible [61, 62]. An elegant method for the synthesis of chiral acetate made use of glycolytic enzymes to convert [1-<sup>3</sup>H]-glucose via its glycolysis intermediates to [3-<sup>2</sup>H,<sup>3</sup>H]-lactate, which is subsequently oxidized by chromiumtrioxide (Figure 2.6) [219]. The chirality of the resulting acetate can be controlled by the solvents (D<sub>2</sub>O or H<sub>2</sub>O) used during the enzymatic reactions. The chiral acetate can be used for the synthesis of e.g. [*methyl*-<sup>2</sup>H,<sup>3</sup>H]-methionine.



**Figure 2.5.: Top:** Non-radical SAM dependent methyl transferases catalyze the nucleophilic transfer of a methyl group from the donor SAM to a nucleophile (Nu; e.g. O,N,C,S). A proton ( $H^+$ ) is usually abstracted through a general base (B), to achieve activation of the nucleophile. The proton is later transferred to the aqueous medium. **Bottom:** The  $S_N2$  reaction proceeds via a single transition state, during which the methyl-carbon is  $sp^2$  hybridized. After transfer of the methyl to the nucleophile the now  $sp^3$  hybridized carbon's configuration is inverted.



**Figure 2.6.:** Bioenzymatic synthesis of chiral acetate, the precursor for the synthesis of SAM carrying an assymetrical methyl, as used by Woodward *et al.* [219]. Ring hydroxyls of hexoses and pentoses are omitted for easier reading. HK – hexokinase, PGI – phosphoglucose isomerase, PFK – phosphofructokinase, ALD – aldolase, TIM – triose phosphate isomerase, G3PDH – glycerinaldehyde-3-phosphate dehydrogenase, PGM – phosphoglycerate mutase, ENO – enolase, Glu – glucose, F1,6BP – fructose-1,6-bisphosphate, 3PG – 3-phosphoglycerate, PEP – phosphoenol pyruvate

In 1980, Woodward *et al.* fed *R*- and *S*-methionine containing an assymetrical methyl group to cultures of *Streptomyces griseus* and found, that the enzymatic transfer of two methyl groups (*N* and *C*-methylation) during the indolmycin biosynthesis proceeded with inversion of the configuration, strongly implying an  $\text{S}_{\text{N}}2$  mechanism [219]. This experiment also demonstrated that, *in vivo*,  $[\text{methyl-}^2\text{H}, 3\text{-}^3\text{H}]\text{-methionine}$  is converted to  $[\text{methyl-}^2\text{H}, 3\text{-}^3\text{H}]\text{-SAM}$  before the methyl group is transferred by a MT. *In vitro* experiments that were conducted using catechol O-methyl transferase (COMT) further supported the hypothesis of an  $\text{S}_{\text{N}}2$  mechanism [218].

The nucleophile attacking the methyl group of SAM is usually activated by abstraction of a proton through a general base (e.g. histidine, lysine) [22, 230, 231] and/or with the help of a lewis acid such as complexed  $\text{Mg}^{2+}$  [108, 200]. In a bimolecular reaction, the rate is highly dependent on the concentration of both compounds.

$$\text{rate} = k[A][B]$$

Low concentrations mean a low rate. This is in part due to the fact, that bimolecular reactions are entropically disfavoured. MTs (and enzymes in general) strongly in-

crease the effective concentration of each compound and thereby decrease entropy, in that methyl donor and acceptor are bound (immobilized) in close proximity to each other in the active site [60]. The right positioning of the substrates is a major factor for efficient catalysis and enzymes go through great lengths to achieve optimal alignment of substrates. One remarkable example are DNA-MTs, which can flip the target nucleotide out of the DNA-helix to provide the best orientation of SAM and the acceptor nucleophile [103].

Since the reaction catalyzed by MTs is a two substrate reaction, kinetics-driven mechanistic studies have been done on a number of different MTs to show the binding mode of the substrates. It turns out, that the reaction mechanism varies for MTs from different organisms and classes. There is no one mechanism that describes every MT. A random bi-bi binding mechanism is for example exhibited by rat liver COMT [36], CheR protein-L-glutamate O-MT from *Salmonella typhimurium* [179] and the protoporphyrin IX O-MT of *Rhodobacter capsulatus* [166]. The protoporphyrin IX MT from etiolated wheat *Triticum aestivum* on the other hand shows a ping-pong bi-bi mechanism [222], whereas ordered bi-bi mechanisms were shown for the cytosine DNA-MT MSPI and isoprenylated P-MT [12, 176]. Meanwhile, the enzymes exhibiting ordered bi-bi mechanisms were different in that one bound SAM first and released SAH last [176], whereas the other one bound the acceptor molecule first [12]. Competitive product inhibition, especially by SAH, is commonly observed for MTs [8, 12, 176, 179].

The mechanisms of RSMTs are outside the scope of this work, but the interested reader is instead referred to current reviews on the topic [24, 226].

#### 2.2.4 Plant O-methyl transferases (O-MTs)

Plant O-methyl transferases (O-MTs) were the prime interest of this work. As such, plant O-MTs represent a large group of plant enzymes that catalyze the transfer of a methyl group to a hydroxyl or carboxyl group of phenylpropanoids, flavonoids or alkaloids. O-methylation greatly effects the (bio)-chemical properties of a compound and can for example have profound influences on reactivity, solubility, bioavailability, antimicrobial or antioxidant activities.

Plant *O*-MTs are subdivided into two groups according to their size and the spatial relationship between three highly preserved motifs (Table 2.1) [91]. Group I members, containing caffeoyl CoA dependent *O*-methyltransferase (CCoAOMT)-like representatives, are usually between 110 and 140 amino acid residues shorter than group II members ( $\approx$ 340–390 amino acids). The distance between motifs A and

**Table 2.1.**: Defining motifs of plant *O*-MTs as described by Joshi *et al.* [91].

motif	consensus	distance to motif ...		
		...	group I	group II
A	(V,I,L)(V,L)(D,K)(V,I)GGXX(G,A)	B	19	52
B	(V,I,F)(A,P,E)X(A,P,G)DAXXXX(W,Y,F)	C	24	30
C	(A,P,G,S)(L,I,V)(A,P,G,S)XX(A,P,G,S)(K,R)(V,I)(E,I)(L,I,V)			

B, and between B and C is also shorter in group I members, than in group II members. In contrast to group II, group I plant *O*-MTs require Mg<sup>2+</sup> for activity. Overall, they are fairly similar to mammalian COMTs [91]. Group II plant *O*-MTs can methylate a variety of substrates, whereas group I plant *O*-MTs are usually very strict in their substrate scope utilizing only a couple of substrates. However, some enzymes from group I are much more relaxed with their acceptance of substrates. For example, phenylpropanoid and flavonoid *O*-methyl transferase (PFOMT) from the ice-plant *Mesembryanthemum crystallinum* and an *O*-MT from chickweed *Stellaria longipes* can utilize several substrates [85, 227].

**Phenylpropanoid and flavonoid *O*-methyl transferase (PFOMT)** is a Mg<sup>2+</sup>-dependent class I plant *O*-MT from the ice plant *M. crystallinum* and was first described in 2003 by Ibdah *et al.* [85]. PFOMT was the first class I MT that provided evidence showing, that methylation of flavonoids is not only restricted to class II plant *O*-MTs. It belongs to a subgroup of class I plant *O*-MTs, that is distinguished from caffeoyl CoA dependent *O*-methyltransferase (CCoAOMT) by a lower sequence homology and a broader substrate promiscuity and regiospecificity. PFOMT methylates a number of flavonoids and phenyl propanoids, preferably at the *meta*-position, provided a catecholic moiety is present. Enzyme purified from its native source *M. crystallinum* is truncated N-terminally by 11 amino

acids, although there is no known signaling sequence [202]. This truncation has deleterious effects on the catalytic efficiency, especially towards substrates such as caffeoyl glucose and caffeoyl-CoA, but also influences the regioselectivity. There is only speculation as to the purpose of this N-terminal truncation *in vivo*, but metabolic regulation is plausible.

The three dimensional structure of PFOMT shows a dimeric protein (pdb: 3C3Y) [108]. Each monomer exhibits a Rossmann  $\alpha/\beta$ -fold consisting of 8  $\alpha$ -helices and 8  $\beta$ -strands. The catalytically important N-terminus is not resolved in the structure. SAH and  $\text{Ca}^{2+}$  were cocrystallized and appear bound in the active site.  $\text{Ca}^{2+}$  is complexed by two aspartate and one asparagine residues with the rest of the coordination spaces occupied by waters.

**SOMT** soy O-methyl transferase (SOMT-2) has been described in the literature to methylate multiple flavonoids at the 4'-position of the B-ring [98, 101, 102]. It has the highest activity towards naringenin, to produce ponciretin (also known as isosakuranetin). Enzymes like SOMT-2, that methylate a *para*-monohydroxylated B ring of flavonoids, either seem to be a rare occurrence or fairly inactive, since descriptions of characterized representatives are scarce in the literature and are only limited to a couple of enzymes [46, 168].

## 2.3 Alkylation and biotransformations

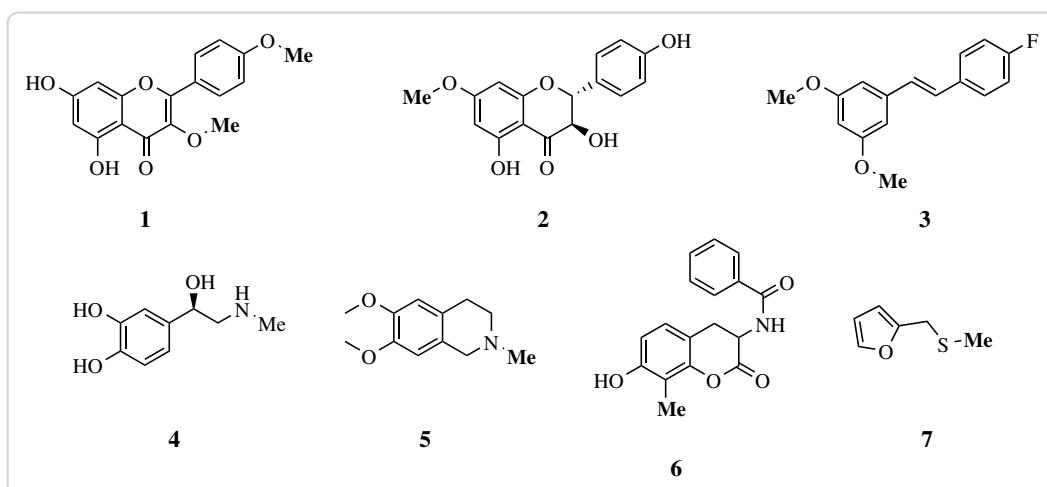
Alkylation reactions are a crucial factor helping nature create highly diverse natural products from a limited number of precursors and as such, these reactions are becoming more and more important in biocatalysis. Methylation, prenylation and glycosylation constitute the major alkylation reactions in nature and can largely influence the (bio)chemical characteristics of a compound. Intra- and intermolecular prenylation is achieved by prenyl transferases, which employ mono- and oligoprenyl diphosphates and are mainly responsible for the over 70 000 terpenoids described today [23]. Glycosyl transferases catalyze the formation of a glycosidic bond using nucleotide- or lipid phospho-sugars (e.g. uridine diphosphate (UDP)-glucose, dolichol phosphate oligosaccharides) as sugar donating substrates [112].

Methylation reactions are catalyzed by MTs using SAM as methyl donor and will be the focus of this section.

The introduction of a methyl group ( $\approx 20 \text{ \AA}$ ) for a hydrogen ( $\approx 5 \text{ \AA}$ ) can have different effects, chemically and biologically. Polar groups (e.g. hydroxyl, amine, carboxyl) are masked by methylation, which majorly alters their chemistry. Possible hydrogen donors are lost and the lipophilicity is increased. The methylation can act as a molecular signal, which might be specifically recognized by other enzymes than the original more polar, hydrogen donating group. This can in turn have dramatic physiological consequences in an organism.

**The industrial potential of methyl transferases** has been demonstrated by several studies [207]. Li and Frost [121] presented an environmentally friendly method to produce vanillin from glucose by genetically modified *E. coli* cells. During the fermentation, methyl transfer was achieved by recombinantly expressed COMT. Recent developments mainly focus on the synthesis of structurally more complex and highly valuable compounds, especially flavonoids, due to their manifold biological effects (Figure 2.7). For example, ermanine (**1**) with a claimed anti-inflammatory and antiviral activity, can be synthesized from its inactive precursor kaempferol (5,7-dihydroxy-3,4'-dimethoxyflavone) by whole-cell biotransformation [15]. Sequential introduction of the two methyl groups was performed by OMT-9 from rice (ROMT-9) and OMT-2 from soybean and resulted in almost quantitative conversion of the substrate. Similarly, transformation of kaempferol and quercetin by an engineered variant of the OMT-7 from *Populus deltoides* gave the 3,7-di-O-methyl products in 58 % and 70 % yield, respectively [88]. The conversion of naringenin (4',5,7-trihydroxyflavanone) to 3-O-methylkaempferol was performed as an enzymatic two-step process involving initial oxidation by flavonol synthase and methylation of the intermediate by ROMT-9 [100].

The most promising approach for the chemoenzymatic production of flavonoids remains the *de novo* synthesis from inexpensive biosynthetic precursors such as p-coumaric acid, which is initially processed into non-methylated flavonoids and subsequently modified by *O*-MT reactions. This way, 7-*O*-methyl-aromadendrin (**2**) and the corresponding flavone genkwanin were obtained from recombinant *E. coli* in



**Figure 2.7:** Natural products synthesized with the help of methyl transferases. *O*-MTs were used for the synthesis of ermanine (**1**), 7-*O*-methyl-aromadendrin (**2**) and 3,5-dimethoxy-4'-fluorostilbene (**3**). *N*-methyl transferase (*N*-MT) catalyze the production of epinephrine (**4**) and *N*-methyltetrahydroisoquinoline (**5**), whereas 3-(benzoylamino)-8-methylumbelliferone (**6**) and furfuryl-methyl-sulfide (**7**) can be produced by *C*-methyl transferase (*C*-MT) and *S*-methyl transferase (*S*-MT)

yields of 2.7 mg/l and 0.2 mg/l, respectively [117, 132]. Similarly, several non-natural (e.g. fluorinated) cinnamic acids can be converted to mono- and dimethylated stilbenes such as 3,5-dimethoxy-4'-fluorostilbene (**3**) using a reconstructed plant pathway in *E. coli*.

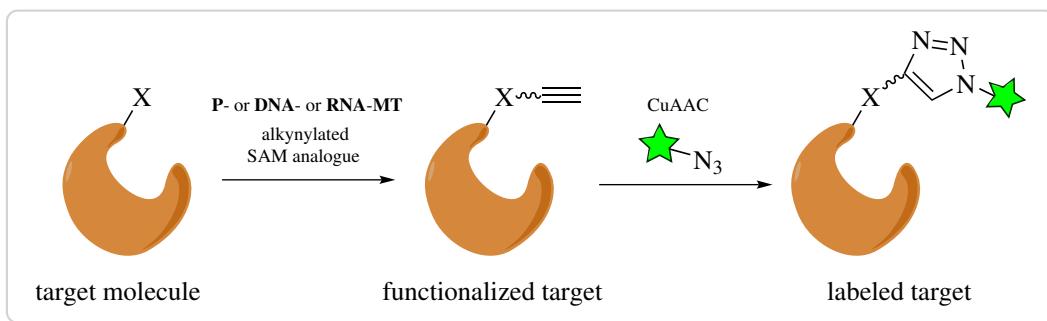
*N*-, *C*- and *S*-methyl transferases have also been successfully used in biocatalytic applications. For example, epinephrine (**4**) and *N*-methyltetrahydroisoquinoline (**5**) were obtained by *in vitro* and *in vivo* biotransformations of their respective precursors [136, 146]. Furthermore, a number of studies describe the production of several other *N*-methylated alkaloid and non-alkaloid compounds [154, 158, 225]. *C*-methyl transferases have been used biocatalytically to modify different phenols [37] and coumarin derivatives, to obtain compounds such as 3-(benzoylamino)-8-methylumbelliferone (**6**) [183]. Also, the composition of the main tocopherol species in plants could be tuned by the introduction of bacterial *C*-MTs [84]. *S*-methyl transferases have only seen a limited number of biocatalytic applications, but a candidate from *Catharanthus roseus* shows some promiscuous activity towards small aliphatic and aromatic thiols (furfuryl-methyl-sulfide (**7**))) [32].

*In vivo* biotransformations for the high-yield methylation of compounds is a feasible method, especially since SAM is a rather expensive cofactor (3000 to 15 000 €/g). However, SAM cannot be easily substituted for artificial analogues *in vivo*.

**Artificial SAM-analogues** have shown tremendous potential in *in vitro* biocatalytic applications. The first description of novel synthetic SAM analogues with extended carbon chains, including *S*-adenosyl-L-ethionine (SAE), allyl and propargyl derivatives, that were also shown to be useful in modifying DNA via the action of several DNA-MTs was provided by Dalhoff, *et al.* [43, 44]. A whole variety of allyl derivatives was examined by different researchers and site-specific introductions of allyl, pent-2-en-4-ynyl and even 4-propargyloxy-but-2-enyl moieties into proteins (i.e. histones) was demonstrated using P-MTs [156, 205]. However, the larger substrate analogues were not necessarily accommodated by the native P-MTs making engineering efforts for the accommodation of larger substrates inevitable [205]. The specific introduction of alkene functionalized groups made it then possible to use click chemistry for further functionalization and/or detection of the labelled proteins, DNA or RNA (Figure 2.8) [148, 156, 171, 205, 211].

In 2012 Bothwell and Luo even described the exchange of the sulfonium with a selenonium center, which afforded *Se*-adenosyl selenomethionine (SeAM) analogues that have since then been described as substrates for several P-MTs [18, 19]. SeAM analogues have the advantage of being more resistant to chemical decomposition than their sulfur counterparts, but also show enhanced transmethylation reactivity [18].

There have been some reports on the use of SAM analogues by small molecule MTs. In 2009 Stecher *et al.* reported the use of the C-MTs NovO and CouO along with synthetic SAM analogues to accomplish biocatalytic Friedel-Crafts alkylations of some aminocoumarine antibiotics [183]. Lee *et al.* were the first ones to describe the transfer of a keto-group from an SAM derivative by means of the small molecule MTs catechol O-methyl transferase (EC 2.1.1.6) and thiopurine S-methyl transferase (EC 2.1.1.67) [115]. Furthermore the work done on the O-MTs RebM and RapM, which modify the antitumor active natural products rebeccamycin and rapamycin respectively, shows the general feasibility of using SAM analogues in combination



**Figure 2.8:** Labelling of macromolecules by using a combination of novel alkyne-derivatized SAM analogues and Cu<sup>I</sup>-catalyzed azide-alkyne 1,3-dipolar cycloaddition (CuAAC). Depending on the type of label used, it can be employed for detection (e.g. through fluorophores, coupled assays) or affinity purification (e.g. biotin). This technique is also feasible for use in activity based protein profiling (ABPP) approaches.

with MTs to modify small molecules [113, 181, 224]. However, no bioactivity data has been reported that shows the biological activity of the newly produced compounds.

## 2.4 Motivation

Hier muss noch was geschehen!

### **3 Material And Methods**

Within this section percentages refer to volume per volume (v/v) percentages unless otherwise specified.

## 3.1 Materials

### 3.1.1 Chemicals

Enzymes and buffers used for molecular cloning were obtained from Life Technologies (Darmstadt, Germany), unless otherwise noted. Flavonoid HPLC standards were purchased from Extrasynthese (Genay, France). Deuterated solvents were acquired from Deutero GmbH (Kastellaun, Germany). Solvents, purchased from VWR (Poole, England), were distilled in-house before use.

All other chemicals were obtained from either Sigma-Aldrich (Steinheim, Germany), Applichem (Darmstadt, Germany), Carl Roth (Karlsruhe, Germany) or Merck (Darmstadt, Germany).

### 3.1.2 Commonly used solutions and buffers

50× 5052	25 % glycerol, 2.5 % (w/v) glucose, 10 % (w/v) $\alpha$ -lactose
binding buffer	50 mM Tris/HCl, 500 mM NaCl, 10 % glycerol, 2.5 mM imidazole pH 7
elution buffer	50 mM Tris/HCl, 500 mM NaCl, 10 % glycerol, 250 mM imidazole pH 7
ysis buffer	50 mM Tris/HCl, 500 mM NaCl, 10 % glycerol, 2.5 mM imidazole, 0.2 % Tween-20 pH 7
1 M MMT pH 4 (10×)	26.8 g/l L-malic acid, 78.1 g/l MES, 26.8 g/l Tris, 2.1 % 10 M HCl
1 M MMT pH 9 (10×)	26.8 g/l L-malic acid, 78.1 g/l MES, 26.8 g/l Tris, 6.7 % 10 M NaOH
20× NPS	1 M $\text{Na}_2\text{HPO}_4$ , 1 M $\text{KH}_2\text{PO}_4$ , 0.5 M $(\text{NH}_4)_2\text{SO}_4$
1 M SSG pH 4 (10×)	14.8 g/l succinic acid, 60.4 g/l $\text{NaH}_2\text{PO}_4 \cdot \text{H}_2\text{O}$ , 32.8 g/l glycine, 0.4 % 10 M NaOH

1 M SSG pH 10 (10×)	14.8 g/l succinic acid, 60.4 g/l $\text{NaH}_2\text{PO}_4 \cdot \text{H}_2\text{O}$ , 32.8 g/l glycine, 10.3 % 10 M NaOH
5× SDS sample buffer	10 % (w/v) SDS, 10 mM $\beta$ -mercaptoethanol, 20 % glycerol, 0.2 M Tris/HCl pH 6.8, 0.05 % (w/v) bromophenolblue
1000× trace elements	50 mM $\text{FeCl}_3$ , 20 mM $\text{CaCl}_2$ , 10 mM $\text{MnCl}_2$ , 10 mM $\text{ZnSO}_4$ , 2 mM $\text{CoCl}_2$ , 2 mM $\text{CuCl}_2$ , 2 mM $\text{NiCl}_2$ , 2 mM $\text{Na}_2\text{MoO}_4$ , 2 mM $\text{Na}_2\text{SeO}_3$ , 2 mM $\text{H}_3\text{BO}_3$

### Preparation of natural deep eutectic solvent (NADES)

Natural deep eutectic solvent (NADES) were prepared by adding each component in a round-bottom flask with a stirrer and stirring the mixture at 50 °C with intermittent sonication treatments until a clear solution was obtained.

**Table 3.1.: NADES-mixtures used within this work.**

name	composition	mole ratio	mass fraction (w/w)
PCH	propane-1,2-diol	1:1:1	0.326
	choline chloride		0.597
	water		0.077
GCH	L-glucose	2:5:5	0.314
	choline chloride		0.608
	water		0.078

### 3.1.3 Culture media used to grow bacteria

LB-medium	10 g/l NaCl, 10 g/l tryptone, 5 g/l yeast extract, pH 7.5
LB-agar	LB + 1.5 % (w/v) agar-agar
TB-medium	12 g/l tryptone, 24 g/l yeast extract, 0.4 % glycerol, 72 mM $\text{K}_2\text{HPO}_4$ , 17 mM $\text{KH}_2\text{PO}_4$
ZY	10 g/l tryptone, 5 g/l yeast extract
ZYP-5052	volume fraction (v/v): 0.928 ZY, 0.05 20× NPS, 0.02 50× 5052, 0.002 1 M $\text{MgSO}_4$ , 0.0002 1000× trace elements

### 3.1.4 Bacterial strains

#### *E.coli*

BL21(DE3)	F <sup>-</sup> <i>ompT hsdSB(r<sub>B</sub><sup>-</sup>,m<sub>B</sub><sup>-</sup>) gal dcm λ(DE3)</i> Invitrogen, Karlsruhe (Germany)
C41(DE3)	F <sup>-</sup> <i>ompT hsdSB(r<sub>B</sub><sup>-</sup>,m<sub>B</sub><sup>-</sup>) gal dcm λ(DE3)</i> Lucigen, Wisconsin (USA)
C43(DE3)	F <sup>-</sup> <i>ompT hsdSB(r<sub>B</sub><sup>-</sup>,m<sub>B</sub><sup>-</sup>) gal dcm λ(DE3)</i> Lucigen, Wisconsin (USA)
DH5 $\alpha$	F <sup>-</sup> $\Phi 80 lacZ \Delta M15 \Delta(lacZYA-argF) U169 recA1 endA1 hsdR17(r_K^-, m_K^+) phoA supE44 \lambda^- thi-1 gyrA96 relA1$ Invitrogen, Karlsruhe (Germany)
JM110	<i>rpsL thr leu thi lacY galK galT ara tonA tsx dam dcm glnV44 Δ(lac-proAB) e14- [F' traD36 proAB<sup>+</sup> lacI<sup>q</sup> lacZΔM15] hsdR17(r<sub>K</sub><sup>-</sup>m<sub>K</sub><sup>+</sup>)</i> Martin-Luther-University Halle-Wittenberg
JW1593 (BW25113 derivative)	<i>rrnB ΔlacZ4787 HsdR514 Δ(araBAD)568 rph-1 ΔydgG (Kan<sup>R</sup>)</i> Keio Collection, National Institute of Genetics (Japan)
MG1655	F <sup>-</sup> $\lambda^- ilvG^- rfb-50 rph-1$ DSMZ, Hamburg (Germany)
One Shot TOP10	F <sup>-</sup> $\Phi 80 lacZ \Delta M15 \Delta(mrr-hsdRMS-mcrBC) recA1 endA1 mcrA \Delta lacX74 araD139 \Delta(ara-leu)7697 galU galK rpsL (\text{Str}^R) \lambda^- nupG$ Invitrogen, Karlsruhe (Germany)
Origami(DE3)	$\Delta(ara-leu)7697 \Delta lacX74 \Delta phoA Pvull phoR araD139 ahpC galE galK rpsL F'[lac + lacI q pro]$ (DE3) <i>gor522::Tn10 trxR (Kan<sup>R</sup>, Str<sup>R</sup>, Tet<sup>R</sup>)</i> Novagen, Wisconsin (USA)
Rosetta(DE3)	F <sup>-</sup> <i>ompT hsdSB(r<sub>B</sub><sup>-</sup>,m<sub>B</sub><sup>-</sup>) gal dcm λ(DE3) pRARE (Cam<sup>R</sup>)</i> Novagen, Wisconsin (USA)
Rosetta(DE3) pLysS	F <sup>-</sup> <i>ompT hsdSB(r<sub>B</sub><sup>-</sup>,m<sub>B</sub><sup>-</sup>) gal dcm λ(DE3) pLysSRARE (Cam<sup>R</sup>)</i> Novagen, Wisconsin (USA)

T7 Express                    *fhuA2 lacZ::T7 gene1 [lon] ompT gal sulA11 R(mcr-73::miniTn10-Tet<sup>S</sup>)2 [dcm] R(zgb-210::Tn10-Tet<sup>S</sup>) endA1 Δ(mcrC-mrr)114::IS10*  
NEB, Massachusetts (USA)

### ***Agrobacterium tumefaciens***

GV3101                    chromosomal background: C58, marker gene: *rif*, Ti-plasmid:  
cured, opine: nopaline  
Sylvestre Marillonet, IPB

### **3.1.5 Plasmids**

**Table 3.3:** Plasmids used in this work.

name	supplier/source
pACYCDuet-1	Merck, Darmstadt (Germany)
pCDFDuet-1	Merck, Darmstadt (Germany)
pET-20b(+)	Merck, Darmstadt (Germany)
pET-28a(+)	Merck, Darmstadt (Germany)
pET-32a(+)	Merck, Darmstadt (Germany)
pET-41a(+)	Merck, Darmstadt (Germany)
pQE30	QIAGEN, Hilden (Germany)
pUC19	Invitrogen, Karslruhe (Germany)

### **3.1.6 Oligonucleotides and synthetic genes**

Oligonucleotides and primers were ordered from Eurofins Genomics (Ebersberg, Germany). The purity grade was *high purity salt free* (HPSF). Synthetic genes or gene fragments were obtained from GeneArt® (Life Technologies, Darmstadt, Germany) or Eurofins Genomics (Ebersberg, Germany).

**Table 3.4.:** Primers used in this work. Recognition sites for endonucleases are underlined. Positions used for site directed mutagenesis are in lower case font.

name	sequence (5'→3')	cloning site
somt1	TTG <u>AAG ACA</u> AAA TGG CTT CTT CAT TAA ACA ATG GCC G	BpiI
somt2	TTG <u>AAG ACA</u> AGG ACA CCC CAA ATA CTG TGA GAT CTT CC	BpiI
somt3	TTG <u>AAG ACA</u> AGT CCT TAG GAA CAC CTT TCT GGG AC	BpiI
somt4	TTG <u>AAG ACA</u> AAA GCT CAA GGA TAG ATC TCA ATA AGA GAC	BpiI
pfomt1.fw	CAG AGA GGC cTA TGA GAT TGG CTT GC	
pfomt1.rv	GCA AGC CAA TCT CAT AgG CCT CTC TG	
pfomt2.fw	<u>CAT ATG</u> GAT TTT GCT GTG ATG AAG CAG GTC	NdeI
pfomt2.rv	<u>GAA TTC</u> AAT AAA GAC GCC TGC AGA AAG TG	EcoRI
pRha1.fw	CTC TAG <u>CAG ATC</u> TCG GTG AGC ATC ACA TCA CCA CAA TTC	BglII
pRha1.rv	CAA TTG <u>AGG ATC</u> CCC ATT TTA ACC TCC TTA GTG	BamHI
pUC1.fw	GCG TAT TGG Gag aTC TTC CGC TTC CTC	
pUC1.rv	GAG GAA GCG GAA GAt ctC CCA ATA CGC	

### 3.1.7 Instruments

CD-spectrometer	Jasco J-815 (Eaton, USA)
electrophoresis (horizontal)	Biometra Compact XS/S (Göttingen, Germany)
electrophoresis (vertical)	Biometra Compact M (Göttingen, Germany) Biometra Minigel-Twin (Göttingen, Germany)
FPLC	ÄKTA purifier (GE Healthcare, Freiburg, Germany)
GC/MS	GC-MS-QP2010 Ultra (Shimadzu, Duisburg, Germany)
HPLC	VWR-Hitachi LaChrom Elite (VWR, Darmstadt, Germany)
ITC	MicroCal iTC200 (Malvern, Worcestershire, UK)
plate-reader	SpectraMax M5 (Molecular Devices, Biberach, Germany)
NMR-spectrometer	Varian Unity 400 (Agilent, Böblingen, Germany) Varian VNMRS 600 (Agilent, Böblingen, Germany)

photospectrometer	Eppendorf Biophotometer Plus (Hamburg, Germany) JASCO V-560 (Eaton, USA)
centrifuges	Colibri Microvolume Spectrometer (Biozym, Hess. Oldendorf, Germany) Eppendorf 5424 (Hamburg, Germany) Hettich Mikro 120 (Kirchlengern, Germany)
centrifuge rotors	Beckman Avanti J-E, Beckman Allegra X-30R (Krefeld, Germany) Beckman JA-10, JA-16.250, JS-4.3 (Krefeld, Germany)

### 3.1.8 Software

All mathematical and statistical computations and graphics were done with the R software (versions 3.1.X, <http://cran.r-project.org/>) [33]. Visualizations of macromolecules were arranged using the PyMol Molecular Graphics System, version 1.7.0.0 (Schrödinger, New York, USA).

Physicochemical calculations and calculations of different molecular descriptors were performed using Marvin Beans 15.4.13.0 (ChemAxon, Budapest, Hungary) and Molecular Operating Environment 2008.10 (Chemical Computing Group, Montreal, Canada). Special software used for X-ray crystal structure solution is discussed separately in the corresponding section (3.5).

## 3.2 Molecular Biology

Basic molecular biology methods like polymerase chain reaction (PCR), DNA restriction/ligation, DNA gel electrophoresis, preparation of competent cells and transformation were performed based on the protocols summarized by Sambrook and Russell [163]. Plasmid DNA was isolated using the QIAprep® Spin Miniprep Kit (QIAGEN, Hilden, Germany) according to the manufacturer's instructions. In vitro site-directed mutagenesis was set-up according to the protocol of the *QuikChange™ Site-Directed Mutagenesis* kit [3] offered by Agilent Technologies (Santa Clara, USA). Nucleotide fragments obtained by PCR, restriction/ligation procedures or excision from electrophoresis gels were purified and concentrated

using the *Nucleospin Gel and PCR Clean-up* kit provided by Machery-Nagel (Düren, Germany) according to the instructions provided by the manufacturer.

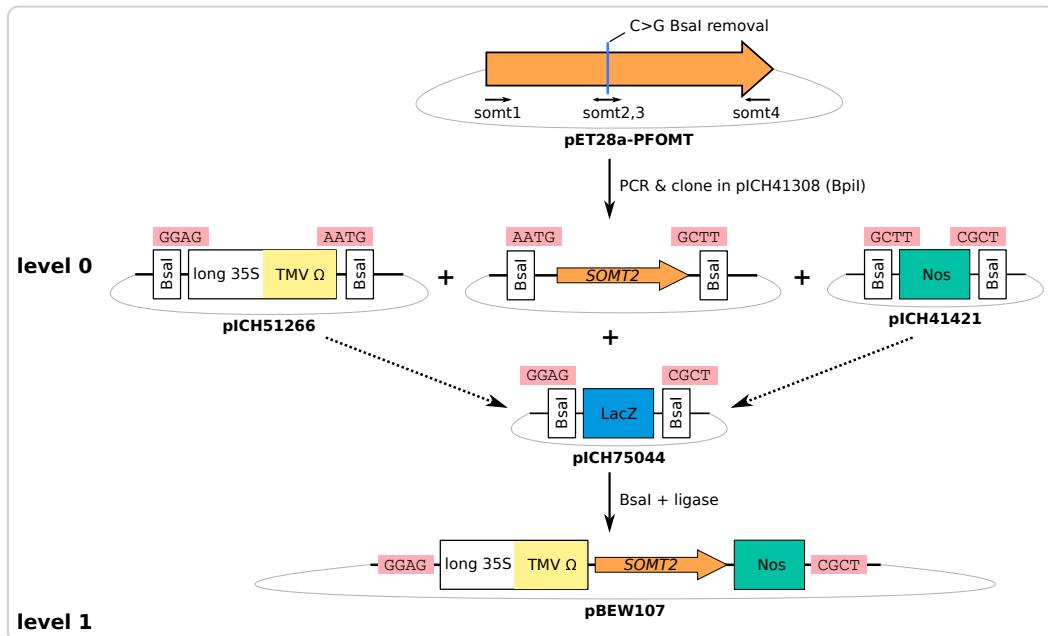
### 3.2.1 Golden Gate Cloning

The Golden Gate cloning procedure is a one-pot method, meaning the restriction digestion and ligation are carried out in the same reaction vessel at the same time [55, 105]. Consequently PCR-fragments, destination vector, restriction endonuclease and ligase are added together in this reaction. The methodology employs type II restriction enzymes, which together with proper design of the fragments allow for a ligation product lacking the original restriction sites. For digestion/ligation reactions of fragments containing BpiI sites, 20 fmol of each fragment or vector, together with 5 U of BpiI and 5 U of T4 ligase were combined in a total volume of 15 µl 1× ligase buffer. For fragments to be cloned via BsaI sites, BpiI in the above reaction was substituted by 5 U BsaI. The reaction mixture was placed in a thermocycler and incubated at 37 °C for 2 min and 16 °C for 5 min. These two first steps were repeated 50 times over. Finally, the temperature was raised to 50 °C (5 min) and 80 °C (10 min) to inactivate the enzymes.

The *SOMT2* gene was amplified from pET28a-SOMT using primers *somt1-somt4* (Table ??), cloned into vector pICH41308 (level 0) using BpiI and consequently subcloned into the level 1 module pICH75044, alongside 35S promoter and nopaline synthase (nos)-terminator, using BsaI (Figure 3.1). The resulting construct was denoted as pBEW107.

### 3.2.2 Subcloning of genes

All subcloning procedures were performed according to section 3.2 and specifically subsection 3.2.1. Specific steps for the subcloning of any genes discussed can be found in the appendix (p.129). The *pfromt* gene was subcloned from the pQE-30 vector kindly provided by Thomas Vogt (Leibniz-Institute of Plant Biochemistry (IPB), Halle, Germany) into the pET-28a(+) vector. The *somt-2* gene was subcloned from the pQE-30 vector kindly provided by Martin Dippe (IPB, Halle, Germany) into the pET-28-MC vector.



**Figure 3.1:** Golden Gate cloning scheme for soy *O*-methyl transferase (*SOMT-2*)

### 3.2.3 Transformation of electrocompetent *Agrobacterium tumefaciens* cells

A 50 µl aliquot of electrocompetent *A. tumefaciens* cells was thawed on ice. (50 to 100) ng of plasmid were added, the solution was mixed gently and transferred to a pre-cooled electroporation cuvette. After pulsing (2.5 kV, 200 Ω) 1 ml of lysogeny broth (LB)-medium was added, the mixture transferred to a 1.5 ml tube and incubated for (3 to 4) hours at 28 °C. The culture was centrifuged (10 000 × g, 1 min) and 900 µl supernatant were discarded. The pellet was resuspended in the remaining liquid, plated onto LB-agar plates supplemented with 40 µg/ml rifampicin and 50 µg/ml carbencillin and incubated at 28 °C for (2 to 3) days.

## 3.3 Treatment of plant material

### 3.3.1 Infiltration of *Nicotiana benthamiana*

Before infiltration *N. benthamiana* plants were pruned, such that only leaves to be infiltrated remained with the plant (Figure 3.2). 5 ml cultures of transformed *A. tumefaciens* in LB-medium (with 40 µg/ml rifampicin and 50 µg/ml carbencillin) were grown over night at 28 °C and 220 rpm. OD<sub>600</sub> of the culture was measured and adjusted to 0.2 by dilution with infiltration buffer (10 mM MES/NaOH, 10 mM MgSO<sub>4</sub> pH 5.5). When multiple *A. tumefaciens* transformed with different constructs/plasmids were used for infiltration, the cultures were mixed and diluted using infiltration buffer, such that OD<sub>600</sub> of each culture in the mix was 0.2. The solution was infiltrated into the abaxial side of *N. benthamiana* leaves using a plastic syringe. The leaf material was harvested after 7 days.

#### Infiltration of *N. benthamiana* for *in vivo* biotransformation using SOMT-2

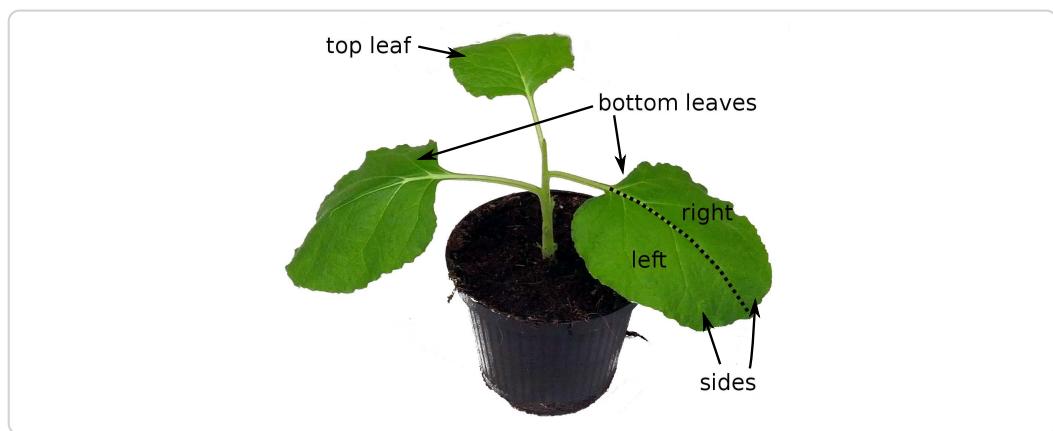
Both sides of *N. benthamiana* leaves were infiltrated with different samples (Figure 3.2). The left side was infiltrated with *A. tumefaciens* cultures transformed with pAGM10733 (phenylalanine ammonia-lyase (PAL)), pAGM10406 (chalcone synthase (CHS)) and pBEW107 (SOMT-2). For the right side the *A. tumefaciens* culture containing pBEW107 was replaced by a control: *A. tumefaciens* transformed with the empty vector pICH75044.

### 3.3.2 Plant material harvest

Infiltrated/Infected areas of *N. benthamiana* leaf material were cut out and grouped by plant number, leaf position (top/bottom) and leaf side (right/left). The grouped clippings were weighed, frozen in liquid nitrogen, ground to a powder, freeze-dried and stored at -80 °C.

### 3.3.3 Extraction of flavonoids from *N. benthamiana* leaves

Two tips of a small spatula of freeze-dried material ( $\approx$ 6 mg), were weighed exactly and extracted with 500 µl 75 % aqueous methanol containing 1 mM ascorbic acid,



**Figure 3.2.:** Pruned *N. benthamiana* plant, with two bottom and one top leaf, ready to be infiltrated.

0.2 % formic acid and 0.1 mM flavone (internal standard). Therefore the suspension was vortexed for 30 s, rotated on an orbital shaker for 10 min and vortexed again for 30 s. The suspension was centrifuged ( $20\,000 \times g$ , 4 °C, 10 min) and the supernatant transferred to a new tube, to remove the insoluble plant material. The supernatant was centrifuged again ( $20\,000 \times g$ , 4 °C, 10 min) and the resulting supernatant was transferred to a HPLC-vial and stored at –20 °C until analysis.

## 3.4 Protein biochemistry

Stock solutions of antibiotics, IPTG or sugars were prepared according to the pET System Manual by Novagen [152], unless otherwise noted.

### 3.4.1 Determination of protein concentration

Protein concentrations were estimated using the absorption of protein solutions at 280 nm, which is mainly dependent on the amino acid composition of the protein studied [70]. Extinction coefficients of proteins were calculated from the amino acid sequence using the ExpPASy servers's ProtParam tool [66].

**Table 3.6.:** Calculated extinction coefficients of proteins used in this work.

protein/enzyme	$\epsilon_{280\text{nm}}^{1\text{ g/l}}$ in $\text{ml mg}^{-1} \text{cm}^{-1}$
PFOMT (reduced)	0.714
PFOMT Y51K N202W (reduced)	0.852
SOMT-2 (oxidized)	1.263
SOMT-2 (reduced)	1.247
COMT	

### 3.4.2 Protein production test (expression test)

The heterologous production of proteins in *E. coli* was assessed in a small scale protein production test, henceforth called expression test. Single colonies of *E. coli* transformed with the constructs to be studied were used to inoculate a 2 ml starter culture in LB-medium containing the appropriate antibiotics. The working concentrations of antibiotics used was as follows: 200 µg/ml ampicillin, 150 µg/ml kanamycin, 50 µg/ml chloramphenicol, 20 µg/ml tetracycline. The starter culture was allowed to grow at 37 °C and 200 rpm over night. A 5 ml sampling culture of the medium to be studied containing the appropriate antibiotics was prepared. The media tested included LB, terrific broth (TB) and auto-induction media like ZYP-5052. The sampling culture was inoculated to an OD<sub>600</sub> of 0.075 using the starter culture and incubated at different temperatures and 200 rpm in a shaking incubator. 1 mM isopropyl-D-thiogalactopyranosid (IPTG) was added when the OD<sub>600</sub> reached 0.6-0.8, if appropriate for the studied construct. 1 ml samples were removed after different times of incubation (e.g. 4, 8, 12 hours), subfractionated (3.4.3) and analyzed via SDS-polyacrylamide gel electrophoresis (PAGE) (3.4.6).

Exact specifications of growth conditions (e.g. temperature, time, constructs) are discussed in the individual sections.

### 3.4.3 Protein subfractionation

The protein subfractionation procedure described herein was adapted from the protocol described in the pET Manual [152]. Overall 5 protein subfractions can be obtained, including *total cell protein*, *culture supernatant (medium) protein*, *periplasmic protein*, *solute cytoplasmic protein* and *insoluble protein*. The OD<sub>600</sub> of the culture sample was measured and the cells harvested by centrifugation at 10 000 × *g*, 4 °C for 5 minutes. The protein in the supernatant medium was concentrated by precipitation with trichloro acetic acid (TCA) (3.4.4) for SDS-PAGE analysis. The periplasmic protein was prepared (3.4.5) and also concentrated by TCA precipitation for SDS-PAGE. Cells were lysed by resuspending the cell pellet in (OD<sub>600</sub> × V × 50) µl of bacterial protein extraction reagent (B-PER) and vortexing vigorously for 30 s. The suspension was incubated at room temperature (RT) for 30 min to assure complete lysis. To separate insoluble protein and cell debris from the soluble cytosolic protein, the suspension was centrifuged at 10 000 × *g* and 4 °C for 10 min. Soluble cytoplasmic protein was contained in the supernatant, whereas insoluble protein remained in the pellet. For SDS-PAGE analysis of the insoluble protein, the pellet was resuspended in the same volume of B-PER. To obtain only the total cell protein fraction, the preparation of periplasmic and soluble cytosolic protein was omitted. Sample volumes of 10 µl of each fraction were used for SDS-PAGE analysis.

### 3.4.4 Protein sample concentration by TCA precipitation

Diluted protein samples were concentrated by TCA precipitation in microcentrifuge tubes. Therefore 0.1 volume (V) of 100 % (w/v) TCA in water was added to the clarified sample, which was then vortexed for 15 s and placed on ice for a minimum of 15 min. The sample was centrifuged at 14 000 × *g*, 4 °C for 15 min. The supernatant was discarded and the pellet was washed twice with 0.2 V ice-cold acetone. The acetone was removed and the pellet set to air-dry in an open tube. After drying, the protein pellet was resuspended in 0.1 V phosphate buffered saline (PBS) containing 1 × SDS-sample buffer by heating to 85 °C and vigorous vortexing, to achieve a

10× concentration. After resuspension the sample was analyzed by SDS-PAGE or stored at -20 °C until use.

### 3.4.5 Preparation of periplasmic protein

Target proteins may be directed to the periplasmic space by N-terminal signal sequences like *pelB* or *DsbA/C* [131]. The periplasma is, other than the cytosol, an oxidizing environment and often used for the production of proteins containing disulfide linkages. The preparation of periplasmic protein was accomplished by an osmotic shock protocol modified from Current Protocols in Molecular Biology [7]. The cell pellet was resuspended in the same volume as the culture sample of 30 mM tris-HCl, 20 % (w/v) sucrose, pH 8 and 1 mM ethylenediaminetetraacetic acid (EDTA) was added. The suspension was stirred for 10 min at RT and the cells were collected by centrifugation at 10 000 × *g*, 4 °C for 10 min. The supernatant was discarded and the cell pellet was resuspended in the same volume of ice-cold 5 mM MgSO<sub>4</sub>. The suspension was stirred for 10 min on ice, while the periplasmic proteins were released into the solution. The cells were collected by centrifugation as before. Periplasmic proteins were contained in the supernatant.

### 3.4.6 Discontinuous SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

The analysis of samples via SDS-PAGE was realized via the discontinuous system first described by Laemmli, which allows separation of proteins based on their electrophoretic mobility, which in turn depends on their size [111]. The SDS-PAGE procedure was carried out according to standard protocols described by Sambrook and Russell [163]. Very dilute and/or samples with high ionic strength were concentrated and/or desalting by the TCA precipitation procedure described in subsection 3.4.4. Generally a 10 % (acrylamide/bisacrylamide) running gel combined with a 4 % stacking gel was used. Reducing SDS-PAGE sample buffer was added to the protein sample to be analyzed, whereafter the sample was heated to 95 °C for 5 min, to allow for total unfolding of the protein. After cooling to RT the samples were transferred into the gel pockets for analysis. The *PageRuler™ Prestained*

*Protein Ladder* (Life Technologies GmbH, Darmstadt, Germany) was used as a molecular weight (MW) marker and run alongside every analysis as a reference. Gels were stained using a staining solution of 0.25 % Coomassie Brilliant Blue G-250 (w/v) in water:methanol:acetic acid (4:5:1) and destained by treatment with water:methanol:acetic acid (6:3:1).

### 3.4.7 Buffer change of protein samples

The buffer in protein samples was exchanged either by dialysis, or by centrifugal filter concentrators (Amicon® Ultra Centrifugal Filter; Merck, Darmstadt, Germany). Large volumes of highly concentrated protein solutions were preferably dialyzed. Respectively, very dilute samples were concentrated and rebuffered using centrifugal concentrators. Dialysis was carried out at least twice against a minimum of 100 times the sample volume. Dialysis steps were carried out at RT for 2 hours, or over-night at 4 °C. Centrifugal concentrators were used according to the manufacturers instructions.

### 3.4.8 Production of recombinant protein

#### Heterologous production of PFOMT

PFOMT was produced as a N-terminally (His)<sub>6</sub>-tagged fusion protein. A 2 ml starter culture of LB containing 100 µg/ml kanamycin was inoculated with a single colony of *E. coli* BL21(DE3) transformed with pET28-pfomt and incubated at 37 °C, 220 rpm for 6 hours. The main culture (N-Z-amino, yeast extract, phosphate (ZYP-5052) containing 200 µg/ml kanamycin) was inoculated with the starter culture such that OD<sub>600</sub> was 0.05. The culture was incubated in a shaking incubator at 37 °C, 220 rpm over night ( $\approx$ 16 h). Due to the autoinducing nature of the ZYP-5052 medium, addition of IPTG was not necessary. Cells were harvested by centrifugation at 10 000  $\times g$ , 4 °C for 10 min and the supernatant discarded. The pellet was resuspended in 50 mM Tris/HCl, 500 mM NaCl, 2.5 mM imidazole, 10 % glycerol pH 7 using a volume of  $\approx$ 10 ml/g of cell pellet. The cells were lysed by sonication (70 % amplitude, 1 s on-off-cycle) for 30 seconds, which was repeated twice. The crude lysate was clarified

by centrifugation at  $15\,000 \times g$ , 4 °C for 15 minutes followed by filtration through a 0.45 µm filter. Consequently, the His-tagged PFOMT was purified by immobilized metal affinity chromatography (IMAC) (3.4.10). The eluted PFOMT protein was dialyzed (3.4.7) against 25 mM HEPES, 100 mM NaCl, 5 % glycerol pH 7 and stored at –20 °C until use.

### Heterologous production of SOMT-2

SOMT-2 was produced as a fusion protein with an N-terminal His-tag. A starter LB-culture ( $\approx 2$  ml) containing 100 µg/ml kanamycin was inoculated with a single colony of *E. coli* BL21(DE3) transformed with pET28MC-somt and incubated at 37 °C, 220 rpm for 6 hours. The starter culture was used to inoculate the main culture (LB-medium containing 100 µg/ml kanamycin), such that  $OD_{600} \approx 0.05$ . The culture was incubated at 37 °C, 220 rpm in a shaking incubator until  $OD_{600} \approx 0.6$ . Expression was induced by addition of 1 mM IPTG. Incubation continued at 37 °C, 220 rpm for 6 hours. Cells were harvested by centrifugation ( $10\,000 \times g$ , 4 °C, 10 min) and used, or stored at –20 °C until use. SOMT-2 was produced in inclusion bodies (IBs), which were prepared as laid out in subsection 3.4.9.

### 3.4.9 Preparation of inlusion bodies (IBs)

Often, when recombinant protein is produced in high levels in *E. coli* it is accumulated in so-called inlusion bodies (IBs) [47]. The accumulating IBs consist mainly of the overproduced target protein, which is inherently quite pure already. IBs can be selectively recovered from *E. coli* cell lysates and can consequently be refolded. IBs were prepared according to a modified protocol by Palmer [155]. The cells were resuspended in 5 ml/g<sub>cells</sub> IB lysis buffer (100 mM Tris/HCl, 1 mM EDTA pH 7), 0.5 mM phenylmethylsulfonylfluoride (PMSF) was added as protease inhibitor. The solution was homogenized using a tissue grinder homogenizer (Ultra Turrax®; IKA®-Werke GmbH & Co. KG, Staufen, Germany). 200 µg/ml lysozyme was added to aid in the breakage of cells and the cells were lysed by sonicating thrice at 70 % amplitude (1 s on-off-cycle) for 30 seconds. DNase I (10 µg/ml) was added and the solution was incubated on ice for 10 min. The lysate was clarified by

centrifuging for 1 h at  $20\,000 \times g$ , 4 °C. The supernatant was discarded and the pellet was resuspended in 5 ml/g<sub>cells</sub> IB wash buffer I (20 mM EDTA, 500 mM NaCl, 2 % (w/v) Triton X-100 pH), followed by thorough homogenization. The solution was centrifuged (30 min at  $20\,000 \times g$ , 4 °C), the supernatant discarded and the pellet was washed twice more. To remove detergent, the pellet was washed twice again with IB washing buffer II (20 mM EDTA, 100 mM Tris/HCl pH 7). The IBs were resuspended in IB solubilization buffer (100 mM Tris/HCl, 5 mM DTT, 6 M GdmCl pH 7), such that the protein concentration was about 25 mg/ml and stored at -20 °C until use.

### 3.4.10 Purification of His-tagged proteins using immobilized metal affinity chromatography (IMAC)

N- or C-terminal oligo-histidine tags (His-tags) are a common tool to ease purification of recombinantly produced proteins. The free electron pairs of the imidazol nitrogens of histidines can complex divalent cations such as Mg<sup>2+</sup> or Ni<sup>2+</sup>, which are usually immobilized on a matrix of nitrilo triacetic acid (NTA)-derivatives. The affinity of the His-tag is correlated with its length and tagged proteins can simply be eluted by increasing the concentration of competing molecules (e.g. imidazole). His-tagged protein was purified by fast protein liquid chromatography (FPLC) via Ni<sup>2+</sup>- (HisTrap FF crude) or Co<sup>2+</sup>-NTA (HiTrap Talon FF crude) columns, obtained from GE Healthcare (Freiburg, Germany), following modified suppliers instructions. First the column was equilibrated with 5 column volumes (CV) of binding buffer (50 mM Tris/HCl, 500 mM NaCl, 10 % glycerol, 2.5 mM imidazole pH 7). The sample (generally clarified lysate) was applied to the column using a flow of 0.75 ml/min. Unbound protein was removed by washing with 3 CV binding buffer. Unspecifically bound proteins were washed away by increasing the amount of elution buffer (50 mM Tris/HCl, 500 mM NaCl, 10 % glycerol, 250 mM imidazole pH 7) to 10 % (constant for 3 to 5 CV). Highly enriched and purified target protein was eluted with 6 to 10 CV of 100 % elution buffer.

### 3.4.11 Refolding of SOMT-2 on a micro scale using design of experiments (DoE)

Design of experiments (DoE) and fractional factorial design (FrFD) have been successfully used to optimize the refolding conditions of several proteins [4, 11, 210]. Thus, an approach using fractional factorial design (FrFD) was used to find optimal refolding conditions for SOMT-2. Factors studied were pH (buffer), arginine, glycerol, divalent cations, ionic strength, redox system, cyclodextrin and co-factor addition. The experimental matrix was constructed using the FrF2 package (<http://cran.r-project.org/web/packages/FrF2/index.html>) in the R software.

**Table 3.7.:** Factors used in the construction of the FrFD.

factor	symbol	setting		unit
		-1	+1	
pH	A	5.5	9.5	-
arginine	B	0	0.5	M
glycerol	C	0	10	% (v/v)
divalent cations <sup>1</sup>	D	no	yes	-
ionic strength <sup>2</sup>	E	low	high	-
redox state <sup>3</sup>	F	reducing	redox-shuffling	-
$\alpha$ -cyclodextrin	G	0	30	mM
SAH	H	0	0.5	mM

**Table 3.8.:** Experimental design matrix for the FrFD.

Experiment	A	B	C	D	E	F	G	H
1	+	+	+	-	-	-	-	+
2	-	-	-	-	-	-	-	-

<sup>1</sup>no: 1 mM EDTA; yes: 2 mM CaCl<sub>2</sub>, MgCl<sub>2</sub>

<sup>2</sup>low: 10 mM NaCl, 0.5 mM KCl; high: 250 mM NaCl, 10 mM KCl

<sup>3</sup>reducing: 5 mM DTT; redox-shuffling: 1 mM glutathione (GSH), 0.2 mM glutathione disulfide (GSSG)

Experiment	A	B	C	D	E	F	G	H
3	+	-	+	+	-	+	+	-
4	-	+	+	-	+	+	+	-
5	+	+	-	-	+	+	-	-
6	-	+	-	+	+	-	+	+
7	+	+	-	+	-	-	+	-
8	-	-	+	-	+	-	+	+
9	+	-	+	+	+	-	-	-
10	-	-	-	-	-	+	+	+
11	+	-	-	+	+	+	-	+
12	-	+	+	+	-	+	-	+

The buffers were mixed from stock solutions and prepared in 1.5 ml microcentrifuge tubes immediately prior to the experiment. 50 µl of solubilized SOMT-2 (1 mg/ml) in IB solubilization buffer was added to 1 ml of each buffer followed by a short vortex boost for rapid mixing. The final protein concentration in the refolding reaction was 50 µg/ml, whereas the remaining GdmCl concentration was ≈286 mM. The refolding reactions were incubated at RT for 1 hour, followed by an overnight incubation at 4 °C. After incubation the refolding reactions were centrifuged (10 000 × *g*, 4 °C, 10 min) to separate insoluble and soluble protein fractions. The supernatant was transferred to a new tube, whereas the pellet was washed twice with 200 µl acetone and once with 400 µl methanol/acetone (1:1). The pellet was resuspended in 100 µl PBS with 20 µl SDS-PAGE sample buffer and 10 µl were used for SDS-PAGE analysis. 100 µl of the supernatant were concentrated using TCA precipitation (3.4.4) and analyzed by SDS-PAGE. The remaining supernatant was rebuffered into 50 mM 2-[Bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)propane-1,3-diol (BisTris) pH 7.5 using Amicon® Ultra 0.5 ml centrifugal filters (Merck, Darmstadt, Germany) according to the manufacturers instructions. The pre-weighed collection tubes were re-weighed after recovery and the volume of recovered liquid calculated ( $\rho \approx 1 \text{ g/cm}^3$ ). The sample was filled up to 100 µl using 50 mM BisTris pH 7.5 and the protein concentration was assessed using the Roti®-Quant protein quantification solution (Carl Roth, Karlsruhe, Germany) according to the man-

factors description. 50 µl of each refolded sample was used for an activity test using naringenin as substrate (3.6.3). The reactions were incubated over night and stopped by the extraction method. However, before the actual extraction 1 µl of anthracene-9-carboxylic acid (AC-9) was added as internal standard. The samples were analyzed by high-performance liquid chromatography (HPLC).

### Assessment of refolding performance

The performance of each buffer on the refolding of SOMT-2 was examined by comparing the SDS-PAGE results, as well as the amount of soluble protein and the conversion of naringenin over night (see subsection 3.6.3). Main effects were analyzed qualitatively using main effects plots [20].

### Upscaling of refolding reactions

Refolding reactions were scaled up to 50 ml. Therefore 2.5 ml solubilized SOMT-2 (1 mg/ml) were added over 10 minutes to 50 ml of refolding buffer while stirring at RT. The refolding reaction was allowed to complete over night at 4 °C.

### 3.4.12 Enzymatic production of SAM and SAE

S-adenosyl-L-methionine (SAM) and S-adenosyl-L-ethionine (SAE) were prepared according to the method described by Dippe, et. al [49].

Preparative reactions (20 ml) were performed in 0.1 M Tris/HCl, 20 mM MgCl<sub>2</sub>, 200 mM KCl pH 8.0 and contained 7.5 mM adenosine triphosphate (ATP), 10 mM D,L-methionine or D,L-ethionine, for the production of SAM or SAE respectively, and 0.2 U S-adenosylmethionine synthase (SAMS) variant I317V. The reaction was stopped by lowering the pH to 4 using 10 M acetic acid after 18 h of incubation at 30 °C, 60 rpm. After 10 min incubation on ice the solution was centrifuged (15 000 × g, 10 min) to remove insoluble matter. The supernatant was transferred to a round bottom flask, frozen in liquid nitrogen and lyophilized. Crude products were extracted from the pellet using 73 % ethanol and purified using ion exchange chromatography (IEX). IEX was performed on a sulfopropyl sepharose matrix (25 ml) via isocratic elution (500 mM HCl). Before injection, the crude extract was

acidified to 0.5 M HCl using concentrated hydrochloric acid. After elution, the product containing fractions were dried via lyophilization. The amount of product was determined by UV/VIS-spectroscopy at 260 nm using the published extinction coefficient of SAM ( $\epsilon_0 = 15\,400\,M^{-1}\,cm^{-1}$ ) after resuspension in water [175].

## 3.5 Crystallographic Procedures

### 3.5.1 Crystallization of proteins

Commercially available crystallization screens were used to find initial crystallization conditions. The tested screens included kits available from Hampton Research (Aliso Viejo, USA) and Jena Bioscience (Jena, Germany). Crystallization screens were processed in 96-well micro-titer plate (MTP)s, where each well possessed 4 subwells aligned in a  $2\times 2$  matrix. The subwells were divided into 3 shallow wells for sitting drop vapour diffusion experimental setups and a fourth subwell, which was deep enough to act as buffer reservoir. This way the performance of each crystallization buffer could be assessed using three different protein solutions with varying concentrations, effectors etc. A pipetting robot (Cartesian Microsys, Zinsser-Analystik; Frankfurt, Germany) was used to mix 200 nl of each, protein and buffer solution, for a final volume of 400 nl. The crystallization preparations were incubated at 16 °C and the progress of the experiment was documented by an automated imaging-system (Desktop Minstrel UV, Rigaku Europe, Kent, UK). Furthermore, fine screens (e.g. for refinement of crystallization conditions) were set up by hand in 24-well MTPs using the hanging drop vapour diffusion method.

#### PFOMT

PFOMT protein was concentrated to (6 to 8) mg/ml and rebuffered to 10 mM Tris/HCl pH 7.5 using Amicon® Ultracel centrifugal concentrators (10 kDa MWCO). The concentrated protein solution was centrifuged at  $14\,000 \times g$ , 4 °C for 10 min to remove any insoluble material or aggregates. Flavonoids and phenylpropanoid substrates were added to the protein solution from 10 mM stock solution in dimethyl sulfoxide (DMSO). Crystallization screens were set up as described above.

*apo*-PFOMT was crystallized using the following conditions – 2 M  $(\text{NH}_4)_2\text{SO}_4$ , 20 %glycerol. The protein solution contained 0.25 mM SAE, 0.25 mM  $\text{MgCl}_2$ , 0.25 mM eriodictyol and 7.53 mg/ml (0.262 mM) PFOMT .

### Crystallization of proteins using NADES

NADES have the potential to be excellent solvents for hydrophobic compounds such as flavonoids or cinnamic acids [42] and in addition they are able to stabilize and activate enzymes [81].

Four different model proteins (bovine trypsin, hen-egg white lysozyme, proteinase K and *Candida cylindrica* lipase B) were used to assess the capability of NADES for protein crystallization. PCH was tested in a full factorial grid layout using PCH concentrations of (20, 30, 40 and 50) % combined with buffers of different pH. The buffers included 0.1 M sodium acetate pH (4.5 and 5.5), 0.1 M sodium citrate pH 6.5, 0.1 M 2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid (HEPES)/NaOH pH (7 and 7.5) and 0.1 M Tris/HCl pH 8.5. Thus, the full factorial design had a size of  $4 \times 6 = 24$  different conditions. Protein solutions were prepared from lyophilized protein and were as follows: 90 mg/ml trypsin in 10 mg/ml benzamidine, 3 mM  $\text{CaCl}_2$ ; 75 mg/ml lysozyme in 0.1 M sodium acetate pH 4.6; 24 mg/ml proteinase K in 25 mM Tris/HCl pH 7.5 and 6 mg/ml lipase B in water. For crystallization 2  $\mu\text{l}$  enzyme solution and 1  $\mu\text{l}$  reservoir buffer were mixed and set up in a hanging drop experiment on a 24-well MTP. The experiments were set up at 4 °C.

### 3.5.2 Data collection and processing

Crystallographic data were collected at the beamline of the group of Professor Stubbs (MLU, Halle, Germany). The beamline was equipped with a rotating anode X-ray source MicroMax007 (Rigaku/MSC, Tokio, Japan), which had a maximum power of 0.8 kW (40 kV, 20 mA) and supplied monochromatic  $\text{Cu-K}\alpha$ -radiation with a wavelength of 1.5418 Å. Diffraction patterns were detected with a Saturn 944+ detector (CCD++, Rigaku/MSC, Tokio, Japan).

Indexing and integration of the reflexes via Fourier transformation (FT) was accomplished using *XDS* [92, 93, 94] or *MOSFLM* [157]. *Scala* [57], which is integrated in the Collaborative Computational Project No. 4 (CCP4)-Suite, was used for scaling of the intensities.

### 3.5.3 Structure solution

For the determination of the electron density  $\rho(\mathbf{r})$ , where  $\mathbf{r}$  is the positional vector, from the diffraction images by FT two terms are necessary as coefficients; the *structure factor amplitudes*,  $F_{\text{obs}}(\mathbf{h})$  and the *phase angles* or *phases*,  $\alpha(\mathbf{h})$ , where  $\mathbf{h}$  is the reciprocal index vector. The structure factor amplitudes can be directly determined from the measured and corrected diffraction intensities of each spot. However, the phase information is lost during the detection of the diffracted photons and there is no direct way to determine the phases. This constitutes the so-called *phase problem*. Thus, additional phasing experiments are necessary in order to obtain the phases. A variety of phasing experiments are available, which include *marker atom substructure methods*, *density modification* and *molecular replacement* (MR) techniques [193]. Phases of the structures herein were exclusively determined by MR [161, 162].

MR was performed using the software *Phaser* [139, 140], which is included in the CCP4-Suite [212]. A previously published PFOMT structure (PDB-code: 3C3Y [108]) was used as a template during MR procedure for the PFOMT structure solution. For the MR of the lysozyme structure the PDB-entry 4NHI was used.

### 3.5.4 Model building, refinement and validation

Macromolecular model building and manipulation, as well as real space refinement and Ramachandran idealization were performed using the Crystallographic Object-Oriented Toolkit (*Coot*) software [54]. Structure refinement was done using the software REFMAC5 [150, 198] as part of the CCP4-suite or the Phyton-based Hierarchial Environment for Integrated Xtallography (PHENIX) [1]. Validation of the structures was carried out using the web service MolProbity (<http://molprobity>.

[biochem.duke.edu/](http://biochem.duke.edu/)) [30]. Structure visualization and the preparation of figures was performed using PyMOL (Schrödinger, New York, USA).

### 3.5.5 ***In silico* substrate docking**

*In silico* molecular docking studies were performed using the AutoDock Vina 1.1.2 or AutoDock 4.2.6 software in combination with the AutoDockTools-Suite (<http://autodock.scripps.edu/>) [82, 147, 195]. Substrates were docked into the PFOMT structure with the PDB-code 3C3Y. The grid box, which determines the search space, was manually assigned to center at 1.581, 5.196 and 25.718 (x, y, z) and had size of (22, 20 and 25) Å (x, y, z). The exhaustiveness of the global search for AutoDock Vina was set to 25, whereas the rest of the input parameters were kept at their defaults.

## 3.6 Analytics

### 3.6.1 Recording of growth curves

Starter cultures ( $\approx$ 2 ml) of the transformed *E. coli* cells were prepared in the medium to be studied, containing the appropriate antibiotics. The cultures were incubated at 37 °C, 200 rpm over night and harvested by centrifugation (5000  $\times g$ , 4 °C, 5 min). The pellet was resuspended in 15 ml PBS and the suspension centrifuged (5000  $\times g$ , 4 °C, 5 min). The supernatant was discarded and the washing step repeated once more. The washed pellet was resuspended in 2 ml of the medium to be studied with the appropriate antibiotics and the OD<sub>600</sub> was measured. Three independent 50 ml cultures of the medium containing the appropriate antibiotics were inoculated such that OD<sup>600</sup>  $\approx$  0.05 using the washed cell suspension. The cultures were incubated at the conditions to be studied and sampled at appropriate intervals of time ( $\approx$ 1 h). One ml samples were kept on ice until all samples were acquired. 100 µl aliquots of the samples were transferred into a clear MTP and the OD<sub>600</sub> was measured. Green fluorescent protein (GFP) fluorescence was measured accordingly, but the MTP used was opaque. Exitation ( $\lambda^{\text{ex}}$ ) and emission ( $\lambda^{\text{em}}$ ) wavelengths were (470 and 510) nm respectively.

### 3.6.2 *In vitro* determination of glucose

The glucose concentration in clarified, aqueous samples was determined by a modified version of the glucose assay kit procedure provided by Sigma-Aldrich [178]. Glucose oxidase (GOD) oxidizes D-glucose to gluconic acid, whereby hydrogen peroxide is produced. The hydrogen peroxide can be detected and quantified by horseradish peroxidase (HRP), which reduces the produced H<sub>2</sub>O<sub>2</sub> and thereby oxidizes its chromogenic substrate *o*-dianisidine via consecutive one-electron transfers. The oxidized diimine form of *o*-dianisidine can then be measured photospectrometrically [31].

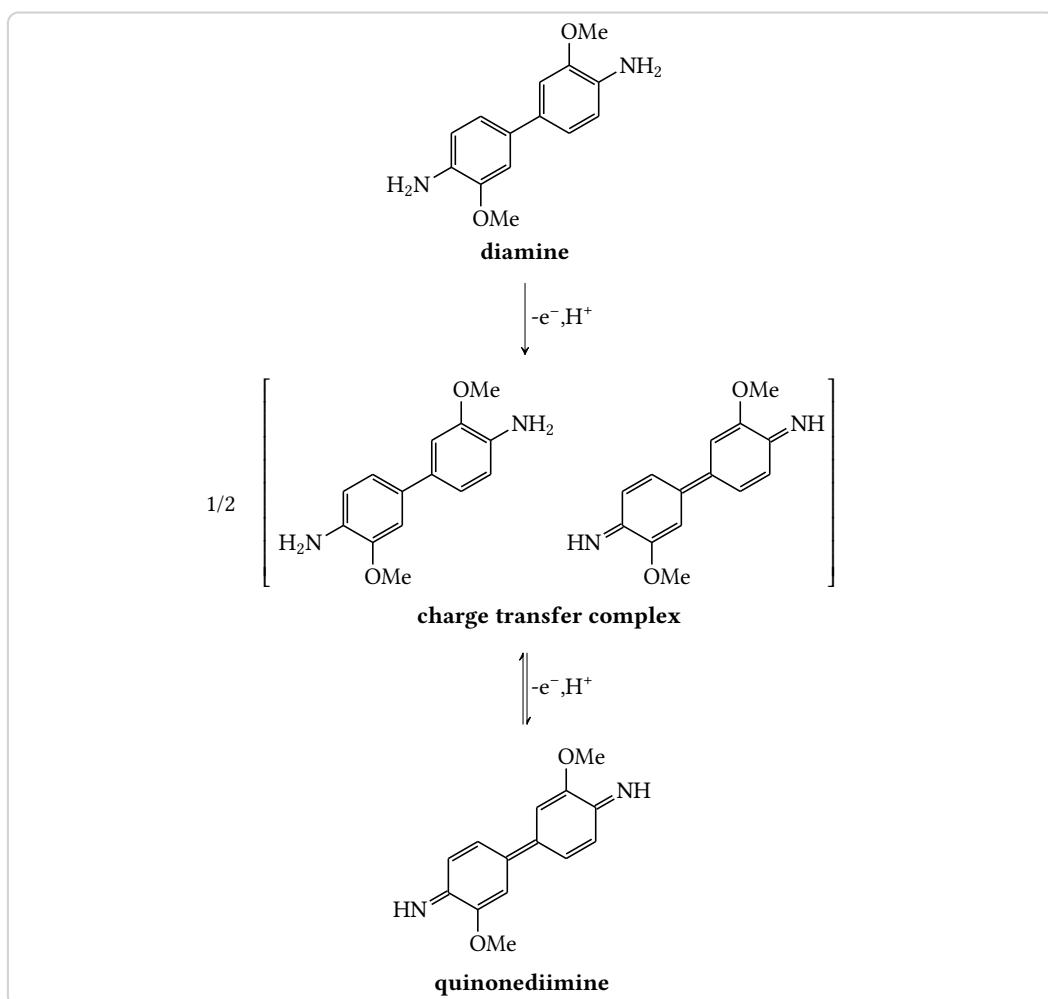
The methodology employs a coupled photospectrometric assay using GOD and HRP with *o*-dianisidine as reporter substrate. The assay was prepared in MTP-format. A reaction solution containing 12.5 U/ml GOD, 2.5 U/ml HRP and 0.125 mg/ml *o*-dianisidine dihydrochloride in 50 mM sodium acetate pH 5.1 was prepared.

Sample solutions from culture supernatants were typically diluted in 9 volumes of water. The reaction was started, by adding 50 µl reaction solution to 25 µl of sample and was incubated at 37 °C and 200 rpm for 30 min in a shaking incubator. 50 µl 6 M sulfuric acid was added to stop the reaction and achieve maximum color development (full oxidation of any *o*-dianisidine charge transfer complexes) (Figure 3.3). The developed pink color was measured at 540 nm in a MTP-reader. A calibration curve of a standard D-glucose solutions (0 to 100 µg/ml), that was always part of the experiments, was used to quantify the sample measurements.

### 3.6.3 *In vitro* O-methyl transferase (O-MT) assay

O-methyl transferase (O-MT) assays were conducted in a total volume of (50 to 100) µl. The standard assay buffer was 100 mM Tris/HCl, 2.5 µM GSH pH 7.5. 1 mM MgCl<sub>2</sub>, which was otherwise omitted, was added for reactions using cation dependent O-MTs (e.g. PFOMT). Reactions contained 0.5 mM alkyl donor (e.g. (S,S)-SAM) and 0.4 mM flavonoid or cinnamic acid substrate. Enzymatic reactions were started by addition of enzyme (usually 0.2 mg/ml) and incubated at 30 °C.

Reactions were stopped by addition of 500 µl ethyl acetate containing 2 % formic acid and vortexed for 15 s to extract the hydrophobic phenylpropanoids and flavonoids.



**Figure 3.3.:** Oxidation of the reporter substrate o-dianisidine. Consecutive one-electron transfers lead to the fully oxidized diimine form of o-dianisidine. The first electron transfer is believed to produce a charge transfer complex intermediate. [31, 90]

After centrifugation ( $10\,000 \times g$ , 4 °C, 10 min) the organic phase was transferred into a new tube. The reaction was extracted once more with 500 µl ethyl acetate, 0.2 % formic acid and the pooled organic phases were evaporated using a vacuum concentrator (Concentrator 5301; eppendorf, Hamburg, Germany). The residue was dissolved in methanol and centrifuged at  $10\,000 \times g$  for 10 min to remove unsoluble matter. The supernatant was transferred into a HPLC vial and analyzed by HPLC (3.6.8).

When detection of hydrophobic (e.g. flavonoids) and hydrophilic compounds (e.g. SAM, SAH) was performed simultaneously reactions were stopped by addition of 0.3 volumes 10 % (w/v) TCA in 50 % acetonitrile. The mixture was vortexed for complete mixing and incubated on ice for at least 30 min. After centrifugation ( $10\,000 \times g$ , 4 °C, 10 min) the supernatant was transferred into HPLC-sample vials and analyzed (see 3.6.8).

### Measurement of activity/pH profiles

Assays to measure activity over larger pH ranges were set up in 50 mM L-malic acid/MES/Tris (MMT)- (pH 4 to 9) or succinate/sodium phosphate/glycine (SSG)-buffer (pH 4 to 10) to keep the concentrations of buffer salts constant for each pH [151].

The protein of interest was first extensively dialyzed against the reaction buffer (e.g. MMT, SSG) at pH 7 with added EDTA (5 mM) and then against the same buffer without EDTA. Standard reaction conditions were 50 mM buffer, 0.4 mM alkyl acceptor (e.g. caffeic acid), 0.5 mM SAM, 2.5 µM GSH and 0.2 mg/ml enyzme. MgCl<sub>2</sub> was either omitted or added at 10 mM to assess influences of divalent cations. Assays were stopped as described in 3.6.3 and analyzed accordingly.

### Estimation of product concentration

Product concentrations were estimated from HPLC runs. The automatically integrated peaks of SAM and SAH provided the area under the curve (AUC). From the AUC of both peaks the concentrations were estimated as follows.

Under the assumption, that

$$AUC^{\text{SAH}} + AUC^{\text{SAM}} = 1 \sim c_0^{\text{SAM}}, x$$

the fraction and concentration of one (e.g. SAH) can be estimated by

$$x^{\text{SAH}} = \frac{AUC^{\text{SAH}}}{AUC^{\text{SAH}} + AUC^{\text{SAM}}}$$

and

$$c^{\text{SAH}} = x^{\text{SAH}} \times c_0^{\text{SAM}}.$$

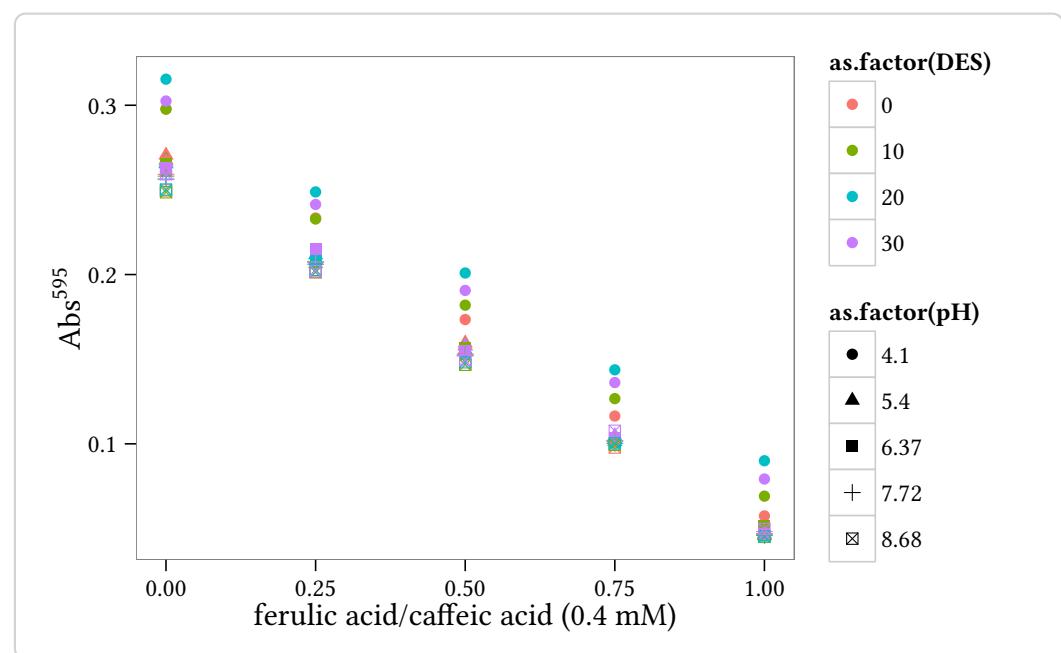
Enzymatic activities can be calculated from the concentrations by standard procedures.

### 3.6.4 Photospectrometric assay for the methylation of catecholic moieties

Catecholic moieties can form stable complexes in the presence of heavy metals such as copper or iron [141, 172]. Hence, caffeic acid can complex ferric ( $\text{Fe}^{3+}$ ) ions and form a colored complex with  $\lambda_{\text{max}} = 595 \text{ nm}$  [48]. Since the complex formation is specific for caffeic acid and methylated derivatives (i.e. ferulic and iso-ferulic acid) cannot complex  $\text{Fe}^{3+}$ , this can be used as a measure for methylation reactions. O-MT assays were prepared as before (3.6.3). However, the reactions were stopped by addition of 0.1 volumes 1 M Tris/HCl pH 8, immediately followed by 0.5 volumes catechol reagent (2 mM  $\text{FeCl}_3$  in 10 mM HCl). The complex formation reaction was allowed to equilibrate for 5 min at RT and the absorbance at 595 nm was measured.

### 3.6.5 Concentration of SOMT-2 using hydrophobic interaction chromatography (HIC)

After refolding using rapid dilution protein samples are very dilute and a concentration step is required. Refolded SOMT-2 was concentrated directly from the



**Figure 3.4.:** Calibration curves of different relative compositions of ferulic acid to caffeic acid, that were taken as described in 3.6.4. The total concentration was always 0.4 mM. At lower pH values around 4, the method seems to overestimate the concentration of caffeic acid. However, the slope of the curves stays the same.

refolding buffer using hydrophobic interaction chromatography (HIC). The ammonium sulfate concentration of the protein sample was brought to 1 M using a 2 M  $(\text{NH}_4)_2\text{SO}_4$  solution and the pH was adjusted to 7 using 5 M NaOH. The sample was centrifuged ( $20\,000 \times g$ , 4 °C, 30 min) to remove insoluble material and the clarified supernatant was applied to a 1 ml HiTrap Phenyl FF (Low Sub) (GE Healthcare, Freiburg, Germany), which had been equilibrated with high salt buffer (1 M  $(\text{NH}_4)_2\text{SO}_4$ , 50 mM HEPES pH 7). The target protein was eluted using a stepwise gradient ((1, 0.8, 0.6, 0.4, 0.2 and 0) M  $(\text{NH}_4)_2\text{SO}_4$ , 50 mM HEPES pH 7; 5 CV each) to remove the ammonium sulfate. The column was washed using 20 % ethanol. Before SDS-PAGE analysis the eluted high salt fractions were desalting using TCA precipitation (3.4.4).

### 3.6.6 Analytical gel filtration

Analytical gel filtration was done using a Superdex 200 10/300 GL column (GE Healthcare, Freiburg, Germany) in combination with a FPLC system according to the manufacturers instructions. The column was equilibrated using an appropriate buffer (e.g. 0.1 M Tris/HCl pH 7.5) and 100 µl of sufficiently concentrated ( $\geq 1$  mg/ml) protein sample were injected. The Gel Filtration Standard by Bio-Rad (München, Germany) was run separately to assess the size of the proteins in the analyzed sample.

### 3.6.7 Binding experiments using Isothermal Titration Calorimetry (ITC)

Isothermal Titration Calorimetry (ITC) can be used to directly characterize the thermodynamics of an observed process, be this a binding interaction or an enzymatic reaction [63].

ITC measurements to describe the interaction between PFOMT and its substrates/-effector were performed using a MicroCal iTC200 device (Malvern, Worcestershire, UK). PFOMT protein was extensively dialyzed against 50 mM MMT-buffer pH 7 prior to ITC experiments. The solution was subsequently centrifuged ( $14\,000 \times g$ , 4 °C, 10 min), to remove insoluble matter and aggregates. The dialysate was stored

at 4 °C and used to prepare substrate and effector solutions. Generally 50 µM protein was provided in the ITC cell and the effectors/substrates to be titrated were loaded into the syringe. The substance concentration in the syringe was ten times higher than the protein solution. Experiments were carried out at 20 °C unless otherwise stated. The stirring speed was set to 500 rpm. The injection volume was set to (2 to 4) µl, amounting to a total of 10 to 19 injections.

### 3.6.8 High-performance liquid chromatography (HPLC) analytics

Due to their aromaticity, methanolic extracts of flavonoids exhibit two major absorption peaks in the UV/VIS region of the light spectrum in the range of (240 to 400) nm [130]. However, even the more simple phenyl propanoids (e.g. cinnamic acids) show absorption of light in the UV/VIS-region.

Methanolic extracts of flavonoids and phenyl propanoids were analyzed by HPLC using a photo diode array (PDA)-detector, which was set to record in the range of (200 to 400) nm. HPLC runs were performed on a reverse-phase C-18 end-capped column (YMC-Pack ODS-A; YMC Europe, Dinslaken, Germany) with a pore size of 120 Å. The mobile phase was aqueous acetonitrile supplemented with 0.2 % formic acid. The flow was kept constant at 0.8 ml/min. 10 µl *O*-MT enzyme assay extract (3.6.3) were injected and analyzed using an acetonitrile gradient starting with 5 % acetonitrile (4 min). The acetonitrile content was increased to 100 % in 21 min and was kept at 100 % for 5 min. Peaks were integrated from the 280 nm trace using the software provided by the manufacturer of the device.

### 3.6.9 Liquid chromatography-tandem mass spectrometry (LC-MS/MS) measurements

The positive and negative ion high resolution electrospray ionization (ESI) and collision induced dissociation (CID) MS<sub>n</sub> spectra as well as higher-energy collisional dissociation (HCD) MS/MS spectra were obtained from an Orbitrap Elite mass spectrometer (Thermo Fisher Scientific, Bremen, Germany) equipped with an

heated-electrospray ionization (H-ESI) ion source (positive spray voltage 4.5 kV, negative spray voltage 3.5 kV, capillary temperature 275 °C, source heater temperature 250 °C, fourier transform mass spectrometry (FTMS) resolving power (RP) 30 000). Nitrogen was used as sheath and auxiliary gas. The MS system was coupled with an ultra-high performance liquid chromatography (UHPLC) system (Dionex UltiMate 3000, Thermo Fisher Scientific), equipped with a RP-C18 column (particle size 1.9 µm, pore size 175 Å, 50 x 2.1 mm inner diameter, Hypersil GOLD, Thermo Fisher Scientific, column temperature 30 °C) and a photodiode array detector ((190 to 400) nm, Thermofisher Scientific). For the UHPLC a gradient system was used starting from H<sub>2</sub>O:CH<sub>3</sub>CN 95:5 (each containing 0.2 % formic acid) raised to 0:100 within 10 min and held at 0:100 for further 3 min. The flow rate was 150 µl/min.

The mass spectra (buffer gas: helium) were recorded using normalized collision energies (NCE) of (30 to 45) % and (75 to 100) % for CID and HCD mass spectra respectively (see Appendix). The instrument was externally calibrated using the Pierce® LTQ Velos ESI positive ion calibration solution (product number 88323, ThermoFisher Scientific, Rockford, IL, 61105 USA) and the Pierce® LTQ Velos ESI negative ion calibration solution (product number 88324, ThermoFisher Scientific, Rockford, IL, 61105 USA) for positive and negative ionization mode respectively.

## **4 Engineering of phenylpropanoid and flavonoid O-methyl trans- ferase (PFOMT)**

## Evaluation of PFOMT towards the acceptance of long-chain SAM analogues

Benjamin Weigel<sup>1,a</sup>, Martin Dippe, Ludger A. Wessjohann<sup>1,c</sup>

Contact: bweigel@ipb-halle.de<sup>a</sup>, law@ipb-halle.de<sup>c</sup>

Affiliation: Leibniz-Institute of Plant Biochemistry, Department of Bioorganic Chemistry<sup>1</sup>

Keywords: methyl transferase, pfomt, SAM

### Abstract

The cation dependent phenylpropanoid and flavonoid O-methyl transferase (PFOMT) from the ice plant, *Mesembryanthemum crystallinum*, methylates a number of flavonoids and phenyl propanoids. A newly solved crystal structure of the protein without any bound ligand shows the fully resolved N-terminus, which acts as a lid to close off the active site. Binding of co-substrates (analogues) (e.g. S-adenosyl-L-homocysteine (SAH), S-adenosyl-L-methionine (SAM), S-adenosyl-L-ethionine (SAE)) is more entropically driven as the chain length increases. However, even though the ethyl-analogue of SAM – SAE – was shown to bind to the enzyme, no conversion of the model substrate caffeic acid was observed for the wild-type and several engineered variants.

---

### 4.1 Introduction

Small changes to molecules can have profound influences on their chemical, physical and biological properties. For example, butyric acid esters differing only by a few methylene groups already exhibit quite divergent smells. However, not only the macroscopically qualitative properties can differ. The quantifiable psychotomimetic effect of methylated and ethylated lysergic acid amids differ by at least an order of magnitude [78, 177]. There are many more of these so-called structure activity relationship (SAR) and quantitative structure activity relationship (QSAR) studies on any number of compounds [5, 133, 164].

Methylation reactions are one of the key tailoring steps during natural product biosynthesis and can in consequence greatly affect a molecules bio- and physico-chemical behavoir [119, 184]. Methyl transferases (MTs) catalyze the transfer of a methyl group from the co-substrate SAM to an activated atom of the acceptor molecule [184].

Between the highly complex core structures of natural products, which are produced by a plethora of enzymes (e.g. poly ketide synthases (PKSs), non-ribosomal peptide synthases (NRPSs), terpene cyclases), and the rather simple alkyl-modification introduced by methylation, nature is missing some medium-sized modifaction options that proceed as elegantly as the methylation by MTs. Thus, natural products containing longer chain alkyl modifications like ethyl or propyl moieties on O, N or S-centers have rarely, if ever been observed.<sup>1</sup>

It has recently been shown however, that a wide array of SAM analogues are used as co-substrates by a variety of MTs [184]. The majority of the work so far has been done on protein methyl transferases (P-MTs) and DNA methyl transferases (DNA-MTs) (Figure A.1), since epi-genetics and finding regions of gene-regulation is of great interest. However, small molecule methyl transferases (*sm*MTs) have also been shown to accept different SAM analogues [113, 115, 181, 183, 224]. There have been a great many of SAM analogues synthesized, both chemically and enzymatically, that were consequently studied with the help of MTs [43, 181, 184].

The *O*-methyl transferase (*O*-MT) PFOMT is a highly promiscuous enzyme with regards to its flavonoid substrates and has extensively been characterized [22, 85, 108, 202]. However, the promiscuity towards different SAM analogues has net yet been described. Combination of both, substrate and co-substrate promiscuity in the small molecule MT PFOMT could provide a powerful tool towards the biosynthetic production of novel small molecules with potentially new and promising biological activities. Functionalization/Detection of substrates could furthermore provide a means of finding new compounds/substrates in complex (e.g. biological) samples analogous to activity based protein profiling (ABPP) approaches.

---

<sup>1</sup>Reaxys searches for natural product isolates with a molecular mass between (150 and 1500) containing the substructures methyl, ethyl or propyl connected to a heteroatom return 66759, 2797 and 52 results respectively. However, it stands to note that 70 % of the propyl results were either esters or otherwise activated moieties. [53]

In this work we show, that PFOMT binds the co-substrate analogues SAH, SAM and SAE with similar affinities. A newly developed crystal structure of the *apo*-enzyme shows the fully resolved N-terminus is lodged in a cleft atop the active site, closing it off. Although semi-rationally designed enzyme variants could not afford enzymatic ethylation of substrates, the regio-selectivity of the methylation reaction was altered.

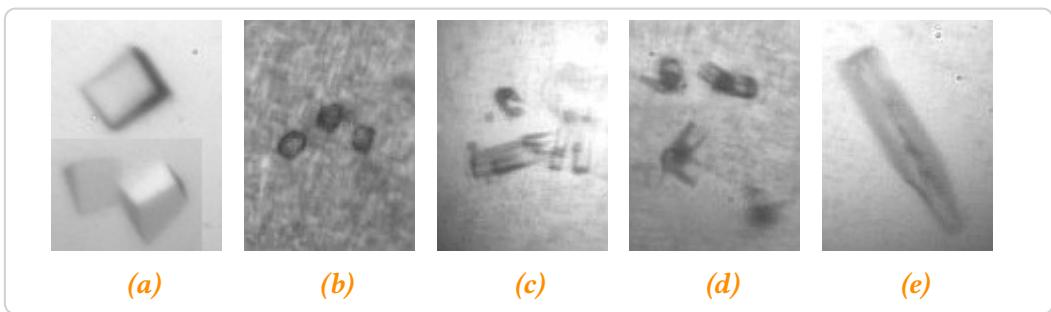
## 4.2 Crystallization of PFOMT

The crystal structure of PFOMT was published in 2008, however binding of substrates could not be accomplished [108]. Nonetheless, the demethylated co-substrate SAH was cocrystallized. The first goal of this study was to crystallize the *apo*-form of the enzyme, to obtain a system that allows for the soaking of substrates. At the same time, PFOMT was to be cocrystallized along with an acceptor substrate and the co-substrate analogs SAE and SAH.

At first the already available crystallization procedures were evaluated [108]. However, reproduction of these results could not be accomplished and new crystallization conditions had to be found.

Several commercially available buffer solutions (see section 3.5) were screened in combination with different protein solutions (e.g. solutions containing co-substrates and acceptor substrates or not) to obtain protein crystals co-crystallized with substrates or of the *apo*-form. Crystals were obtained in various wells after a few days. The crystal shape varied from very smooth and almost cubic (high ammonium sulfate) over sphreulites and intergrown crystals ( $\text{CaCl}_2$ , PEG-4000) to brittle and ragged needles ( $\text{LiCl}$ , PEG-6000) (Figure 4.1).

Crystals that were large enough ( $\geq 50 \mu\text{m}$ ), where screened for diffraction at the home-source after cryoprotection. A rough estimate of the resolution, cell parameters and the space group was acquired, if the diffraction images could be indexed. The screened crystals all had similar cell parameters and belonged to the same space group,  $P2_12_12_1$ , as the previously published structure (pdb: 3C3Y)[108]. However, the unit cell of crystals that grew out of high ammonium sulfate concentrations ( $\geq 1.8 \text{ M}$ ) was approximately four times as large as that of the published



**Figure 4.1:** Some crystal and pseudo-crystal shapes that were observed during the crystallization screen. a – high  $(\text{NH}_4)_2\text{SO}_4$ , b-c –  $\text{CaCl}_2$ , PEG-4000, e –  $\text{LiCl}$ , PEG-6000

structure. Several datasets were collected of crystals from high  $(\text{NH}_4)_2\text{SO}_4$ , since these seemed to be promising candidates to find differences in the bound substrates. Datasets of crystals that grew from other conditions were insufficient for structure solution.

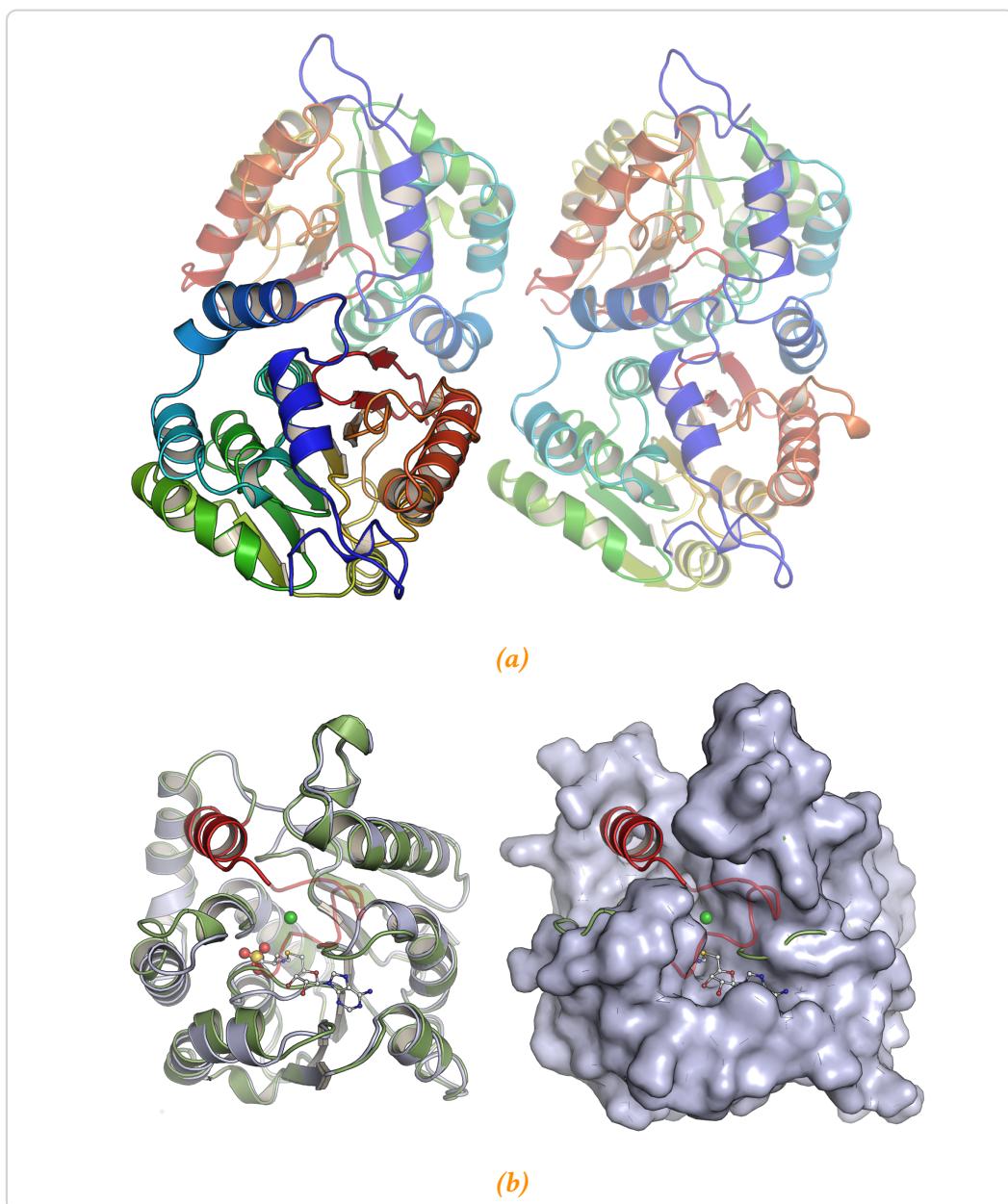
### The crystal structure of *apo*-PFOMT

PFOMT crystallized without any bound substrates under conditions of high  $(\text{NH}_4)_2\text{SO}_4$ . One dataset was solved to completion to obtain a complete structure of this novel *apo*-PFOMT at a resolution of 1.95 Å (Table B.4). The assymmetric unit of *apo*-PFOMT contained two homodimers (4 monomers) (Figure 4.2a), rather than just one homodimer (3C3Y). The active site of each monomer was found to be empty except for a sole sulfate ion, which was positioned where the amino- and carboxylate groups of the SAH residue in the 3C3Y structure (Figure 4.2b). Shifts in the structure of some loops were observed and contrary to the previously published structure the entire N-terminus was resolved up to and including the His-tag.

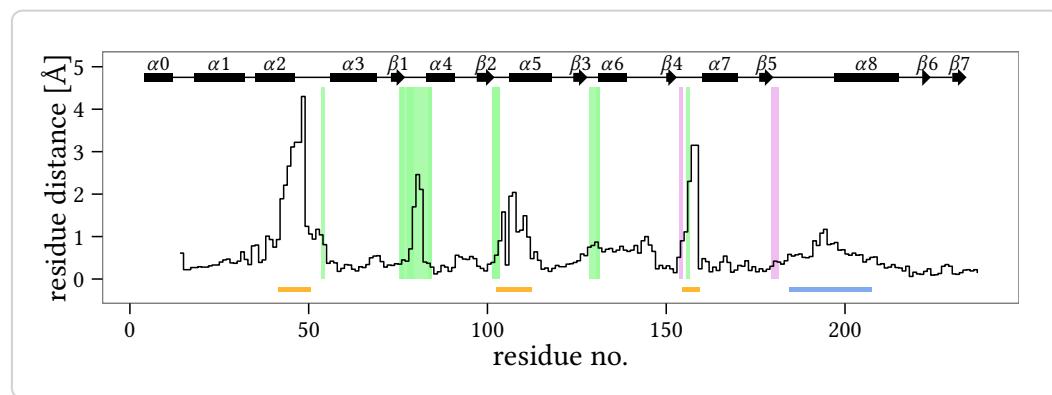
The resolved N-terminus contained another N-terminal  $\alpha$ -helix, which was positioned in a cleft on the surface, where substrates may be bound [108]. This interaction extends up to the His-tag. Considerable movement was observed in different parts of the protein, when no substrate was bound, some of which can be attributed to SAM and metal ion binding residues (Figure 4.3 and Figure A.2) as is obvious for the loop region between  $\beta$ -sheet 1 and  $\alpha$ -helix 4. Nonetheless, most of the movement seemed to be restricted to areas, which are not directly involved in

the binding of either SAM or metal ions. However, all of the regions that moved are located at or near the active site.

Unfortunately soaking of these “*apo*”-crystals did not afford binding of substrates.



**Figure 4.2.**: An overview of the features in the apo-PFOMT structure. **a** – The assymmetric unit of apo-PFOMT consists of two homodimers (4 monomers). Individual monomers are rainbow colored from N-terminus (blue) to C-terminus (red). **b** – Comparison of 3C3Y (steelblue) and apo-PFOMT (green). The N-terminus of apo-PFOMT was resolved up to the N-terminus (red) and even the His-tag (red, transparent) was partly resolved. The N-terminus fits into a cleft on the surface of the 3C3Y structure, shown as a surface model on the right. SAH (white ball-and-sticks) and  $\text{Ca}^{2+}$  (green sphere) are featured in the published structure, whereas a sulphate ion (red/yellow spheres) was bound in the newly solved structure.

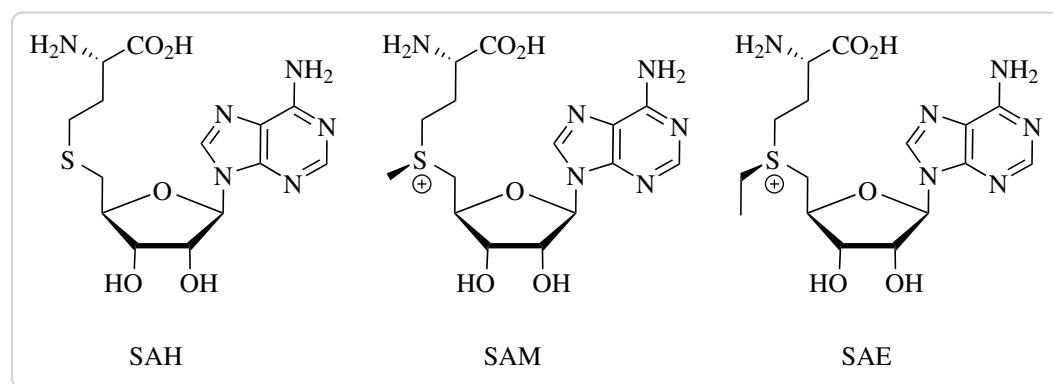


**Figure 4.3.**: Positional differences between the individual residues of the solved apo-PFOMT and the structure with bound SAH (pdb: 3C3Y). The diffraction precision indicator [45] (DPI) of the structures was (0.137 and 0.064) Å respectively. The overall rmsd amounted to 0.9034 Å. The secondary structure of apo-PFOMT is displayed at the top. Helices are displayed as rectangles and sheets are shown as arrows. Graphical background annotations are used to display the binding sites of SAH (green) and the metal ion (plum). The orange bars indicate regions, where much movement seems to happen upon binding or release of the co-substrate. The blue bar shows the region that was annotated as "insertion loop" in previous studies [108].

### 4.3 Substrate binding studies using ITC

The binding of different substrates to PFOMT was examined by Isothermal Titration Calorimetry (ITC), to determine whether the enzyme can bind non-natural SAM analogues. The homologues SAH, SAM and SAE were selected to also study the influence of the alkyl chain length on binding (Figure 4.4). Furthermore the binding of the substrate caffeic acid and the influence of Mg<sup>2+</sup> addition on substrate binding was investigated.

The  $K_D$  values of SAH, SAM and SAE were all in the low micromolar range, around 2 μM. However, the binding enthalpy clearly decreased with the length of the aliphatic chain connected to the sulfur atom (Figure 4.5a). The binding of SAH, gave off more heat than the binding of SAM, which in turn gave off more heat than the binding of SAE (Table 4.1). Thus, the entropic influence must get larger with



**Figure 4.4:** The binding of different SAM analogues was measured via ITC.

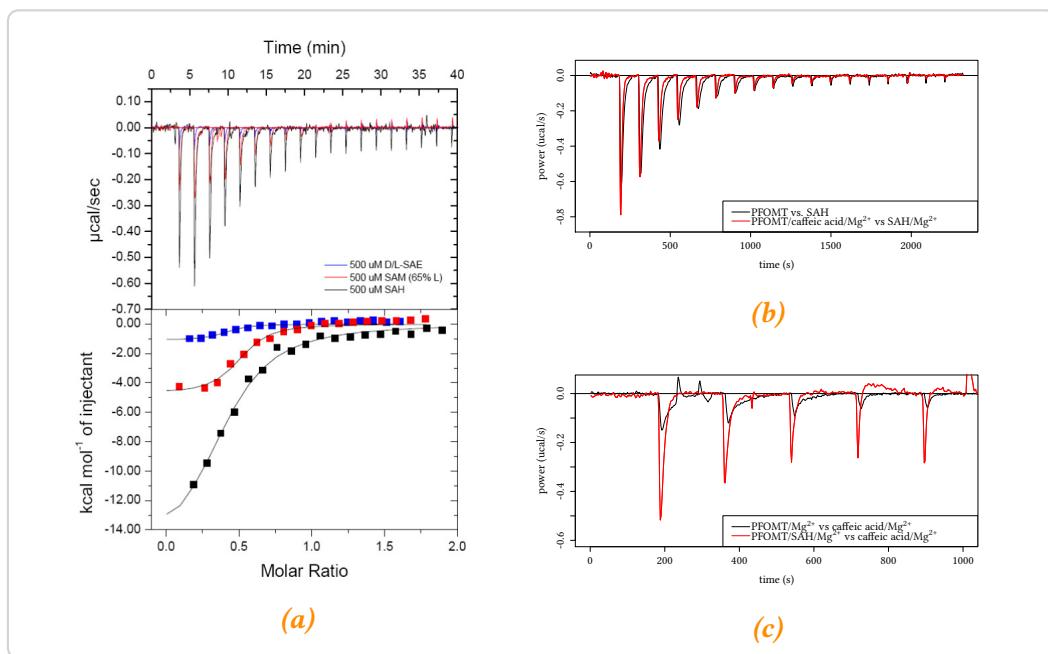
increasing chain length in order for equations (4.1) and (4.2) to still hold true.

$$\Delta G = \Delta H - T\Delta S \quad (4.1)$$

$$\Delta G = \Delta G^0 - RT \ln K \quad (4.2)$$

Indeed, the value for  $\Delta S$  was negative for binding of SAH, but positive for the binding of SAM and SAE (Table 4.1). This relationship between the change of entropy and the change of enthalpy has been found for many biological systems and is called enthalpy-entropy compensation (EEC) [52, 71, 174]. The stoichiometry for the binding process is given by the parameter  $N$ . For all the ligands SAH, SAM and SAE this value was found to be about 0.5, which corresponds to one bound molecule ligand per dimer of PFOMT (Table 4.1).

Upon titration of caffeic acid to PFOMT small amounts of released heat were detected for the system (Figure 4.5c). When the enzyme was incubated with SAH prior to addition of caffeic acid the released heat was slightly increased. The slope of the ITC profile also got steeper. However, the data obtained could not be fitted to afford a sensible solution. When caffeic acid and  $Mg^{2+}$  were incubated with PFOMT prior to addition of SAH, the process of heat production as observed by ITC had a steeper slope (Figure 4.5b). Nonetheless, the thermodynamic parameters did not differ significantly.  $Mg^{2+}$ , in the form of an  $MgCl_2$  solution, titrated to the enzyme solution did not cause signals during the ITC experiments.



**Figure 4.5:** ITC measurements of PFOMT:effector binding. **a** – Binding of SAH, SAM and SAE to PFOMT. **b** – SAH is injected into a PFOMT solution, with (red) or without (black) addition of  $Mg^{2+}$  and caffeic acid. When  $Mg^{2+}$  and caffeic acid were already present, the binding process seems to happen quicker, but is less enthalpic. **c** – Upon addition of caffeic acid to the protein heat is produced, however no sensible binding curve could be obtained.

**Table 4.1.**: Results of fitting a simple one-site binding model to the data obtained from ITC experiments.

	$K_D$ [ $\mu\text{M}$ ]	$\Delta H$ [cal $\text{mol}^{-1}$ ]	$\Delta S$ [cal $\text{mol}^{-1} \text{K}^{-1}$ ]	N
SAH	$2.06 \pm 4.27$	$-10\,380 \pm 1025$	-9.41	$0.505 \pm 0.038$
SAM	$1.08 \pm 3.50$	$-4606 \pm 242$	11.6	$0.492 \pm 0.018$
SAE	$2.22 \pm 3.79$	$-1338 \pm 190$	21.3	$0.513 \pm 0.050$

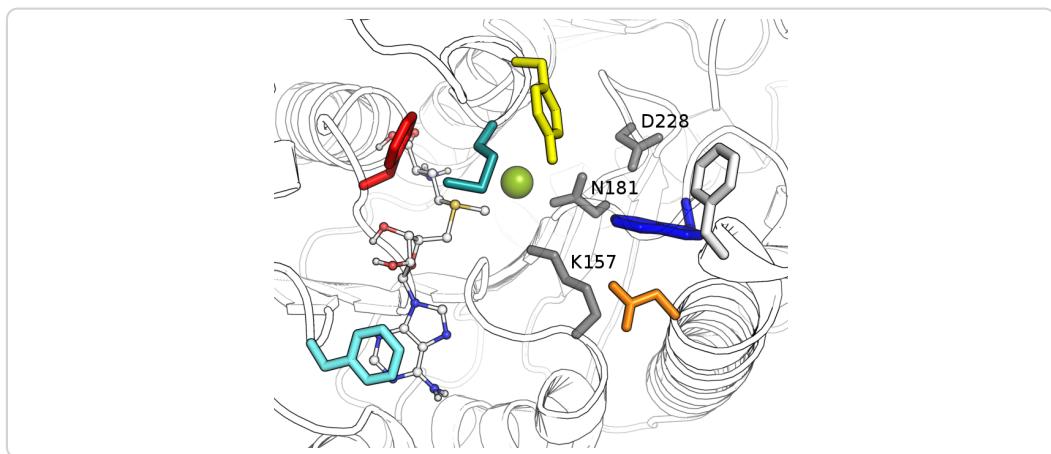
## 4.4 Study of variants for long-chain alkylations

Since the ability to bind the elongated analogue SAE was present in wild-type PFOMT, the activity of the PFOMT protein towards SAE was tested. Activity tests were performed with caffeic acid as substrate under standard reaction conditions. Unfortunately no ethylation of the substrate by PFOMT was observed, even after extended incubation times.

Consequently enzyme variants were prepared to achieve a PFOMT variant with an ethylation activity, since a number of groups were able to accomplish transalkylation with larger substrates by expanding the available space in the active site [205]. The available crystal structures of PFOMT were consulted to select suitable residues. Residues that were exchanged were selected based upon their position in the active site and in relation to the substrate(s) (Figure 4.6). The residues were exchanged to the non-spaceous alanine, as well as amino acids frequently observed at homologous positions in other class I O-MTs.

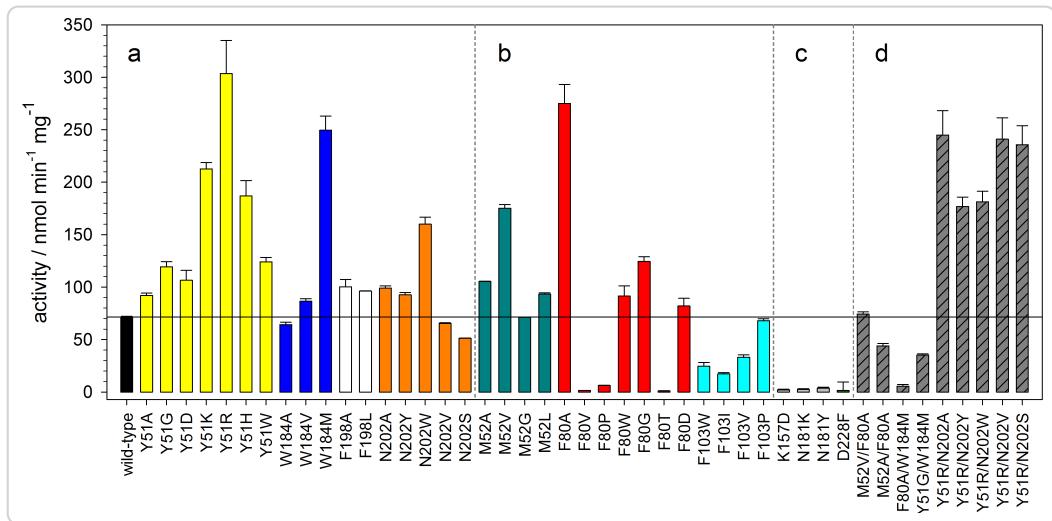
Over 20 enzyme variants were prepared to assess, whether PFOMT ethylation activity would improve over the wild-type. However, no ethylation activity was observed for either variant. Some of the new variants however displayed an increased methylation activity with the substrates caffeic acid and SAM (Figure 4.7). The methylation activity of some of the variants increased by over 4-fold. Interestingly most amino acid substitutions proved as beneficial.

Methylation activity benifited greatly from the replacement of bulky hydrophobic residues by smaller and/or charged residues in the vicinity of the acceptor substrates (Tyr51, Trp184 and Phe198). However, this was not a general trend since the



**Figure 4.6.:** The active site of PFOMT (pdb: 3C3Y). The outline of the protein backbone is displayed, with active site residues portrayed as colored sticks (cyan – F103, red – F80, turquoise – M52, yellow – Y51, white – F198, blue – W184, orange – N202, grey – as labelled). The co-substrate SAM (ball-and-stick model) was docked into the structure.

substitutions N202W and Y51W also improved methylation activity. Looking more closely at residue Tyr51, the activity enhancing effect was greatest, when the tyrosine was substituted by the basic amino acids lysine or arginine. In addition to an enhanced activity the selectivity for the hydroxyl position to be methylated was also altered in these variants. This was not apparent, when caffeic acid was used as a substrate. However when a flavonoid, especially eriodictyol, was used not only the 3' hydroxyl, but to some extent the 4' hydroxyl was methylated (Figure A.3). This effect was improved in some double variants, where also position 202 was altered. For example the variant Y51R N202W almost exclusively methylated flavonoid substrates at the 4' position. A detailed discussion of the results was published in a peer reviewed journal.



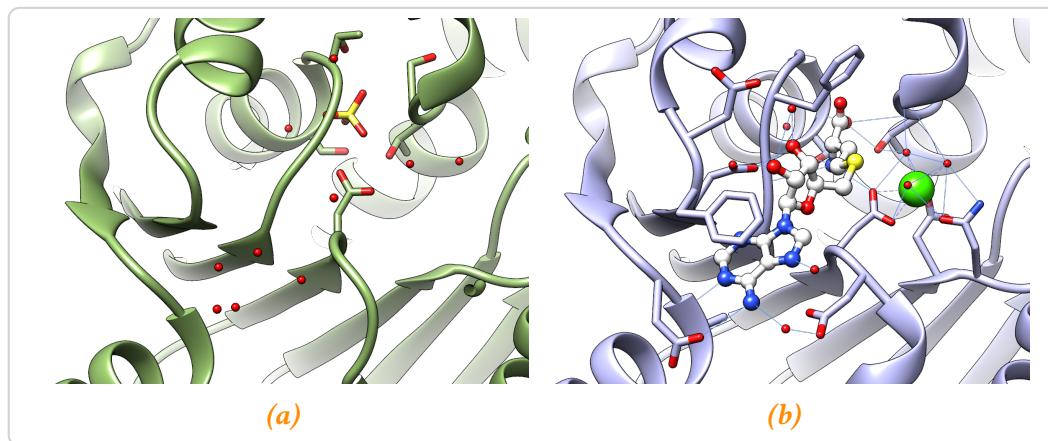
**Figure 4.7.:** Activities of different PFOMT variants towards caffeic acid methylation. Colorations correspond to the ones used in Figure 4.6.

## 4.5 Conclusion/Discussion

Whereas the binding of SAH was solely dependent on the large negative enthalpy, the binding of SAE was almost entirely driven by entropy, since  $\Delta H$  was close to 0 (Table 4.1). Entropy gain can be a major driving force for ligand-protein interactions and in some cases ligand binding can be entirely attributed this gain in entropy [122]. Displacement of protein-bound water molecules contributes strongly to the entropic gain. There were some waters present in the active site of PFOMT in the crystal structure developed herein. However, no metal ion was present in the active site in the *apo*-PFOMT structure. Furthermore Mg<sup>2+</sup> titration via ITC did not afford significant signals, suggesting the notion, that the metal is only bound along with the co-substrate (Figure 4.8). It has been suggested, that the entropy cost to transfer one water molecule from bulk to the protein-bound state can be up to 7 cal mol<sup>-1</sup> K<sup>-1</sup> [51]. The replacement of ordered waters from the active site or from a hydrated metal ion by a growing aliphatic chain could therefore explain the gain in entropy, and SAH is positioned in a way to warrant exactly that (Figure 4.8). Also, the hydrogen and metal complexing bonds consequently lost could explain the less negative enthalpy. However, this is purely hypothetical since more evident data

is missing. Additional insight might be gained by expanding the ITC experiments to even longer SAM analogues. The limited space in the active site, which forces the growing side chain to expel water and possibly the metal ion might also be the reason for the inactivity of PFOMT towards SAE. If the metal ion is blocked from its complexing moieties, activation of the substrate hydroxyl would be hindered.

Comparison of the novel *apo*-PFOMT and the published structure (pdb: 3C3Y) suggests that the movement (upon ligand binding) along multiple parts of the backbone proximal to the active site pocket is a main contributor to the overall rmsd of 0.9 Å (Figure 4.3).



**Figure 4.8.: Comparison of the active sites of **a** – the solved apo-structure (green) and **b** – the ligand-bound structure (steelblue; pdb: 3C3Y). Waters are represented as small red spheres, calcium as a green sphere (complexing bonds are dashed) and SAH is displayed as a white ball-and-stick model. A possible hydrogen bond network (blue lines) for the ligand-bound state is displayed.**

The N-terminus of PFOMT seems to act as a lid, which is closed in the *apo*-form, but highly flexible and therefore unresolved in the ligand bound form. Furthermore, the native enzyme has been shown to be truncated, starting only at residue 12 and being less catalytically efficient than the full length protein [108, 202]. The work presented here consequently supports the notion that the N-terminus plays an important role on the regulation of the enzymatic activity.

During our studies, transethylation activities could not be observed for any of the prepared PFOMT variants. However, some of the variants showed higher

methylation activities towards caffeic acid and even different regioselectivities ( $3' \rightarrow 4'$ ) than the wild-type.

Given the fact that only residues in the active site and therefore in direct contact with the substrates were prepared, the laid out findings provide novel hints for indirect proximal regions in the PFOMT structure that might be studied using site-directed mutagenesis, gene-shuffling or similar approaches in order to work towards a variant that can in fact employ SAE for transalkylation reactions. Furthermore variation of these regions might provide variants with altered substrate specificities which are of high interest.

## 4.6 Contributions

Benjamin Weigel wrote the manuscript, prepared figures, sub-cloned, produced and crystallized PFOMT, solved the *apo*-structure and conducted the ITC experiments. Dr. Martin Dippe prepared most of the PFOMT variants and ethylation activity tests. Dr. Christoph Partier (group of Prof. Dr. Milton T. Stubbs, MLU Halle-Wittenberg) helped collect X-ray datasets.

## **5 Tandem mass-spectrometry studies of flavonoids**

# Comparative CID and HCD MS/MS studies for the characterization of flavanoid aglycones

**Benjamin Weigel<sup>1,a</sup>, Annegret Laub<sup>1,b</sup>, Jürgen Schmidt<sup>1,c</sup>, Ludger A. Wessjohann<sup>1,d</sup>**

**Contact:** bweigel@ipb-halle.de<sup>a</sup>, alaub@ipb-halle.de<sup>b</sup>, law@ipb-halle.de<sup>d</sup>

**Affiliation:** Leibniz-Institute of Plant Biochemistry, Department of Bioorganic Chemistry<sup>1</sup>

**Keywords:** tandem mass spectrometry, LCMS, flavonoids

## Abstract

Flavonoids are an important class of natural compounds and make up a large part of the world's biomass. Due to their anti-inflammatory and anti-oxidant properties, many health benefits are associated with flavonoids and there is a growing interest to use flavonoids in medicinal and dietary contexts. The availability of methods that provide for a quick and reliable identification of flavonoids from different sources is therefore essential. In this work a range of flavonoids was studied using liquid chromatography coupled mass-spectrometry (LC/MS). Two modes of activation, namely CID and HCD, were evaluated to study fragmentation of flavonoids from their  $[M+H]^+$  molecular ions. It was found, that HCD outperformed CID in the ring-fragmentations of methylated flavonoids. Together, both methods provide complementary information that can be used to distinguish different types of flavonoids.

---

## 5.1 Introduction

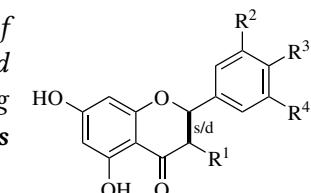
Liquid chromatography-tandem mass spectrometry (LC-MS/MS) has been widely used for the on-line identification of compounds from complex samples, such as crude mixtures from plant or bacterial extracts and is an unexpendable method in the field of metabolomics [56, 123, 127, 165].

Ionization of samples in LC-MS/MS instruments is usually achieved by mild methods operating at atmospheric pressure, such as electrospray ionization (ESI) [208] or atmospheric pressure chemical ionisation (APCI) [79]. However, small molecules rarely produce fragment ions under these conditions and usually only the *m/z* of the molecular ion is observed. A range of different approaches has been used to circumvent this draw-back. The most direct approach is to use electron impact (EI), where the analytes are bombarded with electrons, for ionization. However, EI is a vacuum method and the coupling with liquid chromatography (LC)-systems is not trivial [199]. In order to still generate fragments in liquid chromatography coupled mass-spectrometry (LC/MS) MS/MS methods such as collision induced dissociation (CID) or surface-induced dissociation (SID) were developed [182].

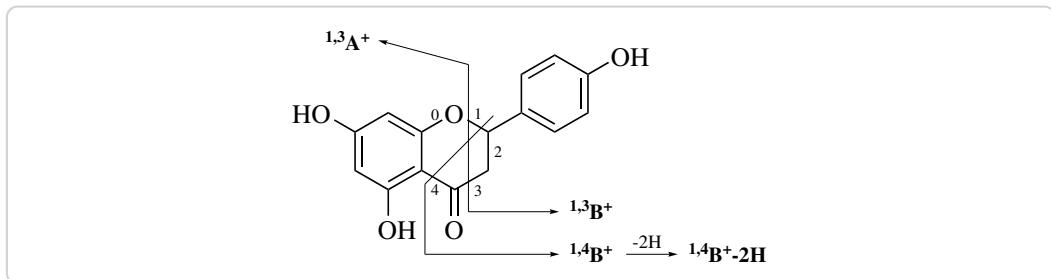
Flavonoids comprise a huge chemical space, with millions of theoretical structures [209]. Due to their biological activities and associated health benefits, applications to quickly identify and characterize these compounds are of special interest. Already, a number of studies have been published that show how MS/MS-approaches using CID can aid in the structural characterization of flavonoids [28, 40, 58, 67, 83, 116, 120, 129, 134, 135]. Researchers have reported that specific patterns of fragmentation along the C-ring can be observed for different classes of flavonoids and can help differentiate between them [40, 129]. However, it was found that the cleavage of the C-ring is less commonly observed for flavonoids methylated at the B-ring, while the loss of small molecules becomes predominant [40, 129].

Fragments of flavonoid aglycones can be represented by a systematic nomenclature first proposed by Ma *et al.* [129]. The labels  $i,jA^+$  and  $i,jB^+$  refer to fragments containing an intact A or B ring, with the superscripts *i* and *j* denoting the bonds of the C-ring that were broken (Figure 5.1). In this work the complementarity of two activation methods, CID and higher-energy collisional dissociation (HCD), for the structural characterization of flavonoids (Table 5.1), especially those methylated at the B-ring, in positive ionization mode was evaluated.

**Table 5.1.**: Substrates studied in this work. Three classes of flavonoids were tested: flavanones (**1-5**), flavones (**6-10**) and flavonols (**11-15**). The topology of the bond in the C-ring specifying flavanones or flavones/flavonols is denoted with **s** (single) or **d** (double), respectively.



	name	[M+H] <sup>+</sup>	s/d	R <sup>1</sup>	R <sup>2</sup>	R <sup>3</sup>	R <sup>4</sup>
<b>1</b>	naringenin	273	s	H	H	OH	H
<b>2</b>	eriodictyol	289	s	H	OH	OH	H
<b>3</b>	ponciretin	287	s	H	H	OCH <sub>3</sub>	H
<b>4</b>	hesperetin	303	s	H	OH	OCH <sub>3</sub>	H
<b>5</b>	homoeriodictyol	303	s	H	OCH <sub>3</sub>	OH	H
<b>6</b>	apigenin	271	d	H	H	OH	H
<b>7</b>	luteolin	287	d	H	OH	OH	H
<b>8</b>	acacetin	285	d	H	H	OCH <sub>3</sub>	H
<b>9</b>	diosmetin	301	d	H	OH	OCH <sub>3</sub>	H
<b>10</b>	chrysoeriol	301	d	H	OCH <sub>3</sub>	OH	H
<b>11</b>	kaempferol	287	d	OH	H	OH	H
<b>12</b>	quercetin	303	d	OH	OH	OH	H
<b>13</b>	myricetin	317	d	OH	OH	OH	OH
<b>14</b>	kaempferide	301	d	OH	H	OCH <sub>3</sub>	H
<b>15</b>	isorhamnetin	317	d	OH	OCH <sub>3</sub>	OH	H



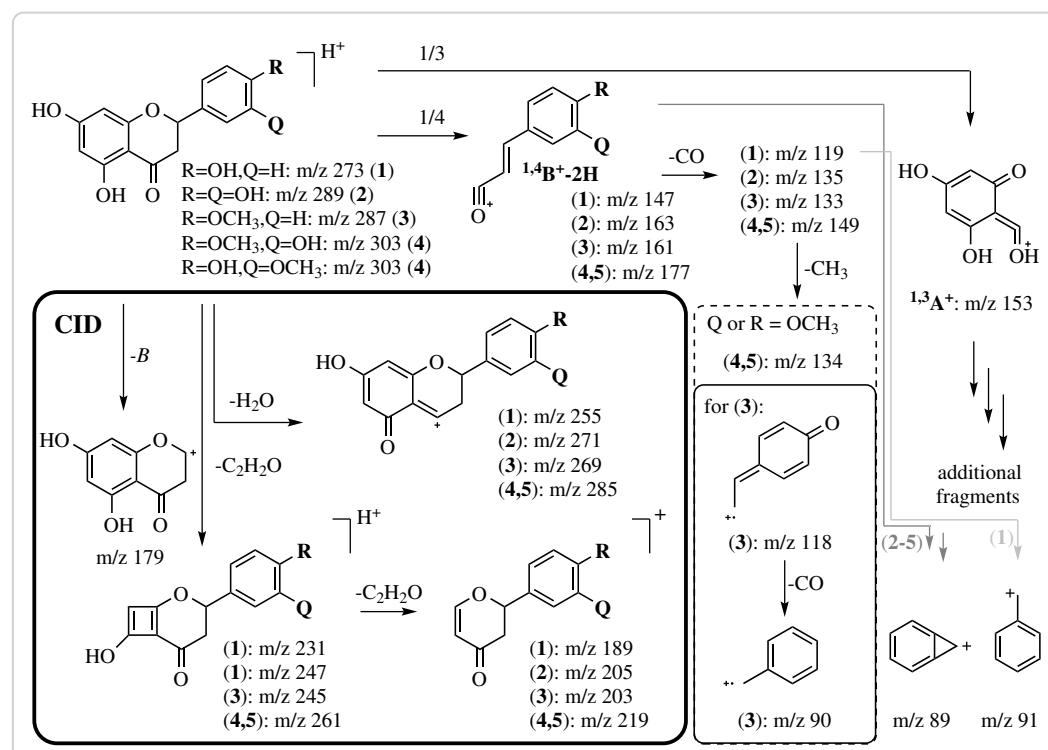
**Figure 5.1:** Ion fragment nomenclature of flavonoid aglycones as proposed by Ma et al., illustrated on naringenin. Ions are labelled according to the ring they contain and the positions of the C ring that were broken. Thus  $^{1,3}A^+$ , contains the ring A and bonds 1 and 3 of the C ring were broken.

## 5.2 Fragmentation of flavanones

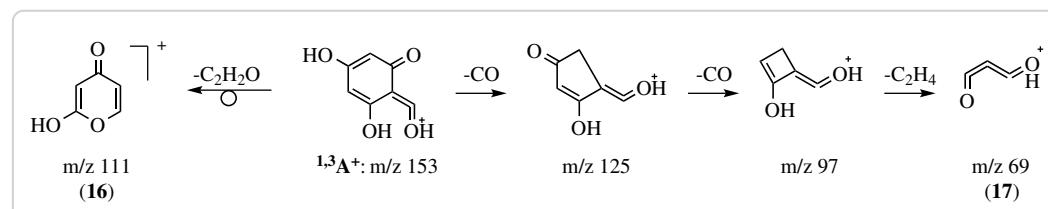
Positive ionization MS<sup>2</sup> spectra of flavanones (Table B.11) are mostly characterized by a base peak at  $m/z$  153, which corresponds to the A-ring of the flavonoid skeleton (Figure 5.2). Even when  $m/z$  153 was not the base peak, it was still dominant in the spectrum with intensities ranging between 20 % and 77 %. Peaks corresponding to the molecular ions  $[M+H]^+$  were not observed for any of the flavanones. The structure of the ion  $^{1,3}A^+$  corresponding to  $m/z$  153 is the same for all compounds (1) to (5) (Figure 5.2). Peaks corresponding to mass-to-charge ratio ( $m/z$ ) values of the respective  $(^{1,4}B^+-2H)$  ions are also present in the mass spectra of each flavanone. Apart from the ions  $^{1,3}A^+$  and  $(^{1,4}B^+-2H)$ , the CID- and HCD-mass spectra of the flavanones differ significantly. CID mainly triggers neutral losses directly from the molecular ion. Losses of water (18 Da) and one or two ketene units ( $C_2H_2O$ , 42 Da) are predominant and afford ions of relatively high masses (Figure 5.2) [95].

Fragment ions from cleavage of the C-ring ( $^{1,3}A^+$  and  $(^{1,4}B^+-2H)$ ) further decomposed under the higher energy conditions in HCD experiments. Thus, the resulting HCD spectra generally display smaller  $m/z$  than the CID spectrum (Figure 5.4). Increasing the normalized collision energy (NCE) from 75 to 100 % in HCD experiments further increased fragmentation. This is made clear by the increasing intensities of smaller fragments upon raising the NCE (Figure 5.4).

Further fragmentation of ion  $(^{1,4}B^+-2H)$  seems to depend on the substituents of the B-ring. Only  $(^{1,4}B^+-2H)$  from eriodictyol (2) loses a water, as suggested by a peak



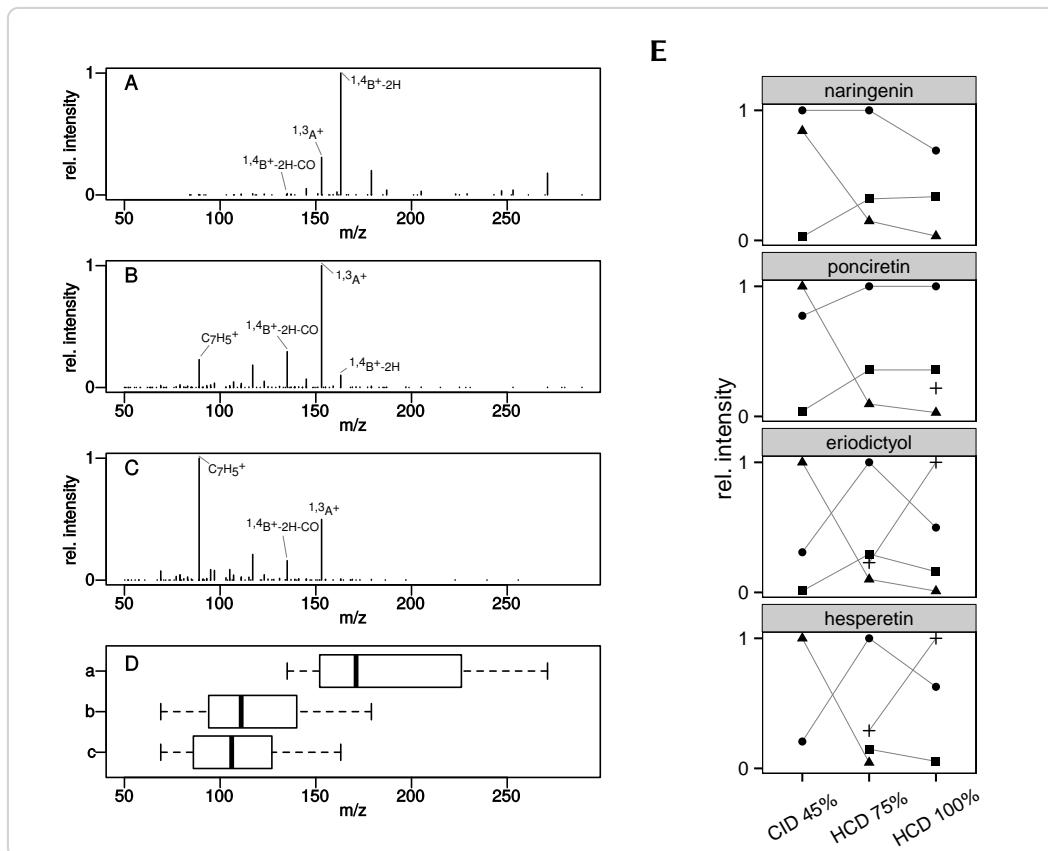
**Figure 5.2:** Major fragmentation pathways of flavanones. Activation using CID conditions at 45 % NCE mainly results in neutral losses of  $\text{H}_2\text{O}$  and ketene ( $\text{C}_2\text{H}_2\text{O}$ ) from the molecular ion  $[\text{M}+\text{H}]^+$  (bold frame). These neutral losses are scarcely observed when HCD with a NCE of 75 % or 100 % is used for activation. Here, C-ring cleavages followed by neutral losses from the cleavage fragments are dominant.



**Figure 5.3:** Proposed  $MS^2$  fragmentation of  $^{1,3}A^+$  after HCD activation. In high energy  $MS^2$  experiments,  $^{1,3}A^+$  might lose two  $\text{CO}$  followed by an unusual  $\text{C}_2\text{H}_4$ . A single loss of ketene ( $\text{C}_2\text{H}_2\text{O}$ ) to afford m/z 111 is also sensible.

at  $m/z$  145. However, the loss of  $\text{CO}$  is the most prominent decomposition of  $(^{1,4}\text{B}^{+}\text{-}2\text{H})$ . The intensities of the peaks corresponding to the  $(^{1,4}\text{B}^{+}\text{-}2\text{H-CO})$  fragment were as high as 36 % in HCD experiments (Figure 5.4). Naringenin (1) seems to

sequentially lose two CO in HCD mode to afford  $m/z$  91 (intensities at 75 and 100 % NCE at 24 and 100 %, respectively). This  $m/z$  is a strong indicator of a benzyl cation or tropyl cation (Figure 5.2). Decay of ( $^{1,4}\text{B}^+ \text{-} 2\text{H}$ ) of the other flavanones likely



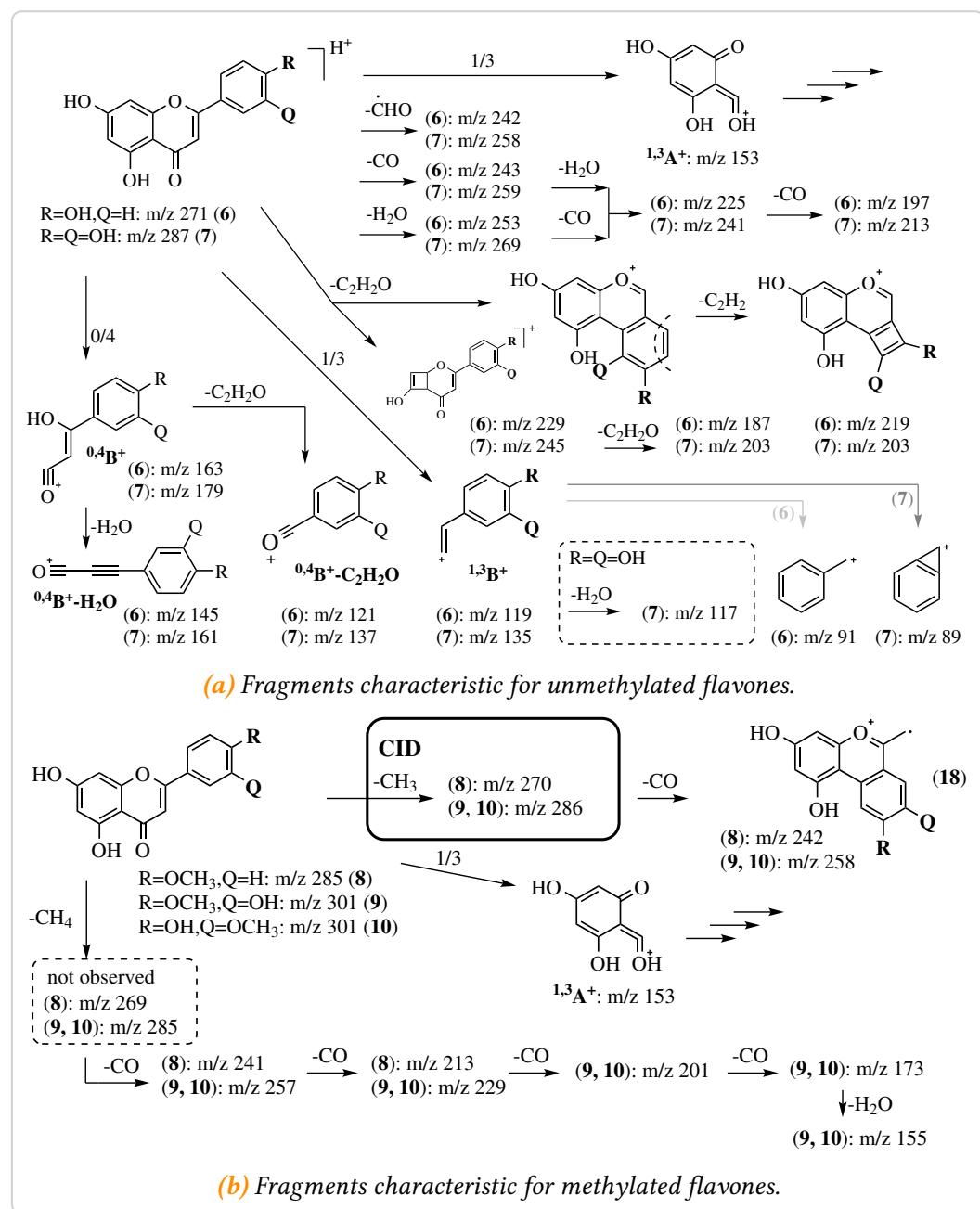
**Figure 5.4.: Comparison of CID and HCD  $\text{MS}^2$  spectra of eriodictyol (2). A – CID at 45 % NCE. B – HCD at 75 % NCE. C – HCD at 100 % NCE. Four different prominent peaks are annotated in each spectrum. D – The shift to smaller masses in HCD spectra and with increasing NCE is illustrated by the boxplot of the distribution of peaks with relative intensities above 1 % in each of the above spectra. E – Relationship between the activation method and the intensity of four fragments (●  $^{1,3}\text{A}^+$ , ▲  $^{1,4}\text{B}^+ \text{-} 2\text{H}$ , ■  $^{1,4}\text{B}^+ \text{-} 2\text{H-CO}$ , +  $\text{C}_7\text{H}_5^+$ ) of different flavanones.**

leads to a stable bicyclo[4.1.0]heptatrienyl cation as the high intensity of peak  $m/z$  89 in HCD mode suggests. Methylated flavanones (3), (4) and (5) show a loss of CO followed by a loss of a methyl radical ( $^{1,4}\text{B}^+ \text{-} 2\text{H-CO-CH}_3\cdot$ ), as suggested by the respective  $m/z$  values of 118 and 134. Another CO loss from this fragment is

possible for ponciretin (**3**) to produce an ion  $m/z$  90, which is at 49 % intensity in the HCD spectrum recorded with NCE of 75 %. The evidence suggests, that this ion's structure is best described by a benzylum/tropylium radical cation (Figure 5.2). It is proposed, that ion  $^{1,3}\text{A}^+$  can decompose via two different pathways under HCD conditions (Figure 5.3). A loss of ketene from  $^{1,3}\text{A}^+$  results in  $m/z$  111. Pyranone (**16**) is suggested as a structure for this ion. Sequential losses of two CO and a  $\text{C}_2\text{H}_4$  could afford ion (**17**). However, further  $\text{MS}^n$  experiments are necessary to confirm these proposals.

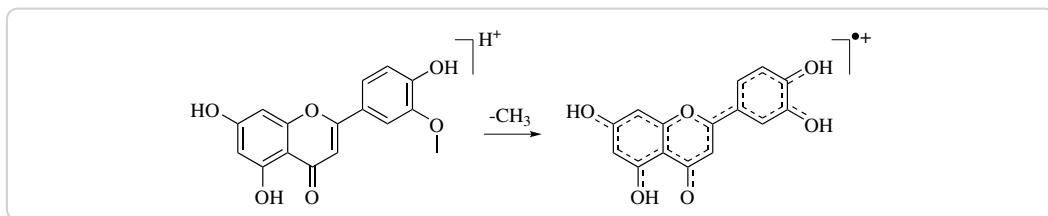
### 5.3 Fragmentation of flavones

The principle fragmentation of flavone aglycones apigenin (**6**), luteolin (**7**), acacetin (**8**) and chrysoeriol (**10**) in positive mode CID tandem mass spectrometry was discussed previously [109, 129]. Non-methylated (**6**, **7**) and methylated flavones (**8 – 10**) show significantly different  $\text{MS}^2$  spectra (Table B.12). Apigenin (**6**) and luteolin (**7**)  $\text{MS}^2$  spectra show a characteristic  $m/z$  153, corresponding to the  $^{1,3}\text{A}^+$  ion, as a base peak in CID mode and at low activation energies in HCD mode (Figure 5.5). Contrary to the flavanones, the  $\text{MS}^2$  of unmethylated flavones show the peak corresponding to the molecular ion  $[\text{M}+\text{H}]^+$ , which is strongest in HCD at NCE of 75 %. Characteristic neutral losses of water, CO and ketene ( $\text{C}_2\text{H}_2\text{O}$ ) were also observed for (**6**) and (**7**) (Figure 5.5, Table B.12). MS-peaks corresponding to a loss of a formyl radical, resulting in  $[\text{M}+\text{H}-\text{CHO}]^{•+}$  were also observed for (**6**) and (**7**). Loss of ketene is proposed to proceed via two different pathways, such that further neutral losses of another ketene, or  $\text{C}_2\text{H}_2$  might be explained (Figure 5.5). Besides the characteristic  $^{1,3}\text{A}^+$  fragment, apigenin (**6**) and luteolin (**7**)  $\text{MS}^2$  spectra also present peaks corresponding to the B-ring fragments  $^{1,3}\text{B}^+$  ( $m/z$  119 and 135) and  $^{0,4}\text{B}^+$  ( $m/z$  163 and 179). From the mass differences of these fragments, the substitution on the B-ring can be deduced. The  $^{0,4}\text{B}^+$  ion might further degrade by neutral losses of ketene (32 Da) or water (18 Da). The base peaks at a NCE of 100 % in HCD,  $m/z$  91 (**6**) and  $m/z$  89 (**7**), are most likely due to a further decomposition of  $^{1,3}\text{B}^+$  in a fashion similar to the flavanones to afford a benzylum or bicycloheptatrienyl cation respectively (Figure 5.5).



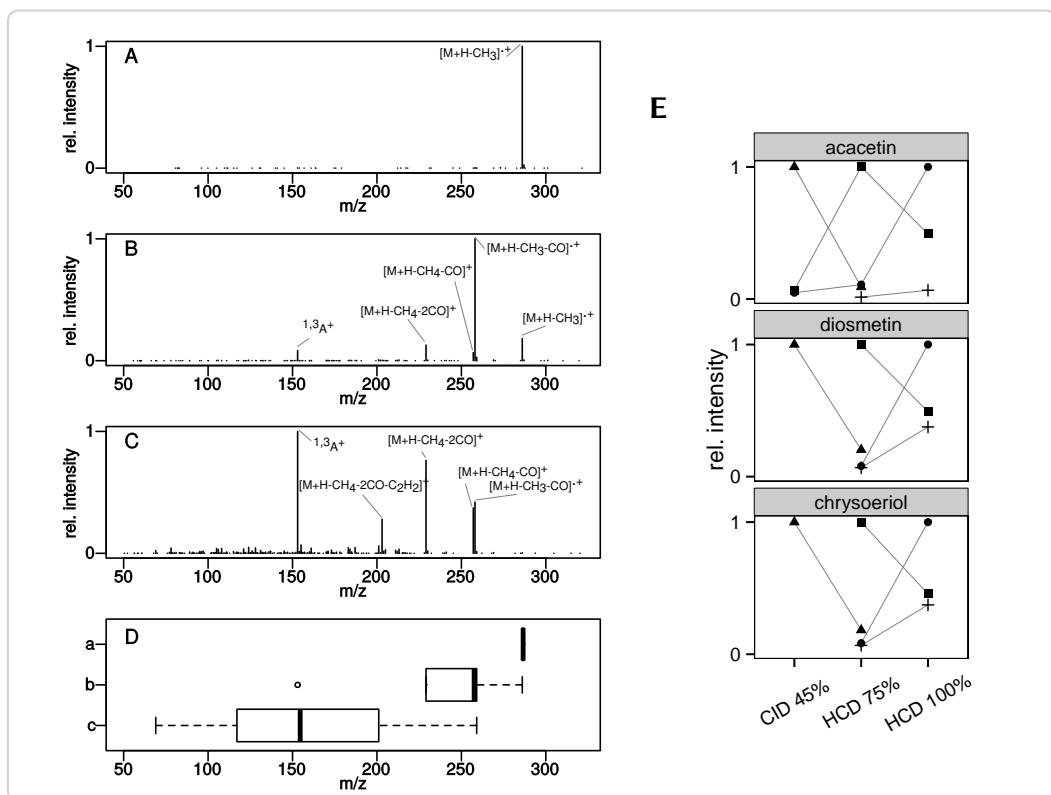
**Figure 5.5.: Major fragmentation pathways of unmethylated and methylated flavones.** Multiple neutral losses of small molecules (e.g. CO, water or ketene) and 0/4 and 1/3 C ring cleavages are predominant in the MS<sup>2</sup> spectra of unmethylated flavones. Methylated flavones loose a methyl group in CID experiments, but only in HCD experiments do other fragmentation reaction become obvious.

The most notable difference between the methylated and unmethylated representatives is the almost complete lack of any fragmentation of the methylated flavones other than a methyl loss, in CID experiments (Table B.12, Figure 5.7). A relatively stable radical cation is formed after the loss of a methyl group, due to the fact that the whole system is essentially conjugated (Figure 5.6). Any other loss

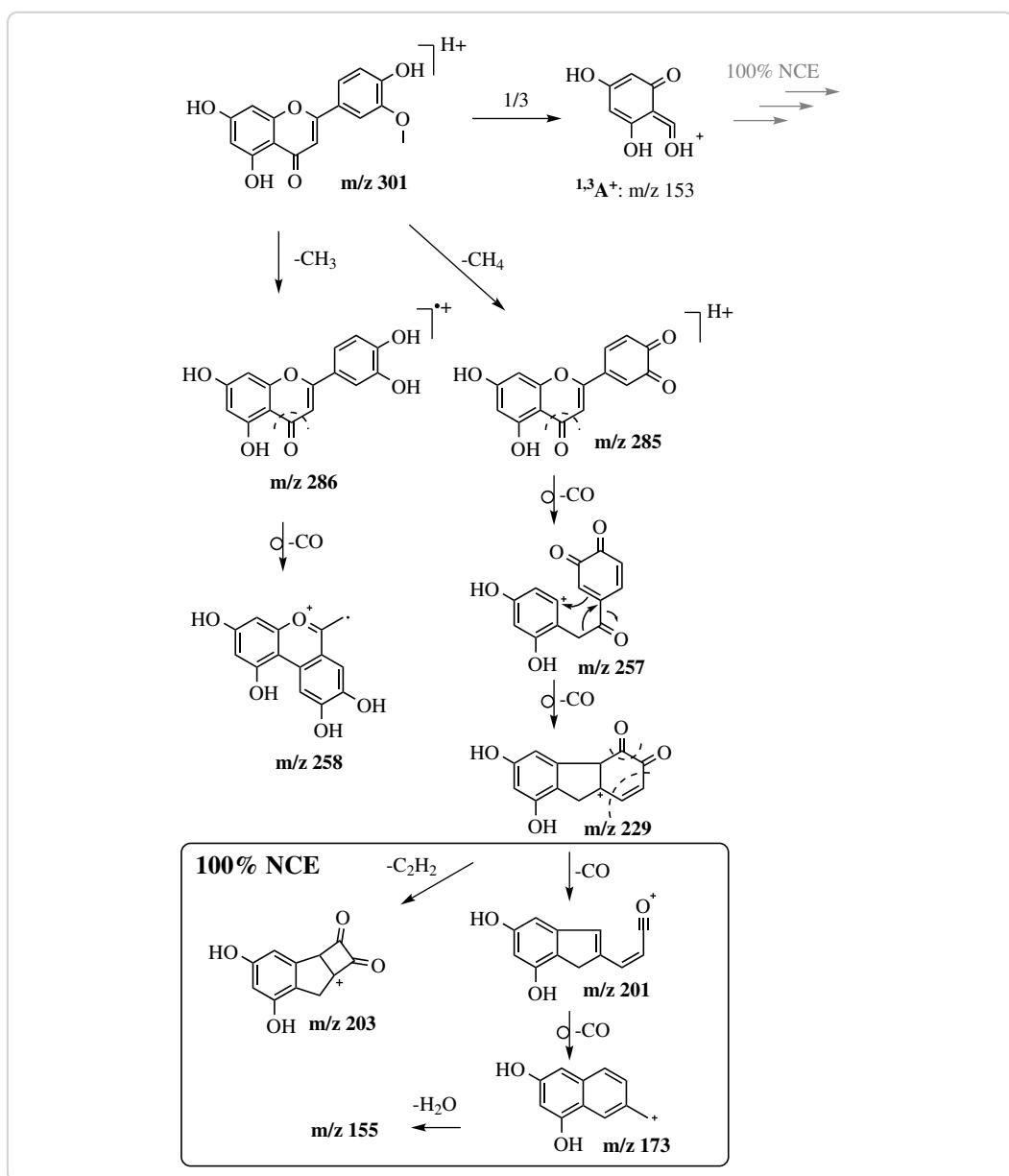


**Figure 5.6.:** Stability of the  $[M+H-CH_3]^{•+}$  ion of flavones. The  $[M+H-CH_3]^{•+}$  ion of methylated flavones like diosmetin is highly stabilized by resonance, explaining the high intensity of the corresponding peak and limiting its fragmentation at low activation energies.

would break this conjugation and therefore requires a higher activation energy. HCD experiments at NCE of (75 to 100) % were suitable to fragment the methylated flavones (**8–10**). The base peak in the HCD spectra of (**8**) ( $m/z$  242) and (**9, 10**) ( $m/z$  257) at 75 % NCE was attributed to another loss of CO from the  $[M+H-CH_3]^{•+}$  ion, while the base peak  $m/z$  153 at 100 % NCE likely corresponds to the  $^{1,3}A^+$  ion (Figure 5.7). Further losses from  $[M+H-CH_3-CO]^{•+}$ , with the proposed structure of a benzochromenylium radical (**20**), were not observed (Figure 5.5, Table B.12). Mass-to-charge ratios of 241 (**8**) and 257 (**9, 10**) were attributed to a neutral loss of methane ( $CH_4$ ), followed by a loss of CO (Figure 5.5, compound 5.8). Interestingly, the abundance of a peak corresponding to a  $[M+H-CH_4]^+$  ion was below 1 % in all spectra, illustrating its susceptibility for additional losses. The fragment  $[M+H-CH_4-CO]^+$  on the other hand might undergo further neutral losses of up to three CO (compounds **10** and **9**) as is illustrated for chrysoeriol in compound 5.8. However, instead of additional CO losses, fragment  $[M+H-CH_3-2CO]^{•+}$  of (**10**) or (**9**) might as well lose a  $C_2H_2$  (compound 5.8), as suggested by the  $MS^2$  spectra (Table B.12). The only C-ring fragmentation of the methylated flavones (**8–10**) occurs at positions 1/3, as the observed  $m/z$  153 ( $^{1,3}A^+$ ) suggests. The higher energy  $MS^2$  spectra suggest, that the  $^{1,3}A^+$  fragment might deteriorate further in the same manner as described for the flavanones (Figure 5.3). Numerous minor peaks in the  $MS^2$  HCD



**Figure 5.7.: Comparison of CID and HCD  $MS^2$  spectra of chrysoeriol (10). A – CID at 45 % NCE. B – HCD at 75 % NCE. C – HCD at 100 % NCE. Four different prominent peaks are annotated in each spectrum. D – The shift to smaller masses in HCD spectra and with increasing NCE is illustrated by the boxplot of the distribution of peaks with relative intensities above 1 % in each of the above spectra. E – Relationship between the activation method and the intensity of four fragments (●  $^{1,3}A^+$ , ▲  $[M+H-CH_3]^{+}$ , ■  $[M+H-CH_3-CO]^{+}$ , +  $[M+H-CH_4-CO]^{+}$ ) of the studied methylated flavones.**



**Figure 5.8:** Proposed pathway of fragmentation of (**10**) after HCD activation. Losses of  $\text{CH}_3$  and  $\text{CH}_4$ , followed by loss of  $\text{CO}$  are the major fragmentations observed in the corresponding MS spectra. However, multiple losses of  $\text{CO}$  only occur after a loss of methane ( $\text{CH}_4$ ), possibly due to the relative stability of the benzochromenylium radical (**19**). At 100 % NCE even higher order fragmentations were observed.

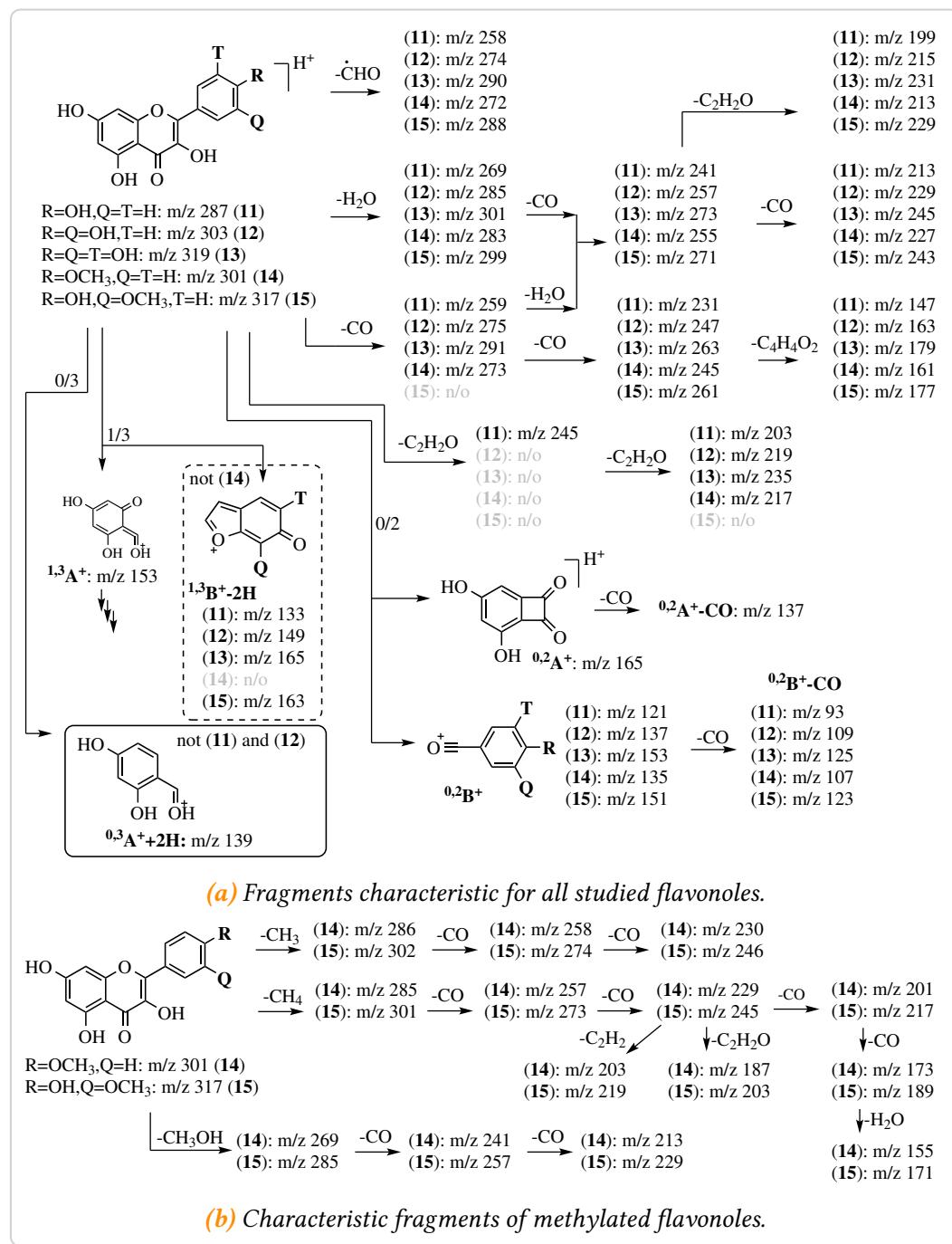
spectra of compounds (**8–10**) could not be assigned a fragment or structure, but many even numbered *m/z* values suggest quite complex rearrangements.

The general trend of smaller sized fragments at higher activation energies is also true for flavones (Figure 5.7).

## 5.4 Fragmentation of flavonols

The principle fragmentation pathways of kaempferol (**11**), quercetin (**12**), myricetin (**13**) and isorhamnetin (**15**) in CID tandem mass spectrometry have been previously reported [129, 134, 216]. Other than flavones, methylated and non-methylated flavonols share similar fragment(ation)s. Whereas in CID methylated flavons hardly showed any fragmentation beyond a methyl loss, methylated flavonoles kaempferide (**14**) and isorhamnetin (**15**) exhibited the same losses as their unmethylated counterparts, albeit at a much lower level (Table B.13, Figure 5.9 and 5.10). These observations are in full agreement with previous reports [129] and hold true in CID as well as HCD measurements. The observed losses from the molecular ion  $[M+H]^+$  are essentially the same as those that were described for the flavones (**6, 7**) (compare Figure 5.9 and 5.5). Lots of high intensity peaks presented in the  $MS^2$  spectra of flavonoles and the base peaks changed between compounds. The base peak of (**11**) in the CID spectra was at *m/z* 165, which corresponds to the  $^{0,2}A^+$  fragment (Figure 5.9). The signals *m/z* 257 and 273 corresponding to the  $[M+H-H_2O]^+$  ions were the base peak in the CID- $MS^2$  spectra of (**12**) and (**13**) respectively. The  $[M+H-CH_3]^{•+}$  ions were highly abundant in the CID experiments of (**14**) and (**15**). The base peak of (**15**) *m/z* 302 corresponds to this fragment. Fragment  $(^{0,3}A^+ + 2H)$  fits the *m/z* 139, which was the base peak in the CID spectrum of (**14**). The MS signal *m/z* 153 corresponding to fragment  $^{1,3}A^+$  was at low abundance in CID spectra, especially for the methylated falvonols (Figure 5.10). However, in HCD experiments *m/z* 153 was the base peak of all flavonols, except kaempferide (**14**) where *m/z* 229 was at 100 % relative intensity.

Neutral losses of CO, water or a formyl radical are suggested by the collected spectra (Figure 5.9, Table B.13). Only for kaempferol (**11**), a neutral loss of 42 Da corresponding ketene was observed. However,  $MS^2$  spectra of all flavonols, except

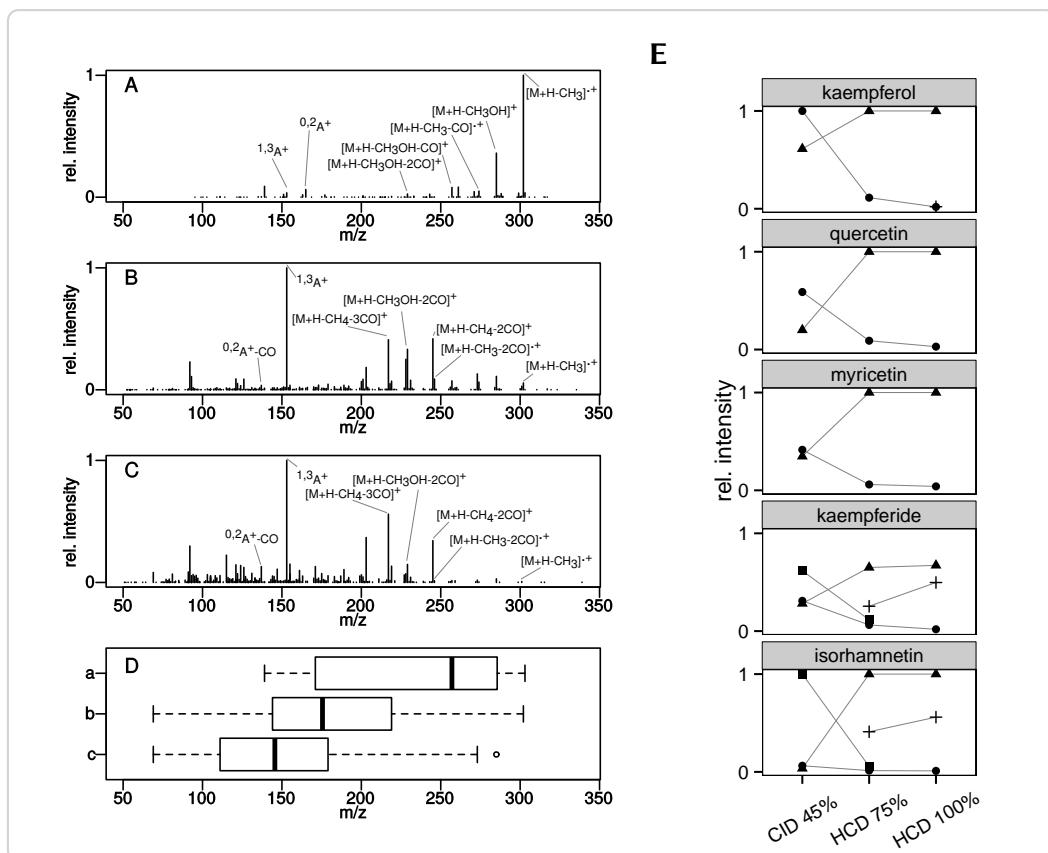


**Figure 5.9.: Major fragmentation pathways of flavonoles.** Unlike flavones, methylated and unmethylated flavonoles share common fragmentations, albeit signals corresponding to small molecule losses are typically small for methylated analogues. Ring fragments observed typically correspond to the cleavage along bonds 0/3 or 0/2. Methylated flavonols shared common fragments with the methylated flavones. However, loss of methanol and a couple CO was also observed. n/o – not observed (relative intensity <1 %).

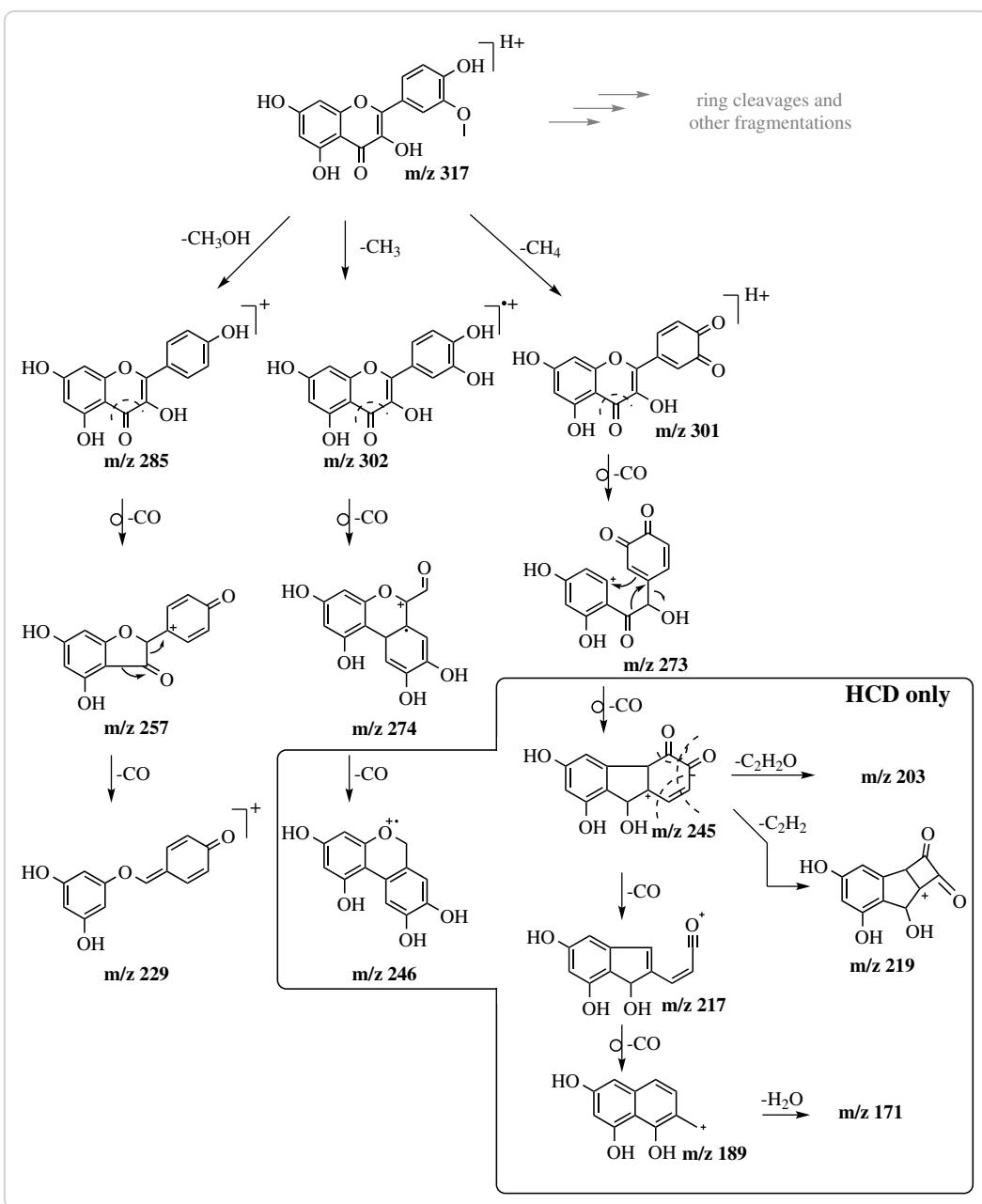
(15), contained signals that could be assigned to the ion  $[M+H-2C_2H_2O]^+$ , suggesting a loss of two ketene units. This advocates the notion that the  $[M+H-C_2H_2O]^+$  ion of flavonols might be highly unstable. Other than the flavones, flavonoles can loose two sequential CO and another  $C_4H_4O_2$ , confirming previously published data [129]. The spectra furthermore suggest, that the  $[M+H-H_2O-CO]^+$  fragment of flavonols can loose another 42 Da ( $C_2H_2O$ ), which was not spotted previously. The data also clearly show, that neutral losses off of the molecular ion are most abundant in CID experiments, whereas the shift to smaller masses in HCD experiments is obvious (Table B.13, Figure 5.10).

The studied flavonoles all displayed an MS signal at  $m/z$  153 corresponding to the  $^{1,3}A^+$  fragment, just as the flavanones and flavones with a 5,7-dihydroxy-substitution of the A-ring did. This further highlights the diagnostic nature of the  $^{1,3}A^+$  fragment of flavonoids in MS/MS spectra. At higher energies,  $^{1,3}A^+$  can further decompose in a manner discussed in the previous sections (Figure 5.3). Characteristic ring cleavage fragments of flavonols include  $^{0,2}A^+$ ,  $^{0,2}B^+$  and  $^{1,3}B^+-2H$  [129, 134], all of which were confirmed in the present study. Overall, the intensity of the  $^{0,2}A^+$  and  $^{1,3}B^+-2H$  fragments decreased in HCD over CID experiments, whereas the intesity of ions  $^{0,2}A^+-CO$ ,  $^{0,2}B^+$  and  $^{1,3}A^+$  increased (Figure 5.10).

Apart from the discussed fragmentations,  $MS^2$  spectra of the methylated flavonols (14) and (15) also showed fragmentations typical of methyl esters, namely methyl, methane and methanol loss. Methyl and methane loss followed by sequential losses of carbon monoxide were already shown for flavones (8–10) and are postulated to proceed in a similar manner in flavonols (14) and (15) (Figure 5.11). Because of the extra hydroxyl at the C-ring, methylated flavonols such as isorhamnetin can loose two CO instead of just one after loss of a methyl radical (compare Figure 5.11 and (compound 5.8)). Other than flavones, spectra of methylated flavonols (14) and (15) also showed signals ( $m/z$  269 and  $m/z$  285) corresponding to a loss of methanol. The data suggests, that these  $[M+H-CH_3OH]^+$  fragments can loose up to two CO, similar to the loss of water and CO (Figure 5.11 and 5.9). The peaks with  $m/z$  301, 273, 245, 217 and 189 in the HCD spectra of isorhamnetin (15), suggest a loss of up to four CO after the initial loss of methane (Figure 5.11). As mentioned before, the smaller mass fragments corresponding to multiple neutral losses are more pronounced at



**Figure 5.10:** Comparison of CID and HCD  $MS^2$  spectra of isorhamnetin (15). **A** – CID at 45 % NCE. **B** – HCD at 75 % NCE. **C** – HCD at 100 % NCE. Four different prominent peaks are annotated in each spectrum. **D** – The shift to smaller masses in HCD spectra and with increasing NCE is illustrated by the boxplot of the distribution of peaks with relative intensities above 1 % in each of the above spectra. **E** – Relationship between the activation method and the intensity of four fragments (●  $0.2A^+$ , ▲  $1.3A^+$ , ■  $[M+H-CH_3]^+$ , +  $[M+H-CH_4-3CO]^+$ ) of the studied methylated flavones.



**Figure 5.11:** Proposed pathways of fragmentation of isorhamnetin (15). Isorhamnetin might lose methyl, methane or methanol upon activation. A similar fragmentation pathway was proposed for the analogous chrysoeriol (compound 5.8). Some fragmentations were observed in HCD mode only (box).

higher activation energies and were thus limited to HCD experiments at a NCE of 100 % (Figure 5.10, Figure 5.11).

## 5.5 Conclusions and Discussion

This comprehensive study shows that, taken together, data from CID and HCD experiments can be complementary to give a much deeper understanding of structural features of flavonoids. Mass errors were calculated for each postulated fragment and ranged from (0.4 to 10) ppm, highlighting the accuracy of the instrument which also allowed for the accurate determination of molecular formulas from MS signals.

The complementary nature of CID and HCD is especially striking, when comparing spectra of (**9**) and (**10**). CID fragmentation of these B-ring methylated flavones afforded MS spectra, where a methyl loss was by far the dominant fragmentation. HCD on the other hand provided higher order fragmentations combined with a higher signal-to-noise ratio, for a deeper insight into structural features. These higher order fragmentations were accelerated by increasing the activation energy, but interpretability of the corresponding spectra was limited. However, with the help of *in silico* methods for the interpretation of MS/MS spectra [14, 215] and the computing power available today, the information contained in highly complex spectra might become more easily accessible. Nonetheless, fine-tuning of the activation energy is an option to optimize fragmentation intensities, especially of the C-ring fragmentations.

Flavones and flavonols share similar patterns of fragmentation and display a loss of a CHO radical, which distinguishes their  $\text{MS}^2$  spectra from those of the flavanones. Distinguishing characteristics between  $\text{MS}^2$  spectra of flavones and flavonols are the C-ring fragmentations, where the  ${}^{0,4}\text{B}^+$  fragment was typically limited to flavones, whereas a (strong)  ${}^{0,2}\text{A}^+$  fragment was only observed for (unmethylated) flavonols. While methylated flavanones did not differ in their fragmentations from their non-methylated analogues, MS spectra of methylated and non-methylated flavones and flavonols showed significant differences. Noticeable loss of  $\text{CH}_3\cdot$  or  $\text{CH}_4$ , followed by losses of CO were typical signs of methylated flavones or flavonols. Loss of methanol was observed in methylated flavonols and

in small amounts at 100 % NCE in flavones, not however in the MS<sup>2</sup> spectra of flavanones. Under the right conditions, all of the studied 5,7-dihydroxy substituted flavonoids presented a <sup>1,3</sup>A<sup>+</sup> ion, with a characteristic *m/z* 153. This information might be of value for studies that want to determine the position of a derivatization of the flavonoid core. To the authors knowledge, a pathway for the decomposition of <sup>1,3</sup>A<sup>+</sup> at high activation energies was proposed for the first time in this work and is universal for all studied compounds. A signal *m/z* 91, stemming from the decay of the <sup>1,4</sup>B<sup>+</sup> or <sup>1,3</sup>B<sup>+</sup> ion, might be a hint for a *para*-monohydroxylated B-ring on flavanones and flavones respectively. Conversely, a peak *m/z* 89 can point in the direction of multiple substitutions on the B-ring.

In summary, the complementary nature of the studied activation methods CID and HCD provides more thorough data for the study of flavonoids. Key ions might only present themselves in the spectra of either method, and together with differences and similarities in the MS/MS spectra, can be used to gain additional insights into the structural characteristics of a studied compound.

## 5.6 Contributions

Benjamin Weigel prepared substances, analyzed mass spectral data and prepared manuscript. Annegret Laub and Jürgen Schmidt conducted LC/MS measurement runs. Through helpful discussions, Jürgen Schmidt helped tremendously with the preparation of the manuscript.

## **6 Enzymatic methylation of Non-catechols**

# Enzymatic methylation of non-catecholic aromatic hydroxyls using class I and class II methyl transferases

Benjamin Weigel<sup>1,a</sup>, Martin Dippe<sup>2,b</sup>, Annegret Laub<sup>1,b</sup>, Ludger A. Wessjohann<sup>1,d</sup>

Contact: bweigel@ipb-halle.de<sup>a</sup>, mdippe@ipb-halle.de<sup>b</sup>, alaub@ipb-halle.de<sup>c</sup>, law@ipb-halle.de<sup>d</sup>

Affiliation: Leibniz-Institute of Plant Biochemistry, Department of Bioorganic Chemistry<sup>1</sup>

Keywords: methyl transferase, SAM, biocatalysis

## Abstract

Phenylpropanoid and flavonoid O-methyl transferase (PFOMT) and soy O-methyl transferase (SOMT-2) are *S*-adenosyl-L-methionine (SAM)-dependent methyl transferases (MTs), belonging to classes I (23–27 kDa, cation-dependent) and II (38–43 kDa, cation-independent) respectively. Methylation of non-catecholic aromatic hydroxyls (phenolic, 3'-hydroxy-4'-methoxy (3O4M), 4'-hydroxy-3'-methoxy (4O3M)) exemplified by different compound classes was achieved by both enzymes, although this has never been described for PFOMT. Active SOMT-2 could not be obtained for *in vitro* experiments, although soluble enzyme was obtained by optimizing refolding conditions using fractional factorial design (FrFD) and design of experiments (DoE). The activity of PFOMT towards non-catechols is increased at high pH. Adjusting the pH to more basic conditions can also partly remedy the negative effect of missing Mg<sup>2+</sup> for class I enzyme PFOMT.

---

## 6.1 Introduction

Non-catechols in nature (biosynthesis, mode of action?), chemical methylation???  
phenol vs methoxyphenols vs catechols

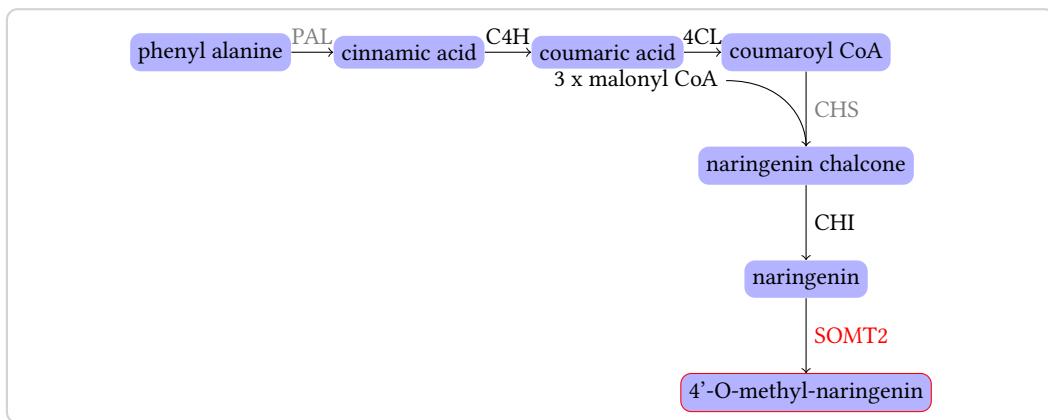
The 4'-hydroxyl of naringenin is non-catecholic in nature. However it is much more acidic than the 4'-hydroxyl of eriodictyol ( $pK_a$  9.8 vs 12.7) and thus about equally as acidic as the 3'-hydroxyl of eriodictyol ( $pK_a$  9.7)<sup>1</sup> [173].

Because SOMT-2 was already well characterized in the literature and acted on flavonoids as well as isoflavonoids, it was selected as a model candidate for enzymes that can methylate 4'-hydroxyls of non-catecholic flavonoids.

## 6.2 SOMT-2

### 6.2.1 *In vivo* biotransformation in *Nicotiana benthamiana*

The group of Sylvestre Marillonet (IPB) established an efficient system to clone and assemble multi enzyme pathways in *Nicotiana benthamiana*, using a modular cloning toolbox, which has already been used to produce flavonoids [106]. All the genes required to establish the pathway up to naringenin in *N. benthamiana* in theory had been cloned previously (Figure 6.1). Only the *SOMT2* gene needed to be cloned into a suitable vector to be transiently expressed in *N. benthamiana*. Infiltrated plants were harvested after 7 days (see section 3.3). The average weight



**Figure 6.1:** Semi-synthetic pathway to naringenin and 4'-O-methyl naringenin in *N. benthamiana*. Enzymes not endogenous to *N. benthamiana* are in gray. PAL - phenylalanine ammonia lyase, C4H - cinnamic acid 4-hydroxylase, 4CL - 4-coumaric acid:CoA ligase, CHS - chalcone synthase, CHI - chalcone isomerase, SOMT2 - soy O-methyl transferase 2

<sup>1</sup> $pK_a$  values were calculated using ChemAxon's MarvinBeans 15.2.16.0

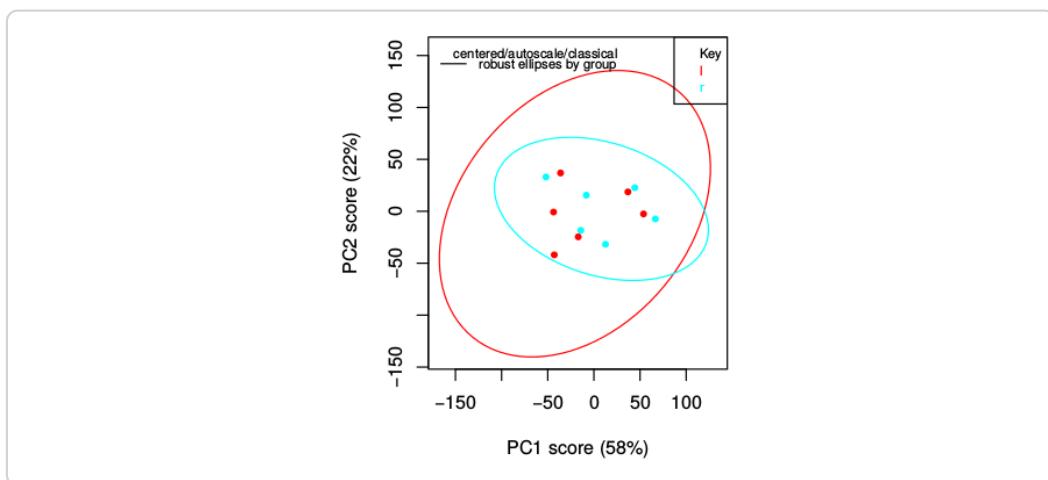
loss after freeze drying was 87.5 %.

The dried material was extracted and analyzed via high-performance liquid chromatography (HPLC) to determine whether ponciretin or the “down-stream” glycosylated products (poncirin, didymine) were produced (Table 6.1). However, through comparison with authentic standards it was apparent, that none of the expected compounds were detected. This finding suggest, that neither naringenin, nor any “down-stream” flavonoids (ponciretin, poncirin, didymin) were present in detectable amounts in the plant tissue at the time of harvest. Although unlikely, it cannot be excluded that higher amounts of the compounds of interest were present at some point in the tissue.

**Table 6.1:** Naringenin and 4'-methylated derivatives that were inquired for in the plant samples via HPLC. The core structure of the compounds is displayed on the left.

	R <sup>1</sup>	R <sup>2</sup>	name
	H	H	naringenin
	CH <sub>3</sub>	H	ponciretin
	CH <sub>3</sub>	rutinose <sup>1</sup>	poncirin
	CH <sub>3</sub>	neohesperidose <sup>2</sup>	didymine

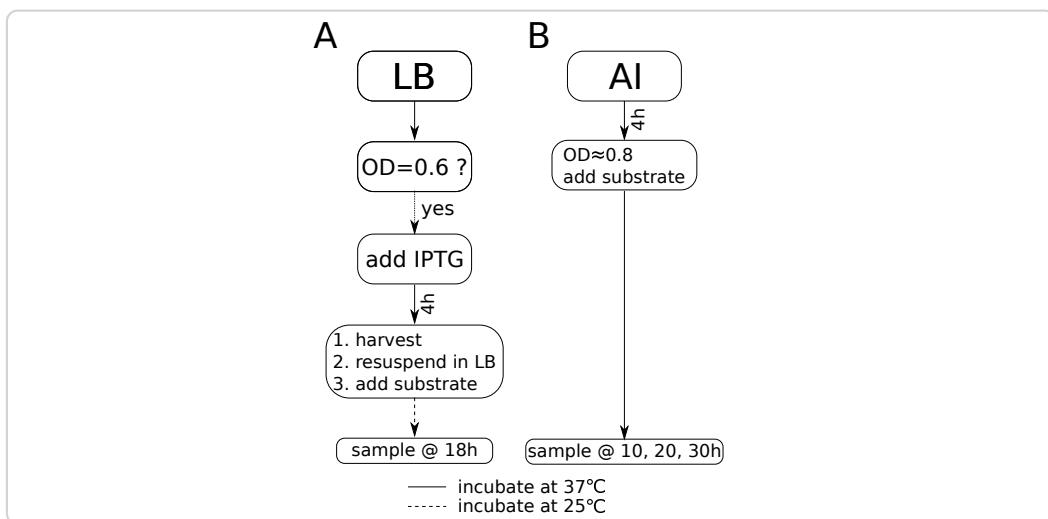
The HPLC chromatograms were analyzed by principal component analysis (PCA) after the data were aligned, centered and scaled, to assess whether the collected plant material samples were different from one another (Figure 6.2 and A.4). The PCA-plot shows that the samples of the different leaf sides do not separate, indicating no difference between infiltration with the *SOMT* gene and vector control between the first two principal components, which account for 80 % of the variance. However, there is a slight separation between top and bottom leaves in the second principal component and between plant 3 and plants (1 and 2) in the first principal component (appendix, Figure A.4). This suggest, that the chemical composition as detected by HPLC is slightly different in the top and bottom leaves, as well as between the different plants.



**Figure 6.2.:** Scatterplot of the first two principal components from the PCA of the HPLC data obtained from leaf material extracts. The samples are colored by leaf side (left/SOMT-2: red, right/vector control: cyan).

### 6.2.2 *In vivo* biotransformation in *E. coli*

Kim *et al.* [98, 101, 102] showed, that SOMT-2 could be used for the biotransformation of different flavonoids in *E. coli* live cultures. The *SOMT-2* gene was cloned into the pET28a(+) and pET41a(+) vectors, to obtain constructs for the production of SOMT-2 without and with a N-terminal Glutathion S-transferase (GST)-tag, respectively since both have been used successfully [98, 101, 102]. However, methylated flavonoids were not detected when biotransformations were prepared according to the methods of the aforementioned authors (Figure 6.3a). Thus, the biotransformation medium was changed to auto-induction (AI)-medium with 0.05 % glucose [185]. Measured growth curves showed, that the glucose present in the medium was depleted after about 5 hours into growth (appendix, Figure A.5). Expression of the *SOMT-2* gene is expected to begin at this time, because the catabolite repression on the *lac* promoter would be relieved and the promoter would be activated by the lactose present in the AI-medium. Thus, 0.1 mM of flavonoid substrate were added at 4 hours to minimize its influence on growth and possible degradation. Although sodium dodecylsulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) samples were prepared throughout the course of the experiment, accumulating



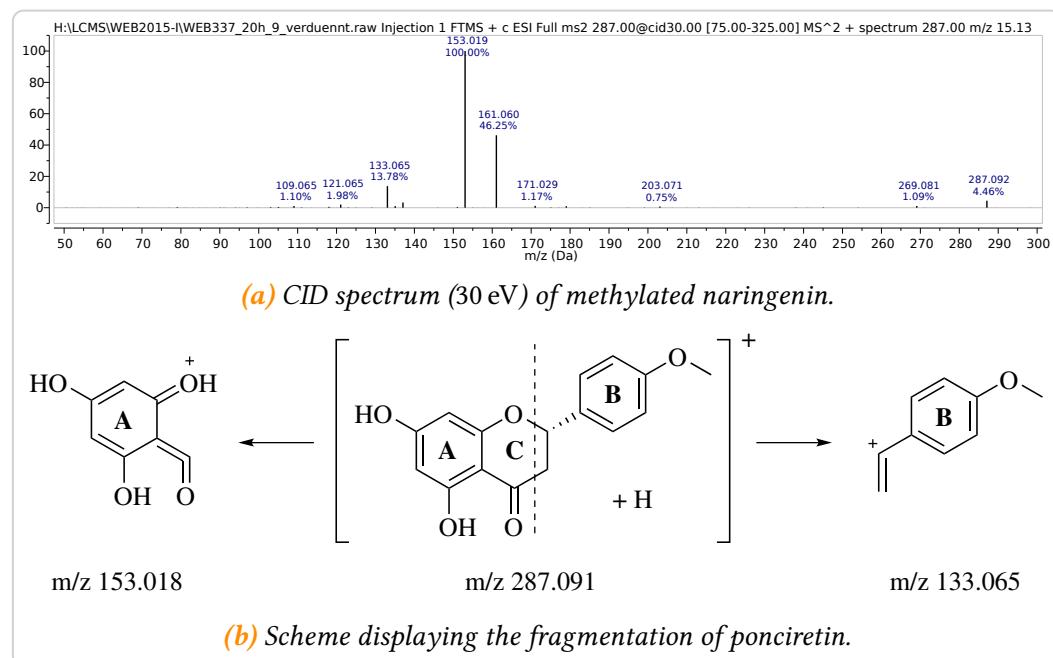
**Figure 6.3.: Biotransformation methods as described by Kim et al. (A) and developed in this work (B). OD – optiocal density at 600 nm, LB – LB-medium, AI – AI-medium.**

SOMT-2 could not be clearly distinguished from endogenous *E. coli* protein in the SDS-PAGE gels (appendix, Figure A.6). Nonetheless, methylation of some of the tested substrates was observed over a course of 30 hours (Table 6.2). Therefore, the sampled medium was extracted using acidified ethyl acetate. Liquid chromatography coupled mass-spectrometry (LC/MS) was employed to determine the site of methylation, since this method is highly sensitive and numerous structural studies on flavonoids using tandem-mass spectrometry experiments have highlighted the feasibility of this approach (see chapter 5) [58, 116]. Collision induced dissociation (CID) was used to obtain structural information about the target molecules, since soft ionization techniques (e.g. electrospray ionization (ESI)) used in LC/MS instruments primarily produce protonated and deprotonated molecular ions, but rarely yield fragments [182]. The CID method collides the precursor ions with a neutral target gas while increasing the energy to induce fragmentation. The produced fragments vary depending on the energy chosen for fragmentation. Flavonoids follow certain different fragmentation pathways [58, 116]. The fragmentation of interest in this work, was the one along the C-ring, which produces two fragments (A- and B-ring) (Figure 6.4b). The mass of the A- and B-ring fragments gives strong evidence for the position (ring) at which methylation occurred. Using the CID technique, an

**Table 6.2.**: In vivo biotransformation of different flavonoids, phenylpropanoids and anthraquinones by SOMT-2 in *E. coli*. Conversion ratios were calculated for samples taken after 30 hours. Multiple substrates containing a 4'-hydroxyl were methylated. Calculation of conversion percentages are only rough estimates, because of the nature of crude medium extracts. Products were determined by liquid chromatography-tandem mass spectrometry (LC-MS/MS). † dihydro cinnamic ketone.

substrate	class	4'-OH conversion product		
alizarin	anthraquinone	✗	✗	-
purpurin	anthraquinone	✗	✗	-
apigenin	flavone	✓	✓( $\geq 54\%$ )	4'-O-methyl apigenin
chrysin	flavone	✗	✗	-
genistein	isoflavone	✓	✓( $< 1\%$ )	Biochanin A
galangin	flavonol	✗	✗	-
kaempferol	flavonol	✓	✓( $\geq 6\%$ )	kaempferide
naringenin	flavanone	✓	✓( $\geq 55\%$ )	ponciretin
eriodictyol	flavanone	✓	✓( $\geq 40\%$ )	hesperetin
homoeriodictyol	flavanone	✓	✓( $> 6\%$ )	3',4'-( <i>O,O</i> )-dimethyl eriodictyol
hesperetin	flavanone	✗	✗	-
phloretin	chalcone	✓	✗	-
resveratrol	stilbene	✓	✓( $\geq 86\%$ )	4'-O-methyl resveratrol
<i>p</i> -coumaric acid	cinnamic acid	✓	✗	-
caffeic acid	cinnamic acid	✓	✗	-
reosmin	cinnamic acid†	✓	✗	-

energy of 30 eV proved sufficient to fragment most flavonoids along the C-ring as is shown here for the methylated naringenin (Figure 6.4). The molecular ion  $[M+H]^+$  of the methylated naringenin has a mass-to-charge ratio (*m/z*) of 287.092. The fragments helping to derive structural information are *m/z* 133 and *m/z* 153, which can only be explained if the B-ring was methylated (Figure 6.4b). If the A-ring was methylated, the expected fragment ions of A and B-ring would have *m/z*-values of 167 and 119 respectively. The LC/MS results suggest, that methylation occurred exclusively at the 4'-hydroxyl, as there was no conversion detected, when the 4'-hydroxyl was absent (Table 6.2). A free 4'-hydroxyl seems therefore necessary for a substance to be a substrate for SOMT-2, which confirms the previous results [98, 101]. Conversion was observed for 4'-hydroxylated (iso)flavonoids and the stilbene resveratrol, however conversion rates of the isoflavone genistein were very low. No



**Figure 6.4.: The masses resulting from the fragmentation into A- and B-ring along the C-ring (dashed line, b) are evidence, that the 4'-hydroxyl on the B-ring is methylated by SOMT-2**

conversion of anthraquinones, cinnamic acid derivates or chalcones was detected, which is also in accordance with previously published data [98, 101]. SOMT-2 acts on phenolic, catecholic as well as (4-hydroxy-3-methoxy-phenyl)-moeities, as is suggested by the assay results that showed methylation of naringenin, eriodictyol and homoeriodictyol respectively. The methylation of (4-hydroxy-3-methoxy-phenyl)-moieties and of stilbenes are properties of SOMT-2 that have not been described before.

The conversion ratios were assessed, but are beset with large errors due to the nature of *E. coli* rich medium extracts. The highest conversions were observed for flavanones and flavones (up to  $\geq 55\%$ ). The tested isoflavones and flavonols showed much lesser conversion ratios (less than 10%). The conversion ratios of apigenin ( $\geq 54\%$ ) and naringenin ( $\geq 55\%$ ) are comparable to the ones published previously [98, 101]. However, genistein only showed minute conversions, which is in disagreement with the data previously published [98, 101]. Conversion of

eriodictyol, homoeriodictyol and kaempferol were not reported before.

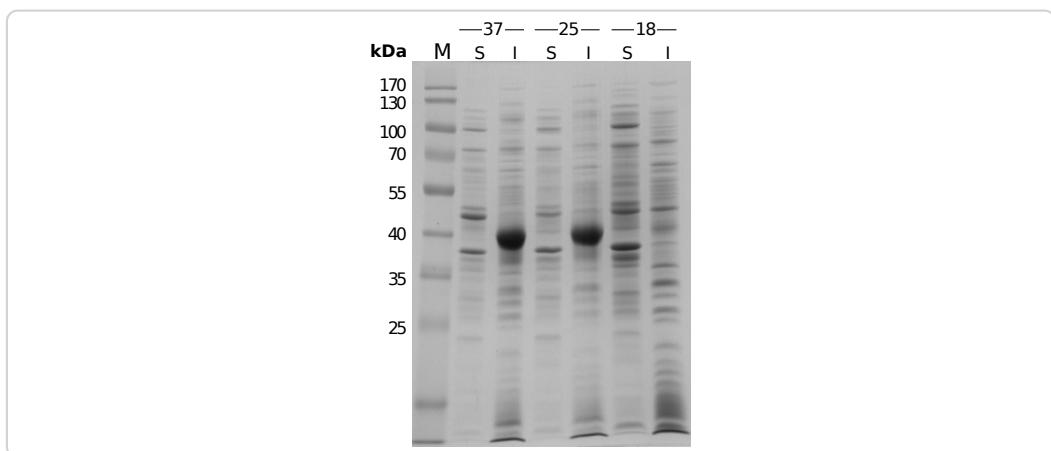
The biotransformation of resveratrol to 3,5-dihydroxy-4'-methoxy-stilbene showed a conversion ratio of  $\geq 86\%$  in 30 hours. This is roughly double the conversion which was recently reported for *in vivo* biotransformations using the specific resveratrol *O*-methyl transferase (*O*-MT) sbCOM1, which only achieved 42 % conversion in 36 hours [96].

### 6.2.3 *In vitro* studies using recombinantly produced SOMT-2

*In vivo* biotransformations are an important tool for the primary characterization of enzymes. However, because live organisms are used and lots of variables are unknown, these systems can cause large errors and are not fit to thoroughly characterize an enzyme. Initially, SOMT-2 was to be purified to homogeneity to be later thoroughly characterized *in vitro*, since the recombinant production of SOMT-2 in *E. coli* as a fusion protein with an N-terminal T7-tag was previously shown. However, the recombinant enzyme had not been characterized [98].

#### Protein production test

Initial protein production tests were carried out using *SOMT-2* cloned into pET28a(+) with an N-terminal His<sub>6</sub>-tag. However, SOMT-2 was not produced in soluble form (Figure 6.5). Numerous systems were tested for the expression of *SOMT-2*. *E. coli* strains used for the trials included BL21(DE3), Rosetta(DE3), Origami(DE3), C41(DE3), C43(DE3), C41(DE3) pLys, C43(DE3) pLys and DH5 $\alpha$ . The *SOMT-2* gene was cloned into multiple other vectors, including pET20b for periplasmic protein production, pET32 for expression with an Trx-tag and vectors that carry promoters for induction by rhamnose. Multiple media, including terrific broth (TB), lysogeny broth (LB) and autoinduction media were used along with different inducers (e.g. lactose, rhamnose, isopropyl-D-thiogalactopyranosid (IPTG)) at different temperatures. Nonetheless SOMT-2 could not be produced in a soluble form and expression only resulted in inclusion bodies (IBs).



**Figure 6.5.:** SDS-PAGE of *pET28a(+)* SOMT-2 expressed in *E. coli* BL21(DE3) in autoinduction medium at different temperatures (shown above). The insoluble fractions show a protein band the same height as the 40 kDa marker band, which corresponds to the SOMT-2 protein (40 425 Da). M – protein size marker, S – soluble fraction, I – insoluble fraction

### In vitro protein refolding

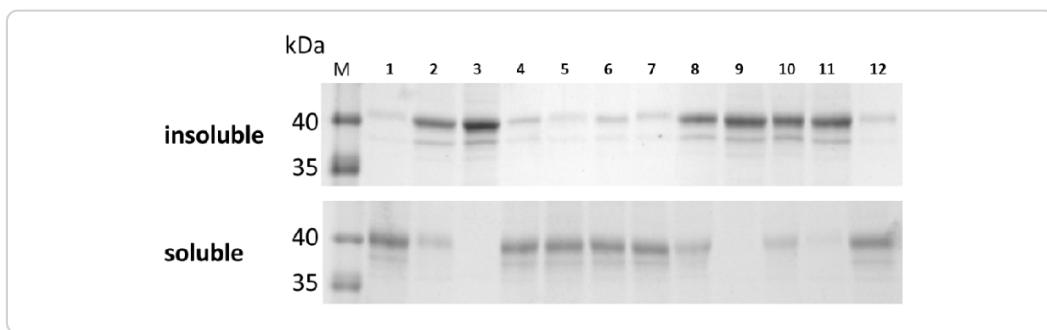
Since the SOMT-2 protein could not be obtained in soluble formen, when recombinantly expressed in *E. coli*, the IBs were prepared [47] and used for *in vitro* refolding studies. For *in vitro* protein refolding, IBs are solubilized using denaturants such as GdmCl or urea. The native tertiary structure of the protein is then restored by removal of the denaturant under the “right” conditions (e.g. pH, salt, additives, etc.). However, this is no trivial task since initially the “right” conditions have to be found by trial and error. The refolding process competes with misfolding and aggregation processes and refolding buffers have to be optimized in order to obtain an efficient refolding system with the best possible results [47, 210, 221]. Refolding efficiency is best measured via biological activity, but even with adequate assays refolding studies are a time-consuming process. The number of experiments required to even test only four variables, for example pH, salt, temperature and protein concentration with 3 states each (e.g. low, medium, high) in all possible combinations results in  $3^4 = 81$ . An experimental setup, which accounts for all possible variable (factor) combinations is also called a *full factorial design*. These setups capture main effects, as well as higher level interaction effects [20, 145].

However, for screening purposes only a fraction of the experiments can be run. The objective of these fractional factorial design (FrFD) experiments is to identify the variables, which have large effects and are worth expanding the experimental investigation upon. FrFDs have been successfully used for a number of protein refolding trials [192, 201, 210].

The following factors were studied for the *in vitro* refolding of SOMT-2: pH, arginine addition, glycerol addition, addition of divalent cations, ionic strength, redox system, cyclodextrin addition and effector (*S*-adenosyl-L-homocysteine (SAH)) addition. These factors were used, because all have been shown to influence refolding success [4, 6, 11, 29, 65, 196, 201, 210, 221]. Two factor levels were used in a twelve-run design. This is sufficient to find some main effects, however no statement about interaction effects can be made. For a complete listing of the buffers and conditions the reader is referred to the materials and methods chapter (Tables 3.7 and 3.8).

Big differences between soluble and insoluble fractions of different refolding buffers could already be noticed from the SDS-PAGE gels (Figure 6.6, see subsection 3.4.11). Refolding buffers 2,3 and 8–11 mainly produced insoluble protein, whereas the majority of the protein in refolding buffers 1, 4–7 and 12 was in soluble form in these buffers after an overnight refolding reaction (Tables 3.7 and 3.8, Figure 6.6). After rebuffering the cleared refolding reactions into a unified buffer the protein concentrations were estimated by BRADFORD-assays [21]. The protein concentrations measured by BRADFORD’s method were consistent with the observations from the SDS-PAGE gels (compare Figure 6.7a and 6.6). Soluble protein was obtained for buffers 1, 4–7 and 12. The highest amount of soluble protein was present, when the refolding reaction took place in buffers 5 or 7. The common denominator of those buffers is that all of them contained arginine, whose addition has proven beneficial for many refolding applications [29, 65, 196].

**Main effects plots (ME-plots)** illustrate the difference between level means for each factor. The ME-plot for protein concentration shows, that arginine seems to be an important factor for the refolding of SOMT-2 (Figure 6.7b). Furthermore, the addition of SAH or glycerin and the pH seem to have an influence, whereas the other factors seem to play only a minor role to achieve high concentrations of



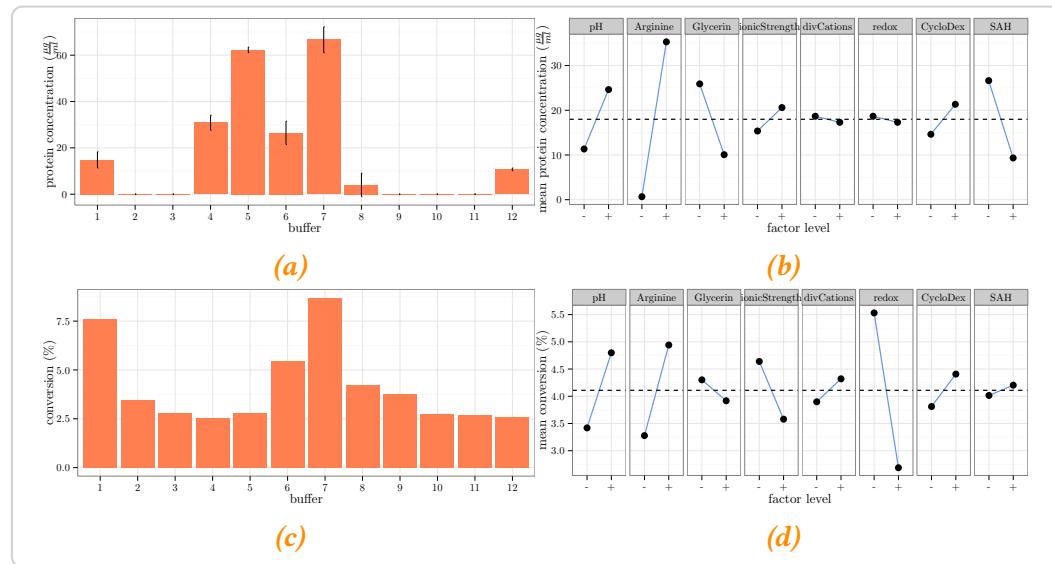
**Figure 6.6.:** SDS-PAGE of the insoluble and soluble fractions of the refolding reactions. Refolding reactions 2,3,8-11 seem to mainly produce misfolded insoluble protein, while the other refolding buffers (1,4-7,12) produce soluble protein.

soluble protein after refolding. However, the Analysis of Variance (ANOVA) test ( $p$ -value 0.0158 for arginine addition) suggests that only arginine addition has a significant influence on refolding, when the significance level is set to 5 % (appendix, Table A.1). The other  $p$ -values are all higher than 0.05, which suggests they have no influence on the yield of soluble protein. Only the  $p$ -value for SAH addition (0.0897) would show significance, if the significance level was raised to 10 %.

Activity tests were conducted with the refolded protein samples to check for naringenin conversion. The conversion of naringenin to ponciretin as follows:

$$\text{conversion} = \frac{\text{AUC}_{\text{ponciretin}}}{\text{AUC}_{\text{naringenin}} + \text{AUC}_{\text{ponciretin}}}$$

The protein activity in the refolded samples was generally very low, as suggested by the low conversions after an overnight activity assay. The maximum conversion of about 8.7 % was observed for the refolded protein sample 7. The activity of the samples did not correlate well with the amount of soluble protein. This becomes clear from the samples 4 and 5, where the amount of soluble protein was high but the observed activity was at a baseline level. The ME-plot also suggests that the main effects have changed. Most notably the redox state of the refolding reaction seemed to have a big influence on the protein activity. Indeed, the ANOVA test suggests that using reducing refolding conditions (DTT) over a redox-shuffling system (GSH:GSSG, oxidizing) has a significant influence on methylation activity



**Figure 6.7.:** Results of *in vitro* protein refolding trials. Measured data (left) is presented alongside the main effects plots (ME-plots) (right). **a** – Protein concentration after refolding and rebuffering into a universal buffer. The highest yield of soluble protein was achieved in buffers 5 and 7. The ME-plots (**b**) illustrate the connection between a factor and the measured protein concentration, illustrating that high pH and arginine concentration might be beneficial in the refolding reactions. **c** – Calculated conversion of naringenin to ponciretin by the refolded protein fractions. Protein refolded in buffers 1 and 7 seem to afford the most active protein by conversion (~volume activity). The ME-plots for the conversion (**d**) show that the redox state (reducing) of the refolding environment is important to achieve active protein.

(*p*-value 0.0218; appendix, Table A.2). However, there is the possibility for SOMT-2 to form intramolecular disulfide bridges, as the modelled structure suggests (Figure A.7). There are also reports, which showed that intermolecular disulfide bridges can contribute to the stability of, mainly archeal, MTs and have no influence on the enzymatic activity [68, 75, 89]. Nevertheless most MTs are only active under reducing conditions and literature suggests, that sometimes assays of MTs are explicitly conducted under reducing conditions [85, 231].

The *p*-value for arginine addition is 0.0824, which also suggests a significant contribution when using a significance level of 10 %. This is plausible, since there cannot be any activity when no soluble protein is present. Judging from the ANOVA test, the remaining factors have no significant impact on the protein activity after

refolding.

### Scaleup of *in vitro* refolding reactions

Buffer 7 (50 mM borate/NaOH, 0.5 M arginine, 2 mM CaCl<sub>2</sub>, 2 mM MgCl<sub>2</sub>, 10 mM NaCl, 0.5 mM KCl, 30 mM  $\alpha$ -cyclodextrin, 5 mM DTT pH 8.5) was used to scale up the refolding reaction from a total volume of 1.05 ml to a volume of 50 ml. After refolding and concentration the protein solution was analyzed by gel filtration, circular dichroism (CD) spectrometry and activity tests were conducted.

The retention time for SOMT-2 during the gel filtration run was 14.5 min (Figure A.8). From the gel filtration standard measurements, this corresponds to a molecular weight of approximatly 57 853 Da, which is roughly the weight of one SOMT-2 monomer (40 kDa). The CD spectrum (Figure A.9) suggest, that the refolded SOMT-2 possesses a secondary structure and is not present as an unfolded random coil. The secondary structure was estimated from the measured CD-spectrum by the K2D3 web service [125]. According to the calculations, the secondary structure elements consist of 12.39 %  $\alpha$ -helix and 32.51 %  $\beta$ -sheet. However, the calculated protein model (A.7) suggests the helix content is much higher (52.3 %), whereas the sheet content is accordingly lower (15.4 %). This finding suggests, that the refolded protein is not in a native state. Also, the enzyme showed no activity for naringenin methylation, which was further evidence that the scaled-up refolding was unsuccessful. Even over the course of many trials a successful large scale refolding of SOMT-2 yielding active protein could not be achieved.

These results display that DoE combined with FrFD can be a valuable tool for the identification of main factors during protein refolding. However, there still exists a discrepancy between small scale refolding reactions and the process of upscaling, which might not be trivial.

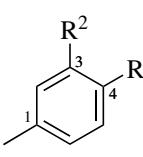
## 6.3 PFOMT

### 6.3.1 Phenolic hydroxyls

Phenolic hydroxyl groups have  $pK_a$ -values of around 10 as demonstrated by four *p*-cresole derivatives (Table 6.3). Catecholic systems have two  $pK_a$ s. The 3-hydroxyl in this example has a much smaller  $pK_a$  than the 4-hydroxyl. This is in part due to the mesomeric (+M) and inductive (+I/-I) properties the substituents display. The M and I-effects let the 3-OH be deprotonated first, which in turn really lowers the acidity and thus increases the  $pK_a$  of the 4-OH. (4-hydroxy-3-methoxy)- and (3-hydroxy-4-methoxy)-derivatives again have a similar  $pK_a$ , with the *meta*-position slightly more acidic due to the +I-effect of the methyl substituent. The nucleophilicity of these phenolic groups happens to coincide with their BRØNSTED acidity. Chemically speaking the hydroxyl with the lower  $pK_a$  always reacts first with an electrophile.

However, different enzymes are able to regioselectively methylate the 3- or the 4-OH of such catecholic systems. Enzyme's active sites create a "microclimate", which can selectively raise or lower the effective  $pK_a$  of functional groups and allows for the efficient manipulation of the macroscopically observed regioselectivity.

**Table 6.3.:**  $pK_a$ -values of phenolic hydroxyl groups exemplified by *p*-cresole derivatives. Substituent positions on the aromatic ring are arbitrary and do not reflect conventions of the International Union of Pure and Applied Chemistry (IUPAC).

	R <sup>1</sup>	R <sup>2</sup>	$pK_a^{-R^1}$	$pK_a^{-R^2}$
	OH	H	10.36	—
	OH	OH	13.1	9.55
	OH	O-Me	10.34	—
	O-Me	OH	—	10.08

Previous studies have established that PFOMT is a 3'-*O*-methyl transferase, which is not able to methylate substrates that bear either phenolic (e.g. naringenin), (3'-hydroxy-4-methoxy)- (e.g. hesperetin) or (4'-hydroxy-3-methoxy)-moieties (e.g. homo-eriodictyol) [85]. The reactions were all run under the same "standard" conditions. However, the reaction buffer/medium can have a tremendous impact on enzymes and their reactions. Therefore reaction conditions require optimization, just as the enzymes themselves, to augment an enzymatic process [17, 104].

Using PFOMT the reaction conditions were screened, to assess if any would promote the methylation of non-catecholic substrates. Changes in the pH of the medium effect enzymatic activity, especially if charged groups are part of the catalytic mechanism, which was proposed for PFOMT [22]. Furthermore, PFOMT is a magnesium dependent enzyme and the activity is effected by altering the concentration of  $Mg^{2+}$ . Thus, the pH was chosen to be varied along with  $Mg^{2+}$  concentration in order to study the influence of those two factors on the methylation reaction.

### 6.3.2 PFOMT pH-profiles are influenced by $Mg^{2+}$

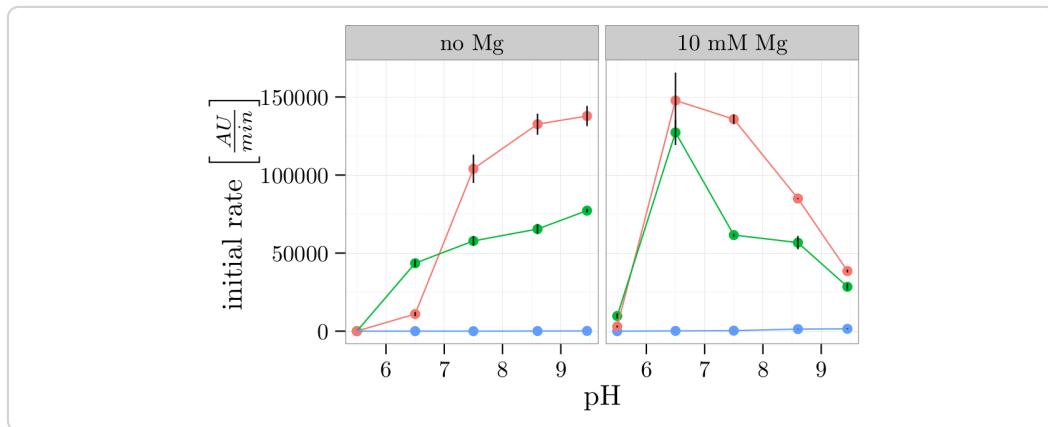
PFOMT was dialyzed extensively against buffer containing 5 mM ethylenediaminetetraacetic acid (EDTA) to obtain enzyme that was virtually free of bound divalent cations. pH-profiles (pH 5.5 – 9.5) of three different substrates (caffeic acid, *iso*-ferulic acid, eriodictyol) were obtained. The pH-profiles were measured without and with the addition of 10 mM  $MgCl_2$ .

The catecholic substrates caffeic acid and eriodictyol were converted by PFOMT much more quickly than *iso*-ferulic acid, which is a (3-hydroxy-4-methoxy)-substituted cinnamic acid (Figure 6.8). The highest rate of *iso*-ferulic acid conversion was two orders of magnitude lower than the highest rate for conversion of the other two substrates. Nonetheless conversion was observed for *iso*-ferulic acid with increasing pH and even an influence of magnesium was observed. Addition of 10 mM  $Mg^{2+}$  increased the rate of *iso*-ferulic acid conversion by 3-fold from 7 pkat/mg to 21 pkat/mg at pH 9.45 (Table 6.4).

The specific activities observed for the conversion of caffeic acid are comparable to published data [49]. For the two catecholic substrates, the pH-optimum shifted from neutral to alkaline pHs, when  $Mg^{2+}$  was omitted. However, the maximum activity remained roughly the same. This was not true for *iso*-ferulic acid, which was hardly converted at all without the addition of  $Mg^{2+}$ .

Linear regression models of the form

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p + \epsilon \quad (6.1)$$



**Figure 6.8.:** Initial rate/pH-profiles for the conversion of three different substrates (red – caffeic acid, green – eriodictyol, blue – iso-ferulic acid) by PFOMT. The non-catecholic substrate iso-ferulic acid is converted much less quickly than the catecholic substrates.

**Table 6.4.:** Maximum specific activity for the conversion of three different substrates with and without addition of magnesium. The pH at which the maximal activity was reached is indicated.

substrate	Mg <sup>2+</sup>	pH	$A_{sp} \left[ \frac{U}{mg} \right]$	$A_{sp} \left[ \frac{pkat}{mg} \right]$
caffeic acid	FALSE	9.45	88	1466
caffeic acid	TRUE	6.50	94	1572
eriodictyol	FALSE	9.45	49	824
eriodictyol	TRUE	6.50	81	1354
iso-ferulic acid	FALSE	9.45	0.4	7
iso-ferulic acid	TRUE	9.45	1.2	21

were fitted to the data in order to make statistically sound inferences regarding a relationship between activity and pH/Mg<sup>2+</sup> addition. However, to simplify this task two subsets were prepared first. The subsets split the data into substrates with catecholic and substrates without catecholic motifs. The *iso*-ferulic acid data was fit to the model

$$\text{activity} = \beta_0 + \beta_1 \times \text{Mg} + \beta_2 \times \text{pH} + \beta_3 \times (\text{Mg} \times \text{pH}), \quad (6.2)$$

which explains about 93.5 % of the variance (Table 6.5). The small p-values of the coefficients for Mg and pH×Mg suggest that there is a strong interaction between Mg<sup>2+</sup> addition and pH. However, the pH as a main effect has little significance.

The pH profile for the catecholic substrates shows that there might be a quadratic relationship between pH and activity. This is true for most enzymatic reactions, where ionizable groups are involved the reaction. A quadratic term was thus included into the linear model to capture this relationship:

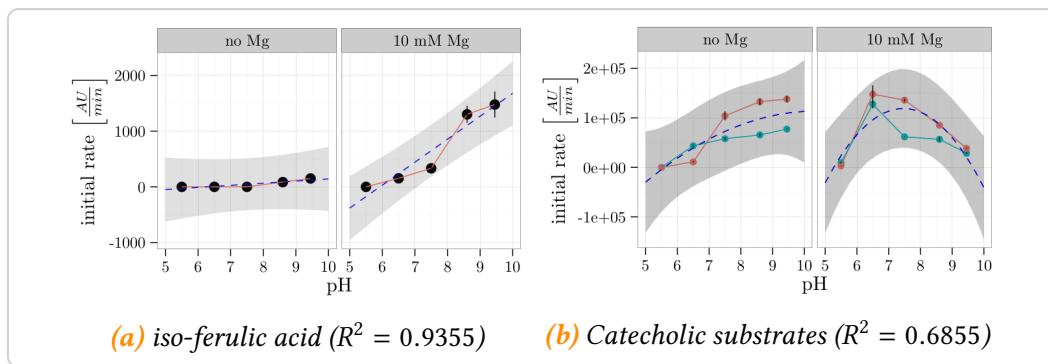
$$\text{activity} = \beta_0 + \beta_1 \times \text{Mg} + \beta_2 \times \text{pH} + \beta_3 \times (\text{Mg} \times \text{pH}) + \beta_4 \times \text{pH}^2 + \beta_5 \times (\text{pH}^2 \times \text{Mg}), \quad (6.3)$$

This model describes the actual data reasonable well, with about 68.6 % of the variance explained (Figure 6.9). Also here the p-values suggest an interaction between Mg<sup>2+</sup> and pH (Table B.5). In addition more complex linear models were prepared and shrunken via the Lasso method and 5-fold cross validation (see Table B.7 and B.9) [191]. The Lasso is a shrinkage method that can shrink coefficients to exactly zero and thus make a model more interpretable. The results of the shrunken models match well with the results obtained by linear modelling and are further statistical evidence that pH and Mg<sup>2+</sup> show main effects and also interaction effects which seem to be associated with the activity towards *iso*-ferulic acid and catechol (eriodictyol, caffeic acid) methylation by PFOMT. Nonetheless, all of these rather simple models can not reflect the reality of such complex systems as enzymes where lots of factors play important roles. The shown results are solely included for purposes of making statistics based inferences in the context of domain knowledge.

To the knowledge of the author, this is the first time the effects of Mg<sup>2+</sup> and pH on methyl transferase activity were systematically analyzed. It was shown, that

**Table 6.5.:** Coefficients of the model (6.2) for activity of iso-ferulic acid methylation. The factor Mg is a categorical variable (addition/no addition) and can therefore only be 0 or 1.

	Estimate	Std. Error	t value	p-value
(Intercept)	-241.4238	420.1485	-0.57	0.5864
pH	38.4239	54.9778	0.70	0.5108
Mg	-2201.3084	594.1797	-3.70	0.0100 *
pH×Mg	373.8131	77.7503	4.81	0.0030 **



**Figure 6.9.:** pH-profiles of substrate conversion. The linear regression models (blue, dashed lines) grasp the general trend of the data reasonably well to draw inferences. 95 % prediction intervals are displayed as shaded areas.

catecholic and non-catecholic substrates could be activated sufficiently by PFOMT at high pHs without the addition of  $Mg^{2+}$ . It is improbable that, if the active site retained the same miromilieu under every reaction condition, an influence on the rate of reaction would be observed. This could be a hint, that the enzyme relays the chemical information of the environment directly to the substrate to aid in activation.

Furthermore, omission of  $Mg^{2+}$  shifts the pH-optimum of the reaction catalyzed by PFOMT to higher pH-values. Addition of  $Mg^{2+}$ , seems to remedy the effect of high pHs on catalysis at least for caffeic acid, since the initial rates observed were not as high. It would be of interest to analyze this behavior with further systematic studies and multiple levels of  $Mg^{2+}$  concentrations.

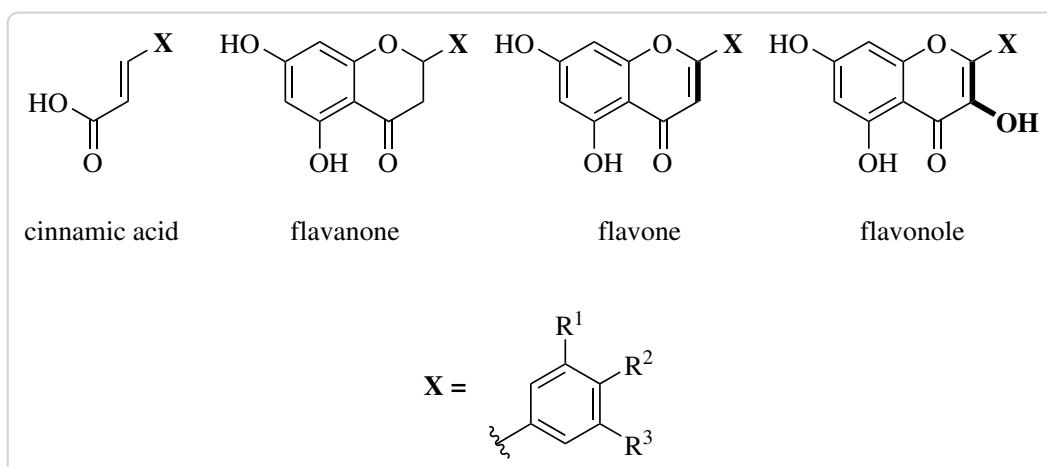
### 6.3.3 Methylation of different chemical motifs

The initial results, that showed the conversion of non-catecholic *iso*-ferulic acid by PFOMT prompted additional experiments with other different non-catechols from multiple flavonoid subgroups (Table 6.6, Figure 6.10). The tested substrates were selected from four different compound groups (cinnamic acids, flavones, flavanones and flavonols) and each group contained each of the structural motifs tested (phenol, catechol, 3O4M, 4O3M), if possible. Each substrate was assessed for conversion with two enzymes (wild-type and 4'-specific variant; Y51R N202W) at four different conditions (pH/ $Mg^{2+}$ : low/no, low/yes, high/no, high/yes). The low and high pH-values were 7.5 and 8.6, respectively. When  $Mg^{2+}$  was added the concentration was 10 mM. The reactions were incubated at 30 °C for 16 h (see section 3.6.3).

There was almost complete conversion of the catecholic substrates under any condition after 16 h of incubation, at least when the wild-type enzyme was used (Table 6.8). This suggests, that the reaction period was chosen too long. Effects on the conversion of catecholic substrates will therefore be disregarded or only discussed superficially. When the wild-type enzyme was employed conversion was observed for all substrates under at least one condition. The highest conversion of non-catecholic substrates was observed at high pH and high  $Mg^{2+}$  conditions

**Table 6.6:** Substrate grid that was tested for methylation with PFOMT. Four different groups of compounds were screened. The groups of flavones, flavanones and cinnamic acids each contained one representative of each motif (phenolic, catecholic, 3'-hydroxy-4'-methoxy (3O4M), 4'-hydroxy-3'-methoxy (4O3M). The substitution patterns refer to Figure 6.10.

substrate	group	motif	R <sup>1</sup>	R <sup>2</sup>	R <sup>3</sup>
naringenin	flavanon	phenolic	H	OH	H
eriodictyol	flavanon	catecholic	OH	OH	H
hesperetin	flavanon	3O4M	OH	OMe	H
homoeriodictyol	flavanon	4O3M	OMe	OH	H
apigenin	flavone	phenolic	H	OH	H
luteolin	flavone	catecholic	OH	OH	H
diosmetin	flavone	3O4M	OH	OMe	H
chrysoeriol	flavone	4O3M	OMe	OH	H
<i>p</i> -coumaric acid	cinnamic acid	phenolic	H	OH	H
<i>m</i> -coumaric acid	cinnamic acid	phenolic	OH	H	H
<i>o</i> -coumaric acid	cinnamic acid	phenolic	H	H	H
caffeic acid	cinnamic acid	catecholic	OH	OH	H
<i>iso</i> -ferulic acid	cinnamic acid	3O4M	OH	OMe	H
ferulic acid	cinnamic acid	4O3M	OMe	OH	H
kaempferol	flavonole	phenolic	H	OH	H
quercetin	flavonole	catecholic	OH	OH	H
myricetin	flavonole	catecholic	OH	OH	OH



**Figure 6.10.:** Substrates from four different groups were screened for methylation by PFOMT under different conditions ( $pH/Mg^{2+}$ ). Figure refers to Table 6.6.

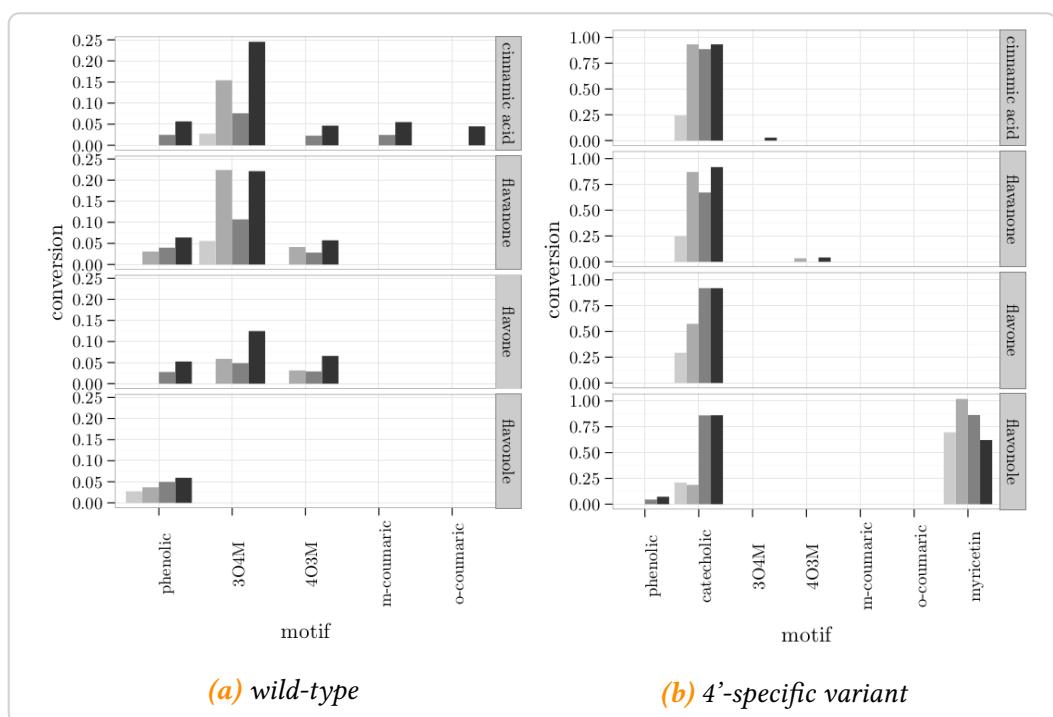
(Figure 6.11). The trend in the data suggests, that methylation efficiency increases with pH, but especially when magnesium is also added. The ME-plot and interaction effects plot (IAE-plot) suggest the same, however there is no statistical evidence to support this notion (see Figure A.10a, A.10b ). Although the statistical tests performed on this data do not back the hypothesis, this does not rule out an influence of pH or  $Mg^{2+}$ . It should be kept in mind, that enzymatic systems are quite complex and simple linear regression models that allow for inferences often cannot describe what is going on accurately and therefore fail. Furthermore this very complex dataset is limited to only about 150 observations.

Overall, methylation of 3O4M motifs was highest, with observed conversions of close to 25 % for the cinnamic acid and flavanone substrates (*iso*-ferulic acid and hesperetin). For these substrates the conversion increased by almost 5-fold upon  $Mg^{2+}$  addition, which is close to the observed increase of the initial rate of *iso*-ferulic acid methylation (6.3.2). Similar results have been shown for SaOMT5, an O-MT from *Streptomyces avermitilis* [223]. Conversion of the somewhat more rigid flavone diosmetin was lower by at least factor two compared to hesperetin and *iso*-ferulic acid. At low pH-values and without addition of  $Mg^{2+}$  barely any conversion of the non-catecholic substrates was observed. The fact, that conversion of 3O4M-moiety bearing substrates is greater than that of the 4'-phenolic and 4O3M

substrates could be due to the fact that the wild-type of PFOMT by and large acts upon 3'-hydroxyls.

The 4'-specific variant for the most part only showed conversion of the catecholic substrates, as expected. Only some conversion of homoeriodictyol ( $\approx 4\%$ ) and iso-ferulic acid ( $\approx 3\%$ ) was observed under high pH/Mg<sup>2+</sup> conditions. However, for the catecholic substrates the same trend – increasing pH/Mg<sup>2+</sup> increases activity – as before holds true. Control experiments without enzyme at high pH and 10 mM revealed, that no substrate conversion took place under these conditions.

The site of methylation was studied by LC-tandem mass-spectrometry (MS/MS). As expected, methylation only took place on the B-ring of the flavonoids. For a throughout discussion refer to chapter 5.



**Figure 6.11:** Conversion of multiple different substrates, catecholic and non catecholic, by PFOMT. pH/Mg<sup>2+</sup> color coded from light to dark:  $\text{X}/\text{X}$ ,  $\text{X}/\checkmark$ ,  $\checkmark/\text{X}$ ,  $\checkmark/\checkmark$ .

To the authors knowledge, this is the first time that methylation of a diverse set of non-catecholic substrates was described for a class I magnesium-dependent methyl transferase. A flavonoid-specific O-MT from *Catharanthus roseus* was described to

**Table 6.8.: Conversion of substrates after 16 hours incubation. Only the maximum conversion is displayed, along with the conditions it was achieved under.**

† – wild-type: substrate conversion was maximal for all pH/Mg<sup>2+</sup> combinations.

‡ – conversion of caffeic acid by the wild-type was set to 100 %.

substrate	wild-type			4'-specific variant		
	conversion %	pH	Mg <sup>2+</sup>	conversion %	pH	Mg <sup>2+</sup>
naringenin	6	↗	✓			
eriodictyol†	94			92	↗	✓
hesperetin	22	↘	✓			
homoeeriodictyol	6	↗	✓	4	↗	✓
apigenin	5	↗	✓			
luteolin†	95			92	↗	
diosmetin	12	↗	✓			
chrysoeriol	7	↗	✓			
<i>p</i> -coumaric acid	6	↗	✓			
<i>m</i> -coumaric acid	6	↗	✓			
<i>o</i> -coumaric acid	4	↗	✓			
caffeic acid†	100‡			93		
iso-ferulic acid	25	↗	✓	3	↗	✓
ferulic acid	5	↗	✓			
kaempferol	6	↗	✓	7	↗	✓
quercetin	93	↗	✗	86	↗	
myricetin	129	↘	✗	102	↘	✓

methylate the 4'-position, when the substrates B-ring possessed a 4O3M substitution [168]. However, said enzyme only showed marginal activities towards catechols and 3O4M derivatives. A class II *O*-MT from wheat, named TaOMT2, is able to sequentially methylate the three hydroxyl-groups on the B-ring of tricetin, in the proposed order 3'-methyl → 3',5'-dimethyl → 3',4',5'-trimethyl [228]. Methylation of dihydroxy-derivatives such as luteolin and eriodictyol by TaOMT2 only gave rise to 3'-mono-methylated products, which is similar to PFOMT. However, there it was also demonstrated that TaOMT2 could methylate tamarixetin, the 4O3M derivative of quercetin, albeit at low activities.

Of the two PFOMT enzymes, 4'-specific variant and wild-type, only the wild-type showed significant methylation of non-catecholic moieties. These findings support the previous results, that could show a pH and magnesium-dependent rate of methylation of *iso*-ferulic acid (subsection 6.3.2). Although methylation of 3'-hydroxyl groups was preferred by the wild-type, a tendency to methylate 4'-hydroxyls, when these were the only ones present, could be demonstrated.

Whereas it is clear, that the N-terminus of PFOMT is important for the function of the enzyme, the role of it *in vivo* is still not fully understood [85, 108, 203]. However, it cannot be ruled out that acts as a signal sequence that can direct the enzyme to different compartments. The findings presented here might give some implications as to the regulation of *O*-MTs, such as PFOMT, since the pH can be quite different in different cell compartments in plants [137].

## 6.4 Conclusion/Discussion

Enzymatic methylation of non-catecholic moieties, was studied using the two methyl transferases PFOMT and SOMT-2, of classes I and II respectively. Therefore multiple different flavonoid and phenylproanoid substrates, displaying either single phenolic, catecholic, 3'-hydroxy-4'-methoxy or 4'-hydroxy-3'-methoxy moieties, were tested. Furthermore, the influence of pH and magnesium addition on PFOMT was systematically studied.

In *in vivo* biotransformation experiments it could be shown, that the class II *O*-methyl transferase SOMT-2 is able to methylate flavonoids and stilbenes at the

4'-OH of the B-ring, regardless the exact moiety (phenolic, catecholic, 4'-hydroxy-3'-methoxy). Although over all the conversions were very low, the conversion of the stilbene resveratrol was superior over all other tested substrates ( $\geq 86\%$  vs.  $\geq 55\%$ ). SOMT-2 exclusively methylated the 4'-OH. Unfortunately, these results are purely based on *in vivo* biotransformations carried out in *E. coli*. SOMT-2 could not be obtained in pure, soluble form for *in vitro* characterization. Nonetheless, using SOMT-2 it was shown that design of experiments and fractional factorial design can be valuable tools for the systematic determination of factors that influence refolding of O-MTs.

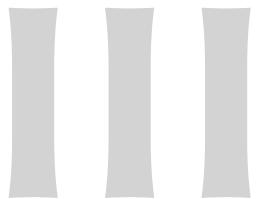
*In vitro* experiments using the class I O-methyl transferase PFOMT, showed that non-catecholic substrates could be methylated. These findings are contrary to the belief, that PFOMT only acted on vicinal aromatic dihydroxyls that are present in compounds such as eriodictyol or caffeic acid. The best conversion of non-catechols was achieved for substrates with 3'-hydroxy-4'-methoxy-moieties (e.g. hesperetin, *iso*-ferulic acid), even though conversion was observed for phenolic (e.g. naringenin) and 4'-hydroxy-3'-methoxy-substrates (e.g. homoeriodictyol), thus demonstrating the ability of PFOMT to methylate both 3'- and 4'-hydroxyls. The best conversions were obtained using the PFOMT wild-type at elevated pH and after Mg<sup>2+</sup> addition. Magnesium addition and pH displayed synergistic effects, meaning the effects of both are not just additive. pH optima shifted under conditions with and without addition of Mg<sup>2+</sup>. When magnesium was omitted, it seemed as though the chemical environment surrounding the enzyme was relayed into the active site. Thus, non-catecholic substrates were methylated at high pH without magnesium, whereas they were hardly methylated at low pH.

These findings show, that the linear stepwise optimization of reaction conditions might not always yield the best overall results, when it comes to such complex systems as enzymes and that synergistic effects need to be considered when looking for the best working conditions.

## 7 Acknowledgements

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetur adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

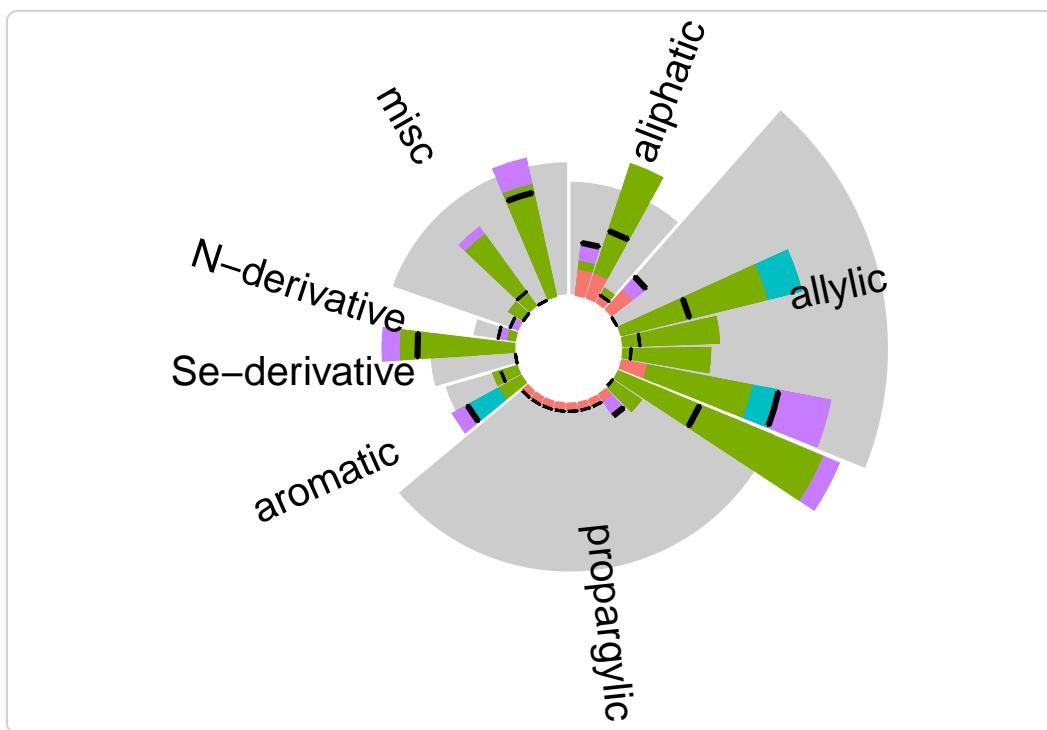
Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetur.



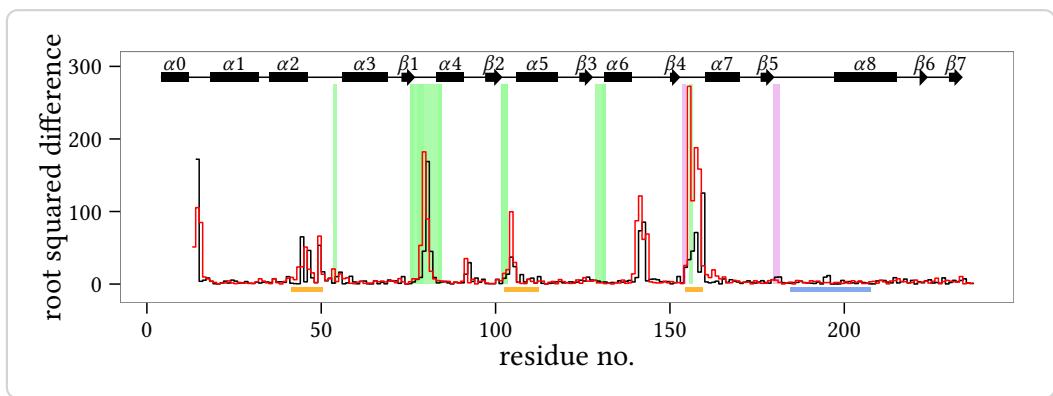
# Appendix

## A Figures

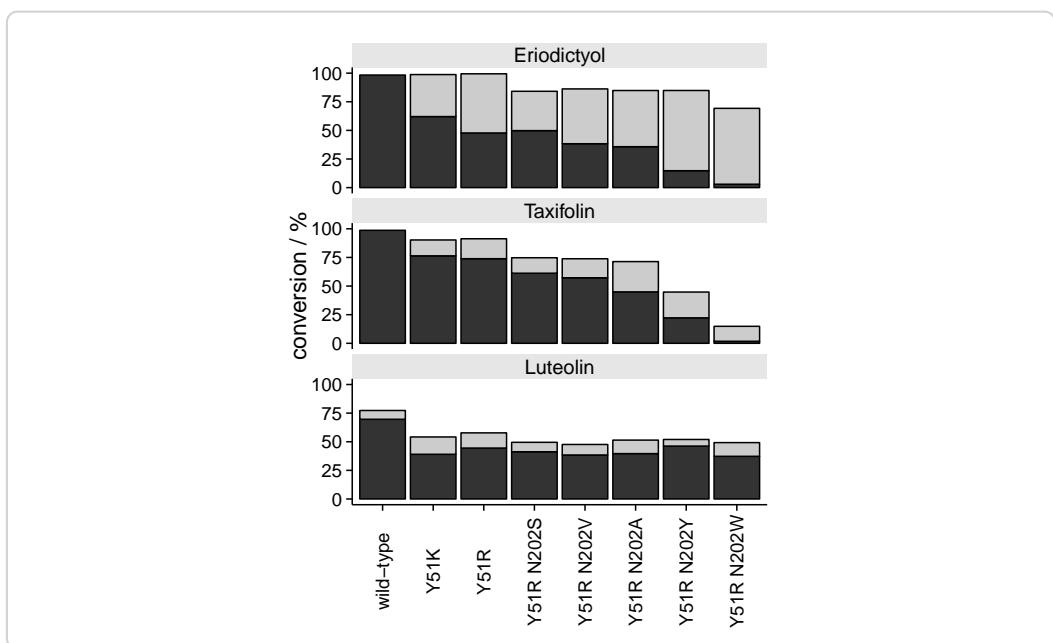
## A.1 Engineering of PFOMT



**Figure A.1:** Graphical representation of the work that has been done on methyl transferases (MTs) in combination with S-adenosyl-L-methionine (SAM) analogues. The grey areas represent individual groups of SAM analogues (aliphatic, allylic, propargylic, aromatic, Se-adenosyl selenomethionine (SeAM) analogues, nitrogen analogues and miscellaneous others). The height of the grey areas represents the number of times a member of the corresponding group has been described as tested in the MT literature. The height of the colored bars represents the times that individual substrate has been tested. The colors represent the different types of MT (red – DNA methyl transferase (DNA-MT), green – protein methyl transferase (P-MT), lilac – small molecule MT, blue – ribonucleic acid (RNA) MT). The black dash across the bar shows the number of times this substrate was actually converted by either enzyme.

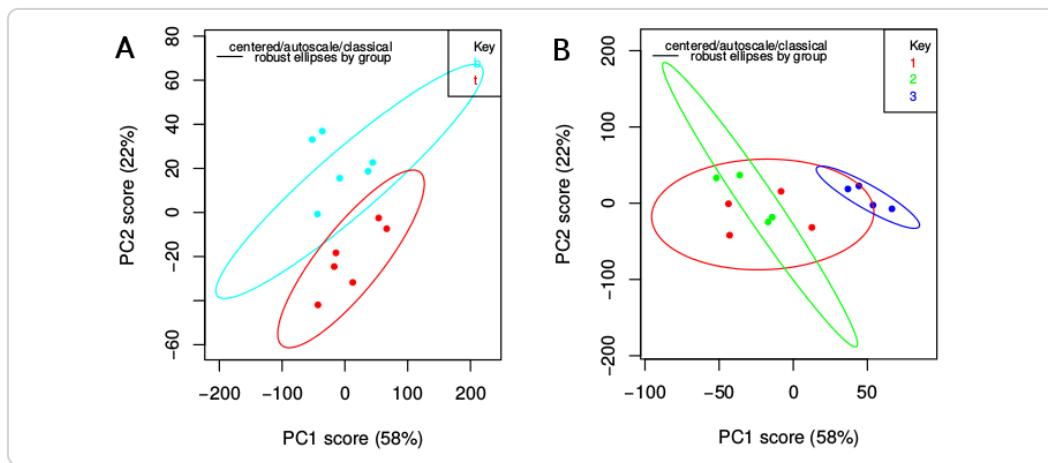


**Figure A.2.:** Differences in the dihedrals  $\psi$  (red) and  $\varphi$  (black) of the solved apo-PFOMT and the structure with bound S-adenosyl-L-homocysteine (SAH) (pdb: 3C3Y). The secondary structure is displayed at the top. Helices are displayed as rectangles and sheets are shown as arrows. Graphical background annotations are used to display the binding sites of SAH (green) and the metal ion (plum). The orange bars indicate regions, where much movement seems to happen upon binding or release of the co-substrate. The blue bar shows the region that was annotated as "insertion loop" in previous studies.



**Figure A.3.:** Differences in the regioselectivity of some phenylpropanoid and flavonoid O-methyl transferase (PFOMT) variants. The products observed in high-performance liquid chromatography (HPLC) and liquid chromatography coupled mass-spectrometry (LC/MS) measurements switched from 3'-methylated (dark grey) to 4'-methylated (light grey) for the displayed variants. The height of the bars corresponds to the total conversion of substrate.

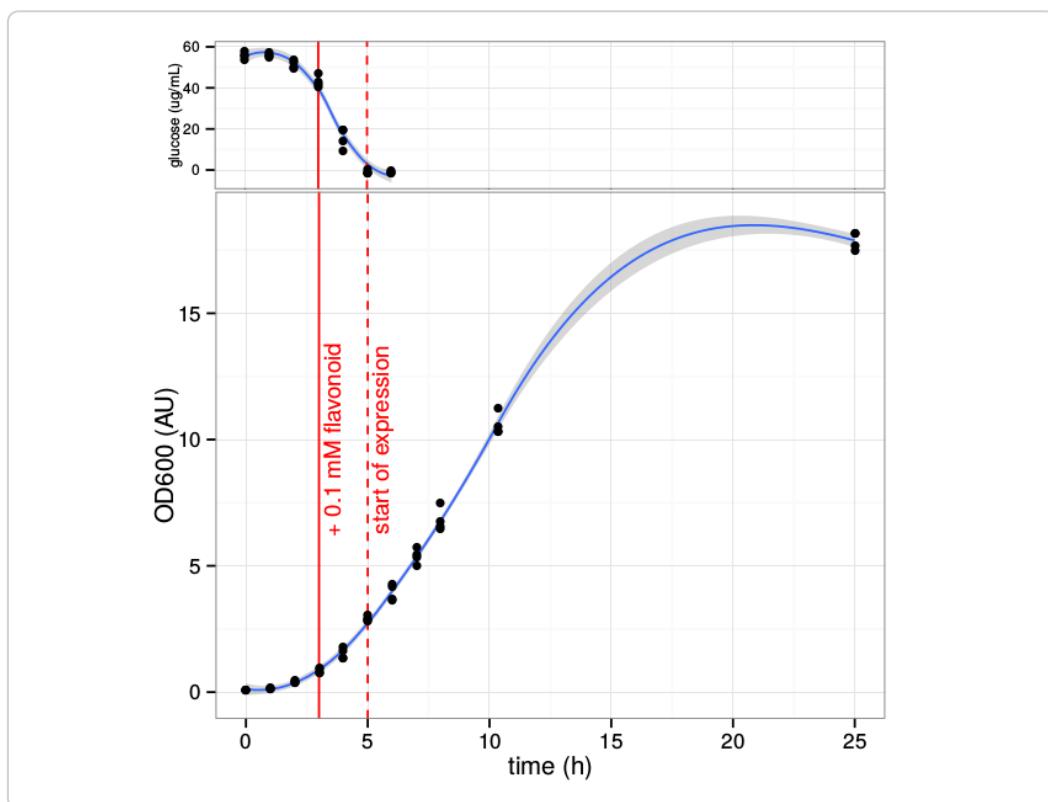
## A.2 Enzymatic methylation of non-catechols



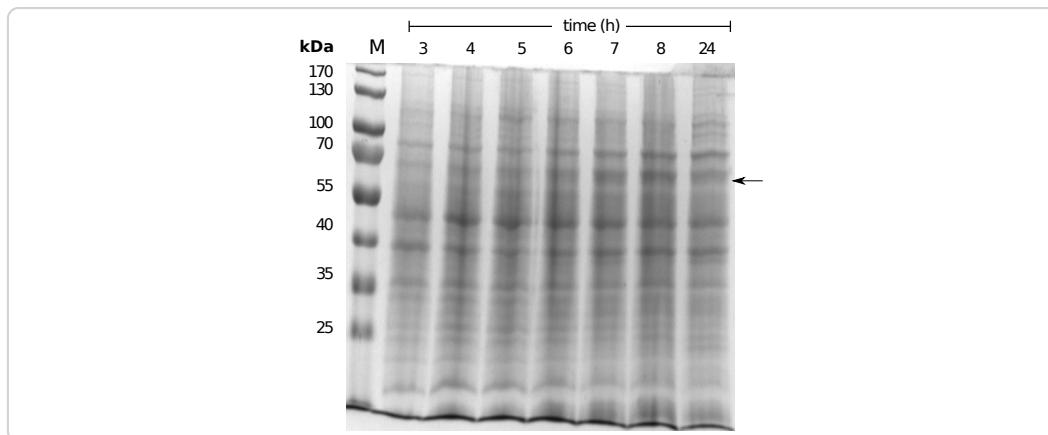
**Figure A.4.:** Additional scatterplots of the principal component analysis (PCA) of HPLC data obtained from *N. benthamiana* leaves infiltrated by *A. tumefaciens* harbouring different constructs. **A** – samples colored by leaf position (top: red; bottom: cyan), **B** – samples colored by plant (plant 1: red; plant 2: green; plant 3: blue)

**Table A.1.:** Results for the Analysis of Variance (ANOVA) of the main effects model describing soluble protein. Significance codes: 0.05 (\*), 0.1 (.)

	df	Sum Sq	Mean Sq	F value	Pr(>F)	
Arginine	1	3595.63	3595.63	24.56	0.0158	*
pH	1	529.87	529.87	3.62	0.1533	
Glycerin	1	752.08	752.08	5.14	0.1083	
ionicStrength	1	82.37	82.37	0.56	0.5077	
divCations	1	5.49	5.49	0.04	0.8588	
redox	1	5.52	5.52	0.04	0.8584	
CycloDex	1	134.67	134.67	0.92	0.4083	
SAH	1	896.83	896.83	6.13	0.0897	.
Residuals	3	439.26	146.42			



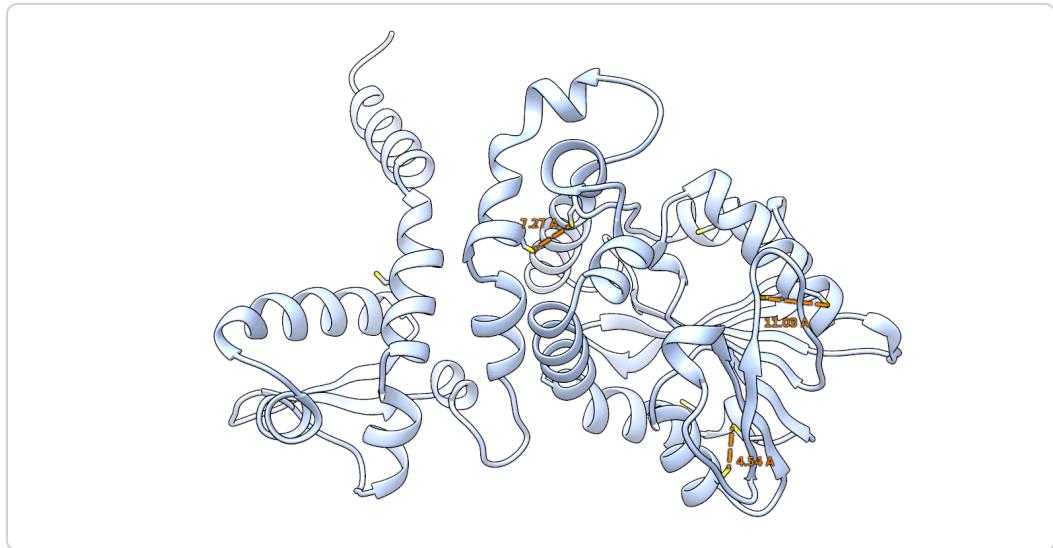
**Figure A.5.:** Growth curve of *E. coli* BL21(DE3) expressing soy O-methyl transferase (SOMT-2) at 37 °C. Glucose is depleted about 5 hours into growth, at which point the start of SOMT-2 expression is expected. The OD<sub>600</sub> after inoculation was about 0.1.



**Figure A.6.:** sodium dodecylsulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) gel of samples acquired during growth curve measurements. The arrow indicated the band that could correspond to the GST-tagged SOMT-2 protein.

**Table A.2.:** Results for the ANOVA of the main effects model describing protein activity.  
Significance codes: 0.05 (\*), 0.1 (.)

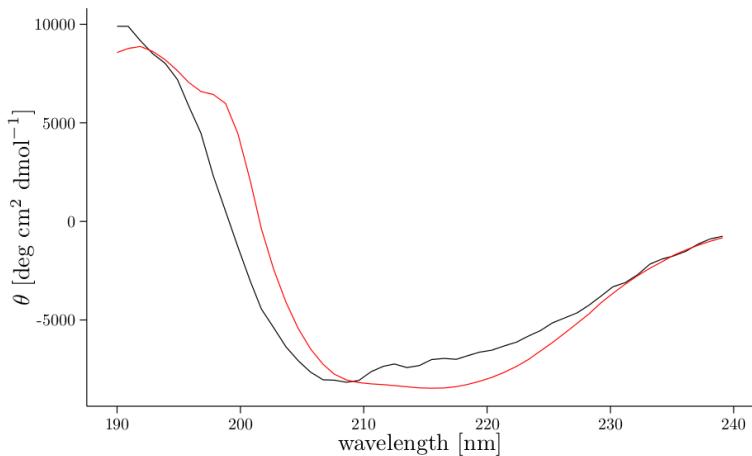
	df	Sum Sq	Mean Sq	F value	Pr(>F)	
Arginine	1	8.31	8.31	6.62	0.0824	.
pH	1	5.71	5.71	4.55	0.1227	
Glycerin	1	0.44	0.44	0.35	0.5945	
ionicStrength	1	3.38	3.38	2.69	0.1997	
divCations	1	0.54	0.54	0.43	0.5605	
redox	1	24.26	24.26	19.31	0.0218	*
CycloDex	1	1.07	1.07	0.85	0.4250	
SAH	1	0.11	0.11	0.09	0.7893	
Residuals	3	3.77	1.26			



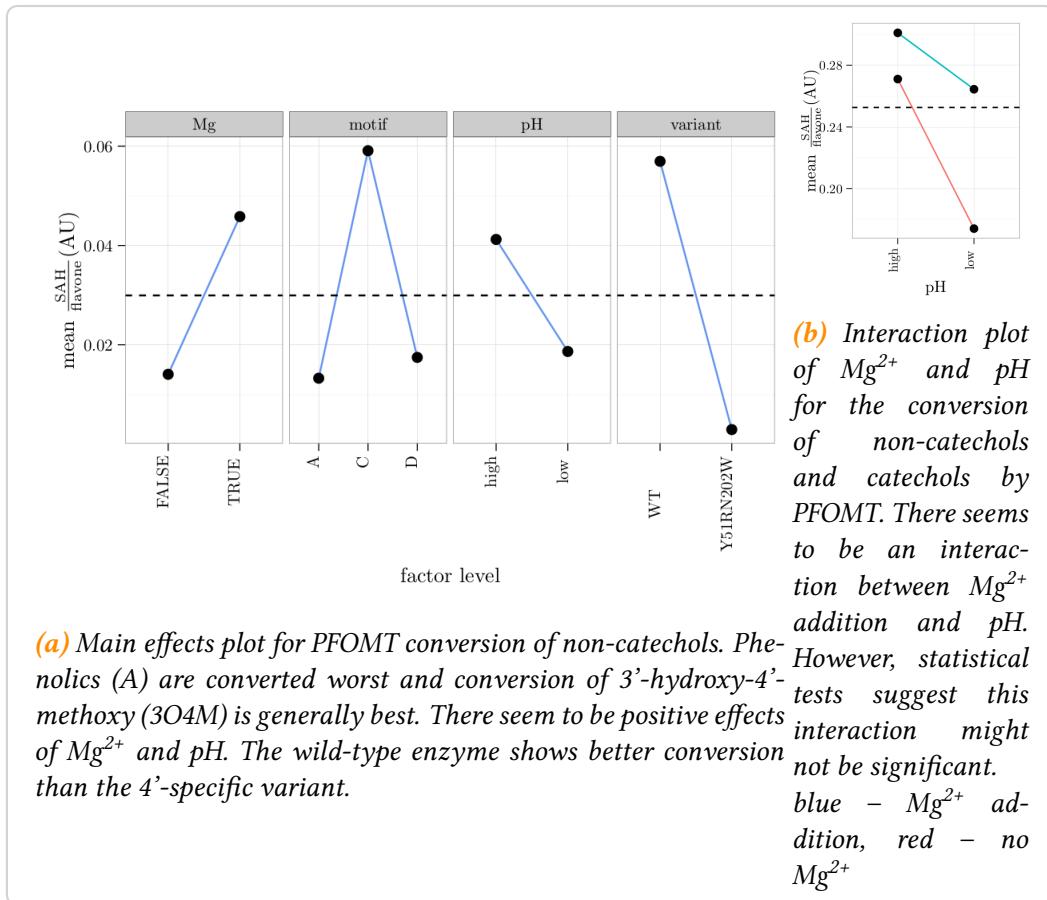
**Figure A.7.:** Graphical representation of a SOMT-2 model obtained from the PHYRE2 web portal (<http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index>) [97]

# Platzhalter

**Figure A.8.:** Chromatogram of the gel filtration analysis of refolded SOMT-2.



**Figure A.9.:** CD-spectrum of refolded SOMT-2 (black) compared to the spectrum that was calculated by the K2D3 web service (<http://cbdm-01.zdv.uni-mainz.de/~andrade/k2d3/index.html>). Secondary structure estimates from the spectrum are 12.39 %  $\alpha$ -helix and 32.51 %  $\beta$ -sheet.

**Figure A.10.**

## B Tables

**Table B.1:** Overview over the constructs produced for the present thesis. Each step during the production of the construct is given in the workflow steps column. Primers (*italic font*) or restriction sites used during each step are displayed in parenthesis.

construct name	description	entry constructs	destination	workflow steps (primers/cloning sites)
pBEW101				
pBEW102	lsrA promoter			
pBEW103	pBEW102 with BamHI cloning site	pBEW102 pBEW4b	pBEW103	amplification ( <i>pRhal1fw/rv</i> ), cloning (BglII, BamHI)
pBEW104	rhaP_BAD promoter			
pBEW106	pICH413038-somt	pET28MC-somt	pICH413038	amplification ( <i>somt1/2/3/4</i> ), cloning (BpiI)
pBEW107		pICH51266, pBEW106, pICH41421	pICH75044	golden gate cloning (BsaI)
pBEW1a				
pBEW1b				
pBEW2a				
pBEW2b				
pBEW3a				
pBEW3b				
pBEW4a				
pBEW4b				
pET28-pfomt	<i>pJomt</i> gene in pET-28a(+), endogenous NdeI site removed	pQE30-pfomt	pET-28a(+)	mutagenesis ( <i>pJomt1fw/rv</i> ), amplification ( <i>pJomt2fw/rv</i> ), cloning (NdeI, EcoRI)
pET20-somt	N-terminal pelB-tag fusion for periplasmic expression			pET20-b(+) pET28-a(+)
pET28-somt				
pET28MC-somt				
pET32-somt				
pET41-somt	N-terminal TrX-tag fusion			pET-32a(+)
pUC19*	N-terminal GST-tag fusion added BglII site	pUC19 lsr-XX-DAS	pET-41a(+)	-
pUCB1	pUC19 derivative with lsrA promoter			mutagenesis ( <i>pUC1.fw/rv</i> ) cloning (NdeI, BglII)
pUCB1-sfGFP-DAS+4				

## B.1 Engineering of PFOMT

**Table B.3.**: SAM analogues that have been used with MTs. Targets: P – peptide/protein, D – DNA, R – RNA, S – small molecule.

analogue	enzyme	target	references
<i>SAM</i>			
-CH <sub>2</sub> -CH <sub>3</sub>	PRMT1, M.TaqI, M.HhaI, M.BcnIB, RebM, RapM	S,P,D	[43, 113, 181, 190] <sup>1</sup>
-CH <sub>2</sub> -CH <sub>2</sub> -CH <sub>3</sub>	PRMT1, M.TaqI, M.HhaI, M.BcnIB	P,D	[43, 190]
-CH <sub>2</sub> -CH <sub>2</sub> -CH <sub>2</sub> -CH <sub>3</sub>	PRMT1	P	[190]
-CH <sub>2</sub> -C <sub>6</sub> H <sub>5</sub>	NovO, CouO,	S,P	[183, 190]
-CH <sub>2</sub> -C(=O)-CH <sub>3</sub>	PRMT1 COMT, TPMT, CazF	S	[115, 213]
-CH <sub>2</sub> -CH=CH <sub>2</sub>	NovO, CouO, RapM, PRMT1, M.TaqI, M.HhaI, M.BcnIB, RebM, Tgs	P,S,D	[43, 113, 171, 181, 183, 188, 190, 204, 205]
-CH <sub>2</sub> -CH=CH-CH <sub>3</sub>	NovO, CouO	S	[183]

<sup>1</sup>Singh *et al.* (2014) published a series of 44 biocatalytically synthesized SAM and Se-adenosyl selenomethionine (SeAM) derivatives, most of which were not tested towards their alkyl donation potential in MT reactions.

analogue	enzyme	target	references
$-\text{CH}_2-\text{C}\equiv\text{CH}$	Dim-5, <i>HsMLL</i> ,	P,R,S	[86, 183, 204, 205, 211]
	TRM1,		
	NovO, CouO,		
	PRMT1		
$-\text{CH}_2-\text{C}\equiv\text{N}$	RebM	S	[181]
$-\text{CH}_2-\text{CH}_2-\text{C}\equiv\text{CH}$	PKMT	P	[86]
$-\text{CH}_2-\text{CH}_2-\text{CH}_2-\text{C}\equiv\text{CH}$	PKMT	P	[86]
$-\text{CH}_2-\text{C}\equiv\text{C}-\text{CH}_3$	NovO, CouO, M.HhaI, M.TaqI, M.BcnIB	S,D	[43, 128, 183]
$-\text{CH}_2-\text{C}\equiv\text{C}-\text{CH}_2-\text{CH}_3$	M.HhaI	D	[128]
$-\text{CH}_2-\text{C}\equiv\text{C}-\text{CH}_2-\text{NH}_2$	M.HhaI	D	[128]
$-\text{CH}_2-\text{C}\equiv\text{C}-\text{CH}_2-\text{NH}-\text{C}(=\text{O})(-\text{CH}_2-)_3-\text{NH}_2$	M.HhaI	D	[128]
$-\text{CH}_2-\text{C}\equiv\text{C}(-\text{CH}_2-)_3-\text{NH}_2$	M.HhaI	D	[128]
$-\text{CH}_2-\text{C}\equiv\text{C}(-\text{CH}_2-)_3-\text{NH}-\text{C}(=\text{O})(-\text{CH}_2-)_3-\text{NH}_2$	M.HhaI	D	[128]
$-\text{CH}_2-\text{C}\equiv\text{C}(-\text{CH}_2-)_3-\text{C}\equiv\text{CH}$	M.HhaI	D	[128]
$-\text{CH}_2-\text{C}\equiv\text{C}(-\text{CH}_2-)_3-\text{N}_3$	M.HhaI	D	[128]
$-\text{CH}_2-\text{CH}=\text{CH}-\text{C}\equiv\text{CH}$	Dim-5, <i>HsMLL</i> ,	P,R	[86, 148, 156, 171, 204, 205, 211]
	TRM1,		
	PRMT1, Tgs		
$-\text{CH}_2-\text{CH}=\text{CH}-\text{CH}_2-\text{C}\equiv\text{CH}$			[86, 204]
$-\text{CH}_2-\text{CH}=\text{CH}-\text{CH}_2-\text{O}-\text{CH}_2-\text{C}\equiv\text{CH}$	PRMT1	P	[204, 205]
<i>SeAM</i>			
$-\text{CH}_3$			
$-\text{CH}_2-\text{C}\equiv\text{CH}$	Dim-5, <i>HsMLL</i> ,	P,R,S	[19, 181, 211, 213]
	TRM1, RebM,		
	CazF		

analogue	enzyme	target	references
<i>N</i> -mustard derivatives			
–CH <sub>2</sub> –CH <sub>2</sub> –I	RebM	S	[224]

**Table B.4.** Crystallographic data, phasing and refinement statistics.

140519_PFOMT	
<i>data collection</i>	
wavelength (Å)	
resolution (Å)	1.95
total reflections	392 368
unique reflections	125 822
completeness (%)	99.12
$I/\sigma(I)$	9.9
$R_{\text{sym}}^a$	
redundancy	
space group	$P2_12_12_1$
cell dimensions (Å)	
<i>a</i>	86.16
<i>b</i>	128
<i>c</i>	129.3
<i>refinement</i>	
$R_{\text{work}}/R_{\text{free}}$	0.21369 / 0.24700
rmsd bond lengths (Å)	0.0199
rmsd bond angles (°)	2.0568
B-values (Å <sup>2</sup> )	21.593
water	
ions	
<i>Ramachandran plot (%)</i>	
favoured	96.82
allowed	2.38
outliers	0.8

## B.2 Enzymatic methylation of non-catechols

**Table B.5.:** Coefficients of the model (6.3) for activity of catecholic methylation. The factor Mg is a categorical variable (addition/no addition) and can therefore only be 0 or 1.

	Estimate	Std. Error	t value	p-value
(Intercept)	-421929.9946	356063.7085	-1.18	0.2557
Mg	-839999.8874	503550.1257	-1.67	0.1175
pH	103271.3345	97739.1728	1.06	0.3086
pH <sup>2</sup>	-4977.7406	6512.6996	-0.76	0.4574
Mg×pH	266920.7964	138224.0638	1.93	0.0740
Mg×pH <sup>2</sup>	-19830.2264	9210.3481	-2.15	0.0492 *

**Table B.7.:** Coefficients obtained for linear regression model suing the iso-ferulic acid subset after shrinkage using the Lasso method and 5-fold cross validation. Only non-zero coefficients (variables actually do have an effect) are retained during the Lasso. Seed was set to 1337.

variable	coefficient
(Intercept)	-509.8385
pH	73.4085
Mg	-1606.1362
pH×Mg	296.0753

**Table B.9.:** Coefficients obtained for linear regression model using the catechols subset after shrinkage using the Lasso method and 5-fold cross validation. Only non-zero coefficients (variables actually do have an effect) are retained during the Lasso. Seed was set to 1336.

variable	coefficient
(Intercept)	-467632.3821
pH	94469.8366
pH×Mg	19068.9540
pH×pH <sup>2</sup>	-381.5863
pH×Mg×pH <sup>2</sup>	-292.3608

### **B.3 Tandem mass-spectrometry studies of flavonoids**

**Table B.11:** Key ions in the positive mode CID and HCD ESI-MS<sup>2</sup> spectra of flavanones.

fragment	CID, 45 % NCE					HCD, 75 % NCE					HCD, 100 % NCE				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
2 [M+H-H <sub>2</sub> O] <sup>+</sup>	255(1)	271(18)	269(1)	285(10)	285(4)										
4 [M+H-C <sub>2</sub> H <sub>2</sub> O] <sup>+</sup>	231(4)	247(3)	245(3)	261(2)	261(2)										
5 [M+H-2C <sub>2</sub> H <sub>2</sub> O] <sup>+</sup>	189(5)	205(3)	203(4)	219(2)	219(1)										
8 AC <sup>+</sup>	179(4)	179(20)	179(5)	179(28)	179(30)	179(1)	179(2)	179(2)	179(2)	179(2)	179(2)	179(1)			
11 1,3A <sup>+</sup>	153(100)	153(31)	153(77)	153(21)	153(57)	153(100)	153(100)	153(100)	153(100)	153(100)	153(69)	153(50)	153(100)	153(63)	153(58)
12 1,3A <sup>+</sup> -CO											125(1)	125(1)			
13 1,3A <sup>+</sup> -C <sub>2</sub> H <sub>2</sub> O											111(2)	111(2)	111(4)	111(2)	111(2)
14 1,3A <sup>+</sup> -2CO						97(3)	97(4)	97(3)	97(4)	97(4)	97(10)	97(8)	97(15)	97(9)	97(9)
15 1,3A <sup>+</sup> -2CO-C <sub>2</sub> H <sub>4</sub>						69(2)	69(2)	69(2)	69(2)	69(2)	69(9)	69(8)	69(13)	69(9)	69(8)
17 1,4B <sup>+</sup> -2H	147(84)	163(100)	161(100)	177(100)	177(100)	147(15)	163(10)	161(10)	177(4)	177(4)	147(3)	163(1)	161(3)		
18 1,4B <sup>+</sup> -2H-H <sub>2</sub> O						145(5)	145(7)	145(7)	145(7)	145(7)					
19 1,4B <sup>+</sup> -2H-CO	119(3)	135(1)	133(4)			119(32)	135(29)	133(36)	149(15)	149(11)	119(34)	135(16)	133(36)	149(5)	149(3)
20 1,4B <sup>+</sup> -2H-CO-CH <sub>3</sub>							118(11)	134(11)	134(7)	91(100)				118(57)	134(20)
22 1,4B <sup>+</sup> -2H-2CO															134(13)
23 1,4B <sup>+</sup> -2H-2CO-CH <sub>3</sub>															
24 1,4B <sup>+</sup> -2H-C <sub>2</sub> H <sub>2</sub> O-H <sub>2</sub> O															
25 1,4B <sup>+</sup> -2H-H <sub>2</sub> O-CO															
26 C <sub>7</sub> H <sub>7</sub> <sup>+</sup>						91(24)	117(18)	91(1)	91(3)	91(2)	91(100)	91(5)	91(7)	91(5)	
27 C <sub>7</sub> H <sub>5</sub> <sup>+</sup>						89(23)		89(29)	89(24)	89(24)	89(100)	89(22)	89(100)	89(100)	

**Table B.12:** Key ions in the positive mode CID and HCD ESI-MS<sup>2</sup> spectra of flavones.

fragment	CID, 45 % NCE					HCD, 75 % NCE					HCD, 100 % NCE				
	(6)	(7)	(8)	(9)	(10)	(6)	(7)	(8)	(9)	(10)	(6)	(7)	(8)	(9)	(10)
1 [M+H] <sup>+</sup>	271 (2)	285 (66)	285 (4)	285 (66)	286 (100)	271 (84)	287 (66)	285 (4)	286 (20)	286 (18)	271 (2)	287 (2)			
2 [M+H-CH <sub>3</sub> ] <sup>•+</sup>	253 (1)	269 (9)	270 (100)	286 (100)	286 (100)	253 (3)	269 (6)	270 (9)	286 (20)	286 (18)					
3 [M+H-H <sub>2</sub> O] <sup>+</sup>	243 (7)	259 (9)				243 (7)									
4 [M+H-CO] <sup>+</sup>	242 (14)	258 (47)				242 (1)	258 (3)								
5 [M+H-CHO] <sup>•+</sup>	229 (21)	245 (13)	243 (1)	242 (7)	229 (4)										
6 [M+H-C <sub>2</sub> H <sub>2</sub> O] <sup>+</sup>						242 (100)	258 (100)	258 (100)							
7 [M+H-CH <sub>3</sub> -CO] <sup>•+</sup>						242 (100)	258 (100)	258 (100)							
8 [M+H-CH <sub>4</sub> -CO] <sup>+</sup>						241 (1)	257 (7)	257 (7)							
9 [M+H-H <sub>2</sub> O-CO] <sup>+</sup>	225 (13)	241 (13)				225 (4)	241 (16)								
14 [M+H-CH <sub>4</sub> -2CO] <sup>+</sup>						197 (4)	213 (7)								
15 [M+H-H <sub>2</sub> O-2CO] <sup>+</sup>						187 (2)	203 (2)								
16 [M+H-2C <sub>2</sub> H <sub>2</sub> O] <sup>+</sup>						199 (1)									
17 [M+H-CH <sub>3</sub> OH-2CO] <sup>•+</sup>															
18 [M+H-CH <sub>4</sub> -2CO-C <sub>2</sub> H <sub>2</sub> ] <sup>+</sup>															
20 [M+H-CH <sub>4</sub> -3CO] <sup>+</sup>															
23 [M+H-CH <sub>4</sub> -4CO] <sup>+</sup>															
24 [M+H-2CO-2C <sub>2</sub> H <sub>2</sub> O] <sup>+</sup>															
27 0,4B <sup>+</sup>	163 (6)	179 (7)				131 (2)	147 (1)								
28 0,4B <sup>+</sup> -H <sub>2</sub> O	145 (13)	161 (12)				163 (8)	179 (3)								
25 0,4B <sup>+</sup> -C <sub>2</sub> H <sub>2</sub> O	121 (6)	137 (7)				145 (17)	161 (16)								
29 1,3A <sup>+</sup>	153 (100)	153 (100)	153 (5)			121 (16)	137 (12)								
30 1,3A <sup>+</sup> -CO						153 (100)	153 (100)	153 (11)	153 (8)	153 (8)	153 (84)	153 (87)	153 (100)	153 (100)	153 (100)
31 1,3A <sup>+</sup> -C <sub>2</sub> H <sub>2</sub> O						125 (1)	125 (2)				125 (3)	125 (2)			
32 1,3A <sup>+</sup> -2CO						111 (2)	111 (2)				111 (4)	111 (3)			
33 1,3A <sup>+</sup> -2CO-C <sub>2</sub> H <sub>4</sub>						97 (2)	97 (2)				97 (9)	97 (9)			
34 1,3B <sup>+</sup>	119 (12)	135 (11)	133 (2)			69 (4)	69 (5)				69 (24)	69 (22)			
35 1,4A <sup>+</sup> +2H						119 (49)	133 (49)	133 (3)			119 (35)				
39 C <sub>7</sub> H <sub>7</sub> <sup>+</sup>	91 (1)					91 (26)	127 (1)				91 (100)	127 (2)			
40 C <sub>7</sub> H <sub>5</sub> <sup>+</sup>							89 (17)				89 (7)	89 (100)	89 (3)		89 (1)

**Table B.13.** Key ions in the positive mode CID and HCD ESI-MS<sup>2</sup> spectra of flavonoles.

fragment	CID, 45 % NCE						HCD, 75 % NCE						HCD, 100 % NCE					
	(11)	(12)	(13)	(14)	(15)	(11)	(12)	(13)	(14)	(15)	(11)	(12)	(13)	(14)	(15)	(11)	(12)	(13)
1 [M+H] <sup>+</sup>							287(25)	303(8)	319(1)	301(9)	286(12)	302(6)						
2 [M+H-CH <sub>3</sub> ] <sup>•+</sup>							286(62)	302(100)			285(5)	301(3)						285(1)
3 [M+H-CH <sub>4</sub> ] <sup>+</sup>							285(2)		285(6)	301(1)								
4 [M+H-H <sub>2</sub> O] <sup>+</sup>	269(32)	285(63)	301(40)	283(2)	299(3)	269(2)	285(6)	301(1)					269(2)	285(3)				
5 [M+H-CO] <sup>+</sup>	259(24)	275(14)	291(6)	273(3)	259(3)	259(3)												
6 [M+H-CHO] <sup>•+</sup>	258(46)	274(20)	290(22)	272(20)	288(3)	258(10)	274(4)	272(2)	272(2)	269(11)	285(11)							
7 [M+H-CH <sub>3</sub> OH] <sup>•+</sup>						269(11)	285(36)											
9 [M+H-C <sub>2</sub> H <sub>2</sub> O] <sup>+</sup>	245(4)																	
10 [M+H-CH <sub>3</sub> -CO] <sup>•+</sup>							258(9)	274(5)			258(58)	274(6)						
12 [M+H-CH <sub>4</sub> -CO] <sup>+</sup>							257(3)				257(17)	273(13)						
13 [M+H-H <sub>2</sub> O-CO] <sup>+</sup>	241(99)	257(100)	273(100)	255(3)	271(5)	241(5)	257(13)	273(6)			241(1)	257(2)						
23 1.(40)	247(37)	263(25)	245(17)	261(8)	231(5)	247(2)							241(2)	257(7)				
14 [M+H-2CO] <sup>+</sup>				241(16)	257(8)	216(2)	232(2)			230(94)	246(9)	216(2)						
15 [M+H-CH <sub>3</sub> OH-CO] <sup>•+</sup>										229(100)	245(42)							
17 [M+H-CH <sub>3</sub> -2CO] <sup>•+</sup>										227(2)								
19 [M+H-CH <sub>4</sub> -2CO] <sup>+</sup>										213(20)	229(49)							
20 [M+H-H <sub>2</sub> O-2CO] <sup>+</sup>										203(2)	235(2)							
21 [M+H-3CO] <sup>+</sup>										219(1)								
22 [M+H-2C <sub>2</sub> H <sub>2</sub> O] <sup>+</sup>	203(7)	219(4)	235(4)	217(2)		231(1)	213(11)	229(33)			203(1)	219(2)						
23 [M+H-H <sub>2</sub> O-CO-C <sub>2</sub> H <sub>2</sub> O] <sup>•+</sup>	199(3)	215(2)	231(2)	213(4)	229(3)	213(4)	229(3)			213(11)	229(33)							
23 [M+H-CH <sub>3</sub> OH-2CO] <sup>•+</sup>										203(3)	219(7)							
24 [M+H-CH <sub>4</sub> -2CO-2C <sub>2</sub> H <sub>2</sub> ] <sup>+</sup>										203(3)	219(8)	201(25)	217(41)	187(2)	203(10)	219(21)	201(50)	217(56)
25 [M+H-H <sub>2</sub> O-2CO-C <sub>2</sub> H <sub>2</sub> ] <sup>+</sup>										201(25)	217(41)						201(50)	217(56)
26 [M+H-CH <sub>4</sub> -3CO] <sup>+</sup>												187(4)	203(18)					
27 [M+H-CH <sub>4</sub> -2CO-C <sub>2</sub> H <sub>2</sub> O] <sup>+</sup>										173(2)	189(4)							
28 [M+H-CH <sub>4</sub> -4CO] <sup>+</sup>												147(8)	163(5)	179(9)				
29 [M+H-2CO-C <sub>2</sub> H <sub>2</sub> O] <sup>+</sup>	147(13)	163(7)	179(7)	161(13)	177(2)	147(9)	163(7)	179(9)	161(4)									
31 0.2A <sup>+</sup>	165(100)	165(59)	165(41)	165(31)	165(6)	165(11)	165(9)	165(6)	165(6)	165(1)	165(2)	165(3)	165(4)	165(1)				
32 0.2A <sup>+</sup> -CO	137(11)	137(23)	137(6)	137(4)		137(14)	137(47)	137(15)	137(7)	137(4)	137(12)	137(38)	137(32)	137(13)				
33 0.2B <sup>+</sup>	121(36)	137(23)	153(35)	155(18)	151(2)	137(47)	153(100)	155(14)	151(2)	121(69)	137(38)	153(100)	135(5)	151(1)				
35 0.3A <sup>+</sup> +2H	153(61)	153(26)	153(35)	139(100)	139(9)	139(3)	139(7)	139(1)	139(1)									
37 1.3A <sup>+</sup>												153(100)	153(100)	153(100)	153(100)	153(100)	153(67)	153(100)
38 1.3A <sup>+</sup> -CO	111(19)	111(7)	111(4)			111(5)	111(6)	111(1)			125(2)	125(1)						
39 1.3A <sup>+</sup> -C <sub>2</sub> H <sub>2</sub> O						97(2)	97(2)	97(3)			111(4)	111(5)						
40 1.3A <sup>+</sup> -2CO						69(7)	69(8)	69(8)	69(3)	69(1)	97(9)	97(10)	97(14)	97(3)				
41 1.3A <sup>+</sup> -2CO-C <sub>2</sub> H <sub>4</sub>						133(3)	149(3)	165(6)	163(1)	133(1)	149(4)	165(4)	147(3)	163(5)				
42 1.3B <sup>+</sup> -2H	133(25)	149(10)	165(41)			163(2)			91(2)	91(4)								
49 C <sub>7</sub> H <sub>7</sub> <sup>+</sup>												89(6)	89(9)	89(7)	89(2)	89(2)		
50 C <sub>7</sub> H <sub>5</sub> <sup>+</sup>																		

## C Affidavit

---

*Appendix C. Affidavit*

---

I hereby declare that this document has been written only by the undersigned and without any assistance from third parties. Furthermore, I confirm that no sources have been used in the preparation of this document other than those indicated in the thesis itself.

Date:....., Location:....., Signature:.....

# Bibliography

- [1] Paul D. Adams et al. "PHENIX: A comprehensive Python-based system for macromolecular structure solution". en. In: *Acta Crystallographica Section D: Biological Crystallography* 66.2 (Feb. 2010), pp. 213–221.
- [2] Giovanni Agati et al. "Flavonoids as antioxidants in plants: location and functional significance." In: *Plant science : an international journal of experimental plant biology* 196 (Nov. 2012), pp. 67–76.
- [3] Agilent Technologies. *QuikChange II Site-Directed Mutagenesis Kit: Instruction Manual*. 2011.
- [4] Neda Akbari et al. "Efficient refolding of recombinant lipase from Escherichia coli inclusion bodies by response surface methodology". In: *Protein Expression and Purification* 70.2 (Apr. 2010), pp. 254–259.
- [5] Martin Alexander and B. K. Lustigman. "Effect of Chemical Structure on Microbial Degradation of Substituted Benzenes". In: *Journal of Agricultural and Food Chemistry* 14.4 (July 1966), pp. 410–413.
- [6] Bernd Anselment et al. "Experimental optimization of protein refolding with a genetic algorithm". In: *Protein Science* 19.11 (2010), pp. 2085–2095.
- [7] F Ausubel et al. "Current Protocols in Molecular Biology". In: (2008), p. 23.
- [8] Peter Ball et al. "Kinetic Properties of a Soluble Catechol 0-Methyltransferase of Human Liver". In: *Eur J Biochem* 26 (1972), pp. 560–569.
- [9] R V Banerjee and R G Matthews. "Cobalamin-dependent methionine synthase." In: *The FASEB journal : official publication of the Federation of American Societies for Experimental Biology* 4.5 (1990), pp. 1450–1459.
- [10] G.M. Barton. "New C-methylflavanones from Douglas-fir". In: *Phytochemistry* 11.1 (Jan. 1972), pp. 426–429.
- [11] Isabelle Benoit et al. "Expression in Escherichia coli, refolding and crystallization of Aspergillus niger feruloyl esterase A using a serial factorial approach". In: *Protein Expression and Purification* 55.1 (Sept. 2007), pp. 166–174.
- [12] Sanjoy K. Bhattacharya and Ashok K. Dubey. "Kinetic mechanism of cytosine DNA methyltransferase MspI". In: *Journal of Biological Chemistry* 274.21 (1999), pp. 14743–14749.
- [13] Amie K Boal et al. "Structural basis for methyl transfer by a radical SAM enzyme." In: *Science (New York, N.Y.)* 332.6033 (2011), pp. 1089–1092.
- [14] S. Bocker and F. Rasche. "Towards de novo identification of metabolites by analyzing tandem mass spectra". In: *Bioinformatics* 24.16 (2008), pp. i49–i55.

- [15] Gyu Kim Bong et al. "Synthesis of ermanin, 5,7-dihydroxy-3,4'-dimethoxyflavone from kaempferol, 3,5,7,4'-tetrahydroxyflavone with two O-methyltransferases expressed in *E. coli*". In: *Bulletin of the Korean Chemical Society* 27.3 (2006), pp. 357–358.
- [16] Ronald T. Borchardt. "Mechanism of alkaline hydrolysis of S-adenosyl-L-methionine and related sulfonium nucleosides". In: *Journal of the American Chemical Society* 101.2 (1979), pp. 458–463.
- [17] U. T. Bornscheuer et al. "Engineering the third wave of biocatalysis". In: *Nature* 485.7397 (May 2012), pp. 185–194.
- [18] Ian R. Bothwell and Minkui Luo. "Large-scale, protection-free synthesis of Se-adenosyl-l-selenomethionine analogues and their application as cofactor surrogates of methyltransferases". In: *Organic Letters* 16.11 (2014), pp. 3056–3059.
- [19] Ian R. Bothwell et al. "Se-adenosyl-L-selenomethionine cofactor analogue as a reporter of protein methylation". In: *Journal of the American Chemical Society* 134.36 (2012), pp. 14905–14912.
- [20] George E. P. Box, J. Stuart Hunter, and William G. Hunter. *Statistics for Experimenters: Design, Innovation, and Discovery*. 2nd ed. New York: Wiley-Interscience, 2005.
- [21] M M Bradford. "A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding." In: *Analytical biochemistry* 72.1-2 (May 1976), pp. 248–254.
- [22] Wolfgang Brandt, Kerstin Manke, and Thomas Vogt. "A catalytic triad – Lys-Asn-Asp – Is essential for the catalysis of the methyl transfer in plant cation-dependent O-methyltransferases". In: *Phytochemistry* 113 (May 2015), pp. 130–139.
- [23] W Brandt et al. "Molecular and structural basis of metabolic diversity mediated by prenyldiphosphate converting enzymes". In: *Phytochemistry* 70.15-16 (Oct. 2009), pp. 1758–1775.
- [24] Joan B Broderick et al. "Radical S - Adenosylmethionine Enzymes". In: (2014).
- [25] W. Butte. "Rapid method for the determination of fatty acid profiles from fats and oils using trimethylsulphonium hydroxide for transesterification". In: *Journal of Chromatography A* 261 (1983), pp. 142–145.

- [26] G. L. Cantoni. "S-Adenosylmethionine; Anew intermediate formed enzymatically from L-methionine and adenosinetriphosphate". In: *The Journal of Biological Chemistry* 204 (1953), pp. 403–416.
- [27] Guohua Cao, Emin Sofic, and Ronald L. Prior. "Antioxidant and Prooxidant Behavior of Flavonoids: Structure-Activity Relationships". In: *Free Radical Biology and Medicine* 22.5 (Jan. 1997), pp. 749–760.
- [28] Hung Ju Chen, Baskaran Stephen Inbaraj, and Bing Huei Chen. "Determination of phenolic acids and flavonoids in Taraxacum formosanum kitam by liquid chromatography-tandem mass spectrometry coupled with a post-column derivatization technique". In: *International Journal of Molecular Sciences* 13.1 (2012), pp. 260–285.
- [29] Jing Chen et al. "Different Effects of L -Arginine on Protein Refolding : Suppressing Aggregates of Hydrophobic Interaction, Not Covalent Binding". In: (2008), pp. 1365–1372.
- [30] Vincent B. Chen et al. "MolProbity: All-atom structure validation for macromolecular crystallography". In: *Acta Crystallographica Section D: Biological Crystallography* 66.1 (Jan. 2010), pp. 12–21.
- [31] a Claiborne and Irwin Fridovich. "of of of". In: *Society* 18 (1979), pp. 2324–2329.
- [32] Heather Coiner et al. "Methylation of sulfhydryl groups: a new function for a family of small molecule plant O-methyltransferases." In: *The Plant journal : for cell and molecular biology* 46.2 (Apr. 2006), pp. 193–205.
- [33] Austria Computing, R Foundation for Statistical Vienna. *R: A language and environment for statistical computing*. Vienna, Austria, 2008.
- [34] Robert a. Copeland. "Molecular pathways: Protein methyltransferases in cancer". In: *Clinical Cancer Research* 19.23 (2013), pp. 6344–6350.
- [35] Marjorie Murphy Cowan. "Plant Products as Antimicrobial Agents". In: *Clin. Microbiol. Rev.* 12.4 (1999), pp. 564–582.
- [36] J K Coward, E P Slixz, and F Y Wu. "Kinetic studies on catechol O-methyltransferase. Product inhibition and the nature of the catechol binding site." In: *Biochemistry* 12.12 (1973), pp. 2291–2297.
- [37] Ivana Crnovčić, Roderich Süssmuth, and Ullrich Keller. "Aromatic C-Methyltransferases with Antipodal Stereoselectivity for Structurally Diverse Phenolic Amino Acids Catalyze the Methylation Step in the Biosynthesis of the Actinomycin Chromophore". In: *Biochemistry* 49.45 (Nov. 2010), pp. 9698–9705.

- [38] Rodney Croteau, Toni M Kutchan, and Norman G Lewis. "Secondary Metabolites". In: *Biochemistry Molecular Biology of Plants*. Ed. by B. Buchanan, W. Gruissem, and R. Jones. Vol. 7. 7. 2000, pp. 1250–1318.
- [39] T.P. Tim Cushnie and Andrew J. Lamb. "Antimicrobial activity of flavonoids". In: *International Journal of Antimicrobial Agents* 26.5 (Nov. 2005), pp. 343–356.
- [40] Filip Cuyckens and Magda Claeys. "Mass spectrometry in the structural analysis of flavonoids". In: *Journal of Mass Spectrometry* 39.1 (2004), pp. 1–15.
- [41] Ermias Dagne et al. "O-Geranylated and O-prenylated flavonoids from Millettia ferruginea". In: *Phytochemistry* 29.8 (1990), pp. 2671–2673.
- [42] Yuntao Dai et al. "Natural deep eutectic solvents as new potential media for green technology". In: *Analytica Chimica Acta* 766 (Mar. 2013), pp. 61–68.
- [43] Christian Dalhoff et al. "Direct transfer of extended groups from synthetic cofactors by DNA methyltransferases." In: *Nature chemical biology* 2.1 (2006), pp. 31–32.
- [44] Christian Dalhoff et al. "Synthesis of S-adenosyl-L-methionine analogs and their use for sequence-specific transalkylation of DNA by methyltransferases." In: *Nature protocols* 1.4 (2006), pp. 1879–1886.
- [45] Daresbury Laboratory. "No Title". In: *Newsletter on protein crystallography* 33 (1997), pp. 25–30.
- [46] Bettina E. Deavours et al. "Functional analysis of members of the isoflavone and isoflavanone O-methyltransferase enzyme families from the model legume *Medicago truncatula*". In: *Plant Molecular Biology* 62.4-5 (2006), pp. 715–733.
- [47] Prasanna K. Devaraneni, Jordan J. Devereaux, and Francis I. Valiyaveetil. "In vitro folding of K vAP, a voltage-gated K + channel". In: *Biochemistry* 50.48 (2011), pp. 10442–10450.
- [48] Martin Dippe et al. "Engineering of a Mg<sup>2+</sup>-dependent O-methyltransferase towards novel regiospecificity". In: *manuscript submitted* (2015).
- [49] M. Dippe et al. "Rationally engineered variants of S-adenosylmethionine (SAM) synthase: reduced product inhibition and synthesis of artificial co-factor homologues". en. In: *Chem. Commun.* 51.17 (Feb. 2015), pp. 3637–3640.

- [50] M M Dixon et al. “The structure of the C-terminal domain of methionine synthase: presenting S-adenosylmethionine for reductive methylation of B12.” In: *Structure (London, England : 1993)* 4.11 (1996), pp. 1263–1275.
- [51] J D Dunitz. “The entropic cost of bound water in crystals and biomolecules.” In: *Science (New York, N.Y.)* 264.5159 (Apr. 1994), p. 670.
- [52] J D Dunitz. “Win some, lose some: enthalpy-entropy compensation in weak intermolecular interactions.” In: *Chemistry & biology* 2.11 (Nov. 1995), pp. 709–712.
- [53] Elsevier. *Reaxys, version 2.19790.2*.
- [54] P. Emsley et al. “Features and development of Coot”. In: *Acta Crystallographica Section D: Biological Crystallography* 66.4 (2010), pp. 486–501.
- [55] Carola Engler, Romy Kandzia, and Sylvestre Marillonnet. “A one pot, one step, precision cloning method with high throughput capability”. In: *PLoS ONE* 3.11 (Jan. 2008), e3647.
- [56] Madeleine Ernst et al. “Mass spectrometry in plant metabolomics strategies: from analytical platforms to data acquisition and processing.” en. In: *Natural product reports* 31.6 (June 2014), pp. 784–806.
- [57] Philip Evans. “Scaling and assessment of data quality”. In: *Acta Crystallographica Section D: Biological Crystallography* 62.1 (Jan. 2006), pp. 72–82. arXiv: S0907444905036693 [doi:10.1107].
- [58] Nicolas Fabre et al. “Determination of flavone, flavonol, and flavanone aglycones by negative ion liquid chromatography electrospray ion trap mass spectrometry”. In: *Journal of the American Society for Mass Spectrometry* 12.6 (2001), pp. 707–715.
- [59] J L Ferrer et al. “Structure of chalcone synthase and the molecular basis of plant polyketide biosynthesis.” In: *Nature structural biology* 6.8 (1999), pp. 775–784.
- [60] Alan Fersht. *Structure and mechanism in protein science*. Ed. by Michelle R. Julet and Georgia L. Hadler. New York: W.H. Freeman and Company, 1999.
- [61] H G Floss and M D Tsai. “Chiral methyl groups.” In: *Advances in enzymology and related areas of molecular biology* 50 (Jan. 1979), pp. 243–302.
- [62] Heinz G. Floss and Sungsook Lee. “Chiral Methyl Groups: Small is Beautiful”. In: *Acc. Chem. Res.* 26.3 (1993), pp. 116–122.

- [63] Matthew W. Freyer and Edwin a. Lewis. "Isothermal Titration Calorimetry: Experimental Design, Data Analysis, and Probing Macromolecule/Ligand Binding and Kinetic Interactions". In: *Methods in Cell Biology* 84.07 (Jan. 2008), pp. 79–113.
- [64] Mendel Friedman et al. "Antimicrobial activities of tea catechins and theaflavins and tea extracts against *Bacillus cereus*." In: *Journal of food protection* 69.2 (2006), pp. 354–361.
- [65] Akiko Fujimoto, Atsushi Hirano, and Kentaro Shiraki. "Ternary system of solution additives with Arginine and salt for refolding of beta-galactosidase". In: *Protein Journal* 29.3 (2010), pp. 161–166.
- [66] Gasteiger E. et al. "Protein Identification and Analysis Tools on the ExPASy Server". In: *The Proteomics Protocols Handbook*, Humana Press. Ed. by John M. Walker. Humana Press, 2005, pp. 571–607.
- [67] Paul J. Gates and Norberto P. Lopes. "Characterisation of Flavonoid Aglycones by Negative Ion Chip-Based Nanospray Tandem Mass Spectrometry". In: *International Journal of Analytical Chemistry* 2012 (2012), pp. 1–7.
- [68] Christine L. Gee et al. "Disulfide-linked dimers of human adrenaline synthesizing enzyme PNMT are catalytically active". In: *Biochimica et Biophysica Acta - Proteins and Proteomics* 1750.1 (June 2005), pp. 82–92.
- [69] Azra Gholami et al. "Natural product biosynthesis in *Medicago* species." In: *Natural product reports* 31.3 (2014), pp. 356–80.
- [70] S C Gill and P H von Hippel. "Calculation of protein extinction coefficients from amino acid sequence data." In: *Analytical biochemistry* 182.2 (Nov. 1989), pp. 319–326.
- [71] Paola Gilli et al. "Enthalpy-Entropy Compensation in Drug-Receptor Binding". In: *Journal of Physical Chemistry* 98.5 (Feb. 1994), pp. 1515–1518.
- [72] R M Goodrich and R J Braddock. "Major By-Products of the Florida Citrus Processing". In: *Processing* (2006), pp. 1–4.
- [73] Carmen H. Gray et al. ".alpha.-Deuterium and carbon-13 isotope effects for a simple, intermolecular sulfur-to-oxygen methyl-transfer reaction. Transition-state structures and isotope effects in transmethylation and transalkylation". In: *Journal of the American Chemical Society* 101.15 (1979), pp. 4351–4358.
- [74] Erich Grotewold. *The Science of Flavonoids The Science of Flavonoids*. Ed. by Erich Grotewold. 1st ed. New York: Springer, 2006, pp. 1–274.

- [75] Amandine Guelorget et al. “Insights into the hyperthermostability and unusual region-specificity of archaeal Pyrococcus abyssi tRNA m1A57/58 methyltransferase”. In: *Nucleic Acids Research* 38.18 (2010), pp. 6206–6218.
- [76] R. Hegnauer. *The Flavonoids*. Ed. by J. B. Harborne. 1st ed. Vol. 15. 2. Chapman and Hall, 1976, p. 357.
- [77] J L Hoffman. “Chromatographic analysis of the chiral and covalent instability of S-adenosyl-L-methionine.” In: *Biochemistry* 25.15 (1986), pp. 4444–4449.
- [78] Albert Hofmann. *Die Mutterkornalkaloide*. Solothurn: Nachtschatten Verlag, 2000.
- [79] E.C. Horning et al. “Liquid chromatograph–mass spectrometer–computer analytical systems”. In: *Journal of Chromatography A* 99 (Jan. 1974), pp. 13–21.
- [80] Yonglin Hu et al. “Crystal structures of a *Populus tomentosa* 4-coumarate:CoA ligase shed light on its enzymatic mechanisms.” In: *The Plant cell* 22.9 (2010), pp. 3093–3104.
- [81] Ze Lin Huang et al. “Deep eutectic solvents can be viable enzyme activators and stabilizers”. In: *Journal of Chemical Technology and Biotechnology* October 2013 (2014).
- [82] Ruth Huey et al. “Software news and update a semiempirical free energy force field with charge-based desolvation”. In: *Journal of Computational Chemistry* 28.6 (Apr. 2007), pp. 1145–1152.
- [83] Richard J. Hughes et al. “A tandem mass spectrometric study of selected characteristic flavonoids”. In: *International Journal of Mass Spectrometry* 210–211 (2001), pp. 371–385.
- [84] Sarah C Hunter and Edgar B Cahoon. “Enhancing vitamin E in oilseeds: unraveling tocopherol and tocotrienol biosynthesis.” In: *Lipids* 42.2 (Mar. 2007), pp. 97–108.
- [85] Mwafaq Ibdah et al. “A Novel Mg<sup>2+</sup>-dependent O-Methyltransferase in the Phenylpropanoid Metabolism of *Mesembryanthemum crystallinum*”. In: *Journal of Biological Chemistry* 278.45 (Nov. 2003), pp. 43961–43972.
- [86] Kabirul Islam et al. “Expanding cofactor repertoire of protein lysine methyltransferase for substrate labeling”. In: *ACS Chemical Biology* 6.7 (2011), pp. 679–684.

- [87] J M Jez et al. “Structure and mechanism of the evolutionarily unique plant enzyme chalcone isomerase.” In: *Nature structural biology* 7.9 (2000), pp. 786–791.
- [88] Eun Ji Joe et al. “Engineering of flavonoid O-methyltransferase for a novel regioselectivity.” In: *Molecules and cells* 30.2 (Aug. 2010), pp. 137–41.
- [89] Julien Jorda and Todd O. Yeates. “Widespread disulfide bonding in proteins from thermophilic archaea”. In: *Archaea* 2011 (2011).
- [90] P.D. Josephy, Thomas Eling, and R.P. Mason. “The horseradish peroxidase-catalyzed oxidation of 3, 5, 3', 5'-tetramethylbenzidine. Free radical and charge-transfer complex intermediates.” In: *Journal of Biological Chemistry* 257.7 (1982), pp. 3669–3675.
- [91] Chandrashekhar P. Joshi and Vincent L. Chiang. “Conserved sequence motifs in plant S-adenosyl-L-methionine-dependent methyltransferases”. In: *Plant Molecular Biology* 37.4 (1998), pp. 663–674.
- [92] Wolfgang Kabsch. “Automatic processing of rotation diffraction data from crystals of initially unknown symmetry land cell constants”. In: *Journal of Applied Crystallography* 26.pt 6 (Dec. 1993), pp. 795–800.
- [93] Wolfgang Kabsch. “Integration, scaling, space-group assignment and post-refinement”. In: *Acta Crystallographica Section D: Biological Crystallography* 66.2 (Feb. 2010), pp. 133–144.
- [94] Wolfgang Kabsch. “Xds”. In: *Acta Crystallographica Section D: Biological Crystallography* 66.2 (Feb. 2010), pp. 125–132.
- [95] Jinguo Kang, Larry a Hick, and William E Price. “A fragmentation study of isoflavones in negative electrospray ionization by MS n ion trap mass spectrometry and triple quadrupole mass spectrometry”. In: *Rapid Communications in Mass Spectrometry* 21.6 (2007), pp. 857–868.
- [96] Sun-Young Kang et al. “Biosynthesis of methylated resveratrol analogs through the construction of an artificial biosynthetic pathway in *E. coli*.” In: *BMC biotechnology* 14.1 (2014), p. 67.
- [97] Lawrence a Kelley et al. “The Phyre2 web portal for protein modeling, prediction and analysis.” In: *Nature protocols* 10.6 (June 2015), pp. 845–858.
- [98] Bong Gyu Kim et al. “Multiple regiospecific methylations of a flavonoid by plant O-methyltransferases expressed in *E. coli*”. In: *Biotechnology Letters* 27.23-24 (Dec. 2005), pp. 1861–1864.

- [99] Bong Gyu Kim et al. “Plant Flavonoid O-Methyltransferases: Substrate Specificity and Application”. In: *Journal of Plant Biology* 53.5 (Sept. 2010), pp. 321–329.
- [100] Bong-Gyu Kim, Eun Ji Joe, and Joong-Hoon Ahn. “Molecular characterization of flavonol synthase from poplar and its application to the synthesis of 3-O-methylkaempferol.” In: *Biotechnology letters* 32.4 (Apr. 2010), pp. 579–84.
- [101] Dae Hwan Kim et al. “Regiospecific methylation of naringenin to ponciretin by soybean O-methyltransferase expressed in *Escherichia coli*”. In: *Journal of Biotechnology* 119.2 (Sept. 2005), pp. 155–162.
- [102] Min Ji Kim, Bong Gyu Kim, and Joong Hoon Ahn. “Biosynthesis of bioactive O-methylated flavonoids in *Escherichia coli*”. In: *Applied Microbiology and Biotechnology* 97.16 (2013), pp. 7195–7204.
- [103] Saulius Klimasauskas et al. “Hhal methyltransferase flips its target base out of the DNA helix”. In: *Cell* 76.2 (1994), pp. 357–369.
- [104] K M Koeller and C H Wong. “Enzymes for chemical synthesis.” In: *Nature* 409.6817 (Jan. 2001), pp. 232–240.
- [105] Youichi Kondou et al. “cDNA Libraries”. In: *Methods* 729 (2011), pp. 183–197.
- [106] Kristin König. “Engineering of the Anthocyanin Biosynthetic Pathway in *Nicotiana benthamiana*”. Master. Martin-Luther-Universität Halle-Wittenberg, 2014.
- [107] Frank Koopman et al. “De novo production of the flavonoid naringenin in engineered *Saccharomyces cerevisiae*.” In: *Microbial cell factories* 11 (2012), p. 155.
- [108] Jakub G. Kopycki et al. “Biochemical and Structural Analysis of Substrate Promiscuity in Plant Mg<sup>2+</sup>-Dependent O-Methyltransferases”. In: *Journal of Molecular Biology* 378.1 (Apr. 2008), pp. 154–164.
- [109] Fabian Kuhn et al. “Differentiation of isomeric flavone/isoflavone aglycones by MS2 ion trap mass spectrometry and a double neutral loss of CO”. In: *Rapid Communications in Mass Spectrometry* 17.17 (2003), pp. 1941–1949.
- [110] Shashank Kumar and Abhay K. Pandey. “Chemistry and Biological Activities of Flavonoids: An Overview”. In: *The Scientific World Journal* 2013.7 (2013), pp. 1–16.
- [111] U K Laemmli. “Cleavage of structural proteins during the assembly of the head of bacteriophage T4.” In: *Nature* 227.5259 (1970), pp. 680–685.

- [112] L L Lairson et al. “Glycosyltransferases: structures, functions, and mechanisms.” In: *Annual review of biochemistry* 77 (2008), pp. 521–555.
- [113] Brian J. C. Law et al. “Site-specific bioalkylation of rapamycin by the RapM 16-O-methyltransferase”. In: *Chem. Sci.* (2015), pp. 2885–2892.
- [114] Gunhild Layer et al. “Crystal structure of coproporphyrinogen III oxidase reveals cofactor geometry of Radical SAM enzymes”. In: *EMBO Journal* 22.23 (2003), pp. 6214–6224.
- [115] Bobby W K Lee et al. “Enzyme-catalyzed transfer of a ketone group from an S-adenosylmethionine analogue: A tool for the functional analysis of methyltransferases”. In: *Journal of the American Chemical Society* 132.11 (2010), pp. 3642–3643.
- [116] Jong Suk Lee et al. “Identification of flavonoids using liquid chromatography with electrospray ionization and ion trap tandem mass spectrometry with an MS/MS library”. In: *Rapid Communications in Mass Spectrometry* 19.23 (2005), pp. 3539–3548.
- [117] Effendi Leonard et al. “Expression of a soluble flavone synthase allows the biosynthesis of phytoestrogen derivatives in *Escherichia coli*.” In: *Applied microbiology and biotechnology* 70.1 (Mar. 2006), pp. 85–91.
- [118] Monica Leopoldini et al. “Iron chelation by the powerful antioxidant flavonoid quercetin”. In: *Journal of Agricultural and Food Chemistry* 54.17 (2006), pp. 6343–6351.
- [119] Jakob P. Ley et al. “Evaluation of bitter masking flavanones from Herba Santa (*Eriodictyon californicum* (H. & A.) Torr., Hydrophyllaceae)”. In: *Journal of Agricultural and Food Chemistry* 53.15 (2005), pp. 6061–6066.
- [120] Chen Li et al. “Tandem mass spectrometric fragmentation behavior of lignans, flavonoids and triterpenoids in *< i>Streblus asper</i>*”. In: *Rapid Communications in Mass Spectrometry* 28.21 (2014), pp. 2363–2370.
- [121] Kai Li and J W Frost. “Synthesis of Vanillin from Glucose”. In: 7863.Table 1 (1998), pp. 10545–10546.
- [122] Liwei Li et al. “PDBcal: A comprehensive dataset for receptor-ligand interactions with three-dimensional structures and binding thermodynamics from isothermal titration calorimetry”. In: *Chemical Biology and Drug Design* 71.6 (June 2008), pp. 529–532.
- [123] Manuel Liebeke et al. “Metabolome Analysis of Gram-Positive Bacteria such as *Staphylococcus aureus* by GC-MS and LC-MS”. In: *Methods in Molecular Biology* 815 (Jan. 2012), pp. 377–398.

- [124] Gordon V. Louie et al. "Structural Determinants and Modulation of Substrate Specificity in Phenylalanine-Tyrosine Ammonia-Lyases". In: *Chemistry and Biology* 13.12 (2006), pp. 1327–1338.
- [125] Caroline Louis-Jeune, Miguel a. Andrade-Navarro, and Carol Perez-Iratxeta. "Prediction of protein secondary structure from circular dichroism using theoretically derived spectra". In: *Proteins: Structure, Function and Bioinformatics* 80.2 (Feb. 2012), pp. 374–381.
- [126] J D Lozada-Ramírez et al. "Cloning, overexpression, purification, and characterization of S-adenosylhomocysteine hydrolase from *Corynebacterium efficiens* YS-314." In: *Biotechnology progress* 24.1 (2008), pp. 120–7.
- [127] Liang Lu et al. "A High-Resolution LC-MS-Based Secondary Metabolite Fingerprint Database of Marine Bacteria". en. In: *Scientific Reports* 4 (Jan. 2014), p. 6537.
- [128] Gražvydas Lukinavičius et al. "Enhanced chemical stability of AdoMet analogues for improved methyltransferase-directed labeling of DNA". In: *ACS Chemical Biology* 8.6 (2013), pp. 1134–1139.
- [129] Y. L. Ma et al. "Characterization of flavone and flavonol aglycones by collision-induced dissociation tandem mass spectrometry". In: *Rapid Communications in Mass Spectrometry* 11.12 (1997), pp. 1357–1364.
- [130] Tom J. Mabry, K. R. Markham, and M. B. Thomas. *The Systematic Identification of Flavonoids*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1970.
- [131] Savvas C Makrides and Savvas C Makrides. "Strategies for Achieving High-Level Expression of Genes in *Escherichia coli*". In: *Microbiological reviews* 60.3 (1996), pp. 512–538.
- [132] Sailesh Malla et al. "Production of 7-O-methyl aromadendrin, a medicinally valuable flavonoid, in *Escherichia coli*." In: *Applied and environmental microbiology* 78.3 (Feb. 2012), pp. 684–94.
- [133] S. V. Mani, D. W. Connell, and R. D. Braddock. "Structure activity relationships for the prediction of biodegradability of environmental pollutants". en. In: *Critical Reviews in Environmental Control* 21.3-4 (Jan. 1991), pp. 217–236.
- [134] Raymond E. March and Xiu Sheng Miao. "A fragmentation study of kaempferol using electrospray quadrupole time-of-flight mass spectrometry at high mass resolution". In: *International Journal of Mass Spectrometry* 231.2-3 (2004), pp. 157–167.

- [135] Raymond March and Jennifer Brodbelt. "Analysis of flavonoids: Tandem mass spectrometry, computational methods, and NMR". In: *Journal of Mass Spectrometry* 43.12 (Dec. 2008), pp. 1581–1617.
- [136] Nektaria Markoglou and Irving W Wainer. "Biosynthesis in an online immobilized-enzyme reactor containing phenylethanolamine N-methyltransferase in single-enzyme and coupled-enzyme formats." In: *Journal of chromatography. A* 948.1-2 (Mar. 2002), pp. 249–56.
- [137] Alexandre Martinière et al. "Development and properties of genetically encoded pH sensors in plants." In: *Frontiers in plant science* 4.December (2013), p. 523.
- [138] M McClelland, M Nelson, and E Raschke. "Effect of site-specific modification on restriction endonucleases and DNA modification methyltransferases." In: *Nucleic acids research* 22.17 (1994), pp. 3640–3659.
- [139] Airlie J. McCoy. "Solving structures of protein complexes by molecular replacement with Phaser". In: *Acta Crystallographica Section D: Biological Crystallography* 63.1 (Jan. 2006), pp. 32–41.
- [140] Airlie J. McCoy et al. "Phaser crystallographic software". In: *Journal of Applied Crystallography* 40.4 (Aug. 2007), pp. 658–674.
- [141] Edoardo Mentasti, Ezio Pelizzetti, and Guido Saini. "Reactions between iron(III) and catechol (o-dihydroxybenzene). Part II. Equilibria and kinetics of the redox reaction in aqueous acid solution". en. In: *Journal of the Chemical Society, Dalton Transactions* 23 (Jan. 1973), p. 2609.
- [142] Diana Metodiewa et al. "Quercetin may act as a cytotoxic prooxidant after its metabolic activation to semiquinone and quinoidal product". In: *Free Radical Biology and Medicine* 26.1-2 (Jan. 1999), pp. 107–116.
- [143] Gurvan Michel et al. "The structure of the RlmB 23S rRNA methyltransferase reveals a new methyltransferase fold with a unique knot". In: *Structure* 10.10 (2002), pp. 1303–1315.
- [144] Ivan Mihel et al. ".alpha.-Deuterium isotope effects and transition-state structure in an intramolecular model system for methyl-transfer enzymes". In: *Journal of the American Chemical Society* 101.15 (July 1979), pp. 4349–4351.
- [145] Douglas Montgomery. *Design and Analysis of Experiments*. 5th ed. New York: John Wiley & Sons, 2008, p. 680.

- [146] Takashi MORISHIGE, Kum-Boo CHOI, and Fumihiko SATO. “In Vivo Bioconversion of Tetrahydroisoquinoline by Recombinant Coclaurine N -Methyltransferase”. In: *Bioscience, Biotechnology and Biochemistry* 68.4 (May 2014), pp. 939–941.
- [147] Garrett M. Morris et al. “Software news and updates AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility”. In: *Journal of Computational Chemistry* 30.16 (Dec. 2009), pp. 2785–2791.
- [148] Yuri Motorin et al. “Expanding the chemical scope of RNA:methyltransferases to site-specific alkynylation of RNA for click labeling”. In: *Nucleic Acids Research* 39.5 (2011), pp. 1943–1952.
- [149] N. Mulinacci et al. “Polyphenolic content in olive oil waste waters and related olive samples”. In: *Journal of Agricultural and Food Chemistry* 49.8 (2001), pp. 3509–3514.
- [150] Garib N. Murshudov, Alexei a. Vagin, and Eleanor J. Dodson. “Refinement of macromolecular structures by the maximum-likelihood method”. In: *Acta Crystallographica Section D: Biological Crystallography* 53.3 (May 1997), pp. 240–255.
- [151] Janet Newman. “Novel buffer systems for macromolecular crystallization”. In: *Acta Crystallographica Section D: Biological Crystallography* 60.3 (2004), pp. 610–612.
- [152] Novagen. *pET System Manual*. 11th ed. Vol. 123. 2 Pt 1. Darmstadt: EMD Chemicals, 2014, p. 369. arXiv: 1007.2303v3.
- [153] Zouhaier B Ouallagui et al. “Valorization of Olive Processing By-Products : Characterization , Investigation of Chemico-Biological Activities and Identification of Active Compounds”. In: 64 (2012), pp. 61–64.
- [154] Javier Palazón et al. “Application of metabolic engineering to the production of scopolamine.” In: *Molecules (Basel, Switzerland)* 13.8 (Jan. 2008), pp. 1722–42.
- [155] Ira Palmer and Paul T. Wingfield. “Preparation and extraction of insoluble (Inclusion-body) proteins from Escherichia coli”. In: *Current Protocols in Protein Science* 1.SUPPL.70 (Nov. 2012), Unit6.3.
- [156] Wibke Peters et al. “Enzymatic site-specific functionalization of protein methyltransferase substrates with alkynes for click labeling”. In: *Angewandte Chemie - International Edition* 49.30 (2010), pp. 5170–5173.

- [157] Harold R. Powell. "The Rossmann Fourier autoindexing algorithm in MOSFLM". In: *Acta Crystallographica Section D: Biological Crystallography* 55.10 (1999), pp. 1690–1695.
- [158] O. Rigbers and S.-M. Li. "Ergot Alkaloid Biosynthesis in Aspergillus fumigatus: OVERPRODUCTION AND BIOCHEMICAL CHARACTERIZATION OF A 4-DIMETHYLALLYLTRYPTOPHAN N-METHYLTRANSFERASE". In: *Journal of Biological Chemistry* 283.40 (Aug. 2008), pp. 26859–26868.
- [159] K D Robertson. "DNA methylation, methyltransferases, and cancer." In: *Oncogene* 20.24 (2001), pp. 3139–3155.
- [160] Keith D Robertson. "DNA methylation and human disease." In: *Nature reviews. Genetics* 6.8 (2005), pp. 597–610.
- [161] M. G. Rossmann and D. M. Blow. "The detection of sub-units within the crystallographic asymmetric unit". In: *Acta Crystallographica* 15.1 (Jan. 1962), pp. 24–31.
- [162] Michael G. Rossmann. "Molecular replacement - Historical background". In: *Acta Crystallographica - Section D Biological Crystallography* 57.10 (Sept. 2001), pp. 1360–1366.
- [163] J Sambrook and D W Russell. *Molecular Cloning: A Laboratory Manual*. 3rd ed. Cold Spring Harbor (NY, USA): Cold Spring Harbor Laboratory Press, 2001.
- [164] Cleydson Breno R Santos et al. "A SAR and QSAR study of new artemisinin compounds with antimalarial activity". In: *Molecules* 19.1 (2014), pp. 367–399.
- [165] Yuji Sawada and Masami Yokota Hirai. *Integrated LC-MS / MS system for plant metabolomics*. en. May 2013.
- [166] Artur Sawicki and Robert D Willows. "S-adenosyl-L-methionine:magnesium-protoporphyrin IX O-methyltransferase from Rhodobacter capsulatus: mechanistic insights and stimulation with phospholipids." In: *The Biochemical journal* 406.3 (2007), pp. 469–478.
- [167] Jason W. Schmidberger et al. "Halomethane biosynthesis: Structure of a SAM-dependent halide methyltransferase from arabidopsis thaliana". In: *Angewandte Chemie - International Edition* 49.21 (2010), pp. 3646–3648.
- [168] Gudrun Schröder et al. "Flavonoid methylation: A novel 4'-O-methyltransferase from Catharanthus roseus, and evidence that partially methylated flavanones are substrates of four different flavonoid dioxygenases". In: *Phytochemistry* 65.8 (2004), pp. 1085–1094.

- [169] H L Schubert et al. “The X-ray structure of a cobalamin biosynthetic enzyme, cobalt-precorrin-4 methyltransferase.” In: *Nature structural biology* 5.7 (1998), pp. 585–592.
- [170] Heidi L. Schubert, Robert M. Blumenthal, and Xiaodong Cheng. “Many paths to methyltransfer: A chronicle of convergence”. In: *Trends in Biochemical Sciences* 28.6 (2003), pp. 329–335.
- [171] Daniela Schulz, Josephin Marie Holstein, and Andrea Rentmeister. “A chemo-enzymatic approach for site-specific modification of the RNA cap”. In: *Angewandte Chemie - International Edition* 52.30 (2013), pp. 7874–7878.
- [172] N. Schweigert, a. J B Zehnder, and R. I L Eggen. “Chemical properties of catechols and their molecular modes of toxic action in cells, from microorganisms to mammals”. In: *Environmental Microbiology* 3.2 (2001), pp. 81–91.
- [173] N. Schweigert, a. J B Zehnder, and R. I L Eggen. “Chemical properties of catechols and their molecular modes of toxic action in cells, from microorganisms to mammals”. In: *Environmental Microbiology* 3.2 (2001), pp. 81–91.
- [174] Mark S. Searle, Martin S. Westwell, and Dudley H. Williams. “Application of a generalised enthalpy-entropy relationship to binding co-operativity and weak associations in solution”. en. In: *Journal of the Chemical Society, Perkin Transactions 2* 1 (Jan. 1995), p. 141.
- [175] S K Shapiro and D J Ehninger. “Methods for the analysis and preparation of adenosylmethionine and adenosylhomocysteine.” In: *Analytical biochemistry* 15.2 (May 1966), pp. 323–333.
- [176] Y. Q. Shi and R. R. Rando. “Kinetic mechanism of isoprenylated protein methyltransferase”. In: *Journal of Biological Chemistry* 267.14 (1992), pp. 9547–9551.
- [177] Alexander Ann Shulgin. *Tihkal: The Continuation*. Ed. by Dan Joy. 1st ed. Berkeley: Transform Press, 1997, p. 804.
- [178] Sigma-Aldrich. *Technical Bulletin no. 2003-03: freezing of microbial samples prior to testing*. Parenteral Drug Association. 2003.
- [179] Simon A. Simms and Kotha Subbaramaiah. “Kinetic Mechanism of the CheR Methyltransferase”. In: *Journal of Biological Chemistry* 266.19 (1991), pp. 12741–12746.

- [180] Henrik T. Simonsen et al. "Methylenedioxy- and methoxyflavones from Melicope coodeana syn. Euodia simplex". In: *Phytochemistry* 60.8 (Aug. 2002), pp. 817–820.
- [181] Shanteri Singh et al. "Facile chemoenzymatic strategies for the synthesis and utilization of S-adenosyl-L-methionine analogues". In: *Angewandte Chemie - International Edition* 53.15 (2014), pp. 3965–3969.
- [182] Lekha Sleno and Dietrich a. Volmer. "Ion activation methods for tandem mass spectrometry". In: *Journal of Mass Spectrometry* 39.10 (2004), pp. 1091–1112.
- [183] Harald Stecher et al. "Biocatalytic Friedel-Crafts alkylation using non-natural cofactors". In: *Angewandte Chemie - International Edition* 48.50 (Jan. 2009), pp. 9546–9548.
- [184] Anna Winona Struck et al. "S-Adenosyl-Methionine-Dependent Methyltransferases: Highly Versatile Enzymes in Biocatalysis, Biosynthesis and Other Biotechnological Applications". In: *ChemBioChem* 13.18 (Nov. 2012), pp. 2642–2655.
- [185] F William Studier. "Protein production by auto-induction in high density shaking cultures." In: *Protein expression and purification* 41.1 (May 2005), pp. 207–234. arXiv: *NIHMS150003*.
- [186] Fusao Takusagawa et al. "Crystal Structure of S-Adenosylmethionine Synthetase \*\*". In: 271.1 (1996), pp. 136–147.
- [187] Loverine P Taylor and Erich Grotewold. "Flavonoids as developmental regulators." In: *Current opinion in plant biology* 8.3 (June 2005), pp. 317–23.
- [188] Martin Tengg et al. "Molecular characterization of the C-methyltransferase NovO of Streptomyces sphaeroides, a valuable enzyme for performing Friedel-Crafts alkylation". In: *Journal of Molecular Catalysis B: Enzymatic* 84 (Dec. 2012), pp. 2–8.
- [189] David J Thomas et al. "Arsenic (+3 Oxidation State) Methyltransferase and the Methylation of Arsenicals". In: *Experimental biology and medicine (Maywood, N.J.)* 232.1 (Jan. 2007), pp. 3–13.
- [190] Marie Thomsen et al. "Chemoenzymatic synthesis and in situ application of S-adenosyl-L-methionine analogs." In: *Organic & biomolecular chemistry* 11.43 (2013), pp. 7606–10.
- [191] Robert Tibshirani. *Regression Selection and Shrinkage via the Lasso*. 1994.

- [192] Dominique a Tobbell et al. “Identification of in vitro folding conditions for procathepsin S and cathepsin S using fractional factorial screens.” In: *Protein expression and purification* 24.2 (Mar. 2002), pp. 242–254.
- [193] Jaime Torres-Corzo, Roberto Rodríguez-Della Vecchia, and Leonardo Rangel-Castilla. *Observation of the ventricular system and subarachnoid space in the skull base by flexible neuroendoscopy: normal structures*. 1st ed. Vol. 141. 2. New York: Garland Science, 2009, pp. 165–168.
- [194] Emmanouil a. Trantafyllidis et al. “When plants produce not enough or at all: metabolic engineering of flavonoids in microbial hosts”. In: *Frontiers in Plant Science* 6.January (2015), pp. 1–16.
- [195] Oleg Trott and Arthur J. Olson. “Software news and update AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading”. In: *Journal of Computational Chemistry* 31.2 (Jan. 2010), pp. 455–461.
- [196] Kouhei Tsumoto et al. “Role of arginine in protein refolding, solubilization, and purification.” In: *Biotechnology progress* 20.5 (2004), pp. 1301–1308.
- [197] P. Urban et al. “Characterization of recombinant plant cinnamate 4-hydroxylase produced in yeast - Kinetic and spectral properties of the major plant P450 of the phenylpropanoid pathway”. In: *European Journal of Biochemistry* 222.3 (1994), pp. 843–850.
- [198] Alexei a. Vagin et al. “REFMAC5 dictionary: Organization of prior chemical knowledge and guidelines for its use”. en. In: *Acta Crystallographica Section D: Biological Crystallography* 60.12 I (Nov. 2004), pp. 2184–2195.
- [199] Jacob L. Venzie et al. “Electron-impact and glow-discharge ionization LC-MS analysis of green tea tincture”. In: *Analytical and Bioanalytical Chemistry* 387.1 (2007), pp. 321–333.
- [200] J Vidgren, L a Svensson, and a Liljas. “Crystal structure of catechol O-methyltransferase.” In: *Nature* 368.6469 (1994), pp. 354–358.
- [201] Renaud Vincentelli et al. “High-throughput automated refolding screening of inclusion bodies.” In: *Protein science : a publication of the Protein Society* 13.10 (2004), pp. 2782–2792.
- [202] Thomas Vogt. “Regiospecificity and kinetic properties of a plant natural product O-methyltransferase are determined by its N-terminal domain”. In: *FEBS Letters* 561.1-3 (Mar. 2004), pp. 159–162.

- [203] Thomas Vogt. “Regiospecificity and kinetic properties of a plant natural product O-methyltransferase are determined by its N-terminal domain”. In: *FEBS Letters* 561.1-3 (Mar. 2004), pp. 159–162.
- [204] Rui Wang et al. “Formulating a fluorogenic assay to evaluate S-adenosyl-L-methionine analogues as protein methyltransferase cofactors”. In: *Molecular BioSystems* 7.11 (2011), p. 2970.
- [205] Rui Wang et al. “Labeling substrates of protein arginine methyltransferase with engineered enzymes and matched S-adenosyl-l-methionine analogues”. In: *Journal of the American Chemical Society* 133.20 (2011), pp. 7648–7651.
- [206] Williard J. Werner et al. “In vitro phosphinate methylation by PhpK from *Kitasatospora phosalacinea*”. In: *Biochemistry* 50.42 (2011), pp. 8986–8988.
- [207] Ludger a. Wessjohann et al. “Alkylating enzymes”. In: *Current Opinion in Chemical Biology* 17.2 (Mar. 2013), pp. 229–235.
- [208] C M Whitehouse et al. “Electrospray interface for liquid chromatographs and mass spectrometers.” In: *Analytical chemistry* 57.3 (Mar. 1985), pp. 675–679.
- [209] Christine a Williams and Renee J Grayer. “Anthocyanins and other flavonoids.” In: *Natural product reports* 21.4 (2004), pp. 539–573.
- [210] Melissa Swope Willis et al. “Investigation of protein refolding using a fractional factorial screen: a study of reagent effects and interactions.” In: *Protein science : a publication of the Protein Society* 14.7 (2005), pp. 1818–1826.
- [211] Sophie Willnow et al. “A Selenium-Based Click AdoMet Analogue for Versatile Substrate Labeling with Wild-Type Protein Methyltransferases”. In: *ChemBioChem* 13.8 (2012), pp. 1167–1173.
- [212] Martyn D. Winn et al. “Overview of the CCP4 suite and current developments”. In: *Acta Crystallographica Section D: Biological Crystallography* 67.4 (Apr. 2011), pp. 235–242.
- [213] Jaclyn M. Winter et al. “Expanding the structural diversity of polyketides by exploring the cofactor tolerance of an inline methyltransferase domain”. In: *Organic Letters* 15.14 (2013), pp. 3774–3777.
- [214] Tomasz Włodarski et al. “Comprehensive structural and substrate specificity classification of the *Saccharomyces cerevisiae* methyltransferome”. In: *PLoS ONE* 6.8 (2011).
- [215] Sebastian Wolf et al. “In silico fragmentation for computer assisted identification of metabolite mass spectra.” In: *BMC Bioinformatics* 11 (2010), p. 148.

- [216] J.-L. Wolfender et al. “Evaluation of Q-TOF-MS/MS and multiple stage IT-MSn for the Dereplication of Flavonoids and Related Compounds in Crude Plant Extracts”. In: *Analisis* 28.10 (Dec. 2000), pp. 895–906.
- [217] Eckhard Wollenweber. “On the occurrence of acylated flavonoid aglycones”. In: *Phytochemistry* 24.7 (1985), pp. 1493–1494.
- [218] R W Woodard et al. “Stereochemical Course of the Transmethylation Catalyzed by Catechol O-Methyltransferase”. In: *The Journal of Biological Chemistry* 255.19 (1980), pp. 9124–9127.
- [219] Ronald W Woodard et al. “Stereochemistry of Indolmycin Biosynthesis. Steric Course of C- and N-Methylation Reactionsa”. In: *Journal of the American Chemical Society* 102.20 (1980), pp. 6314–6318.
- [220] B. Xiao et al. “Structure and catalytic mechanism of the human histone methyltransferase SET7/9”. In: 421.February (2003).
- [221] Hiroshi Yamaguchi and Masaya Miyazaki. “Refolding techniques for recovering biologically active recombinant proteins from inclusion bodies.” In: *Biomolecules* 4.1 (2014), pp. 235–51.
- [222] William C Yee, Susan J Eglsær, and William R Richards. “Confirmation of a ping-pong mechanism for S-adenosyl-L-methionine:magnesium protoporphyrin methyltransferase of etiolated wheat by an exchange reaction.” In: *Biochemical and biophysical research communications* 162.1 (1989), pp. 483–490.
- [223] Yoon Youngdae et al. “Characterization of an O-methyltransferase from Streptomyces avermitilis MA-4680”. In: *Journal of Microbiology and Biotechnology* 20.9 (2010), pp. 1359–1366.
- [224] Changsheng Zhang et al. “Natural product diversification using a non-natural cofactor analogue of S-adenosyl-L-methionine”. In: *Journal of the American Chemical Society* 128.9 (2006), pp. 2760–2761.
- [225] Qi Zhang and Wilfred A van der Donk. “Catalytic promiscuity of a bacterial  $\alpha$ -N-methyltransferase.” In: *FEBS letters* 586.19 (Sept. 2012), pp. 3391–7.
- [226] Qi Zhang, Wilfred A. van der Donk, and Wen Liu. “Radical-Mediated Enzymatic Methylation: A Tale of Two SAMS”. In: *Accounts of Chemical Research* 45.4 (Apr. 2012), pp. 555–564.
- [227] Xing-Hai Zhang and C. C. Chinnappa. “Molecular characterization of a cDNA encoding caffeoyl-coenzyme A 3-O-methyltransferase of *Stellaria longipes*”. In: *Journal of Biosciences* 22.2 (Mar. 1997), pp. 161–175.

- [228] Jian Min Zhou et al. "Sequential O-methylation of tricetin by a single gene product in wheat". In: *Biochimica et Biophysica Acta - General Subjects* 1760.7 (2006), pp. 1115–1124.
- [229] Saijie Zhu et al. "Efficient synthesis of eriodictyol from l-tyrosine in escherichia coli". In: *Applied and Environmental Microbiology* 80.10 (2014), pp. 3072–3080.
- [230] Chloe Zubieta et al. "Structural Basis for the Modulation of Lignin Monomer Methylation by Caffeic Acid / 5-Hydroxyferulic Acid 3 / 5- O -Methyltransferase". In: *The Plant Cell* 14.June (2002), pp. 1265–1277.
- [231] C Zubieta et al. "Structures of two natural product methyltransferases reveal the basis for substrate specificity in plant O-methyltransferases." In: *Nature structural biology* 8.3 (Mar. 2001), pp. 271–279.

## **Acronyms**

**Å** Ångström, 0.1 nm

**3O4M** 3'-hydroxy-4'-methoxy

**4CL** 4-coumarate:CoA ligase

**4O3M** 4'-hydroxy-3'-methoxy

**ABPP** activity based protein profiling

**AC-9** anthracene-9-carboxylic acid

**AI** auto-induction 94, *see* ZYP-5052

**ANOVA** Analysis of Variance

**APCI** atmospheric pressure chemical ionisation

**ATP** adenosine triphosphate

**AUC** area under the curve

**BisTris** 2-[Bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)propane-1,3-diol

**B-PER** bacterial protein extraction reagent

**C4H** cinnamate-4-hydroxylase

**CBD** cobalamin binding domain

**CCoAOMT** caffeoyl CoA dependent O-methyltransferase

**CCP4** Collaborative Computational Project No. 4

**CD** circulary dichroism

**CHI** chalcone isomerase

**CHS** chalcone synthase

**CID** collsion induced dissociation

**C-MT** C-methyl transferase

**CoA** coenzyme A

**COMT** catechol O-methyl transferase

**Coot** Crystallographic Object-Oriented Toolkit

**CV** column volumes

**dAdo** 5'-deoxyadenosyl

**DMSO** dimethyl sulfoxide

**DNA** desoxyribonucleic acid

**DNA-MT** DNA methyl transferase

**DoE** design of experiments

**DTT** dithiothreitol; (2*S*,3*S*)-1,4-bis(sulfanyl)butane-2,3-diol

**EDTA** ethylenediaminetetraacetic acid

**EEC** enthalpy-entropy compensation

**EI** electron impact

**ESI** electrospray ionization

**F3H** flavanone-3-hydroxylase

**FNS** flavone synthase

**FPLC** fast protein liquid chromatography

**FrFD** fractional factorial design

**FT** Fourier transformation

**FTMS** fourier transform mass spectrometry

**GdmCl** guanidinium hydrochloride

**GFP** green fluorescent protein

**GOD** glucose oxidase

**GSH** glutathione,  $\gamma$ -L-glutamyl-L-cysteinylglycine

**GSSG** glutathione disulfide

**GST** Glutathion S-transferase

**HCD** higher-energy collisional dissociation

**HEPES** 2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid

**H-ESI** heated-electrospray ionization

**HIC** hydrophobic interaction chromatography

**HPLC** high-performance liquid chromatography

**HRP** horseradish peroxidase

**IAE-plot** interaction effects plot

**IB** inclusion body

**IEX** ion exchange chromatography

**IFS** isoflavone synthase

**IMAC** immobilized metal affinity chromatography

**IPB** Leibniz-Institute of Plant Biochemistry

**IPTG** isopropyl-D-thiogalactopyranosid

**ITC** Isothermal Titration Calorimetry

**LB** lysogeny broth

**LC** liquid chromatography

**LC/MS** liquid chromatography coupled mass-spectrometry

**LC-MS/MS** liquid chromatography-tandem mass spectrometry

**m/z** mass-to-charge ratio

**ME-plot** main effects plot

**MES** 2-(*N*-morpholino)ethanesulfonic acid

**MLU** Martin-Luther-Universität

**MMT** L-malic acid/MES/Tris

**MR** molecular replacement

**MS/MS** tandem mass-spectrometry

**MT** methyl transferase

**MTP** micro-titer plate

**MW** molecular weight

**MWCO** molecular weight cut-off

**NADES** natural deep eutectic solvent

**NCE** normalized collision energy

**N-MT** *N*-methyl transferase

**nos** nopaline synthase

**NPS** nitrogen, phosphate, sulfate buffer

**NRPS** non-ribosomal peptide synthase

**NTA** nitrilo triacetic acid

**O-MT** *O*-methyl transferase

**PAGE** polyacrylamide gel electrophoresis

**PAL** phenylalanine ammonia-lyase

**PBS** phosphate buffered saline

**PCA** principal component analysis

- PCH** propane-1,2-diol/choline chloride,natural deep eutectic solvent (NADES)-mixture
- PCR** polymerase chain reaction
- PDA** photo diode array
- PDB** Protein Data Base 46, 47
- PFOMT** phenylpropanoid and flavonoid O-methyl transferase
- PHENIX** Phyton-based Hierachial Environment for Integrated Xtallography
- PKS** poly ketide synthase
- PMSF** phenylmethylsulfonylfluoride
- P-MT** protein methyl transferase
- QSAR** quantitative structure activity relationship
- rmsd** root mean squared deviation
- RNA** ribonucleic acid
- RNA-MT** RNA methyl transferase
- ROS** reactive oxygen species
- RP** resolving power
- rRNA** ribosomal ribonucleic acid
- RSMT** radical SAM methyl transferase
- RT** room temperature
- SAE** *S*-adenosyl-L-ethionine, (2*S*)-2-amino-4-[[*(2S,3S,4R,5R)*-5-(6-aminopurin-9-yl)-3,4-dihydroxyoxolan-2-yl]methyl-ethylsulfonio]butanoat
- SAH** *S*-adenosyl-L-homocysteine
- SAM** *S*-adenosyl-L-methionine
- SAMS** *S*-adenosylmethionine synthase
- SAR** structure activity relationship
- SDS** sodium dodecylsulfate
- SeAM** Se-adenosyl selenomethionine
- SET** suvar3-9, enhancer-of-zeste, trithorax
- SID** surface-induced dissociation
- smMT** small molecule methyl transferase
- S-MT** *S*-methyl transferase

**SOMT-2** soy O-methyl transferase

**SPOUT** *SpoU-TrmD*

**SSG** succinate/sodium phosphate/glycine

**TB** terrific broth

**TCA** trichloro acetic acid

**Ti-plasmid** tumor inducing plasmid

**Tris** tris(hydroxymethyl)-aminomethane

**U** enzyme unit; measure for enzymatic activity (1 U = 1 µmole/min = 1/60 µkat)

**UDP** uridine diphosphate

**UHPLC** ultra-high performance liquid chromatography

**UV** ultra violet

**UV/VIS** ultra violet/visible (light spectrum)

**V** volume

**ZYP** N-Z-amine, yeast extract, phosphate 38, 166, 168, *see ZYP-5052*

# Glossary

**GOD** Glucose oxidase is an enzyme....

**His<sub>6</sub>-tag** Hexa-histidine tag commonly used for recombinant protein production.

**Isothermal Titration Calorimetry (ITC)** Fill in description here

**MTP** Micro-titer plate. Small format rectangular plastic plate containing wells to allow for storage of multiple small samples or the containment multiple simultaneous reactions. Typical sizes include 24, 96 and 384-wells

**PFOMT** Phenylpropanoid and flavonoid O-methyl transferase from *Mesembryanthemum crystallinum*, which was first described by Ibdah et al. in 2003 [85]

**T7-tag** Initial 11 amino acids of the T7 gene 10 protein.

**Ti-plasmid** Commonly found plasmids in *A. tumefaciens* and *A. rhizogenes* that confer virulence

**Trx-tag** Thioredoxin tag used to increase solubility and stability of recombinantly expressed proteins.

**ZYP-5052** Autoinduction medium developed by Studier [185]. The naming stems from the components N-Z-amine, yeast extract and phosphate. The numbering designates the composition; e.g. 5052 refers to 0.5 % glycerol, 0.05 % glucose and 0.2 % lactose. 166