Biotransformations from and to methylated flavonoids

Dissertation

zur Erlangung des Doktorgrades der Naturwissenschaften
(Dr. rer. nat)

der

Naturwissenschaftlichen Fakultät II
Chemie, Physik und Mathematik

der Martin-Luther-Universität
Halle-Wittenberg

vorgelegt von

Herrn Diplom-Biochemiker Benjamin Weigel
geb. am 06. April 1986 in Bernburg (Saale)

Halle (Saale), May 3, 2017

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

– *unknown* –

erster Gutachter: Prof. Dr. Ludger A. Wessjohann
zweiter Gutachter: Prof. Dr. Andreas Liese
Tag der öffentlichen Verteidigung: 06.04.2017

Acknowledgements

> The true delight is in the finding out rather than in the knowing.
>
> – *Isaac Asimov* –

First I would like to thank Professor Ludger Wessjohann for giving me the opportunity to work on this challenging problem and become a better scientist, and also for his steady professional support throughout the years.

Thanks go to Annegret Laub, Dr. Jürgen Schmidt, Anja Ehrlich, Christine Kuhnt, Dr. Andrea Porzel, Martina Lerbs and Gudrun Hahn for all their help with analytics measurements and problems. I am grateful to all the people in the IPB and especially the group of Bioorganic Chemistry for a great and always helpful professional and personal atmosphere. Many thanks go to Dr. Danilo Meyer and Dr. Martin Dippe, who were an enrichment to the biocatalysis lab and always provided helpful ideas and discussions.

I would like to thank Dr. Thomas Vogt for his advice. Thanks to Dr. Sylvestre Marillonet and Ramona Grützner for theoretical and practical help with the Golden Gate system and plant expression.

I thank Professor Milton Stubbs for the opportunity to run crystallization and diffraction experiments in his group. Here, a special thanks to Dr. Christoph Parthier for his practical and theoretical help concerning the solution of crystal structures.

I thank Dr. Marco Kloos for the *black-ops* style screening experiments during a long night in Leipzig.

To the people of “IPB Aktuell“ – Steve Ludwig, Robert Berger, Anne-Katrin Bauer, Jeanette Keim, Sebastian Stark, as well as Rainer and Julia Kufka – a very special thank you. Your are like a family <3. Thank you Anne for being an awesome roomie (that

was not annoyed by my messiness) throughout all the years. Thanks Robert for all the time we spent watching TV-series, and the personal, professional and R related conversations. Thanks Steve for the longest consecutive foosball match of (probably) all time … and your mother. Thanks Jeanette for getting me set up when I first started and all the great help, especially in moving. Thanks for the effort we shared writing that manuscript. Thanks Sebastian for the Pflaumenknödel and scientific sceptisism. Thanks Rainer and Julia for always being so positive (when I see you in my head I always see you grinning).

A big shoutout goes to the ThursdayNightBoardGamingAndWhiskyTasting (TNBGAWT)-crew – René Wartner <3, Christian Arlt, Marcus Böhme, sometimes others and „Hexeee! “. I always had a blast and appreciate that we kept it going for all these years. I dearly miss it.

Thanks to the guys and girls, who have attended every biochemistry retreat – Bea, Susi, René, Christian and Marcus. This weekend is always welcome and a great way to meet up again. I hope it keeps on going till the end.

I am truly grateful to my surrogate father, Melvin Allen Vermillion. Thanks for always keeping an open door. You will forever be in my heart.

I thank my family for all their love and support. My grandparents Ingrid and Siegfried, who will always have us over for lunch and prepare the the meal that we so crave. My grandpa Siegfried, whose mindset and ethical views I deeply admire. I thank my parents Heike and René for all they have done in making dreams become reality and their dependability. My brother, Konstantin, for his ability and willingness to get things organized, even when it is high time. To Oma Lilo, I miss you.

I thank my wife Maika, for keeping up and sticking with me over the years. Thank you for condoning my flaws and making me a better person. I am grateful to have you in my life – forever and always.

Contents

Acknowledgements iii

Contents v

List of Figures ix

List of Tables xvii

List of Schemes xxi

I Preface 1

1 Abstracts 3
- 1.1 English Abstract . . . 3
- 1.2 Deutsche Zusammenfassung . . . 4

2 List of Publications 7

II Thesis 9

3 Introduction and Motivation 11
- 3.1 Flavonoids . . . 11
  - 3.1.1 Overview . . . 11
  - 3.1.2 The phenyl propanoid pathway . . . 13
  - 3.1.3 Biological activity . . . 13
- 3.2 Methyl transferases (MTs) . . . 15
  - 3.2.1 Overview . . . 15
  - 3.2.2 *S*-Adenosyl-ʟ-methionine . . . 18
  - 3.2.3 Methyl transferase mechanisms . . . 19

3.2.4 Plant *O*-methyl transferases (*O*-MTs) . . . 22
3.3 Alkylations and biotransformations . . . 23
3.3.1 Overview . . . 23
3.3.2 Methyl transferases for industrial use . . . 24
3.3.3 Artificial SAM analogues . . . 26
3.4 Motivation . . . 27

4 Material And Methods 29
4.1 Materials . . . 29
4.1.1 Chemicals . . . 29
4.1.2 Commonly used solutions and buffers . . . 29
4.1.3 Culture media used to grow bacteria . . . 30
4.1.4 Bacterial strains . . . 31
4.1.5 Plasmids . . . 32
4.1.6 Oligonucleotides and synthetic genes . . . 32
4.1.7 Instruments . . . 33
4.1.8 Software . . . 34
4.2 Molecular Biology . . . 34
4.2.1 Golden Gate Cloning . . . 35
4.2.2 Subcloning of genes . . . 35
4.2.3 Transformation of electrocompetent *Agrobacterium tumefaciens* cells 35
4.3 Treatment of plant material . . . 37
4.3.1 Infiltration of *Nicotiana benthamiana* . . . 37
4.3.2 Plant material harvest . . . 37
4.3.3 Extraction of flavonoids from *N. benthamiana* leaves . . . 37
4.4 Protein biochemistry . . . 38
4.4.1 Determination of protein concentration . . . 38
4.4.2 Protein production test (expression test) . . . 39
4.4.3 Protein subfractionation . . . 39
4.4.4 Protein sample concentration by TCA precipitation . . . 40
4.4.5 Preparation of periplasmic protein . . . 40
4.4.6 Discontinous SDS-polyacrylamide gel electrophoresis (SDS-PAGE) 41
4.4.7 Buffer change of protein samples . . . 41
4.4.8 Production of recombinant protein . . . 42
4.4.9 Preparation of inlusion bodies (IBs) . . . 43
4.4.10 Purification of His-tagged proteins using immobilized metal affinity chromatography (IMAC) . . . 43
4.4.11 Refolding of SOMT-2 on a micro scale using design of experiments (DoE) . . . 44
4.4.12 Enzymatic production of SAM and SAE . . . 46

4.5 Crystallographic Procedures 47
4.5.1 Crystallization of proteins 47
4.5.2 Data collection and processing 48
4.5.3 Structure solution 49
4.5.4 Model building, refinement and validation 49
4.5.5 *In silico* substrate docking 50
4.6 Analytics 50
4.6.1 Recording of growth curves 50
4.6.2 *In vitro* determination of glucose 50
4.6.3 *In vitro* *O*-methyl transferase (*O*-MT) assay 51
4.6.4 Photospectrometric assay for the methylation of catecholic moieties 54
4.6.5 Concentration of SOMT-2 using hydrophobic interaction chromatography (HIC) 54
4.6.6 Analytical gel filtration 55
4.6.7 Binding experiments using Isothermal Titration Calorimetry (ITC) 55
4.6.8 High-performance liquid chromatography (HPLC) analytics . . . 56
4.6.9 Liquid chromatography-tandem mass spectrometry (LC-MS/MS) measurements 56

5 Engineering of phenylpropanoid and flavonoid *O*-methyl transferase (PFOMT) 59
5.1 Introduction 60
5.2 Crystallization of PFOMT 62
5.3 Substrate binding studies using ITC 65
5.4 Study of variants for long-chain alkylations 68
5.5 Conclusion/Discussion 70
5.6 Contributions 72

6 Tandem mass-spectrometry studies of flavonoids 73
6.1 Introduction 74
6.2 Fragmentation of flavanones 77
6.3 Fragmentation of flavones 81
6.4 Fragmentation of flavonols 86
6.5 Conclusions 91
6.6 Contributions 92

7 Enzymatic methylation of non-catechols 93
7.1 Introduction 94
7.2 SOMT-2 95
7.2.1 *In vivo* biotransformation in *Nicotiana benthamiana* 95
7.2.2 *In vivo* biotransformation in *E. coli* 97
7.2.3 *In vitro* studies using recombinantly produced SOMT-2 101

7.3 PFOMT 110
7.3.1 Phenols. 110
7.3.2 PFOMT pH-profiles are influenced by Mg^{2+} 111
7.3.3 Methylation of different chemical motifs. 116
7.4 Conclusion 121

8 Summary 123

Affidavit 125

Curriculum vitæ 126

III Appendix 127

A Engineering of PFOMT 129

B Tandem mass-spectrometry studies of flavonoids 133

C Enzymatic methylation of non-catechols 137
C.1 SOMT expression studies 137
C.2 Conversion of non-catechols by PFOMT 142
C.2.1 Modelling and shrinkage of catechols subset (pH profile) 142
C.3 Identification of products from conversion of non-catechols by PFOMT 143
C.3.1 *p*-Coumaric acid methyl ester 143
C.3.2 *iso*-Ferulic acid esters and caffeic acid dimethylether 144
C.3.3 3’,4’-Dimethyl eriodictyol 144
C.3.4 3’,4’-Dimethyl luteolin 146

D Additional information 149

Bibliography 165

Acronyms 167

Glossary 173

List of Figures

3.1 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/malonate derived C_2-moieties (orange). 12

3.2 Major 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. 14

3.3 Topology plots of the five major structural classes of methyl transferases and radical SAM methyl transferases (RSMTs) (modified and extended from Schubert *et al.* [39]). Helices are depicted as circles and β-strands as triangles. The SAM binding site is depicted as a flag. Radical SAM methyl transferases all share a common “radical SAM”-domain, which contains the iron sulfur cluster (red checker). The individual radical SAM methyl transferase classes are differentiated according to the other domains they contain. 16

3.4 Reactions catalyzed by methyl transferases (MTs). Different shades of gray were used to differentiate between different groups of compounds and methyl transferases. In contrast to other methyl transferases, the group of radical SAM methyl transferases also requires additional co-factors to SAM. 17

3.5 Mechnism of the methyl transfer reaction catalyzed by methyl transferases. Non-radical *S*-adenosyl-L-methionine dependent methyl transferases catalyze the nucleophilic transfer of a methyl group from the donor *S*-adenosyl-L-methionine 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. The S_N2 reaction proceeds via a single transition state, during which the methyl-carbon is sp^2 hybridzed. After transfer of the methyl to the nucleophile the carbon's configuration is inverted. 20

3.6 Bioenzymatic synthesis of chiral acetate, the precursor for the synthesis of SAM carrying an assymetrical methyl, as used by Woodward *et al.* [61]. 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 21

3.7 Natural productes synthesized with the help of methyl transferases. *O*-methyl transferases were used for the synthesis of ermanine (**1**), 7-*O*-methyl-aromadendrin (**2**) and 3,5-dimethoxy-4'-fluorostilbene (**3**). *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) 25

3.8 Labelling of macromolecules by using a combination of novel alkine-derivatized *S*-adenosyl-L-methionine analogues and Cu^I-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 might also be also feasable for use in activity based protein profiling (ABPP) approaches. 27

4.1 Golden Gate cloning scheme for SOMT-2 36

4.2 Pruned *N. benthamiana* plant, with two bottom and one top leaf, ready to be infiltrated. 38

4.3 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. [157, 158] 52

5.1 Some crystal and pseudo-crystal shapes that were observed during the crytsallization screen. a – high $(NH_4)_2SO_4$, b-c – $CaCl_2$, PEG-4000, e – LiCl, PEG-6000 62

5.2 An overview of the features in the *apo*-PFOMT structure. **a** – The assymetric unit of *apo*-PFOMT consists of two homodimers (4 monomers). Individual monomers are rainbow colored from N- (blue) to C-terminus (red). **b** – Comparison of 3C3Y (steelblue) and *apo*-PFOMT (green). The N-terminus of *apo*-PFOMT (red) and even part of the His-tag (red, transparent) was 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^{2+} (green sphere) are featured in the published structure, whereas a sulfate ion (red/yellow spheres) was bound in the newly solved structure. 64

5.3 Positional differences between the individual residues of the solved *apo*-PFOMT and the structure with bound SAH (pdb: 3C3Y). The diffraction precision indicator [172] (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 [33]. 65

5.4 The binding of different SAM analogues was measured via ITC 66

5.5 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. 67

5.6 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, turqouise – 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. 69

5.7 Activities of different PFOMT variants towards caffeic acid methylation. Colorations correspond to the ones used in Figure 5.6. 70

5.8 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. 71

6.1 Comparison of CID and HCD 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 %. Labels a to c correspond to spectra A to C, respectively. **E** – Relationship between the activation method and the intensity of four fragments (● $^{1,3}A^+$, ▲ ($^{1,4}B^+$-2H), ■ ($^{1,4}B^+$-2H-CO), + $C_7H_5^+$) of different flavanones. 80

6.2 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 %. Labels a to c correspond to spectra A to C, respectively. 83

6.3 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 %. Labels a to c correspond to spectra A to C, respectively. 89

7.1 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 96

7.2 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). 97

7.3 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. 98

7.4 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-2 100

7.5 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 102

7.6 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. 104

7.7 Results of *in vitro* protein refolding trials. Measured data (left) is presented alongside the ME-plots (right). The dashed line through the ME-plots illustrates the overall mean. **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, suggesting that high pH and arginine concentration might have been 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 afforded 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 was important to achieve active protein. 108

7.8 CD-spectrum of refolded SOMT-2 (black) compared to the spectrum that was calculated (red) by the K2D3 web service (*http://cbdm-01.zdv.uni-mainz.de/~andrade/k2d3//index.html*) from the SOMT-2 sequence. Secondary structure estimates from the measured spectrum are 12.39 % α-helix and 32.51 % β-sheet. 109

7.9 Specific activity/pH-profiles for the conversion of three different substrates (● caffeic acid [catecholic], ▲ eriodictyol [catecholic], ■ *iso*-ferulic acid [3-hydroxy-4-methoxy]) by PFOMT. The specific activity for the non-catecholic substrate *iso*-ferulic acid was much lower than the specific activity for the catecholic substrates. When magnesium is omitted, the activity is increased by increasing the pH 112

7.10 pH-profiles of substrate conversion along with predicted data. Predicted data from 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. 115

7.11 Conversion of multiple different substrates, catecholic and non-catecholic, by PFOMT wild-type (**a**) and the 4'-specific variant Y51R N202W (**b**). Every individual box represents one substrate *p*-coumaric acid (A.1), . . ., chrysoeriol (C.4). Low pH conditions are displayed on the left side of each panel, high pH conditions on the right. Mg^{2+}-conditions are color-coded as follows: no Mg^{2+} – white bars, 10 mM Mg^{2+} – black bars. 118

A.1 Differences in the dihedral-angles ψ (red) and φ (black) of the solved *apo*-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. 130

A.2 Differences in the regioselectivity of some PFOMT variants. The products observed in HPLC and LC/MS measurements switched from 3'-*O*-methylated (dark grey) to 4'-*O*-methylated (light grey) for the displayed variants. The height of the bars corresponds to the total conversion of substrate. 131

C.1 Additional scatterplots of the 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) 137

C.2 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_{600} after inoculation was about 0.1. 138

C.3 SDS-PAGE gel of samples aquired during growth curve measurements. The arrow indicates the band that could correspond to the GST-tagged SOMT-2 protein. 139

C.4 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*) [227]. Cysteines are shown as sticks. The distance between neighboring cysteines that could by oxidized to disulfide bridges is shown in orange. 140

C.5 Chromatogram of the gel filtration analysis of refolded SOMT-2 (blue). Gel-filtrations standards (red) were used to assess the size of the SOMT-2 protein. The estimated molecular weights for the eluting peaks were 165 kDa (12.47 ml) and 65.5 kDa (14.26 ml). Protein standard: 9.15 ml – thyroglobulin (670 kDa), 12.62 ml – γ-globulin (158 kDa), 15.43 ml – ovalbumin (44 kDa), 17.63 ml – myoglobin (17 kDa), 21.11 ml – vitamin B12 (1.35 kDa) 141

C.6 MS^2 spectra of (**1**) and (**3**). (**a**) negative mode MS^2 of (**1**). (**b**) positive mode MS^2 of (**1**). (**c**) positive mode MS^2 of (**3**). 145

C.7 MS^2 spectra of 3'4'-dimethyl eriodictyol (**4**). (**a**) positive mode MS^2 CID spectrum of (**4**). (**b**) positive mode MS^2 HCD spectrum of (**4**). 146

C.8 MS^2 spectra of 3'4'-dimethyl luteolin (**5**). (**a**) positive mode MS^2 CID spectrum of (**5**). (**b**) positive mode MS^2 HCD spectrum of (**5**). 147

C.9 Product composition after conversion of flavones with PFOMT. Bar chart of the the peak heights of the unidentified (black) and 3',4'-*O*-dimethylated products (gray) in the selected ion chromatograms (HCD at 100 % NCE). The conversion experiments were conducted with the wild-type PFOMT at pH 8.6 with 10 mM Mg^{2+} added. 147

List of Tables

3.1 Defining motifs of plant *O*-MTs as described by Joshi *et al.* [76]. 22

4.1 Natural deep eutectic solvent (NADES)-mixtures used within this work.. 30

4.2 Plasmids used in this work. 32

4.3 Primers used in this work. Recognition sites for endonucleases are underlined. Positions used for site directed mutagenesis are in lower case font. 33

4.4 Calculated extinction coefficients of proteins used in this work.. 39

4.5 Factors and their high and low levels (+/-) used in the construction of the fractional factorial design (FrFD). 44

4.6 Experimental design matrix for the FrFD. 45

5.1 Results of fitting a simple one-site binding model to the data obtained from ITC experiments. 68

7.1 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. 97

7.2 *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 LC-MS/MS. 99

7.3 pK_a-values of phenolic groups examplified by *p*-cresol derivatives. Substituent numberings on the aromatic ring do not reflect conventions of the International Union of Pure and Applied Chemistry (IUPAC). 110

7.4 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 by the column titled “pH”. 113

7.5 Coefficients of the model (Equation 7.1) 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. 114

7.6 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, single phenolic, catecholic, $3^{(')}$-hydroxy-$4^{(')}$-methoxy (3O4M) and $4^{(')}$-hydroxy-$3^{(')}$-methoxy (4O3M). $R^{1/2}$ denote possible methylation positions.. 117

7.8 Conversion of substrates after 16 hours incubation. Only the maximum conversion is displayed, along with the conditions it was achieved under. 119

7.9 Products of the enzymatic methylation of the studied substrates. The products were confirmed by authentic standards or via LCMS. 120

A.1 Crystallographic data, phasing and refinement statistics. 129

B.1 Key ions in the positive mode CID and HCD ESI-MS^2 spectra of flavanones. 134

B.2 Key ions in the positive mode CID and HCD ESI-MS^2 spectra of flavones. 135

B.3 Key ions in the positive mode CID and HCD ESI-MS^2 spectra of flavonoles. 136

C.1 Results for the ANOVA of the main effects model describing soluble protein obtained after refolding. 139

C.2 Results for the ANOVA of the main effects model describing protein activity after refolding.. 140

C.3 Coefficients of the model (C.1) for activity of catechol methylation by PFOMT. The factor Mg is a categorical variable (addition/no addition) and can therefore only be 0 or 1. 142

C.4 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.. 143

D.2 SAM analogues that have been used with MTs. Targets: *P* – peptide/protein, *D* – DNA, *R* – RNA, *S* – small molecule. 149

D.1 Overview over the contructs 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. 152

List of Schemes

6.1 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. 76

6.2 Major fragmentation pathways of flavanones. Activation using CID conditions at 45 % NCE mainly results in neutral losses of H_2O and ketene (C_2H_2O) from the molecular ion $[M+H]^+$ (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. 78

6.3 Proposed MS^2 fragmentation of $^{1,3}A^+$ after HCD activation. In high energy MS^2 experiments, $^{1,3}A^+$ might loose two CO followed by an unusual C_2H_4. A single loss of ketene (C_2H_2O) to afford m/z 111 is also sensible. 79

6.4 Major fragmentation pathways of non-methylated 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^2 spectra of non-methylated flavones. Methylated flavones loose a methyl group in CID experiments, but only in HCD experiments do other fragmentation reaction become obvious. 82

6.5 Stability of the $[M+H-CH_3]^{\bullet+}$ ion of flavones. The $[M+H-CH_3]^{\bullet+}$ 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. 84

6.6 Proposed pathway of fragmentation of (**10**) after HCD activation. Losses of $CH_3^{\bullet}$ and CH_4, followed by loss of CO are the major fragmentations observed in the corresponding MS spectra. However, multiple losses of CO only occur after a loss of methane (CH_4), possibly due to the relative stability of the benzochromenylium radical cation (**iii**). At 100 % NCE even higher order fragmentations were observed. 85

6.7 Major fragmentation pathways of flavonols. Unlike flavones, methylated and non-methylated flavonols 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 %). 87

6.8 Proposed pathways of fragmentation of isorhamnetin (**15**). Isorhamnetin might loose methyl, methane or methanol upon activation. A similar fragmnentation pathway was proposed for the analogous chrysoeriol (Scheme 6.6). Some fragmentations were observed in HCD mode only (box). 90

I

Preface

1 Abstracts

1.1 English Abstract

The present study outlines the useability of two plant *O*-methyl transferases (*O*-MTs), phenylpropanoid and flavonoid *O*-methyl transferase (PFOMT) and soy *O*-methyl transferase (SOMT-2), of classes I and II for the biocatalytic methylation of common structural motifs encountered throughout the group of plant polyphenolic compounds. Class I plant *O*-MTs, such as PFOMT, are magnesium-dependent caffeoyl CoA dependent *O*-methyltransferase (CCoAOMT)-like enzymes comprising 200 to 280 amino acids, generally with a limited substrate scope. Class II plant *O*-MTs (i.e. SOMT-2) are metal-ion independent, 340 to 390 amino acids in length and generally more promiscuous than class I representatives.

Biophysical characterization of PFOMT using Isothermal Titration Calorimetry demonstrated its ability to bind the biosynthetically obtained *S*-adenosyl-L-methionine analogue *S*-adenosyl-L-ethionine (SAE). However, although numerous variants were characterized, conversion of SAE by PFOMT was not observed. Nonetheless, a novel crystal structure of the *apo*-form of PFOMT obtained during the present study gives new insights into the movements and domains involved in substrate binding.

A systematic grid of 15 flavonoids (i.e. flavanones, flavones and flavonoles) with different substitutions at the B-ring was characterized by tandem mass-spectrometry (MS/MS) studies in positive mode. Thereby, the fragmentation patterns produced by activation via collision induced dissociation (CID) and higher-energy collisional dissociation (HCD) were compared. Fragmentation, especially the appearance of characteristic C-ring cleavage fragments such as $^{1,3}A^+$, of methylated members of flavones and flavonoles was easily achieved in HCD experiments. The results obtained demon-

strate the complementary nature of both activation methods and the importance of HCD spectra for the identification of methylated flavones and flavonols.

In vivo studies using SOMT-2 showed its capability to methylate flavonoids and stilbenes at the 4’-position. Design of experiments and fractional factorial design were used successfully to optimize *in vitro* refolding conditions for SOMT-2. However, sufficiently active enzyme could not be obtained. Moreover, using PFOMT it was demonstrated, that the activity of this class I plant *O*-MT could be modulated by pH and magnesium concentration to achieve previously unobserved methylations of non-catecholic moieties such as $3^{(’)}$-hydroxy-$4^{(’)}$-methoxy and $4^{(’)}$-hydroxy-$3^{(’)}$-methoxy. This resulted in the production of e.g. 3’,4’-dimethyl eriodictyol and 3’,4’-dimethyl luteolin by PFOMT from hesperetin and diosmetin, respectively. The knowledge obtained from the MS/MS studies was crucial in identifying these products.

1.2 Deutsche Zusammenfassung

Die vorliegende Arbeit umreißt die Verwertbarkeit zweier pflanzlicher *O*-Methyltransferasen (*O*-MTs), PFOMT und SOMT-2, der Klassen I and II für die biokatalytische Methylierung in polyphenolischen Verbindungen verbreiteter Strukturmotive. Klasse I Planzen-*O*-MTs, wie PFOMT, sind magnesiumabhängige Kaffeoyl CoA-abhängige *O*-Methyltransferase (CCoAOMT)-ähnliche Enzyme von 200 bis 280 Aminosäuren und einem gemeinhin limitierten Substratumfang. Vertreter der Klasse II (d.h. SOMT-2) sind metallionen-unabhängig, 340 bis 390 Aminosäuren lang und im Allgemeinen promiskuitiver im Substratspektrum als Vertreter der Klasse I.

Die biophysikalische Charakterisierung der PFOMT mittels isothermer Titrationskalorimetrie zeigte die Bindung von *S*-Adenosyl-L-ethionin (SAE), ein biosynthetisch gewonnenes *S*-Adenosyl-L-methionin-Derivat, an das Enzym. Eine Umsetzung von SAE durch PFOMT konnte jedoch nicht gezeigt werden, obgleich mehrere Enzymvarianten untersucht wurden. Nichtsdestotrotz wurde eine neue Kristallstruktur des *apo*-Enzyms gelöst, welche neuerliche Einblicke in die Bewegungen innerhalb des Enzyms und an der Substratbindung involvierter Domänen liefert.

Fünfzehn Flavonoide (d.h. Flavanone, Flavone und Flavonole) mit unterschiedlichen Substitutionsmustern am B-Ring wurden mittels Tandem-Massenspektrometrie (MS/MS) im positiven Modus charakterisiert. Dabei wurden zwei Aktivierungsmeth-

oden, *collision induced dissociation* (HCD) und *higher-energy collisional dissociation* (CID), verglichen. HCD erleichterte die Fragmentierung, besonders das Erscheinen durch C-Ring-Spaltung enstandener Fragmente (z.B. $^{1,3}A^{+}$), methylierter Vertreter von Flavonen und Flavolonen. Die Ergebnisse demonstrieren die Komplementarität beider Aktivierungsmethoden und die Bedeutung von HCD-Spektren für die Identifizierung methylierter Flavone und Flavonole.

In vivo Studien der SOMT-2 zeigten die Methylierung von Flavonoiden und Stilbenen an der 4'-Position. Statistische Versuchplanung und ein Teilfaktorplan wurden erfolgreich verwendet, um die Rückfaltungsausbeuten der *in vitro* Rückfaltung der SOMT-2 zu optimieren. Jedoch konnten nur ungenügende Mengen an aktivem Protein erhalten werden. Es wurde weiterhin gezeigt, dass die Aktivität der PFOMT durch Variation von pH-Wert und Magnesium so weit moduliert werden kann, dass vorher unbeschriebene Methylierungen von 3'-Hydroxy-4'-methoxy und 4'-Hydroxy-3'-methoxy-Motiven möglich sind. So konnten unter anderem 3',4'-Dimethyleriodictyol und 3',4'-Dimethylluteolin durch Umsetztung von Hesperetin und Diosmetin mittels PFOMT produziert werden. Für die Identifizierung der genannten Produkte wurde unter anderem auf das in den MS/MS-Studien gewonnene Wissen zurückgegriffen.

2 List of Publications

Parts of this work have been published:

1. Wessjohann, L. A., Keim, J., Weigel, B., & Dippe, M. (2013). Alkylating enzymes. Current Opinion in Chemical Biology, 17(2), 229–235. *http://doi.org/10.1016/j.cbpa.2013.02.016*
2. Dippe, M., Weigel, B., Heinke, R., Vogt, T., & Wessjohann, L. A. (2015). Engineering of a Mg2+-dependent O-methyltransferase towards novel regiospecificity. Manuscript to be submitted.

Publications that are not part of this work:

1. Landgraf, R., Smolka, U., Altmann, S., Eschen-Lippold, L., Senning, M., Sonnewald, S., Weigel, B., … Rosahl, S. (2014). The ABC Transporter ABCG1 Is Required for Suberin Formation in Potato Tuber Periderm. The Plant Cell, 13(August), 3403–3415. *http://doi.org/10.1105/tpc.114.124776*
2. Mullen, D. G., Weigel, B., Barany, G., & Distefano, M. D. (2010). On-resin conversion of Cys(Acm)-containing peptides to their corresponding Cys(Scm) congeners. Journal of Peptide Science, 16(5), 219–22. *http://doi.org/10.1002/psc.1223*

II

Thesis

3 Introduction and Motivation

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 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 phenylpropanoids and flavonoids, as well as their modification, and will therefore mainly focus on these compounds.

3.1 Flavonoids

3.1.1 Overview

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 related pathways and take on various structural (e.g. cell walls) and non-structural roles (e.g. plant defense, flower color) [1]. 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 3.1). They consist of three rings named A, B and C. Ring A and B are of acetate/malonate 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 (Figure 3.1). For example, a phenyl group at C-2 or C-3 gives flavonoids or isoflavonoids respectively, an unsubstituted C-4 means flavan, whereas a carbonyl group at C-4 indicates flavan*ones* *et cetera*. Flavonoids are usually poly-hydroxylated, but can also carry multiple other different substitutions. *O*-methylations are common at all hydroxyl positions, but flavonoids can also be *C*-methylated [2]. 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) [3–6].

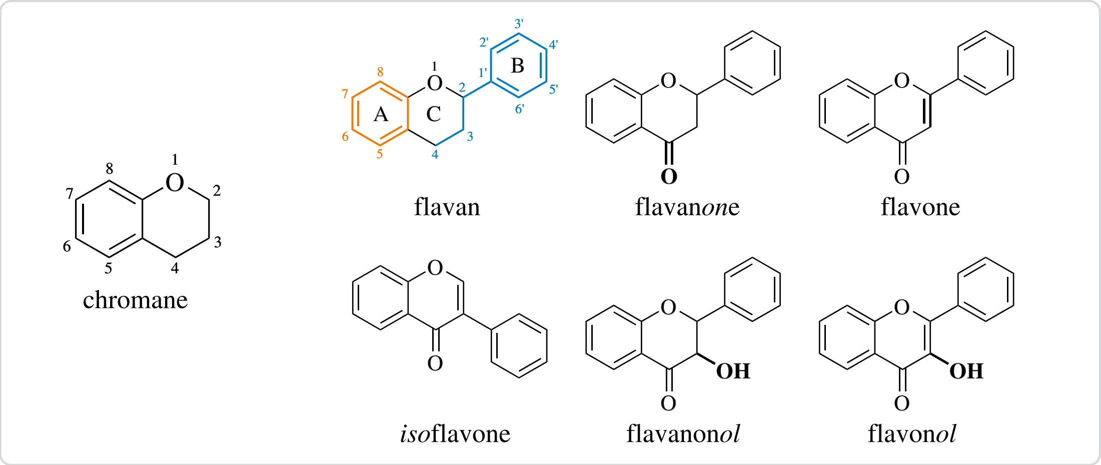


Figure 3.1.: The central feature of the flavonoids is the chromane ring. The names of the different groups of flavonoids are derived from the substitution of this moiety. From a biosynthetic point of view, flavonoids are built up from phenylpropanoid (blue) and acetate/malonate derived C_2*-moieties (orange).*

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 scavange reactive oxygen species (ROS) [7]. Furthermore, flavonoids can act as regulators during plant development [8]. 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 [9].

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 [10–12]. However, many flavonoids are scarce in nature or the production from by-products is not sufficient 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* [13, 14]. For example, eriodictyol was produced from tyrosine in metabolically engineered *E. coli* at levels of up to 107 mg/ml [15], whereas naringenin was produced in *S. cerevisiae* from glucose at levels of 109 mg/ml [16].

3.1.2 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 3.2) [17, 18]. 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 [19, 20]. 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 [21]. Next, the linear chalcone can cyclize spontaneously, or catalyzed by a chalcone isomerase (CHI), via a MICHAEL-type addition to afford the flavanone naringenin [22]. 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 [23]. Dihydroflavonols are again precursors for the biosynthesis flavonols, flavanols and anthocyani(di)nes.

3.1.3 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 scavangers for free ROS or mask metal ions by chelation to suppress the production of radicals [24, 25]. The substitution pattern plays an important role in the antioxidant

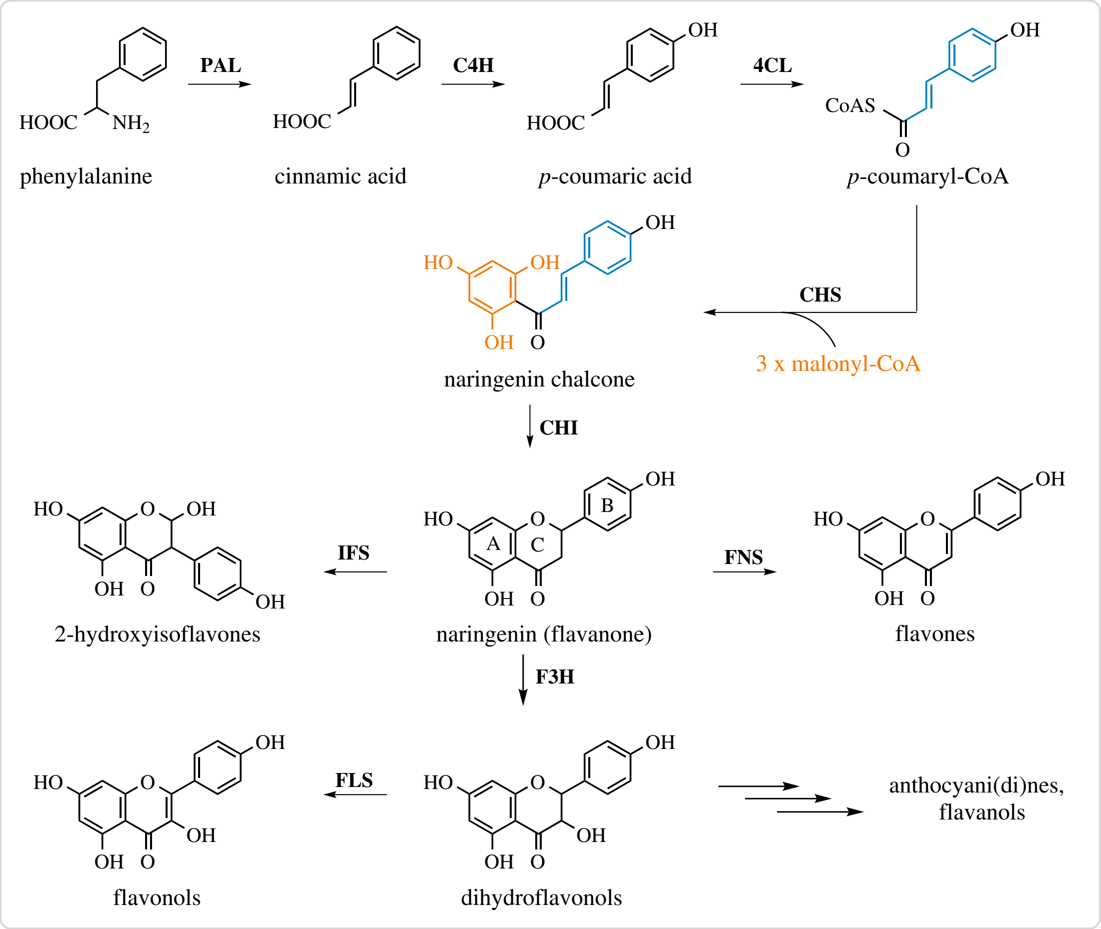

HOOC
NH2
phenylalanine
PAL
HOOC
cinnamic acid
C4H
OH
HOOC
p-coumaric acid
4CL
OH
CoAS
O
p-coumaryl-CoA
CHS
3 x malonyl-CoA
HO
OH
OH
OH
O
naringenin chalcone
CHI
HO
O
OH
OH
O
OH
2-hydroxyisoflavones
IFS
HO
A
B
C
O
OH
OH
O
naringenin (flavanone)
FNS
HO
O
OH
OH
O
flavones
F3H
HO
O
OH
OH
OH
O
flavonols
FLS
HO
O
OH
OH
OH
O
dihydroflavonols
anthocyani(di)nes,
flavanols

Figure 3.2.: Major 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.

activity. Generally, the more free hydroxyls are present, the stronger the antioxidant activity of the flavonoid [24]. Hydroxyls can donate an electron, or a hydrogen atom to free ROS to inactivate such molecules. The resulting flavonoid radicals are stabilized by resonance, however possess prooxidant properties [26].

Numerous flavonoids possess antimicrobial activities [27]. For example, catechins from green and black teas have been shown to be effective against *Bacillus cereus* at nanomolar concentrations [28]. The mechanisms of action for the antimicrobial activity can be the inactivation of enzymes, binding of adhesins, membrane disruption or cell wall complexation [29].

3.2 Methyl transferases (MTs)

3.2.1 Overview

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_N2 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 database (*http://www.uniprot.org*). 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 too (Figure 3.4)[30, 31]. 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) [13, 32–34]. 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 epigentic regulation of gene expression and have been associated with a number of cancers and other diseases [35–37]. In bacteria, DNA-MTs are an essential part of the restriction modification system [38].

According to their structure, MTs can be classified into five main groups (I–V) (Figure 3.3)[39]. 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 β-sheet flanked by α-helices. Most *sm*MTs, DNA-MTs and some P-MTs belong to class I. Even though some of the enzymes belonging to class I share as little as 10 % sequence similarity, there is a pronounced structural conservation [39]. Class II MTs comprise a long anti-parallel β-sheet encompassed by numerous α-helices [40]. In class III MTs the SAM binding site is located between two $\alpha\beta\alpha$ domains [41]. A knot structure at the C-terminus contributes to SAM binding in the *SpoU*-*T*rmD (SPOUT) family of class IV RNA-MTs [42]. Protein lysine MTs make up the largest part of P-MTs and structurally belong to class V MTs containing a *s*uvar3-9, *e*nhancer-of-zeste, *t*rithorax (SET) domain [43]. Interestingly, a recent study of the methyltransferome of baker's yeast (*S. cerevisiae*)

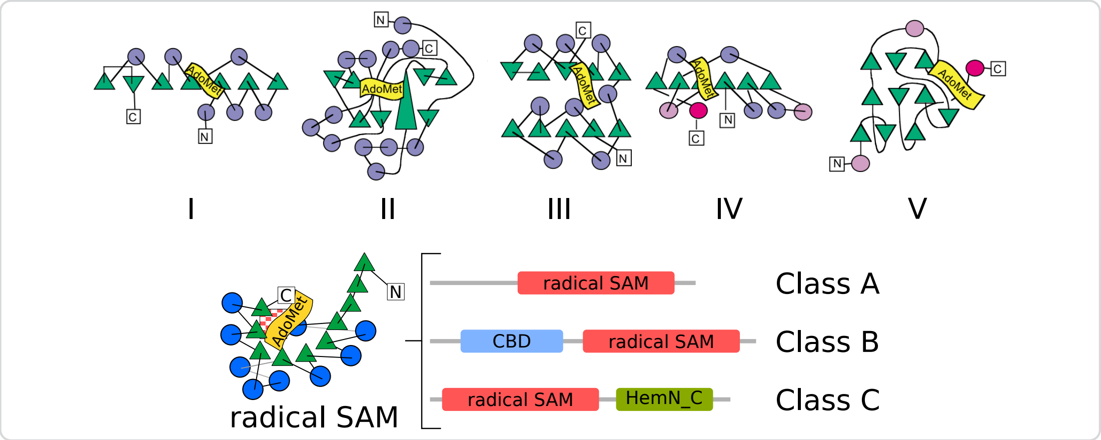


Figure 3.3.: Topology plots of the five major structural classes of methyl transferases and radical SAM methyl transferases (RSMTs) (modified and extended from Schubert et al. [39]). Helices are depicted as circles and β-strands as triangles. The SAM binding site is depicted as a flag. Radical SAM methyl transferases all share a common "radical SAM"-domain, which contains the iron sulfur cluster (red checker). The individual radical SAM methyl transferase classes are differentiated according to the other domains they contain.

argues, that there are four more classes of MTs [44]. Opposite to the studies by Schubert *et al.*, this work mainly relied on bioinformatics 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-β, all-α (RNA/DNA 3-helical bundle) and transmembrane, were postulated.

Radical SAM methyl transferases (RSMTs) comprise another class of resently 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 [45]. 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 [46, 47]. Three classes (A, B, C) with distinct structural and mechanistic characteristics have been recognized within the RSMTs [48]. The centerpiece of RSMTs is the *radical SAM* domain, whose structure was first described in the ribosomal ribonucleic acid (rRNA) methyl transferase RlmN of *E. coli* [46]. This domain consists of an α_6/β_6 partial barrel

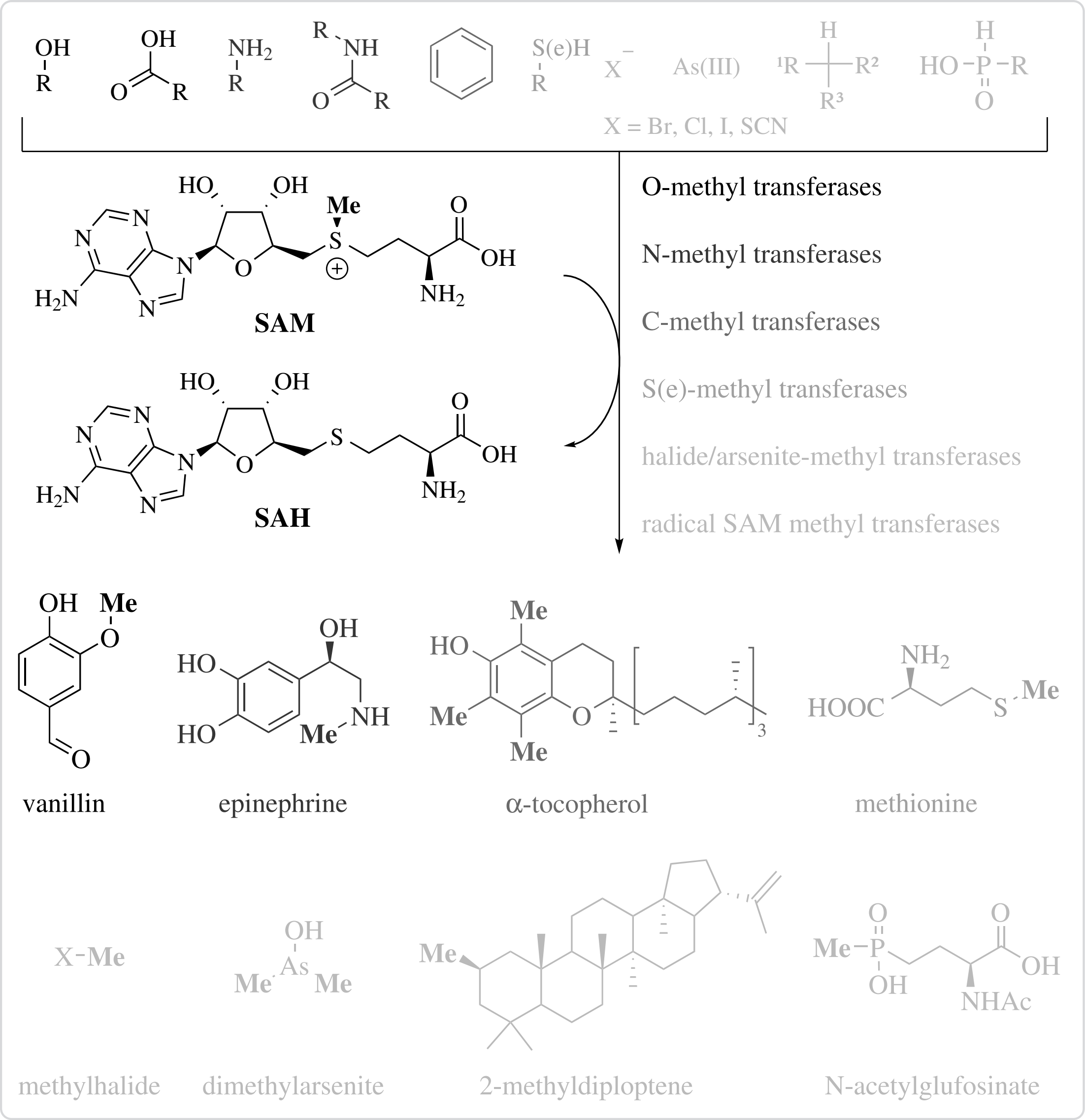

X = Br, Cl, I, SCN
SAM
SAH
O-methyl transferases
N-methyl transferases
C-methyl transferases
S(e)-methyl transferases
halide/arsenite-methyl transferases
radical SAM methyl transferases
vanillin
epinephrine
α-tocopherol
methionine
methylhalide
dimethylarsenite
2-methyldiploptene
N-acetylglufosinate

Figure 3.4.: *Reactions catalyzed by methyl transferases (MTs). Different shades of gray were used to differentiate between different groups of compounds and methyl transferases. In contrast to other methyl transferases, the group of radical SAM methyl transferases also requires additional co-factors to SAM.*

and contains the [4Fe-4S] cluster, as well as the SAM binding site (Figure 3.3). Class A only contains the radical SAM domain and mainly comprises 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 [49]. 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 [47].

3.2.2 *S*-Adenosyl-ʟ-methionine

S-adenosyl-ʟ-methionine (SAM), first described in 1953 [50], 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 [47]. 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. Demethylated SAM is called *S*-adenosyl-ʟ-homocysteine (SAH), which is a good leaving group. Therefore, nucleophilic transfer of the methyl group of SAM is thermodynamically highly favoured ($\Delta G^0 \approx -70\,\mathrm{kJ/mol}$ for SAM + homocysteine ⟶ SAH + methionine) and allows the rapid and selective methylation of a range of substrates [39]. 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 reactiveness, SAM is still quite stable at physiological conditions compared to other sulfonium species like the trimethylsulfonium ion, which quickly reacts with nucleophiles and is often used for derivatization prior to GC analytics [51]. Meanwhile, SAM is readily cleaved into adenine and *S*-ribosylmethionine under alkaline conditions [52] and other deteriorating processes such as racemization and intramolecular cleavage are to be reckoned with [53].

SAM is produced by the enzyme SAM synthetase (EC 2.5.1.6) from methionine and ATP in a two step reaction [54]. 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. 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 [55]. The cobalamine (vitamin B_{12}) dependent methionine synthase (EC 2.1.1.13) can remethylate homocysteine to methionine using N^5-methyltetrahydrofolate as a methyl donor [56]. Taken together, reactions leading from and to SAM are commonly called the activated methyl cycle.

3.2.3 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_N2 mechanism, through an sp^2 hybridized transition state and results in the inversion of configuration (Figure 3.5). The S_N2 mechanism was proposed as early as 1979 [57, 58], but only with the development of the chiral methyl group methodology (Figure 3.6) extended mechanistic studies were made possible [59, 60]. An elegant method for the synthesis of chiral acetate made use of glycolytic enzymes to convert [1-^{3}H]-glucose via its glycolysis intermediates to [3-^{2}H,^{3}H]-lactate, which is subsequently oxidized by chromiumtrioxide (Figure 3.6) [61]. The chirality of the resulting acetate can be controlled by the solvents (D_2O or H_2O) used during the enzymatic reactions. The chiral acetate can be used for the synthesis of e.g. [*methyl*-^{2}H,^{3}H]-methionine.

In 1980, Woodward *et al.* fed *R*- and *S*-methionine containing an assymetrical methyl group to cultures of *Streptomyces griseus*. They 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 S_N2 mechanism [61]. This experiment also demonstrated that, *in vivo*, [*methyl*-^{2}H,^{3}H]-methionine is converted to [*methyl*-^{2}H,^{3}H]-SAM before the methyl group is transferred by a MT. *In vitro* experiments conducted using catechol *O*-methyl transferase (COMT) further supported the hypothesis of an S_N2 mechanism [62].

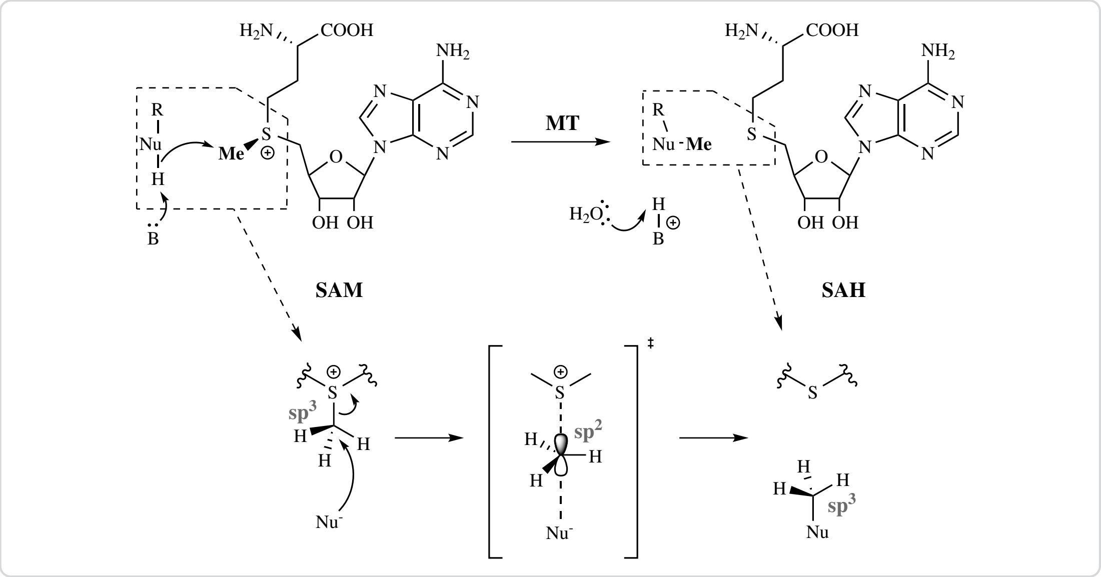


Figure 3.5.: Mechnism of the methyl transfer reaction catalyzed by methyl transferases. Non-radical S-adenosyl-L-methionine dependent methyl transferases catalyze the nucleophilic transfer of a methyl group from the donor S-adenosyl-L-methionine to a nucleophile (Nu; e.g. O,N,C,S). A proton (H^+) is usually abstracted through a general base (B), to achieve activation of the nucleophile. The proton is later transferred to the aqueous medium. The S_N2 reaction proceeds via a single transition state, during which the methyl-carbon is sp^2 hybridzed. After transfer of the methyl to the nucleophile the carbon's configuration is inverted.

The nucleophile attacking the methyl group of SAM is usually activated by abstraction of a proton through a general base (e.g. histidine, lysine) [63–65] or with the help of a LEWIS acid such as complexed Mg^{2+} [33, 66]. In a bimolecular reaction, the rate is highly dependent on the concentration of both compounds:

$$rate = k[A][B].$$

Low concentrations of either result in a low rate. Thus, bimolecular reactions are entropically disfavoured (concentration and entropy are inversely correlated). MTs (and enzymes in general) strongly increase the effective concentration of each reactant and thereby decrease entropy, because methyl donor and acceptor are bound ("immobilized") in close proximity to each other in the active site [67]. 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,

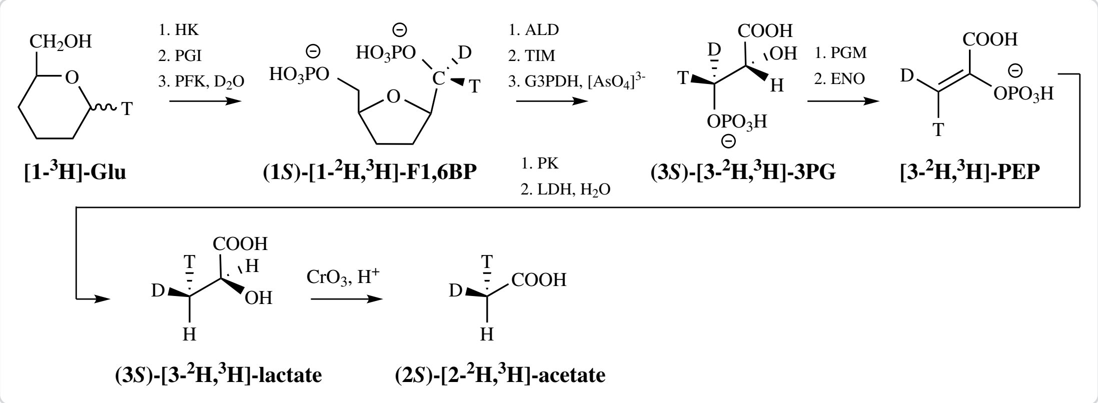

CH2OH
O
T
1. HK
2. PGI
3. PFK, D2O
[1-3H]-Glu
HO3PO
HO3PO
D
C
T
O
(1S)-[1-2H,3H]-F1,6BP
1. ALD
2. TIM
3. G3PDH, [AsO4]3-
1. PK
2. LDH, H2O
COOH
D
OH
T
H
OPO3H
(3S)-[3-2H,3H]-3PG
1. PGM
2. ENO
COOH
D
OPO3H
T
[3-2H,3H]-PEP
COOH
T
H
D
OH
H
(3S)-[3-2H,3H]-lactate
CrO3, H+
T
D
COOH
H
(2S)-[2-2H,3H]-acetate

Figure 3.6.: Bioenzymatic synthesis of chiral acetate, the precursor for the synthesis of SAM carrying an assymetrical methyl, as used by Woodward et al. [61]. 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

which can flip the target nucleotide out of the DNA-helix to provide the best orientation of SAM and the acceptor nucleophile [68].

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 [69], CheR protein-L-glutamate *O*-MT from *Salmonealla typhimurium* [70] and the protoporphyrin IX *O*-MT of *Rhodobacter capsulatus* [71]. The protoporphyrin IX MT from etiolated wheat *Triticum aestivum* on the other hand shows a ping-pong bi-bi mechanism [72], whereas ordered bi-bi mechanisms were shown for the cytosine DNA-MT MSpI and isoprenylated P-MT [73, 74]. Meanwhile, the enzymes exhibiting ordered bi-bi mechanisms were different in that some bound SAM first and released SAH last [74], whereas the others bound the acceptor molecule first [73]. Competetive product inhibition, especially by SAH, is commonly observed for MTs [70, 73–75].

The mechanisms of RSMTs are outside the scope of this work, but the interested reader is referred to current reviews on the topic [47, 48].

3.2.4 Plant *O*-methyl transferases (*O*-MTs)

Plant *O*-methyl transferases (*O*-MTs) were the prime interest of this work. 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 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 3.1) [76]. 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 (≈340–390 amino acids). The distance between motifs A and 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^{2+} for activity. Overall, they are fairly similar to mammalian COMTs [76]. 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 phenyl propanoid derived substrates [77, 78].

Table 3.1.: Defining motifs of plant O-MTs as described by Joshi et al. [76].

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)DAXXXK(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)			

Phenylpropanoid and flavonoid *O*-methyl transferase (PFOMT) is a Mg^{2+}-dependent class I plant *O*-MT from the ice plant *M. crystallinum* and was first described in 2003 by Ibdah *et al.* [78]. 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 CCoAOMT by a lower sequence homology and a broader substrate promiscuity and regiospecificity. PFOMT methylates a number of flavonoids and phenyl propanoids at the *meta*-position of the B ring, 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 [79]. 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.

PFOMT is a biological dimer, as the three dimensonal structure of PFOMT shows (pdb: 3C3Y) [33]. Each monomer exhibits a Rossmann α/β-fold consisting of 8 α-helices and 8 β-strands. The catalytically important N-terminus is not resolved in the structure. SAH and Ca^{2+} were cocrystallized and appear bound in the active site. Ca^{2+} is complexed by two aspartate and one asparagin residues with the rest of the coordination spaces occupied by waters.

Soy *O*-methyl transferase (SOMT-2) has been described in the literature to methylate multiple flavonoids at the 4'-position of the B-ring [80–82]. It has the highest activity towards naringenin, to produce ponciretin (also known as isosakuranetin). No structural data of soy *O*-methyl transferase (SOMT-2) or *in vitro* activity studies have been published to date. Enzymes like SOMT-2, that methylate a *para*-monohydroxylated B ring of flavonoids, either seem to be a rare occurance or fairly inactive, since descriptions of characterized representatives are scarce in the literature and are only limited to a couple of enzymes [83, 84].

3.3 Alkylations and biotransformations

3.3.1 Overview

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 oligo-prenyl diphosphates and are mainly responsible for the over 70 000 terpenoids described today [85]. 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 [86]. 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 ($V \approx 20$ Å^3) for a hydrogen ($V \approx 5$ Å^3) 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 in contrast to the original more polar, hydrogen donating group. This can in turn have dramatic physiological consequences in an organism.

3.3.2 Methyl transferases for industrial use

The industrial potential of MTshas been demonstrated by several studies [13]. Li and Frost [87] 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 3.7). 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 [88]. 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 [89]. 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 [90].

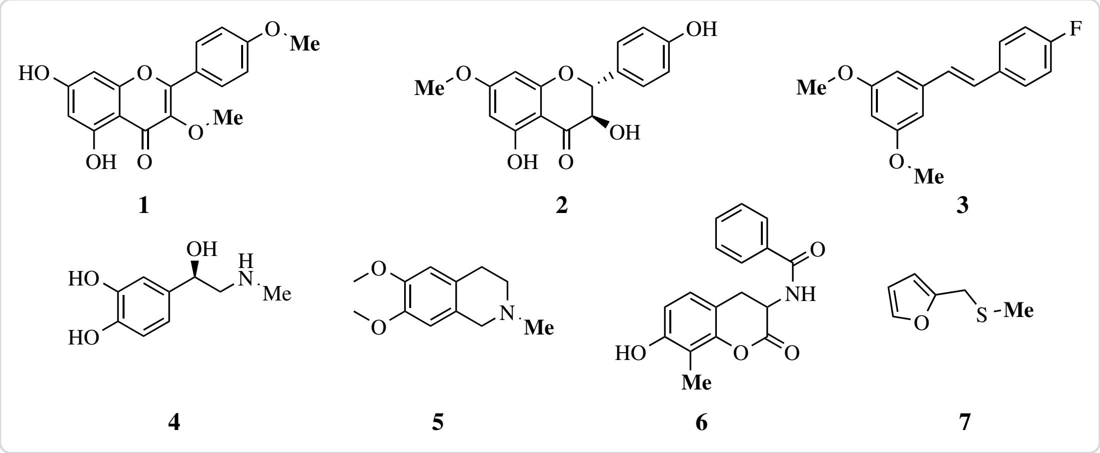


*Figure 3.7.: Natural productes synthesized with the help of methyl transferases. O-methyl transferases were used for the synthesis of ermanine (**1**), 7-O-methyl-aromadendrin (**2**) and 3,5-dimethoxy-4'-fluorostilbene (**3**). N-methyl transferase (N-MT) catalyze the production of epinephrine (**4**) 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)*

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 yields of 2.7 mg/l and 0.2 mg/l, respectively [91, 92]. 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 reconstructued plant pathway in *E. coli.*

N-, *C*- and *S*-methyl transferases have also been successfully used in biocatalytical applications. For example, epinephrine (**4**) and *N*-methyltetrahydro-isoquinoline (**5**) were obtained by *in vitro* and *in vivo* biotransformations of their respective precursors [93, 94]. Furthermore, a number of studies describe the production of several other *N*-methylated alkaloid and non-alkaloid compounds [95–97]. *C*-methyl transferases (*C*-MTs) have been used biocatalytically to modify different phenols [98] and coumarin derivatives, to obtain compounds such as 3-(benzoylamino)-8-methylumbelliferone (**6**)

[99]. Also, the composition of the main tocopherol species in plants could be tuned by the introduction of bacterial *C*-MTs [100]. *S*-methyl transferases (*S*-MTs) 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 and can produce molecules such as furfuryl-methyl-sulfide (**7**) [101].

In vivo biotransformations for the high-yield methylation of compounds is a feasable 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*.

3.3.3 Artificial SAM analogues

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.* [102, 103]. 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 [104, 105]. However, the larger substrate analogues were not neccesarily accomodated by the native P-MTs making engineering efforts for the accomodation of larger substrates inevitable [104]. The specific introduction of alkine functionalized groups made it then possible to use click chemistry for further functionalization and/or detection of the labelled proteins, DNA or RNA (Figure 3.8) [104–108].

In 2012 Bothwell and Luo even described the exchange of the sulfonium with a selenonium center, which afforded *Se*-adenosyl selenomethionine (*Se*AM) analogues that have since then been described as substrates for several P-MTs [109, 110]. *Se*AM analogues have the advantage of being more resistant to chemical decomposition than their sulfur counterparts, but also show enhanced transmethylation reactivity [110]. 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 alkyations of some aminocoumarine antibiotics [99]. 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

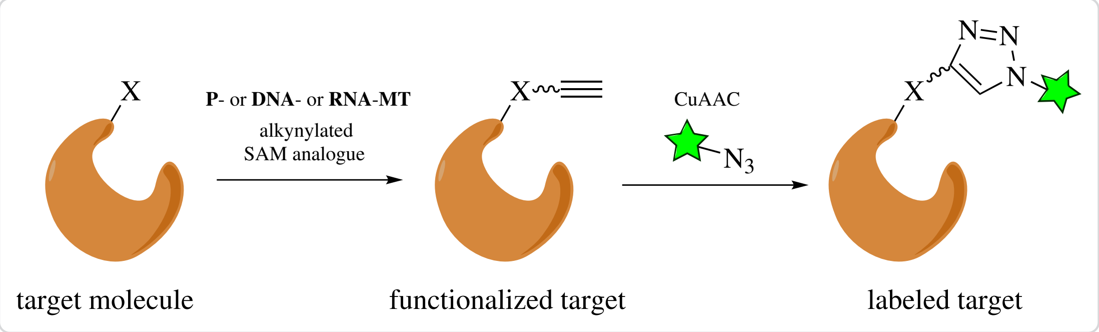


Figure 3.8.: Labelling of macromolecules by using a combination of novel alkine-derivatized S-adenosyl-L-methionine analogues and Cu^{I}-catalyzed azide-alkyne 1,3-dipolar cycloaddition (CuAAC). Depending on the type of label used, it can be employed for detection (e.g. through fluorophores, coupled assays) or affinity purification (e.g. biotin). This technique might also be also feasable for use in activity based protein profiling (ABPP) approaches.

O-methyl transferase (EC 2.1.1.6) and thiopurine *S*-methyl transferase (EC 2.1.1.67) [111]. 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 feasability of using SAM analogues in combination with MTs to modify small molecules [112–114]. However, no bioactivity data has been reported that shows the biological activity of the newly produced compounds.

3.4 Motivation

The motivation of this work was to assess the useability of plant *O*-methyl transferases (*O*-MTs) for the derivatization and functionalization of phenyl propanoid derived phenolics, especially flavonoids.

Phenylpropanoid and flavonoid *O*-methyl transferase (PFOMT) from the ice-plant *Mesembryanthemum crystallinum* was to be used as a model enzyme to study the promiscuity of class I plant *O*-MTs towards the alkyl donor, using the hemisynthetically produced *S*-adenosyl-L-methionine analogue *S*-adenosyl-L-ethionine, (2*S*)-2-amino-4-[[(2*S*,3*S*,4*R*,5*R*)-5-(6-aminopurin-9-yl)-3,4-dihydroxyoxolan-2-yl]methyl-ethylsulfonio]butanoat. Biophysical methods such as x-ray crystallography and Isothermal Titration Calorimetry (ITC) should be used to study the binding of the artificial analogues. The obtained knowledge could be used to aid the development of novel small molecule methyl transferase with desirable catalytic properties.

Furthermore, the enzymatic methylation of different structural motifs encountered throughout the phenyl propanoids was to be studied using class I and class II plant *O*-MTs. Soy *O*-methyl transferase (SOMT-2) and PFOMT are examples of both classes and should thus be used in this work. The results should help in understanding the specific catalytic properties of both classes.

Furthermore, the analytical power of tandem mass-spectrometry (MS/MS) to study substitutions commonly occuring in 5,7-dihydroxylated flavonoids should be assessed using a distinct set of examplary compounds. The insights obtained should provide a reliable and fast method to determine structural properties of unknown flavonoids.

4 Material And Methods

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

4.1 Materials

4.1.1 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 aquired 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).

4.1.2 Commonly used solutions and buffers

50× 5052	25 % glycerol, 2.5 % (w/v) glucose, 10 % (w/v) α-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
lysis 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 Na_2HPO_4, 1 M KH_2PO_4, 0.5 M $(NH_4)_2SO_4$
1 M SSG pH 4 (10×)	14.8 g/l succinic acid, 60.4 g/l $NaH_2PO_4 \cdot H_2O$, 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 $NaH_2PO_4 \cdot H_2O$, 32.8 g/l glycine, 10.3 % 10 M NaOH
5× SDS sample buffer	10 % (w/v) SDS, 10 mM β-mercaptoethanol, 20 % glycerol, 0.2 M Tris/HCl pH 6.8, 0.05 % (w/v) bromophenolblue
1000× trace elements	50 mM $FeCl_3$, 20 mM $CaCl_2$, 10 mM $MnCl_2$, 10 mM $ZnSO_4$, 2 mM $CoCl_2$, 2 mM $CuCl_2$, 2 mM $NiCl_2$, 2 mM Na_2MoO_4, 2 mM Na_2SeO_3, 2 mM H_3BO_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 4.1.: Natural deep eutectic solvent (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

4.1.3 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 yeaxst extract, 0.4 % glycerol, 72 mM K_2HPO_4, 17 mM KH_2PO_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 $MgSO_4$, 0.0002 1000× trace elements

4.1.4 Bacterial strains

E.coli

BL21(DE3)	F$^-$ *ompT hsdSB*(r$_B^-$,m$_B^-$) *gal dcm* λ(DE3) Invitrogen, Karslruhe (Germany)
C41(DE3)	F$^-$ *ompT hsdSB*(r$_B^-$,m$_B^-$) *gal dcm* λ(DE3) Lucigen, Wisconsin (USA)
C43(DE3)	F$^-$ *ompT hsdSB*(r$_B^-$,m$_B^-$) *gal dcm* λ(DE3) Lucigen, Wisconsin (USA)
DH5α	F$^-$ Φ80*lacZ*ΔM15 Δ(*lacZYA-argF*) U169 *recA*1 *endA*1 *hsdR*17(r$_K^-$m$_K^+$) *phoA supE*44 λ^- *thi*-1 *gyrA*96 *relA*1 Invitrogen, Karlsruhe (Germany)
JM110	*rpsL thr leu thi lacY galK galT ara tonA tsx dam dcm glnV*44 Δ(*lac-proAB*) e14- [F' *traD*36 *proAB*$^+$ *lacI*^{q} *lacZ*ΔM15] *hsdR*17(r$_K^-$m$_K^+$) Martin-Luther-University Halle-Wittenberg, AG Bordusa
JW1593 (BW25113 derivative)	*rrnB* Δ*lacZ*4787 *HsdR*514 Δ(*araBAD*)568 *rph*-1 Δ*ydgG* (Kan^R) Keio Collection, National Institute of Genetics (Japan)
MG1655	F$^-$ λ^- *ilvG*$^-$ *rfb*-50 *rph*-1 DSMZ, Braunschweig (Germany)
One Shot TOP10	F$^-$ Φ80*lacZ*ΔM15 Δ(*mrr-hsdRMS-mcrBC*) *recA*1 *endA*1 *mcrA* Δ*lacX*74 *araD*139 Δ(*ara-leu*)7697 *galU galK rpsL* (Str^R) λ^- *nupG* Invitrogen, Karlsruhe (Germany)
Origami(DE3)	Δ(*ara-leu*)7697 Δ*lacX*74 Δ*phoA Pvull phoR araD*139 *ahpC galE galK rpsL* F'[*lac + lacI q pro*] (DE3)*gor*522::*Tn*10 *trxB* (Kan^R, Str^R, Tet^R) Novagen, Wisconsin (USA)
Rosetta(DE3)	F$^-$ *ompT hsdSB*(r$_B^-$,m$_B^-$) *gal dcm* λ(DE3) pRARE (Cam^R) Novagen, Wisconsin (USA)

Rosetta(DE3) pLysS	F^{-} *ompT hsdSB*(r$_{B}^{-}$,m$_{B}^{-}$) *gal dcm* λ(DE3) pLysSRARE (Cam^{R}) Novagen, Wisconsin (USA)
T7 Express	*fhuA*2 *lacZ*::*T*7 *gene*1 *[lon] ompT gal sulA*11 *R*(*mcr*-73::*miniTn*10–Tet^{S})2 *[dcm] R*(*zgb*-210::*Tn*10–Tet^{S}) *endA*1 Δ(*mcrC*-*mrr*)114::*IS*10 NEB, Massachusetts (USA)

Agrobacterium tumefaciens

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

4.1.5 Plasmids

Table 4.2.: 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)

4.1.6 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 4.3.: 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 AAG ACA AAA TGG CTT CTT CAT TAA ACA ATG GCC G	BpiI
somt2	TTG AAG ACA AGG ACA CCC CAA ATA CTG TGA GAT CTT CC	BpiI
somt3	TTG AAG ACA AGT CCT TAG GAA CAC CTT TCT GGG AC	BpiI
somt4	TTG AAG ACA 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	CAT ATG GAT TTT GCT GTG ATG AAG CAG GTC	NdeI
pfomt2.rv	GAA TTC AAT AAA GAC GCC TGC AGA AAG TG	EcoRI
pRha1.fw	CTC TAG CAG ATC TCG GTG AGC ATC ACA TCA CCA CAA TTC	BglII
pRha1.rv	CAA TTG AGG ATC 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	

4.1.7 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) Colibri Microvolume Spectrometer (Biozym, Hess. Oldendorf, Germany)

centrifuges	Eppendorf 5424 (Hamburg, Germany) Hettich Mikro 120 (Kirchlengern, Germany) Beckman Avanti J-E, Beckman Allegra X-30R (Krefeld, Germany)
centrifuge rotors	Beckman JA-10, JA-16.250, JS-4.3 (Krefeld, Germany)

4.1.8 Software

All mathematical and statistical computations and graphics were done with the R software (versions 3.1.X, *http://cran.r-project.org/*) [115]. Visualizations of macromolecules were arranged using the PyMol Molecular Graphics System, version 1.7.0.0 (Schrödinger, New York, USA) or UCSF Chimera version 1.9 (*http://www.cgl.ucsf.edu/chimera*)[116]. 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 seperately in the corresponding section (4.5).

4.2 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 [117]. 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 [118] 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.

4.2.1 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 [119, 120]. 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 inbcubated 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 4.3), 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 4.1). The resulting construct was denoted as pBEW107.

4.2.2 Subcloning of genes

All subcloning procedures were performed according to section 4.2 and specifically subsection 4.2.1. Specific steps for the subcloning of any genes discussed can be found in the appendix (p.152). The *pfomt* 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.

4.2.3 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

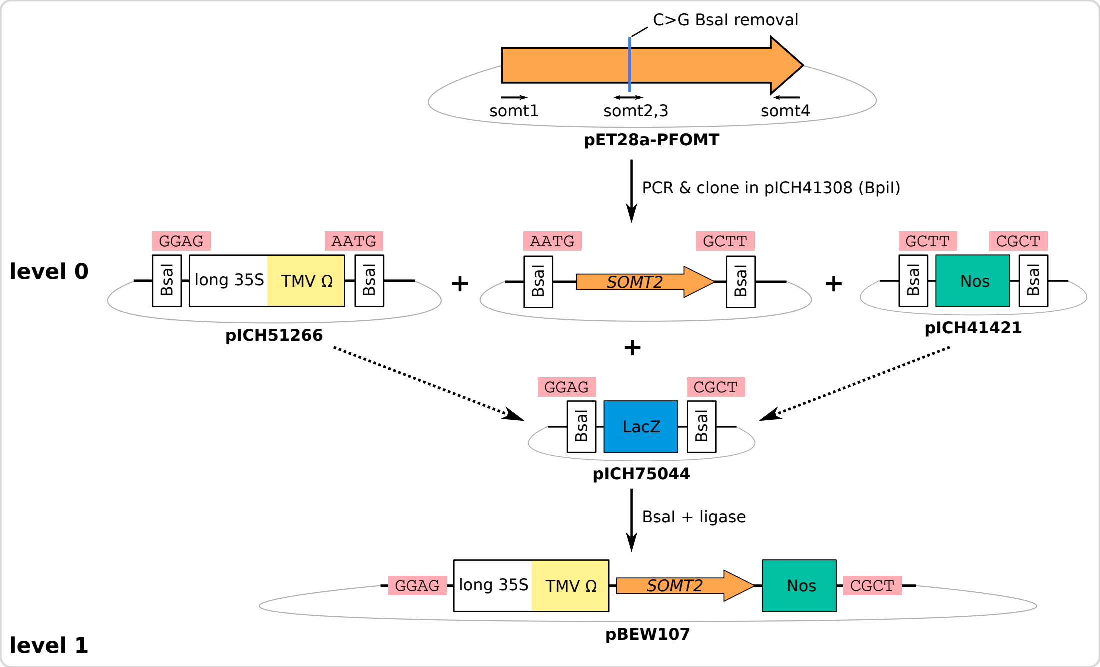

C>G BsaI removal
somt1
somt2,3
somt4
pET28a-PFOMT
PCR & clone in pICH41308 (BpiI)
level 0
GGAG
AATG
BsaI
long 35S
TMV Ω
BsaI
pICH51266
+
AATG
GCTT
BsaI
SOMT2
BsaI
+
GCTT
CGCT
BsaI
Nos
BsaI
pICH41421
+
GGAG
CGCT
BsaI
LacZ
BsaI
pICH75044
BsaI + ligase
GGAG
long 35S
TMV Ω
SOMT2
Nos
CGCT
pBEW107
level 1

Figure 4.1.: Golden Gate cloning scheme for soy O-methyl transferase (SOMT-2)

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.

4.3 Treatment of plant material

4.3.1 Infiltration of *Nicotiana benthamiana*

Before infiltration *N. benthamiana* plants were pruned, such that only leaves to be infiltrated remained with the plant (Figure 4.2). 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_{600} of the culture was measured and adjusted to 0.2 by dilution with infiltration buffer (10 mM MES/NaOH, 10 mM $MgSO_4$ pH 5.5). When multiple *A. tumefaciens* transformed with different contructs/plasmids were used for infiltration, the cultures were mixed and diluted using infiltration buffer, such that OD_{600} 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 4.2). 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.

4.3.2 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.

4.3.3 Extraction of flavonoids from *N. benthamiana* leaves

Two tips of a small spatula of freeze-dried material (≈6 mg), were weighed exactly and extracted with 500 µl 75 % aqueous methanol containing 1 mM ascorbic acid, 0.2 % formic

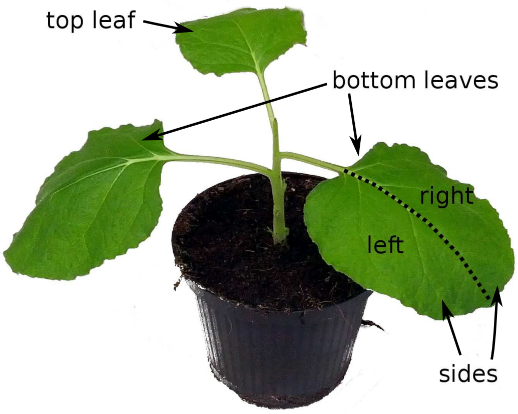


Figure 4.2.: Pruned N. benthamiana plant, with two bottom and one top leaf, ready to be infiltrated.

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 × g, 4 °C, 10 min) and the supernatant transferrerd to a new tube, to remove the insoluble plant material. The supernatant was centrifuged again (20 000 × g, 4 °C, 10 min) and the resulting supernatant was transferred to a HPLC-vial and stored at −20 °C until analysis.

4.4 Protein biochemistry

Stock solutions of antibiotics, IPTG or sugars were prepared according to the pET System Manual by Novagen [121], unless otherwise noted.

4.4.1 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 [122]. Extinction coefficients of proteins were calculated from the amino acid sequence using the ExpPASy servers's ProtParam tool [123].

Table 4.4.: Calculated extinction coefficients of proteins used in this work.

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

4.4.2 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 antiobiotics 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_{600} 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_{600} 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 (4.4.3) and analyzed via SDS-polyacrylamide gel electrophoresis (PAGE) (4.4.6).

4.4.3 Protein subfractionation

The protein subfractionation procedure described herein was adapted from the protocol described in the pET Manual [121]. Overall 5 protein subfractions can be obtained, including *total cell protein, culture supernatant (medium) protein, periplasmic protein, solube cytoplasmic protein* and *insoluble protein*. The OD_{600} 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) (4.4.4) for SDS-PAGE analysis. The periplasmic protein was prepared

(4.4.5) and also concentrated by TCA precipitation for SDS-PAGE. Cells were lysed by resuspending the cell pellet in ($OD_{600} \times V \times 50$) 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.

4.4.4 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.

4.4.5 Preparation of periplasmic protein

Target proteins may be directed to the periplasmic space by N-terminal signal sequences like *pelB* or *DsbA/C* [124]. The periplasma is, other than the cytosol, an oxidizing environment and often used for the production of proteins containing disufide linkages. The preparation of periplasmic protein was accomplished by an osmotic shock protocol modified from Current Protocols in Molecular Biology [125]. 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 $MgSO_4$. 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.

4.4.6 Discontinous SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

The analysis of samples via SDS-PAGE was realized via the discontinous system first described by Laemmli, which allows separation of proteins based on their electrophoretic mobility, which in turn depends on their size [126]. The SDS-PAGE procedure was carried out according to standard protocols described by Sambrook and Russell [117]. Very dilute and/or samples with high ionic strength were concentrated and/or desalted by the TCA precipitation procedure described in subsection 4.4.4. 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).

4.4.7 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.

4.4.8 Production of recombinant protein

Heterologous production of PFOMT

Phenylpropanoid and flavonoid *O*-methyl transferase (PFOMT) was produced as a N-terminally $(His)_6$-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*-amine, *y*east extract, *p*hosphate (ZYP-5052) containing 200 µg/ml kanamycin) was inoculated with the starter culture such that OD_{600} was 0.05. The culture was incubated in a shaking incubator at 37 °C, 220 rpm over night (≈16 h). Due to the autoinducing nature of the ZYP-5052 medium, addition of IPTG was not neccesary. 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 ≈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) (4.4.10). The eluted PFOMT protein was dialyzed (4.4.7) 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 (≈ 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 4.4.9.

4.4.9 Preparation of inlusion bodies (IBs)

Often, when recombinant protein is produced at high levels in *E. coli* it is accumulated in so-called inlusion bodies (IBs) [127]. 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 [128]. The cells were resusupended in 5 ml/g_{cells} 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_{cells} 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.

4.4.10 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^{2+} or Ni^{2+}, 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^{2+}- (HisTrap

FF crude) or Co^{2+}-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. Unound 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.

4.4.11 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 [129–131]. 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 4.5.: *Factors and their high and low levels (+/-) used in the construction of the fractional factorial design (FrFD).*

factor	symbol	setting (level)		unit
		-	+	
pH	A	5.5	9.5	-
arginine	B	0	0.5	M
glycerol	C	0	10	% (v/v)
divalent cations†	D	no	yes	-
ionic strength‡	E	low	high	-
redox state⋆	F	reducing	redox-shuffling	-
α-cyclodextrin	G	0	30	mM
SAH	H	0	0.5	mM

†no: 1 mM EDTA; yes: 2 mM $CaCl_2$, $MgCl_2$
‡low: 10 mM NaCl, 0.5 mM KCl; high: 250 mM NaCl, 10 mM KCl
⋆reducing: 5 mM DTT; redox-shuffling: 1 mM glutathione (GSH), 0.2 mM glutathione disulfide (GSSG)

Table 4.6.: *Experimental design matrix for the FrFD.*

Experiment	A	B	C	D	E	F	G	H
1	+	+	+	-	-	-	-	+
2	-	-	-	-	-	-	-	-
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 over night 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 (4.4.4) 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 ($\rho \approx 1\,\mathrm{g/cm^3}$). 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 manufacturers description. 50 µl of each refolded sample was used for an activity test using naringenin as substrate (4.6.3). 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 4.6.3). Main effects were analyzed qualitatively using main effects plots [132].

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.

4.4.12 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 [133].

Preparative reactions (20 ml) were performed in 0.1 M Tris/HCl, 20 mM $MgCl_2$, 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 ($\varepsilon_0 = 15\,400\,\mathrm{M^{-1}\,cm^{-1}}$) after resuspension in water [134].

4.5 Crystallographic Procedures

4.5.1 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 2×2 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 × 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 $(NH_4)_2SO_4$, 20 %glycerol. The protein solution contained 0.25 mM SAE, 0.25 mM $MgCl_2$, 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 [135] and in addition they are able to stabilize and activate enzymes [136].

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 $CaCl_2$; 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 µl enzyme solution and 1 µl 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.

4.5.2 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 Cu-K_α-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* [137–139] or *MOSFLM* [140]. *Scala* [141], which is integrated in the

Collaborative Computational Project No. 4 (CCP4)-Suite, was used for scaling of the intensities.

4.5.3 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 neccessary 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 [142]. Phases of the structures herein were exclusively determined by MR [143, 144]. MR was performed using the software *Phaser* [145, 146], which is included in the CCP4-Suite [147]. A previously published PFOMT structure (PDB-code: 3C3Y [33]) was used as a template during MR procedure for the PFOMT structure solution.

4.5.4 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 [148]. Structure refinement was done using the software REFMAC5 [149, 150] as part of the CCP4-suite or the Phyton-based Hierarchial Environment for Integrated Xtallography (PHENIX) [151]. Validation of the structures was carried out using the web service MolProbity (*http:// molprobity.biochem.duke.edu/*) [152]. Structure visualization and the preparation of figures was performed using PyMOL (Schrödinger, New York, USA) and UCSF Chimera (*http:// www.cgl.ucsf.edu/ chimera*)[116].

4.5.5 *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/*) [153–155]. 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.

4.6 Analytics

4.6.1 Recording of growth curves

Starter cultures (≈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_{600} was measured. Three independent 50 ml cultures of the medium containing the appropriate antibiotics were inoculated such that $OD^{600} \approx 0.05$ using the washed cell suspension. The cultures were incubated at the conditions to be studied and sampled at appropriate intervals of time (≈1 h). One ml samples were kept on ice until all samples were aquired. 100 µl aliquots of the samples were transferred into a clear MTP and the OD_{600} was measured.

Green fluorescent protein (GFP) fluorescence was measured accordingly, but the MTP used was opaque. Exitation (λ^{ex}) and emission (λ^{em}) wavelengths were (470 and 510) nm respectively.

4.6.2 *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 [156]. 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_2O_2 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 [157].

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 4.3). 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.

4.6.3 *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 $MgCl_2$, 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. After centrifugation (10 000 × *g*, 4 °C, 10 min) the organic phase was transferred into a new tube. The reaction was extraced 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 AG, Hamburg, Germany). The residue was dissolved in methanol and centrifuged at 10 000 × *g* for 10 min to remove unsoluble matter. The supernatant was transferred into a HPLC vial

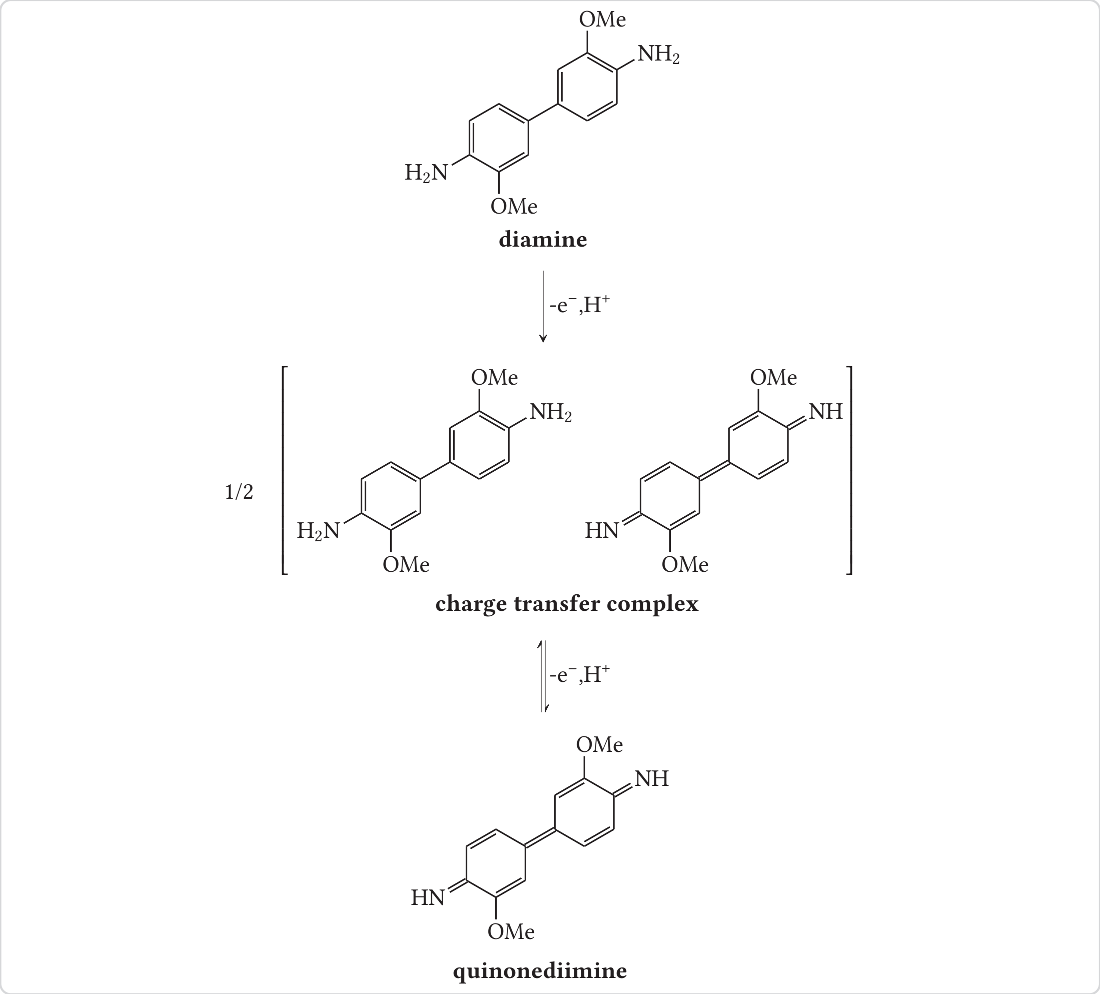

OMe
NH2
H2N
OMe
diamine
-e⁻,H⁺
1/2
OMe
NH2
H2N
OMe
OMe
NH
HN
OMe
charge transfer complex
-e⁻,H⁺
OMe
NH
HN
OMe
quinonediimine

Figure 4.3.: 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. [157, 158]

and analyzed by HPLC (4.6.8). When detection of hydrophobic (e.g. flavonoids) and hydrophilic compounds (e.g. SAM, *S*-adenosyl-L-homocysteine (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 × *g*, 4 °C, 10 min) the supernatant was transferred into HPLC-sample vials and analyzed (see 4.6.8).

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 [159]. 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. $MgCl_2$ was either omitted or added at 10 mM to assess the influence of divalent cations. Assays were stopped as described in 4.6.3 and analyzed accordingly.

Estimation of product concentration and enzymatic activities

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

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

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

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

and

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

The amount n is obtained by multiplying the concentration c by the injection volume V. Enzymatic activities (i.e. initial rates) can be calculated from the concentrations by standard procedures.

4.6.4 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 [160, 161]. Hence, caffeic acid can complex ferric (Fe^{3+}) ions and form a colored complex with $\lambda_{max} = 595$ nm [162]. Since the colored complex is specific for caffeic acid and not observed with methylated derivatives (i.e. ferulic and iso-ferulic acid), this can be used as a measure for methylation reactions. *O*-MT assays were prepared as before (subsection 4.6.3). 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 $FeCl_3$ 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.

4.6.5 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 refolding buffer using hydrophobic interaction chromatography (HIC). The ammonium sulfate concentration of the protein sample was brought to 1 M using a 2 M $(NH_4)_2SO_4$ 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 $(NH_4)_2SO_4$, 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 $(NH_4)_2SO_4$, 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 desalted using TCA precipitation (4.4.4).

4.6.6 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 (≥1 mg/ml) protein sample were injected. The Gel Filtration Standard by Bio-Rad (München, Germany) was run seperately to assess the size of the proteins in the analyzed sample.

4.6.7 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 [163]. ITC measurements to describe the interaction between PFOMT and its substrates/effector where 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 susequently centrifuged (14 000 × 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.

4.6.8 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 [164]. 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 (4.6.3) 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.

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

The positive and negative ion high resolution electrospray ionization (ESI) and collision induced dissociation (CID) MSn 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_2O:CH_3CN 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.

5 Engineering of phenylpropanoid and flavonoid *O*-methyl transferase (PFOMT)

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

Benjamin Weigel[1,a], **Martin Dippe**, **Ludger A. Wessjohann**[1,c]

Contact: bweigel@ipb-halle.de[a], wessjohann@ipb-halle.de[c]

Affiliation: Leibniz-Institute of Plant Biochemistry, Department of Bioorganic Chemistry[1]

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 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.

5.1 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 quantifyable psychotomimetic effect of methylated and ethylated lysergic acid amids differ by at least an order of magnitude [165, 166]. There are many more of these so-called structure activity relationship (SAR) and quantitative structure activity relationship (QSAR) studies on any number of compounds [167–169].

Methylation reactions are one of the key tailoring steps during natural product biosynthesis and can in consequence greatly affect a molecules bio- and physicochemical behavoir [34, 170]. Methyl transferases (MTs) catalyze the transfer of a methyl group from the co-substrate SAM to an activated atom of the acceptor molecule [34].

Between the highly complex core structures of natural products, which are produced by a plethora of enzymes (e.g. polyketide 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.[1]

It has recently been shown however, that a wide array of SAM analogues are used as co-substrates by a variety of MTs [34]. The majority of the work so far has been done on protein methyl transferases (P-MTs) and DNA methyl transferases (DNA-MTs), 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 [99, 111–114]. There have been a great many of SAM analogues synthesized, both chemically and enzymatically, that were consequently studied with the help of MTs [34, 102, 113].

The *O*-methyl transferase (*O*-MT) PFOMT is a highly promiscuous enzyme with regards to its flavonoid substrates and has extensively been characterized [33, 65, 78, 79]. However, the promiscuity towards different SAM analogues has not 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.

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.

---

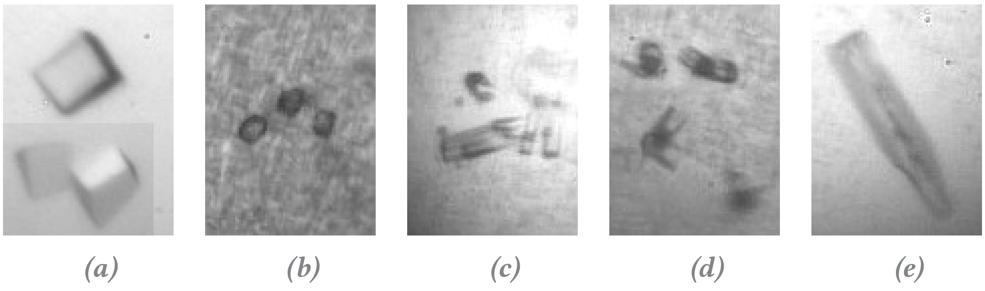
[1] Reaxys searches for natural product isolates with a molecular mass between (150 and 1500) containing the substructures methyl, ethyl or propyl connected to a heteroatom return 66759, 2797 and 52 results respectively. However, it stands to note that 70 % of the propyl results were either esters or otherwise activated moieties. [171]

5.2 Crystallization of PFOMT

The crystal structure of PFOMT was published in 2008, however binding of substrates could not be accomplished [33]. 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 [33]. However, reproduction of these results could not be accomplished and new crystallization conditions had to be found.

Several commercially available buffer solutions (see section 4.5) 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 ($CaCl_2$, PEG-4000) to brittle and ragged needles (LiCl, PEG-6000) (Figure 5.1).



(a) (b) (c) (d) (e)

Figure 5.1.: *Some crystal and pseudo-crystal shapes that were observed during the crytsallization screen. a – high $(NH_4)_2SO_4$, b-c – $CaCl_2$, PEG-4000, e – LiCl, PEG-6000*

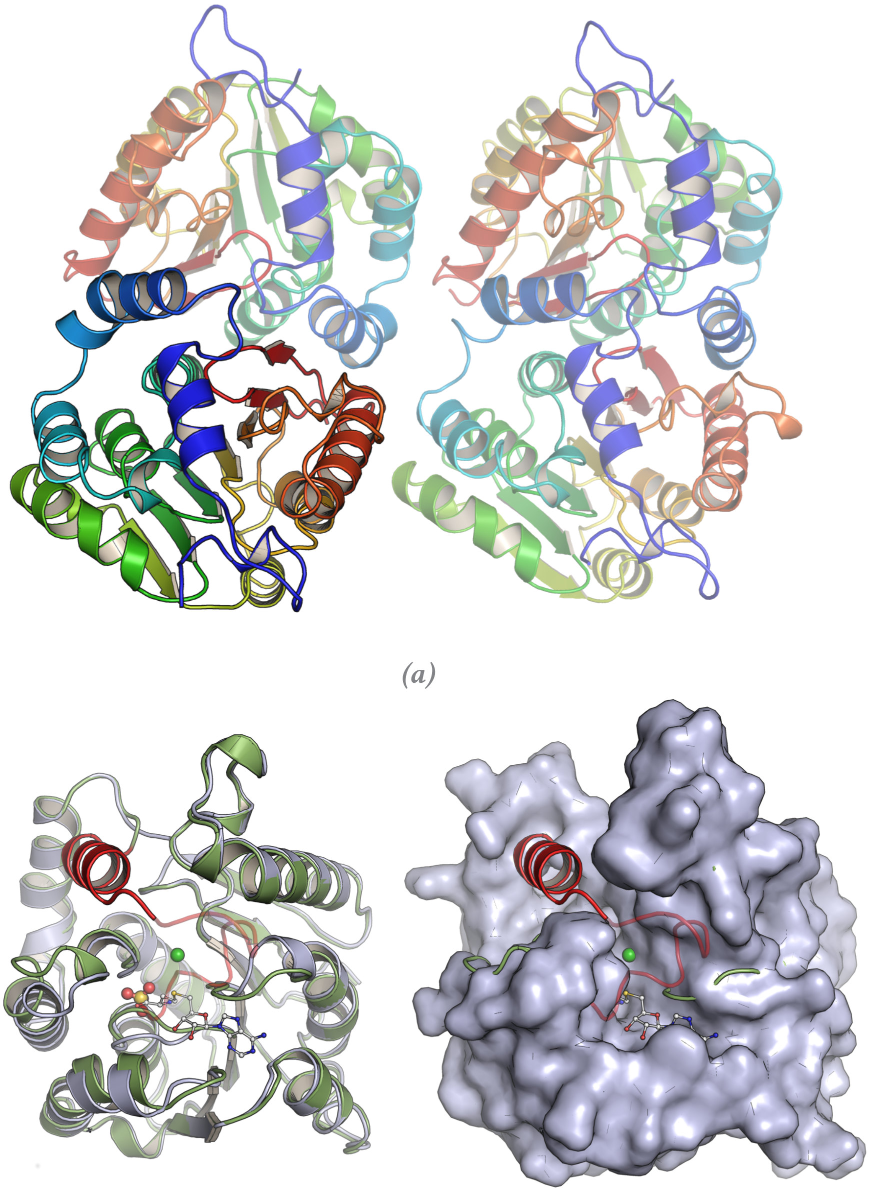
Crystals that were large enough (≥ 50 µm), where screened for diffraction at the home-source after cryoprotection. A rough estimate of the resolution, cell parameters and the space group was aquired, if the diffraction images could be indexed. The screened crystals all had similar cell parameters and belonged to the same space group, $P2_12_12_1$, as the previously published structure (pdb: 3C3Y)[33]. However, the unit cell of crystals that grew out of high ammonium sulfate concentrations (≥1.8 M) was approximately four times as large as that of the published structure. Several datasets were collected of crystals from high $(NH_4)_2SO_4$, 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 $(NH_4)_2SO_4$. One dataset was solved to completion to obtain a complete structure of this novel *apo*-PFOMT at a resolution of 1.95 Å (Table A.1). The assymetric unit of *apo*-PFOMT contained two homodimers (4 monomers) (Figure 5.2a), 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 are located in the 3C3Y structure (Figure 5.2b). 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 α-helix, which was positioned in a cleft on the surface, where substrates may be bound [33]. 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 5.3 and Figure A.1) as is obvious for the loop region between β-sheet 1 and α-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.



(a)

(b)

Figure 5.2.: *An overview of the features in the* apo-*PFOMT structure.* ***a*** *– The assymetric unit of apo-PFOMT consists of two homodimers (4 monomers). Individual monomers are rainbow colored from N- (blue) to C-terminus (red).* ***b*** *– Comparison of 3C3Y (steelblue) and apo-PFOMT (green). The N-terminus of apo-PFOMT (red) and even part of the His-tag (red, transparent) was 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^{2+} *(green sphere) are featured in the published structure, whereas a sulfate ion (red/yellow spheres) was bound in the newly solved structure.*

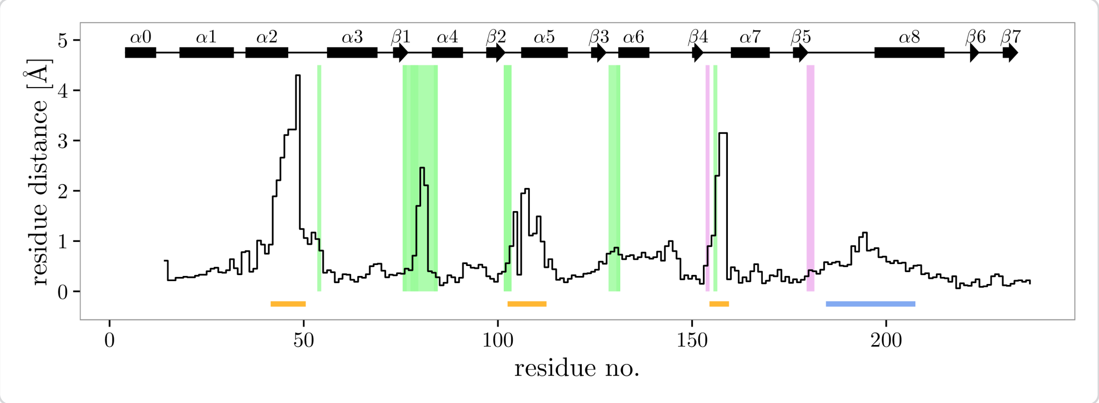


Figure 5.3.: Positional differences between the individual residues of the solved apo-PFOMT and the structure with bound SAH (pdb: 3C3Y). The diffraction precision indicator [172] (DPI) of the structures was (0.137 *and* 0.064) Å *respectively. The overall rmsd amounted to* 0.9034 Å*. The secondary structure 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 [33].*

5.3 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 5.4). Furthermore the binding of the substrate caffeic acid and the influence of Mg^{2+} 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 5.5a). 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 5.1). Thus, the entropic influence must get larger with increasing chain length in order for equations (5.1) and (5.2) to still hold true.

$$\Delta G = \Delta H - T\Delta S \tag{5.1}$$

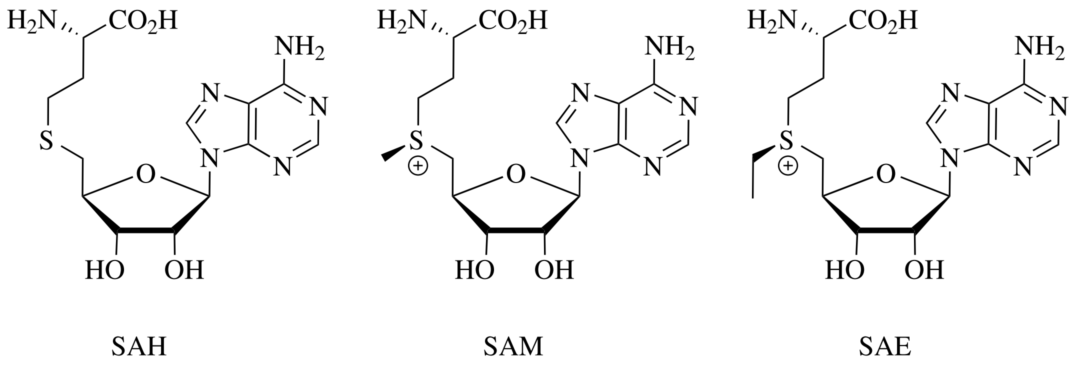
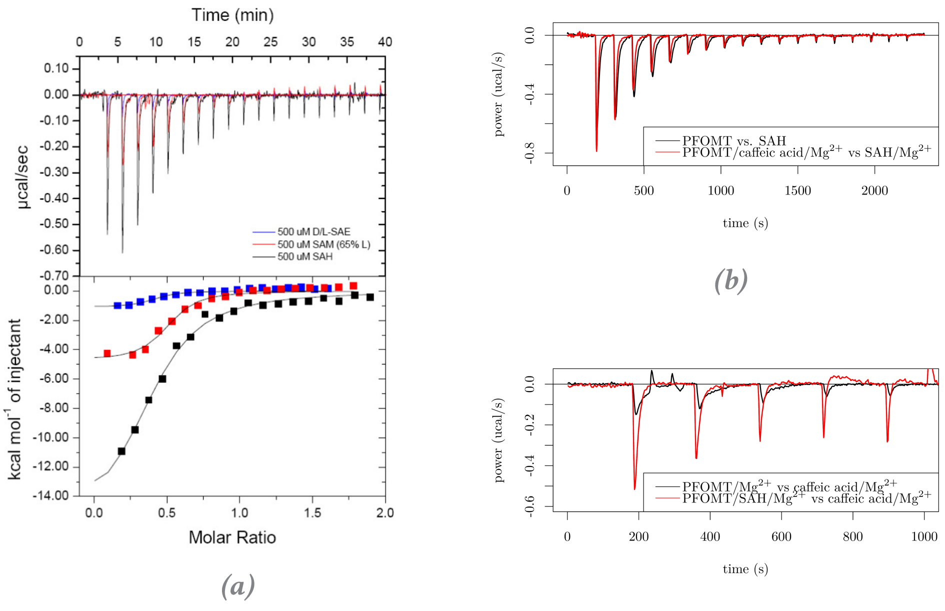


Figure 5.4.: The binding of different SAM analogues was measured via ITC

$$\Delta G = \Delta G^0 - RT \ln K \tag{5.2}$$

Indeed, the value for ΔS was negative for binding of SAH, but positive for the binding of SAM and SAE (Table 5.1). This relationship between the change of entropy and the change of enthalpy has been found for many biological systems and is called enthalpy-enthropy compensation (EEC) [173–175]. 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 ligand molecule per dimer of PFOMT (Table 5.1).

Upon titration of caffeic acid to PFOMT small amounts of released heat were detected for the system (Figure 5.5c). 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 5.5b). 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.



(a)

(b)

(c)

Figure 5.5.: 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 5.1.: Results of fitting a simple one-site binding model to the data obtained from ITC experiments.

	K_D [µM]	ΔH [cal mol^{-1}]	ΔS [cal mol^{-1} K^{-1}]	N
SAH	2.06 ± 4.27	−10 380 ± 1025	−9.41	0.505 ± 0.038
SAM	1.08 ± 3.50	−4606 ± 242	11.6	0.492 ± 0.018
SAE	2.22 ± 3.79	−1338 ± 190	21.3	0.513 ± 0.050

5.4 Study of variants for long-chain alkylations

Since the ability to bind the enlongated 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 [104]. 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 5.6). 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 5.7). 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 or charged residues in the vicinity of the acceptor substrates (Tyr51, Trp184 and Phe198). However, this was not a general trend since the 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

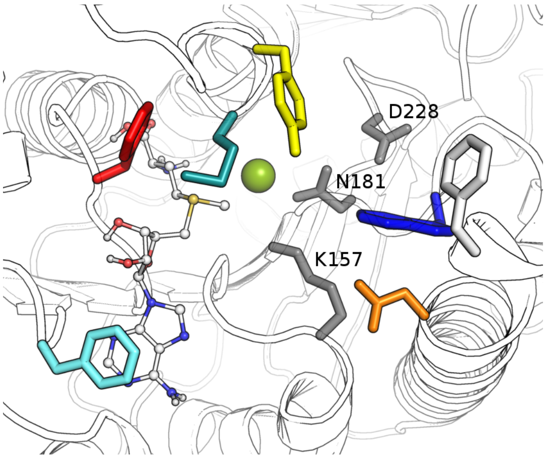


Figure 5.6.: *The active site of PFOMT (pdb: 3C3Y). The outline of the protein backbone is displayed, with active site residues portrayed as colored sticks (cyan – F103, red – F80, turqouise – M52, yellow – Y51, white – F198, blue – W184, orange – N202, grey – as labelled). The co-substrate SAM (ball-and-stick model) was docked into the structure.*

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.2). 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.

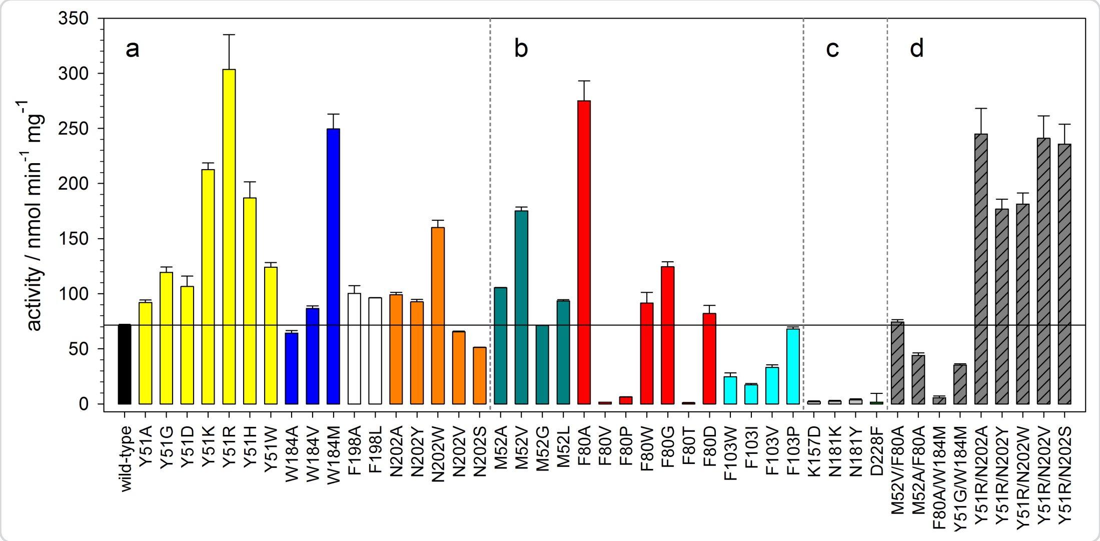
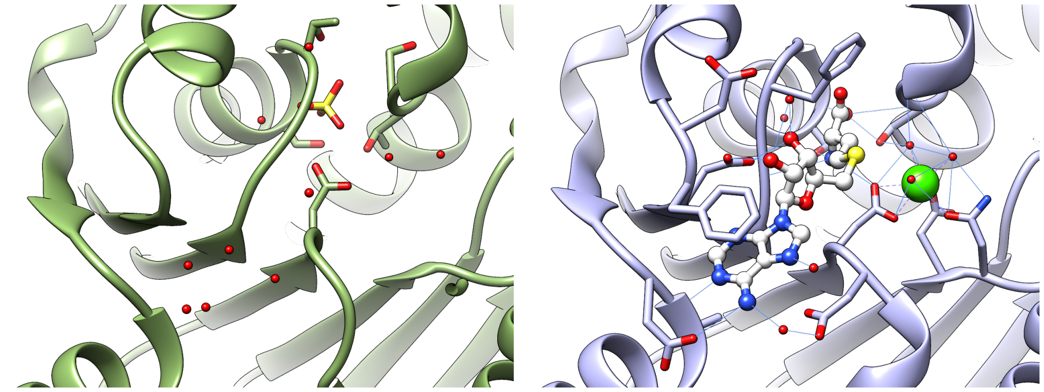


Figure 5.7.: Activities of different PFOMT variants towards caffeic acid methylation. Colorations correspond to the ones used in Figure 5.6.

5.5 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 ΔH was close to 0 (Table 5.1). 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 [176]. 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^{2+} titration via ITC did not afford significant signals, suggesting the notion, that the metal is only bound along with the co-substrate (Figure 5.8). 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 $\mathrm{cal\,mol^{-1}\,K^{-1}}$ [177]. 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 5.8). 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 promixal to the active site pocket is a main contributer to the overall rmsd of 0.9 Å (Figure 5.3).



(a)

(b)

*Figure 5.8.: 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 [33, 79]. 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’⟶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.

5.6 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.

6 Tandem mass-spectrometry studies of flavonoids

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

Benjamin Weigel[1,a], **Annegret Laub**[1,b], **Jürgen Schmidt**[1,c], **Ludger A. Wessjohann**[1,d]

Contact: bweigel@ipb-halle.de[a], alaub@ipb-halle.de[b], jschmidt@ipb-halle.de[c], wessjohann@ipb-halle.de[d]

Affiliation: Leibniz-Institute of Plant Biochemistry, Department of Bioorganic Chemistry[1]

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 availabilty 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.

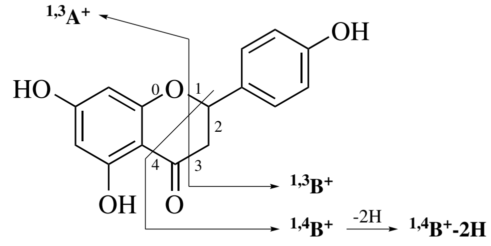
6.1 Introduction

Liquid chromatography-tandem mass spectrometry (LC-MS/MS) has been widely used for the 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 [178–181].

Ionization of samples in LC-MS/MS instruments is usually achieved by soft methods operating at atmospheric pressure, such as electrospray ionization (ESI) [182] or atmospheric pressure chemical ionisation (APCI) [183]. However, small molecules rarely produce fragment ions under these conditions and usually only the M+H]$^+$ or M-H]$^-$ 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 ionization (EI), where the analytes are bombarded with electrons, for ionization. However, EI is operating under high-vacuum and the coupling with liquid chromatography (LC)-systems is not trivial [184]. 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 [185].

Flavonoids comprise a huge chemical space, with thousands of theoretical structures [186]. 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 [187–196]. Researchers have reported that specifc patterns of fragmentation along the C-ring can be observed for different classes of flavonoids and can help differentiate between them [188, 194]. 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 [188, 194].

Fragments of flavonoid aglycones can be represented by a systematic nomenclature first proposed by Ma *et al.* [194]. The labels $^{i,j}A^+$ and $^{i,j}B^+$ 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 (Scheme 6.1). Our group currently works with methyl transferases that act on flavonoids. Identifying the site of methylation is a crucial step in identifying the product of an enzymatic methylation. MS/MS has been shown to be a rather quick and reliable method to identify characteristic key ions of flavonoids, that can help identify the localization of different functional groups [188, 189, 194, 197, 198]. In this work the complementarity of two activation methods, CID and higher-energy collisional dissociation (HCD), for the structural characterization of flavonoids (Table 6.1), especially those methylated at the B-ring, in positive ionization mode was evaluated. An specific array of different flavonoids (Table 6.1) was studied, to get a holistic impression of the



Scheme 6.1: *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.*

fragmentations of these compounds.

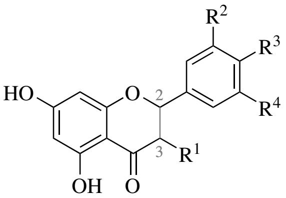


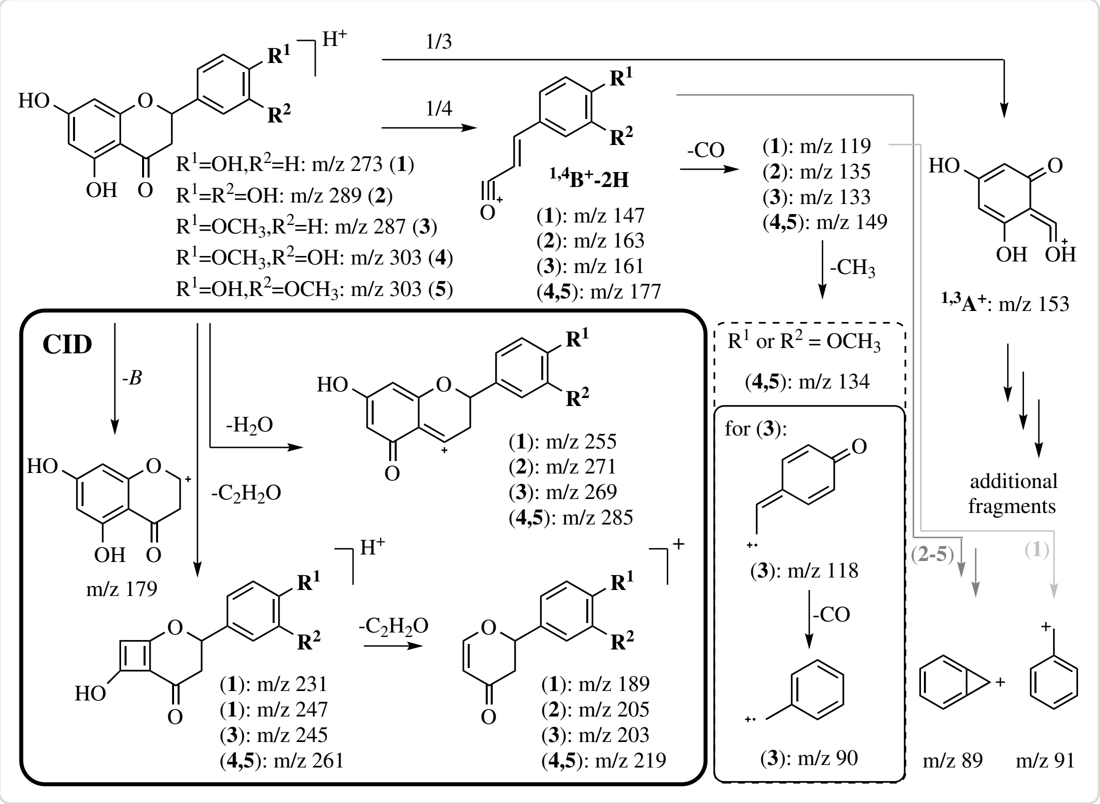
Table 6.1.: *Substrates studied in this work. Three classes of flavonoids were tested: flavanones (**1-5**), flavones (**6-10**) and flavonols (**11-15**). The topology of the bond between C2 and C3 in the C-ring specifying flavanones or flavones/flavonols is denoted with - (single) or = (double), respectively.*

	name	$[M+H]^+$	C2-C3	R^1	R^2	R^3	R^4
1	naringenin	273	-	H	H	OH	H
2	eriodictyol	289	-	H	OH	OH	H
3	ponciretin	287	-	H	H	OCH_3	H
4	hesperetin	303	-	H	OH	OCH_3	H
5	homoeriodictyol	303	-	H	OCH_3	OH	H
6	apigenin	271	=	H	H	OH	H
7	luteolin	287	=	H	OH	OH	H
8	acacetin	285	=	H	H	OCH_3	H
9	diosmetin	301	=	H	OH	OCH_3	H
10	chrysoeriol	301	=	H	OCH_3	OH	H
11	kaempferol	287	=	OH	H	OH	H
12	quercetin	303	=	OH	OH	OH	H
13	myricetin	317	=	OH	OH	OH	OH
14	kaempferide	301	=	OH	H	OCH_3	H
15	isorhamnetin	317	=	OH	OCH_3	OH	H

6.2 Fragmentation of flavanones

Positive ionization MS^2 spectra of flavanones (Table B.1) are mostly characterized by a base peak at *m/z* 153, which corresponds to the A-ring fragment $^{1,3}A^+$ of the flavonoid skeleton (Scheme 6.2). In contrast, negative mode MS^2 spectra of 3,7-dihydroxy flavanones show an *m/z* 151, which correspond to the negatively charged $^{1,3}A^-$ ion [189]. 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**) (Scheme 6.2). 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 (Scheme 6.2) [199].

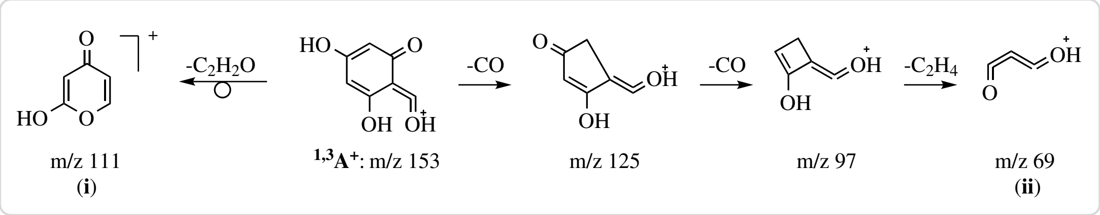
Fragment ions from cleavage of the C-ring ($^{1,3}A^+$ and $^{1,4}B^+$-2H)) are further decomposed under the higher energy conditions in HCD experiments. Thus, the resulting HCD spectra generally display smaller *m/z* than the CID spectrum (Figure 6.1). Increasing the normalized collision energy (NCE) 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 6.1).



H+
R1
R2
HO
O
OH O
1/3
1/4
R1=OH,R2=H: m/z 273 (1)
R1=R2=OH: m/z 289 (2)
R1=OCH3,R2=H: m/z 287 (3)
R1=OCH3,R2=OH: m/z 303 (4)
R1=OH,R2=OCH3: m/z 303 (5)
1,4B+-2H
(1): m/z 147
(2): m/z 163
(3): m/z 161
(4,5): m/z 177
-CO
(1): m/z 119
(2): m/z 135
(3): m/z 133
(4,5): m/z 149
-CH3
R1 or R2 = OCH3
(4,5): m/z 134
for (3):
(3): m/z 118
-CO
(3): m/z 90
1,3A+: m/z 153
additional fragments
(2-5)
(1)
m/z 89
m/z 91
CID
-B
m/z 179
-H2O
(1): m/z 255
(2): m/z 271
(3): m/z 269
(4,5): m/z 285
-C2H2O
(1): m/z 231
(1): m/z 247
(3): m/z 245
(4,5): m/z 261
-C2H2O
(1): m/z 189
(2): m/z 205
(3): m/z 203
(4,5): m/z 219

Scheme 6.2: Major fragmentation pathways of flavanones. Activation using CID conditions at 45 % NCE mainly results in neutral losses of H_2O and ketene (C_2H_2O) from the molecular ion $[M+H]^+$ (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.

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 at



Scheme 6.3: Proposed MS^2 fragmentation of $^{1,3}A^+$ after HCD activation. In high energy MS^2 experiments, $^{1,3}A^+$ might loose two CO followed by an unusual C_2H_4. A single loss of ketene (C_2H_2O) to afford m/z 111 is also sensible.

m/z 145. However, the loss of CO is the most prominent decomposition of (^{1,4}B$^+$-2H). The intensities of the peaks corresponding to the (^{1,4}B$^+$-2H-CO) fragment were as high as 36 % in HCD experiments (Figure 6.1). 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 benzylium or tropylium cation (Scheme 6.2). Decay of (^{1,4}B$^+$-2H) of the other flavanones likely 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}B$^+$-2H-CO-CH3$^{\cdot}$), 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 benzylium/tropylium radical cation (Scheme 6.2). It is proposed, that ion ^{1,3}A$^+$ can decompose via two different pathways under HCD conditions (Scheme 6.3). A loss of ketene from ^{1,3}A$^+$ results in *m/z* 111. Pyranone (**i**) is suggested as a structure for this ion. Sequential losses of two CO and a C_2H_4 could afford ion (**ii**). However, further MS^n experiments are necessary to confirm these proposals.

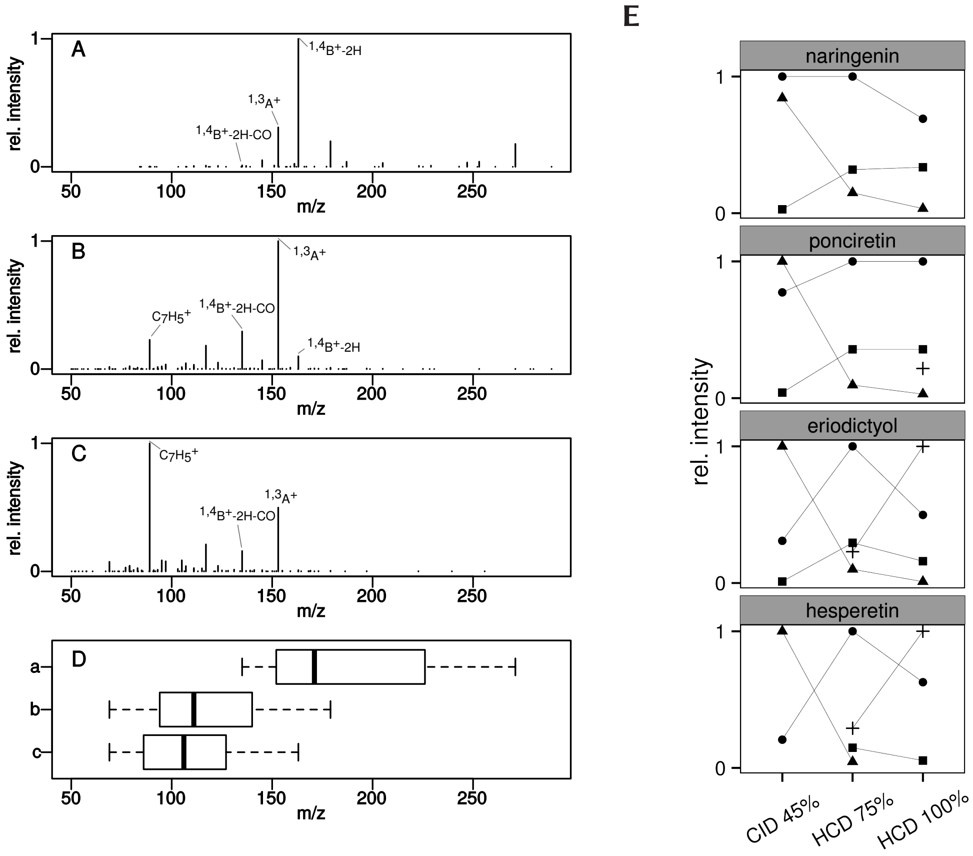
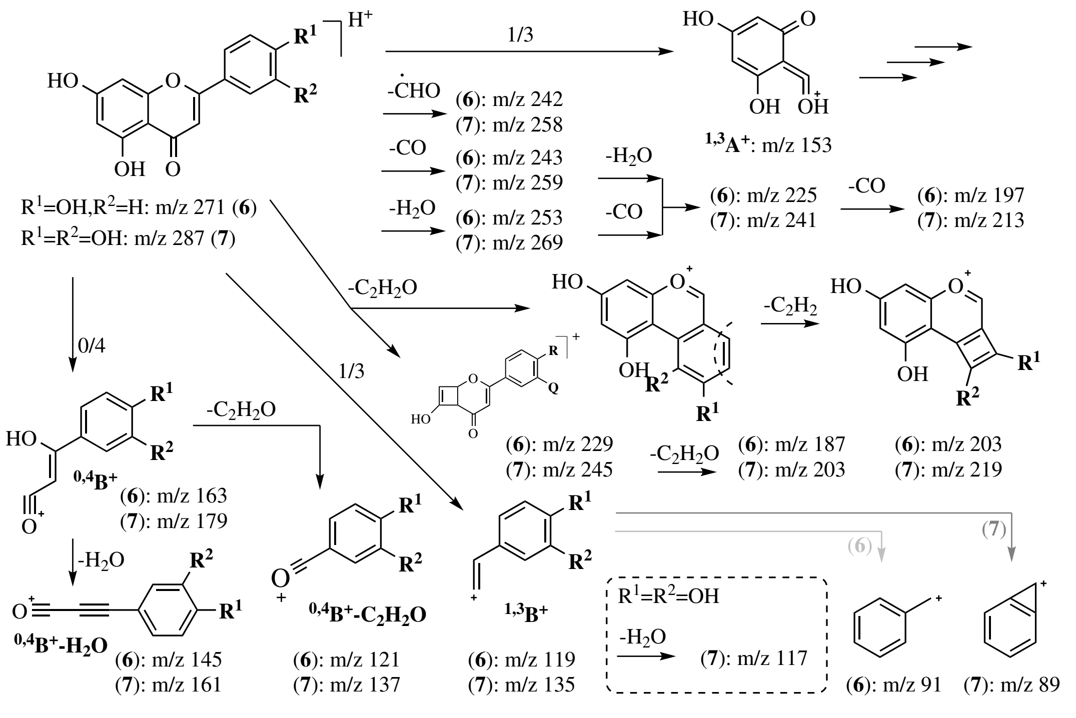


*Figure 6.1.: Comparison of CID and HCD MS^2 spectra of eriodictyol (**2**). **A** – CID at 45 % NCE. **B** – HCD at 75 % NCE. **C** – HCD at 100 % NCE. Four different prominent peaks are annotated in each spectrum. **D** – The shift to smaller masses in HCD spectra and with increasing NCE is illustrated by the boxplot of the distribution of peaks with relative intensities above 1 %. Labels a to c correspond to spectra A to C, respectively.*
***E** – Relationship between the activation method and the intensity of four fragments (● $^{1,3}A^+$, ▲ ($^{1,4}B^+$-2H), ■ ($^{1,4}B^+$-2H-CO), + $C_7H_5^+$) of different flavanones.*

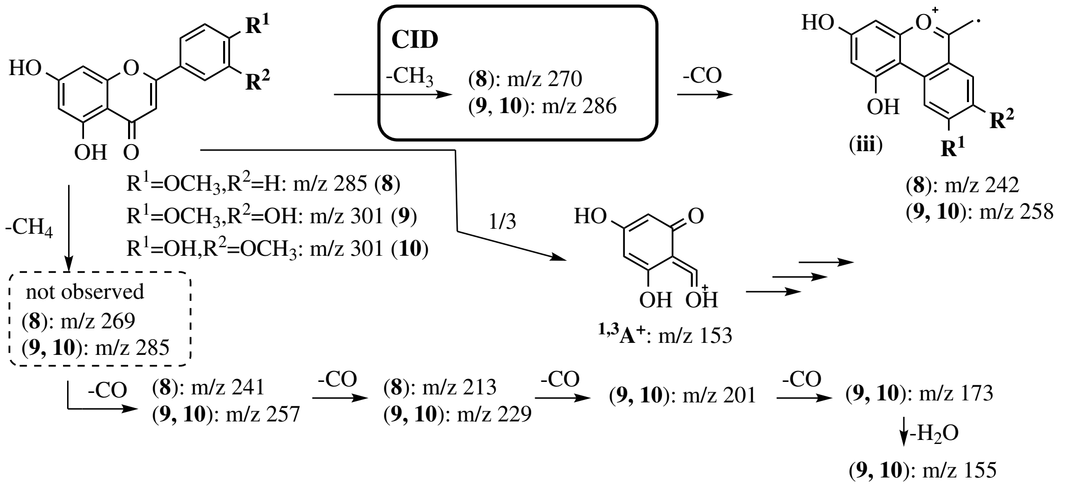
6.3 Fragmentation of flavones

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 [194, 197]. Non-methylated (**6**, **7**) and methylated flavones (**8** – **10**) show significantly different MS^2 spectra (Table B.2). Apigenín (**6**) and luteolin (**7**) MS^2 spectra show a characteristic *m/z* 153, corresponding to the $^{1,3}A^+$ ion, as a base peak in CID mode and at low activation energies in HCD mode (Scheme 6.4). Contrary to the flavanones, the MS^2 of non-methylated flavones show the peak corresponding to the molecular ion $[M+H]^+$, which is strongest in HCD at NCE of 75 %. Characteristic neutral losses of water, CO and ketene (C_2H_2O) were also observed for (**6**) and (**7**) (Scheme 6.4, Table B.2). MS-peaks corresponding to a loss of a formyl radical, resulting in $[M+H\text{-}CHO]^{\bullet+}$ 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 C_2H_2 might be explained (Scheme 6.4). Besides the characteristic $^{1,3}A^+$ fragment, apigenin (**6**) and luteolin (**7**) MS^2 spectra also present peaks corresponding to the B-ring fragments $^{1,3}B^+$ (*m/z* 119 and 135) and $^{0,4}B^+$ (*m/z* 163 and 179). From the mass differences of these fragments, the substitution on the B-ring can be deduced. The $^{0,4}B^+$ 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 $^{1,3}B^+$ in a fashion similar to the flavanones to afford a benzylium or bicycloheptatrienyl cation respectively (Scheme 6.4).

The most noteable difference between the methylated and non-methylated representatives is the almost complete lack of any fragmentation of the methylated flavones other than a methyl loss, in CID experiments (Table B.2, Figure 6.2). 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 (Scheme 6.5). Any other loss 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 $[M+H\text{-}CH_3]^{\bullet+}$ ion, while the base peak *m/z* 153 at 100 % NCE likely corresponds to the $^{1,3}A^+$ ion (Figure 6.2). Further losses from $[M+H\text{-}CH_3\text{-}CO]^{\bullet+}$, with the proposed structure of a benzochromenylium radical cation (**iii**), were not observed


H+
HO
O
OH O
R1
R2
R1=OH,R2=H: m/z 271 (6)
R1=R2=OH: m/z 287 (7)
1/3
HO
O
OH OH
1,3A+: m/z 153
-CHO
(6): m/z 242
(7): m/z 258
-CO
(6): m/z 243
(7): m/z 259
-H2O
(6): m/z 253
(7): m/z 269
-H2O
-CO
(6): m/z 225
(7): m/z 241
-CO
(6): m/z 197
(7): m/z 213
-C2H2O
1/3
0/4
-C2H2
(6): m/z 229
(7): m/z 245
-C2H2O
(6): m/z 187
(7): m/z 203
(6): m/z 203
(7): m/z 219
0,4B+
(6): m/z 163
(7): m/z 179
-C2H2O
0,4B+-C2H2O
(6): m/z 121
(7): m/z 137
1,3B+
(6): m/z 119
(7): m/z 135
-H2O
0,4B+-H2O
(6): m/z 145
(7): m/z 161
R1=R2=OH
-H2O
(7): m/z 117
(6)
(7)
(6): m/z 91
(7): m/z 89

(a) Fragments characteristic for non-methylated flavones.


HO
O
OH O
R1
R2
R1=OCH3,R2=H: m/z 285 (8)
R1=OCH3,R2=OH: m/z 301 (9)
R1=OH,R2=OCH3: m/z 301 (10)
CID
-CH3
(8): m/z 270
(9, 10): m/z 286
-CO
(iii)
(8): m/z 242
(9, 10): m/z 258
-CH4
not observed
(8): m/z 269
(9, 10): m/z 285
1/3
HO
O
OH OH
1,3A+: m/z 153
-CO
(8): m/z 241
(9, 10): m/z 257
-CO
(8): m/z 213
(9, 10): m/z 229
-CO
(9, 10): m/z 201
-CO
(9, 10): m/z 173
-H2O
(9, 10): m/z 155

(b) Fragments characteristic for methylated flavones.

Scheme 6.4: Major fragmentation pathways of non-methylated 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^2 spectra of non-methylated flavones. Methylated flavones loose a methyl group in CID experiments, but only in HCD experiments do other fragmentation reaction become obvious.

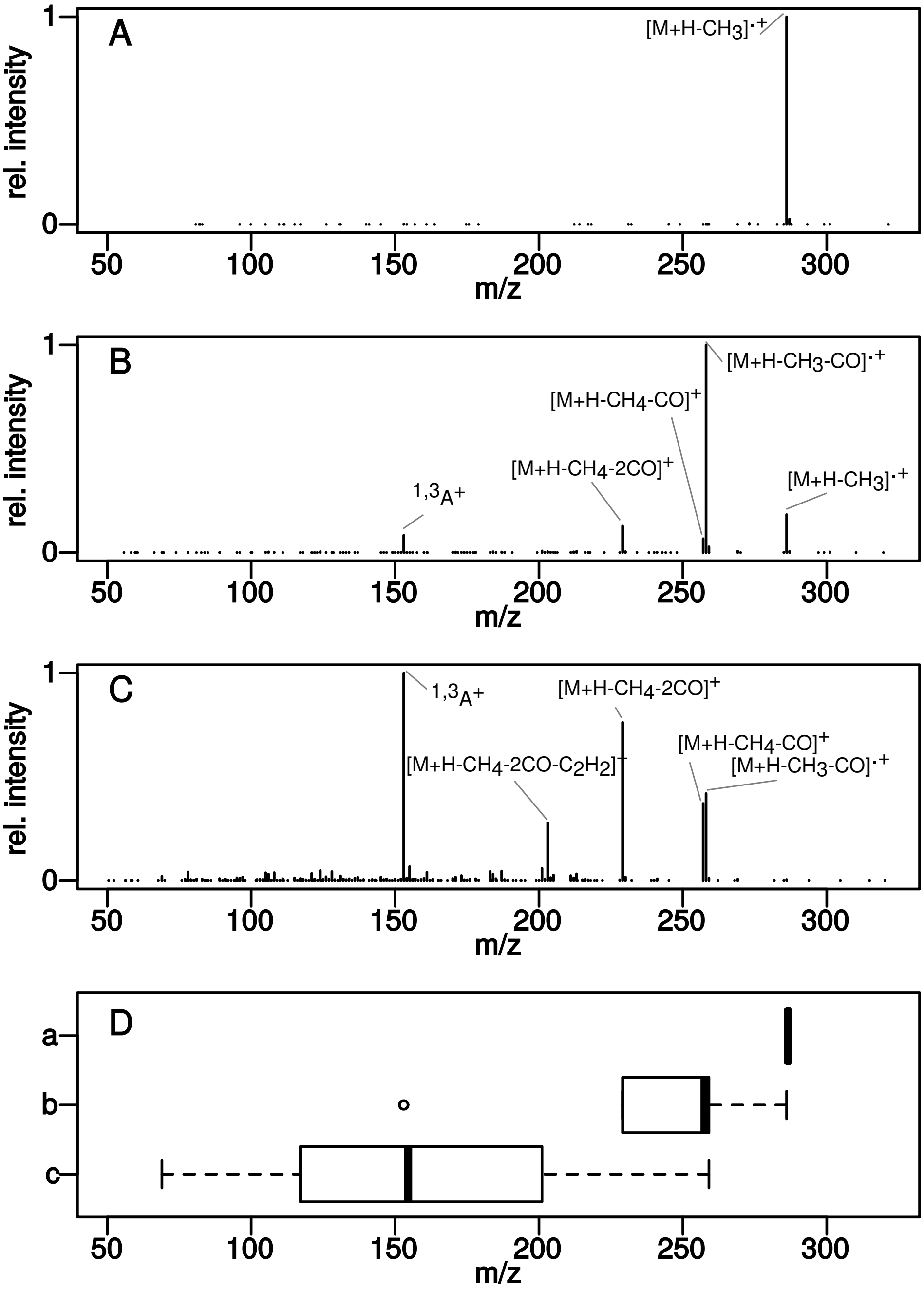
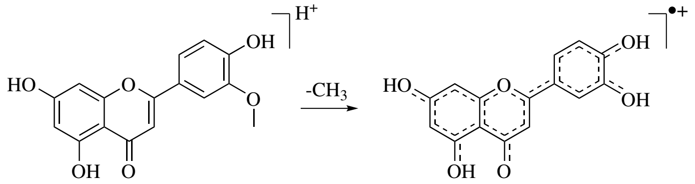


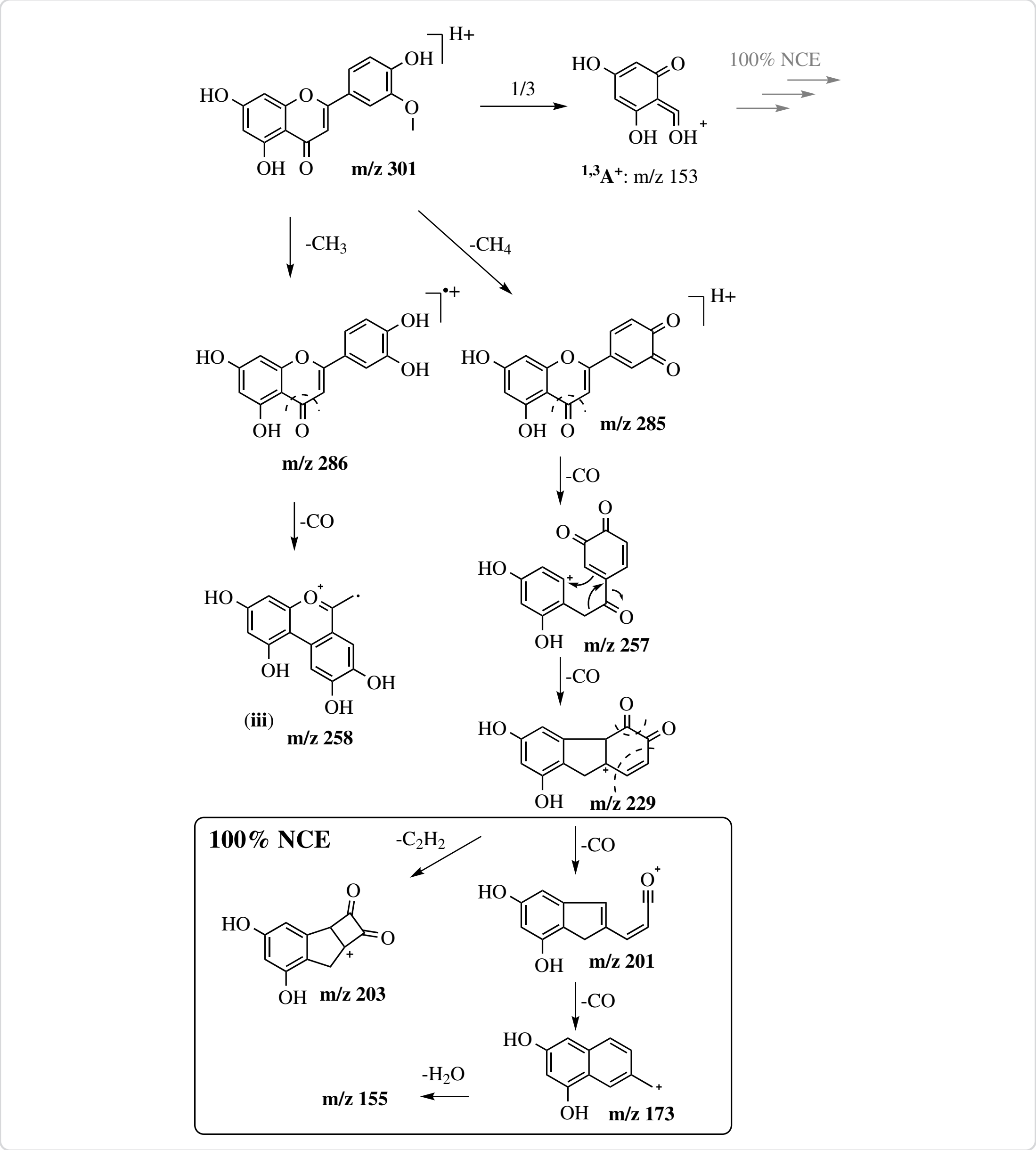
*Figure 6.2.: Comparison of CID and HCD MS^2 spectra of chrysoeriol (**10**). **A** – CID at* 45 % *NCE. **B** – HCD at* 75 % *NCE. **C** – HCD at* 100 % *NCE. Four different prominent peaks are annotated in each spectrum. **D** – The shift to smaller masses in HCD spectra and with increasing NCE is illustrated by the boxplot of the distribution of peaks with relative intensities above* 1 %. *Labels a to c correspond to spectra A to C, respectively.*



Scheme 6.5: Stability of the [M+H-CH_3]$^{\bullet+}$ ion of flavones. The [M+H-CH_3]$^{\bullet+}$ 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.

(Scheme 6.4, Table B.2). 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 (Scheme 6.4, Scheme 6.6). 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 Scheme 6.6. However, instead of additional CO losses, fragment [M+H-CH_3-2CO]$^{\bullet+}$ of (**10**) or (**9**) might as well loose a C_2H_2 (Scheme 6.6), as suggested by the MS^2 spectra (Table B.2). 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 (Scheme 6.3). Numerous minor peaks in the MS^2 HCD 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 6.2).

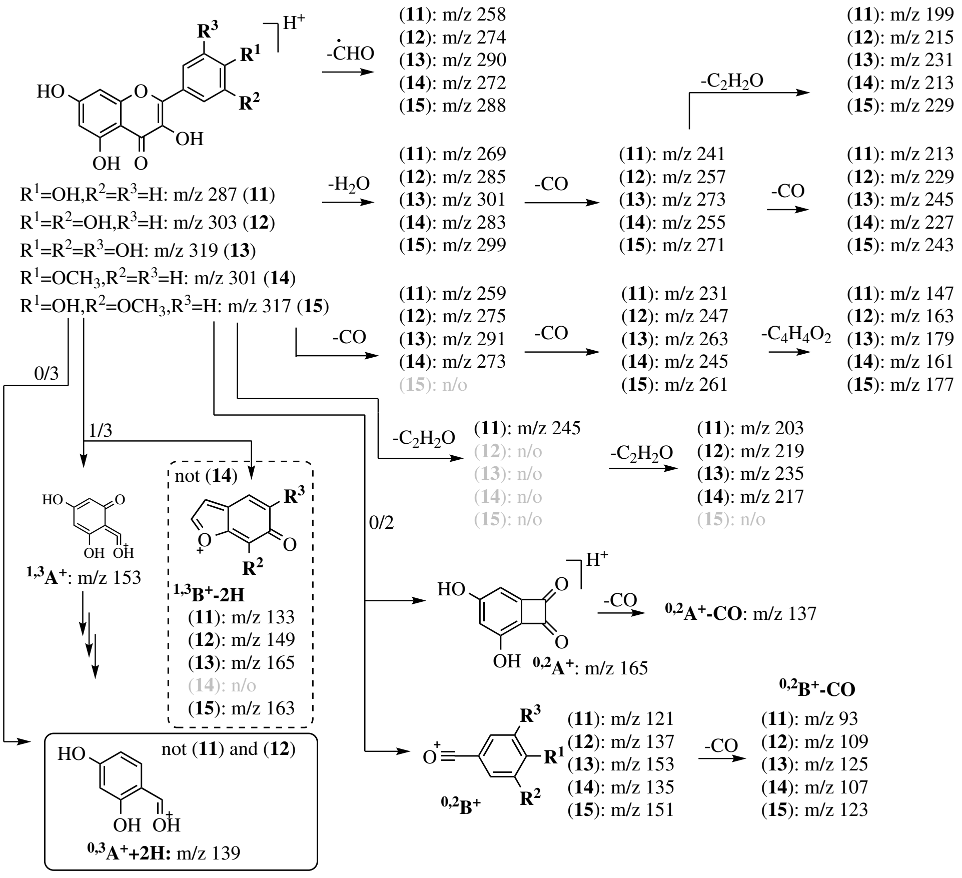

H+
m/z 301
1/3
$^{1,3}A^+$: m/z 153
100% NCE
-CH3
-CH4
•+
m/z 286
H+
m/z 285
-CO
m/z 257
-CO
(iii)
m/z 258
-CO
m/z 229
100% NCE
-C2H2
-CO
m/z 203
m/z 201
-CO
m/z 173
-H2O
m/z 155

*Scheme 6.6: Proposed pathway of fragmentation of (**10**) after HCD activation. Losses of $CH_3^\bullet$ and CH_4, followed by loss of CO are the major fragmentations observed in the corresponding MS spectra. However, multiple losses of CO only occur after a loss of methane (CH_4), possibly due to the relative stability of the benzochromenylium radical cation (**iii**). At* 100 % *NCE even higher order fragmentations were observed.*

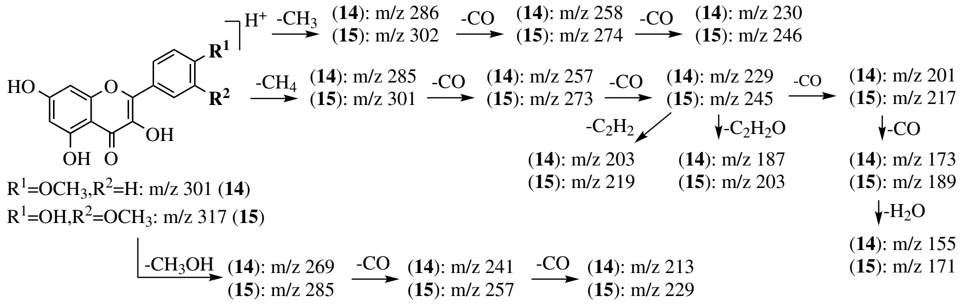
6.4 Fragmentation of flavonols

The principle fragmentation pathways of kaempferol (**11**), quercetin (**12**), myricetin (**13**) and isorhamnetin (**15**) in CID tandem mass spectrometry have been previously reported [194, 195, 200]. Other than flavones, methylated and non-methylated flavonols share similar fragment(ation)s. Whereas in CID, methylated flavones hardly showed any fragmentation beyond a methyl loss, the methylated flavonols kaempferide (**14**) and isorhamnetin (**15**) exhibited the same losses as their non-methylated couterparts, albeit at much lower levels (Table B.3, Scheme 6.7 and 6.3). These observations are in full agreement with previous reports [194] 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 Scheme 6.7 and 6.4). Many high intensity peaks presented in the MS^2 spectra of flavonols 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 (Scheme 6.7). 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 $[M+H-CH_3]^{\bullet+}$ 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 6.3). 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 (Scheme 6.7, Table B.3). Only for kaempferol (**11**), a neutral loss of 42 Da corresponding ketene was observed. However, MS^2 spectra of all flavonols, except (**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, flavonols can loose two sequential CO and another $C_4H_4O_2$, confirming previously published data [194]. 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 from the molecular ion are most abundant in CID experiments, whereas


H+
R3
R1
R2
HO
O
OH
OH O
R1=OH,R2=R3=H: m/z 287 (11)
R1=R2=OH,R3=H: m/z 303 (12)
R1=R2=R3=OH: m/z 319 (13)
R1=OCH3,R2=R3=H: m/z 301 (14)
R1=OH,R2=OCH3,R3=H: m/z 317 (15)
-CHO
(11): m/z 258
(12): m/z 274
(13): m/z 290
(14): m/z 272
(15): m/z 288
-C2H2O
(11): m/z 199
(12): m/z 215
(13): m/z 231
(14): m/z 213
(15): m/z 229
-H2O
(11): m/z 269
(12): m/z 285
(13): m/z 301
(14): m/z 283
(15): m/z 299
-CO
(11): m/z 241
(12): m/z 257
(13): m/z 273
(14): m/z 255
(15): m/z 271
-CO
(11): m/z 213
(12): m/z 229
(13): m/z 245
(14): m/z 227
(15): m/z 243
-CO
(11): m/z 259
(12): m/z 275
(13): m/z 291
(14): m/z 273
(15): n/o
-CO
(11): m/z 231
(12): m/z 247
(13): m/z 263
(14): m/z 245
(15): m/z 261
-C4H4O2
(11): m/z 147
(12): m/z 163
(13): m/z 179
(14): m/z 161
(15): m/z 177
-C2H2O
(11): m/z 245
(12): n/o
(13): n/o
(14): n/o
(15): n/o
-C2H2O
(11): m/z 203
(12): m/z 219
(13): m/z 235
(14): m/z 217
(15): n/o
0/3
1/3
0/2
HO O
OH OH
1,3A+: m/z 153
not (14)
R3
O
R2
1,3B+-2H
(11): m/z 133
(12): m/z 149
(13): m/z 165
(14): n/o
(15): m/z 163
H+
HO
O
O
OH
0,2A+: m/z 165
-CO
0,2A+-CO: m/z 137
HO
not (11) and (12)
OH OH
0,3A++2H: m/z 139
O
R3
R1
R2
0,2B+
(11): m/z 121
(12): m/z 137
(13): m/z 153
(14): m/z 135
(15): m/z 151
-CO
0,2B+-CO
(11): m/z 93
(12): m/z 109
(13): m/z 125
(14): m/z 107
(15): m/z 123

(a) Fragments characteristic for all studied flavonols.


H+
R1
R2
HO
O
OH
OH O
R1=OCH3,R2=H: m/z 301 (14)
R1=OH,R2=OCH3: m/z 317 (15)
-CH3
(14): m/z 286
(15): m/z 302
-CO
(14): m/z 258
(15): m/z 274
-CO
(14): m/z 230
(15): m/z 246
-CH4
(14): m/z 285
(15): m/z 301
-CO
(14): m/z 257
(15): m/z 273
-CO
(14): m/z 229
(15): m/z 245
-CO
(14): m/z 201
(15): m/z 217
-C2H2
(14): m/z 203
(15): m/z 219
-C2H2O
(14): m/z 187
(15): m/z 203
-CO
(14): m/z 173
(15): m/z 189
-H2O
(14): m/z 155
(15): m/z 171
-CH3OH
(14): m/z 269
(15): m/z 285
-CO
(14): m/z 241
(15): m/z 257
-CO
(14): m/z 213
(15): m/z 229

(b) Characteristic fragments of methylated flavonols.

Scheme 6.7: Major fragmentation pathways of flavonols. Unlike flavones, methylated and non-methylated flavonols 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 %).

the shift to smaller masses in HCD experiments is obvious (Table B.3, Figure 6.3).

The studied flavonols 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 (Scheme 6.3). Characteristic ring cleavage fragments of flavonols include $^{0,2}A^+$, $^{0,2}B^+$ and $^{1,3}B^+$-2H [194, 195], 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 6.3).

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**) (Scheme 6.8). 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 Scheme 6.8 and (Scheme 6.6)). 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 (Scheme 6.8 and 6.7). 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 inital loss of methane (Scheme 6.8). As mentioned before, the smaller mass fragments corresponding to multiple neutral losses are more pronounced at higher activation energies and were thus limited to HCD experiments at a NCE of 100 % (Figure 6.3, Scheme 6.8).

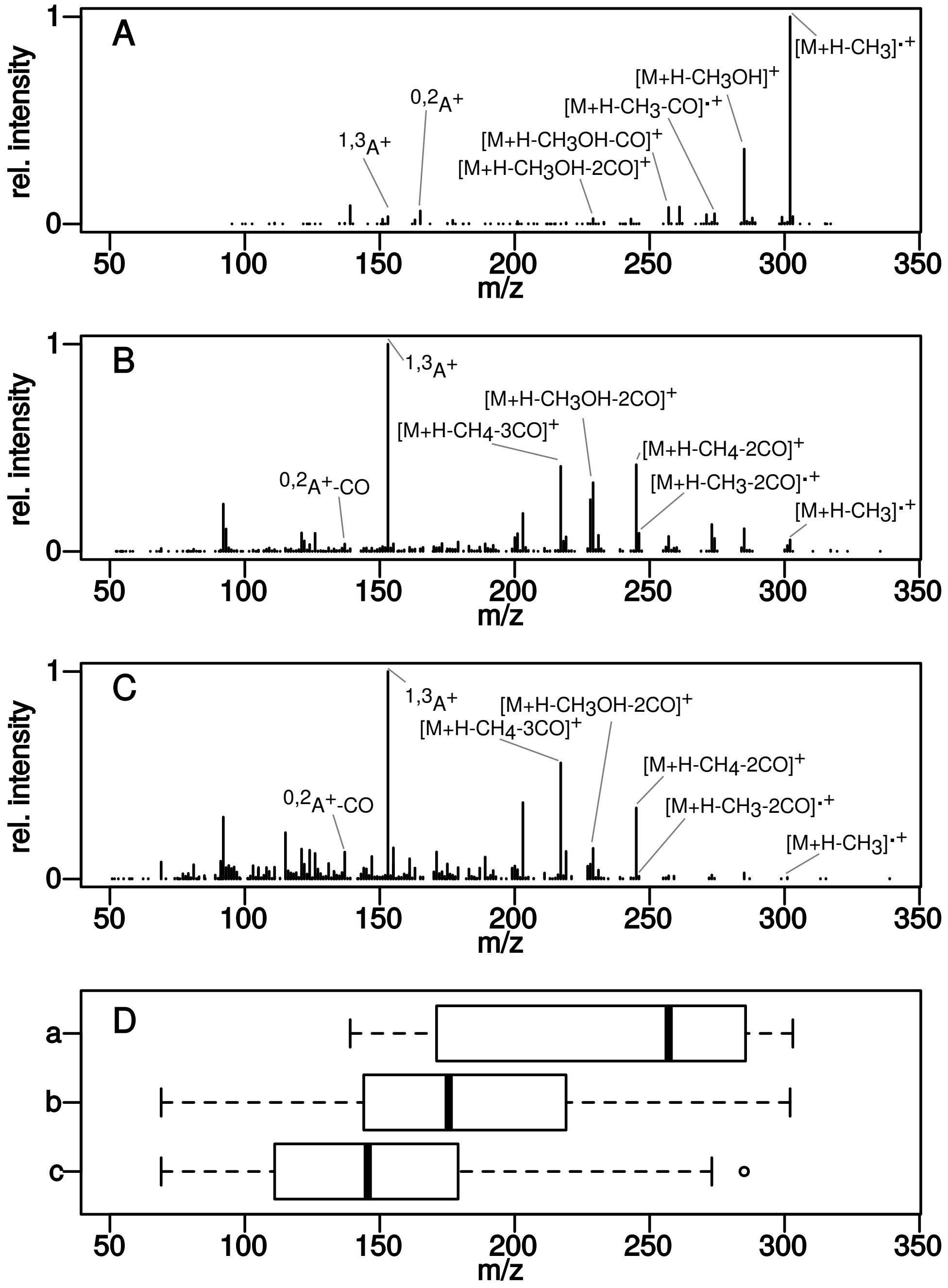
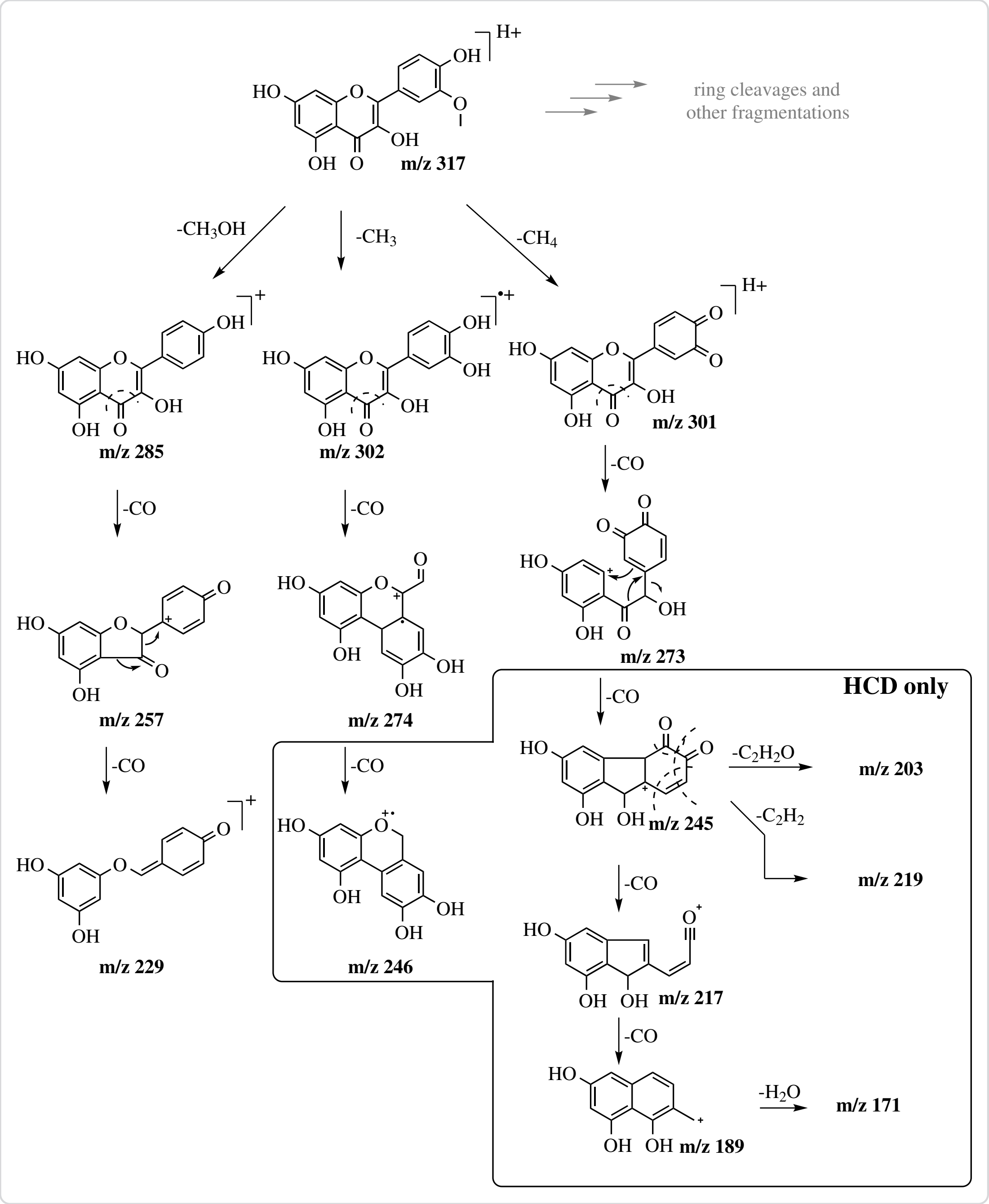


*Figure 6.3.: Comparison of CID and HCD MS^2 spectra of isorhamnetin (**15**). **A** – CID at 45 % NCE. **B** – HCD at 75 % NCE. **C** – HCD at 100 % NCE. Four different prominent peaks are annotated in each spectrum. **D** – The shift to smaller masses in HCD spectra and with increasing NCE is illustrated by the boxplot of the distribution of peaks with relative intensities above 1 %. Labels a to c correspond to spectra A to C, respectively.*


H+
m/z 317
ring cleavages and
other fragmentations
-CH3OH
-CH3
-CH4
m/z 285
m/z 302
m/z 301
-CO
m/z 257
m/z 274
m/z 273
HCD only
m/z 229
m/z 246
m/z 245
-C2H2O
m/z 203
-C2H2
m/z 219
m/z 217
m/z 189
-H2O
m/z 171

*Scheme 6.8: Proposed pathways of fragmentation of isorhamnetin (**15**). Isorhamnetin might loose methyl, methane or methanol upon activation. A similar fragmnentation pathway was proposed for the analogous chrysoeriol (Scheme 6.6). Some fragmentations were observed in HCD mode only (box).*

6.5 Conclusions

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 [201, 202] 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 MS^2 spectra from those of the flavanones. Distinguishing characteristics between MS^2 spectra of flavones and flavonols are the C-ring fragmentations, where the $^{0,4}B^+$ fragment was typically limited to flavones, whereas a (strong) $^{0,2}A^+$ fragment was only observed for (non-methylated) 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. Noticable loss of $CH_3^{\bullet}$ or 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^2 spectra of flavanones. Under the right conditions, all of the studied 5,7-dihydroxy substituted flavonoids presented a $^{1,3}A^+$ 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 $^{1,3}A^+$ 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 $^{1,4}B^{+}$ or $^{1,3}B^{+}$ 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.

6.6 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.

7 Enzymatic methylation of non-catechols

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

Benjamin Weigel[1,a], Martin Dippe[2,b], Annegret Laub[1,b], Ludger A. Wessjohann[1,d]

Contact: bweigel@ipb-halle.de[a], mdippe@ipb-halle.de[b], alaub@ipb-halle.de[c], wessjohann@ipb-halle.de[d]

Affiliation: Leibniz-Institute of Plant Biochemistry, Department of Bioorganic Chemistry[1]

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, $3^{(\prime)}$-hydroxy-$4^{(\prime)}$-methoxy (3O4M), $4^{(\prime)}$-hydroxy-$3^{(\prime)}$-methoxy (4O3M)) examplified by different compound classes was achieved by both enzymes. This is the first time this behavior is described for PFOMT. It is shown, that the activity of PFOMT towards non-catechols is increased at high pH. Adjusting the pH to more basic conditions can also partly remedy the deletrious effect of missing Mg^{2+} for class I enzyme PFOMT. Soluble SOMT-2 enzyme was obtained by optimizing *in vitro* refolding conditions using fractional factorial design (FrFD) and design of experiments (DoE). However, the activity of the refolded SOMT-2 was insufficient for *in vitro* experiments.

7.1 Introduction

Phenyl propanoid derived polyphenols are one of the most abundant plant secondary product in nature. Representatives such as flavonoids and anthocyanidins play important roles in plant development, flower color or in the defense systems combating biotic and abiotic stresses. Lignin, the main component of wood, is mostly comprised of phenyl propanoid alcohols and is responsible for the structure and rigidity in (lignified) vascular plants.

The properties of flavonoids, lignins and other is largely influenced by their derivatization pattern, and thus so-called tailoring enzymes are widespread in nature. Tailoring enzymes such as methyl-, prenyl- and glycosyl-transferases are responsible for hydroxylations, (*C* and *O*)-methylations, -prenylations and -glycosylations, which are common derivatizations of polyphenols, respectively [3–6, 13, 34].

The biosynthetic or biocatalytic production of tailor made natural materials (e.g. lignins) with desireable properties is an important step towards a more ecological and economical industry and has the potential to greatly impact many areas of modern life [63]. Flavonoids and polyphenols such as resveratrol, that exhibit health-beneficial effects (e.g. antioxidative, antimicrobial), are often used in functional foods or to improve the "properties" of food products [29, 170, 203].

It is therefore of great importance to evaluate and develop new methods to biocatalytically produce known and new polyphenols with interesting properties. The aim of this study was to assess two plant *O*-methyl transferases (*O*-MTs), PFOMT and SOMT-2, of classes I and II towards their potential for the methylation of structural motifs commonly found in polyphenols. The main interest was to study methylation of so-called non-catecholic moieties, such as phenolic, $3^{(')}$-hydroxy-$4^{(')}$-methoxy (3O4M) and $4^{(')}$-hydroxy-$3^{(')}$-methoxy (4O3M) structural motifs and to see whether permethoxylated compounds could be produced by these enzymes.

7.2 SOMT-2

SOMT-2 is already well characterized in the literature and acts on flavonoids as well as isoflavonoids[80–82]. However, *in vitro* data of this enzyme has not been published. It was selected as a model candidate as a class II plant *O*-MT, that can already methylate 4'-hydroxyls of non-catecholic flavonoids.

7.2.1 *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 [204]. All the genes required to establish the pathway up to naringenin in *N. benthamiana* in theory had been cloned previously (Figure 7.1). 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 4.3). The average weight loss after freeze drying was 87.5 %.

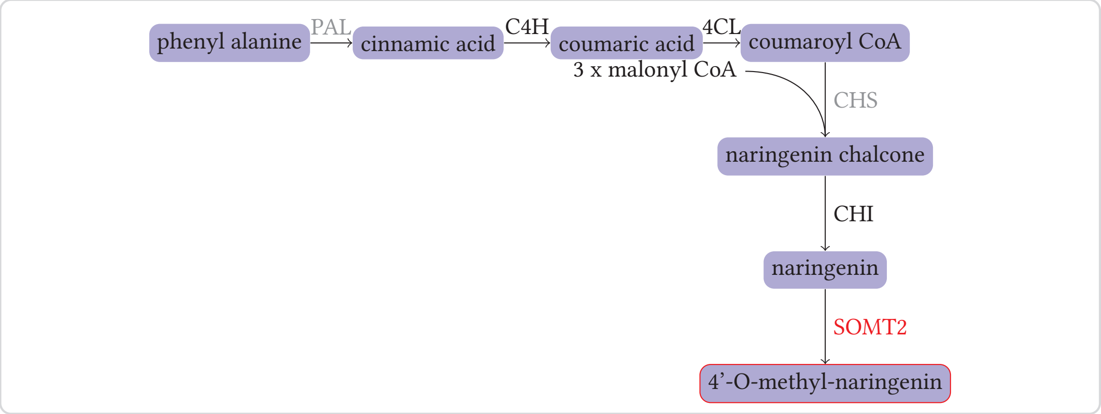
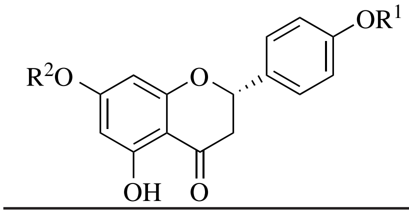


Figure 7.1.: *Semi-synthetic pathway to naringenin and 4'-O-methyl naringenin in N. benthamiana. Enzymes not endogenous to N. benthamiana are in gray. PAL - phenylalanine ammonia lyase, C4H - cinnamic acid 4-hydroxylase, 4CL - 4-coumaric acid:CoA ligase, CHS - chalcone synthase, CHI - chalcone isomerase, SOMT2 - soy O-methyl transferase 2*

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 7.1). 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 exluded that higher amounts of the compounds of interest were present at some point in the tissue.

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 7.2 and C.1). 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

Table 7.1.: 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^1	R^2	name
H	H	naringenin
CH_3	H	ponciretin
CH_3	rutinose [1]	poncirin
CH_3	neohesperidose [2]	didymine

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 C.1). This suggest, that the chemical composition as detected by HPLC is slightly different in the top and botton leaves, as well as between the different plants.

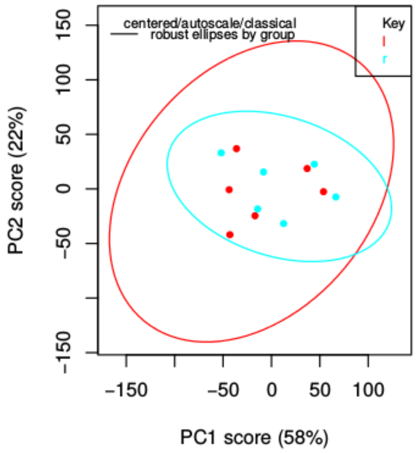


Figure 7.2.: Scatterplot of the first two principal components from the PCA of the HPLC data obtained from leaf material extracts. The samples are colored by leaf side (left/SOMT-2: red, right/vector control: cyan).

7.2.2 *In vivo* biotransformation in *E. coli*

Kim *et al.* [80–82] 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 glutathione *S*-transferase (GST)-tag, respectively since both have been used successfully [80–82]. However, methylated flavonoids were not detected when biotransformations were prepared according to the methods of the aforementioned authors (Figure 7.3a). Thus, the biotransformation medium was changed to auto-induction (AI)-medium with 0.05 % glucose [205]. Measured growth curves showed, that the glucose present in the medium was depleted after about 5 hours into growth (appendix, Figure C.2). 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 SOMT-2 could not be clearly distinguished from endogenous *E. coli* protein in the SDS-PAGE gels (appendix, Figure C.3). Nonetheless, methylation of some of the tested substrates was observed over a course of 30 hours (Table 7.2). 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 feasabil-

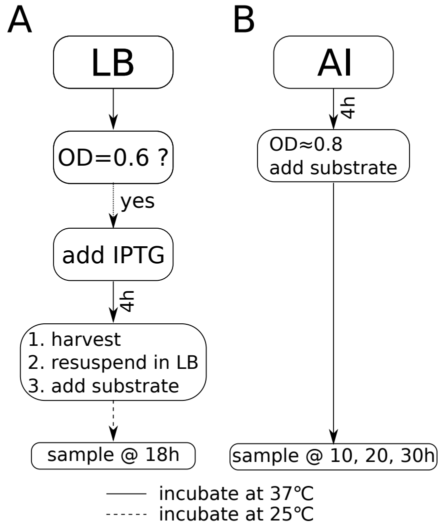


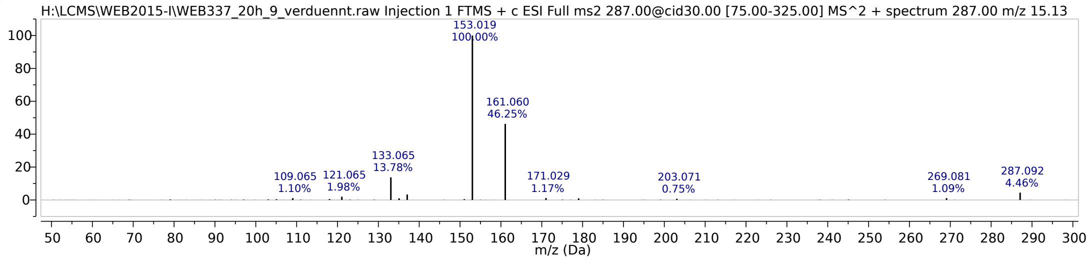
*Figure 7.3.: Biotransformation methods as described by Kim et al. (**A**) and developed in this work (**B**). OD – optiocal density at* 600 nm, *LB – LB-medium, AI – AI-medium.*

Table 7.2.: 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).

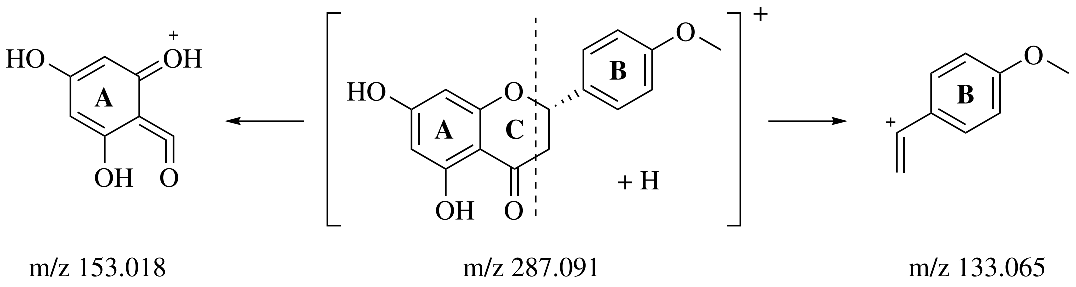
substrate	class	4'-OH	conversion	product
alizarin	anthraquinone	✗	✗	–
purpurin	anthraquinone	✗	✗	–
apigenin	flavone	✓	✓(≥54 %)	4'-*O*-methyl apigenin
chrysin	flavone	✗	✗	–
genistein	isoflavone	✓	✓(<1 %)	Biochanin A
galangin	flavonol	✗	✗	–
kaempferol	flavonol	✓	✓(≥6 %)	kaempferide
naringenin	flavanone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

[†] dihydro cinnamic ketone

ity of this approach (see chapter 6) [189, 192]. 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 [185]. 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 [189, 192]. The fragmentation of interest in this work, was the one along the C-ring, which produces two fragments (A- and B-ring) (Figure 7.4b). The mass of the A- and B-ring fragments gives strong evidence for the position (ring) at which methylation occured. Using the CID technique, an energy of 30 eV proved sufficient to fragment most flavonoids along the C-ring as is shown here for the methylated naringenin (Figure 7.4). 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 7.4b). 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



(a) CID spectrum (30 eV) of methylated naringenin.



(b) Scheme displaying the fragmentation of ponciretin.

*Figure 7.4.: 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*

occured exclusively at the 4'-hydroxyl, as there was no conversion detected, when the 4'-hydroxyl was absent (Table 7.2). A free 4'-hydroxyl seems therefore necessary for a substance to be a substrate for SOMT-2, which confirms the previous results [80, 81]. Conversion was observed for 4'-hydroxylated (iso)flavonoids and the stilbene resveratrol, however, conversion rates of the isoflavone genistein were very low. No conversion of anthraquinones, cinnamic acid derivates or chalcones was detected, which is also in accordance with previously published data [80, 81]. SOMT-2 acts on phenolic, catecholic as well as (4-hydroxy-3-methoxy-phenyl)-moieties, 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 ≥55 %). The tested isoflavones and flavonols showed much lower conversion ratios (less then 10 %). The conversion ratios of apigenin (≥54 %) and naringenin (≥55 %) are comparable to the ones published previously [80, 81]. However, genistein only showed minute conversions, which is in disagreement with the data previously published [80, 81]. 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 ≥86 % in 30 hours. This is roughly double the conversion which was recently reported for *in vivo* biotransformations using the specific resveratrol *O*-MT sbCOM1, which only achieved 42 % conversion in 36 hours [206].

7.2.3 *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 thouroughly characterize an enzyme. Initially, SOMT-2 was to be purified to homogeneity to be later thourougly characterized *in vitro*, since the recombinant production of SOMT-2 in *E. coli* as a fusion protein with an N-terminal T7-tag was previosuly shown. However, the recombinant enzyme had not been characterized [80].

Protein production test

Initial protein production tests were carried out using *SOMT-2* cloned into pET28a(+) with an N-terminal His$_6$-tag. However, SOMT-2 was not produced in soluble form (Figure 7.5). 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 DH5α. The *SOMT-2* gene was cloned into multiple other vectors, inluding pET20b for periplasmic protein proiduction, 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 autoinductiuon 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).

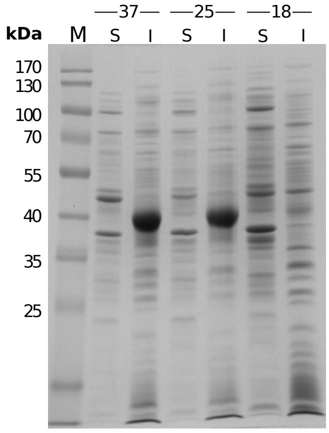


Figure 7.5.: SDS-PAGE of pET28a(+) SOMT-2 expressed in E. coli BL21(DE3) in autoinduction medium at different temperatures (shown above). The insoluble fractions show a protein band the same height as the 40 kDa *marker band, which corresponds to the SOMT-2 protein (*40 425 Da*). M – protein size marker, 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 [127] 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 [127, 129, 207]. 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 [132, 208]. 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 [129, 209, 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 [129–131, 207, 209, 211–214]. 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 4.5 and 4.6).

Big differences between soluble and insoluble fractions of different refolding buffers could already be noticed from the SDS-PAGE gels (Figure 7.6, see subsection 4.4.11). 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 after an overnight refolding reaction (Tables 4.5 and 4.6, Figure 7.6). After rebuffering the cleared refolding reactions into a unified buffer the protein concentrations were estimated by BRADFORD-assays [215]. The protein concentrations measured by BRADFORD's method were consistent with the observations from the SDS-PAGE gels (compare Figure 7.7a and 7.6). 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 [211, 213, 214].

Main effects plots (ME-plots) illustrate the difference between level means for each factor. Therefore, the mean of the measured property (i.e. protein concetration) for each level of every factor, described by + or -, of the used refolding buffers is plotted in relation to the overall mean. For example, the levels of the factor pH are “low” (-) and “high” (+). When x_i^- is the measured concentration from a refolding reaction in buffer i with “low” pH and x_j^+ is the measured concentration from a refolding reaction

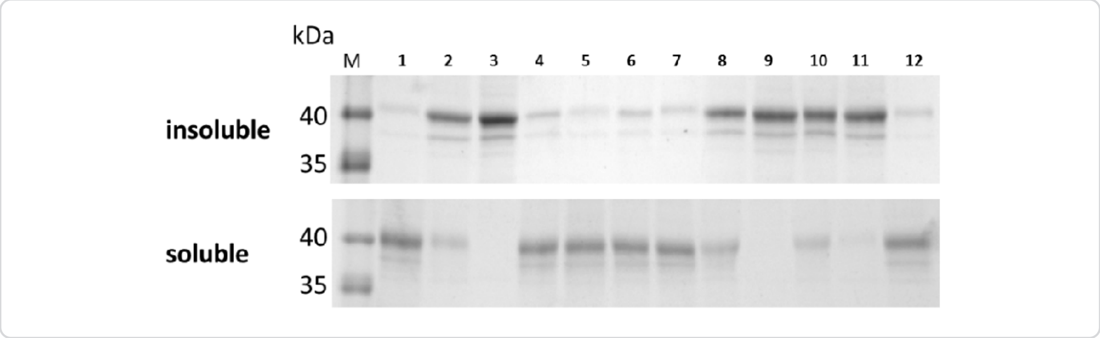


Figure 7.6.: SDS-PAGE of the insoluble and soluble fractions of the refolding reactions. Refolding reactions 2,3,8-11 seem to mainly produce misfolded insoluble protein, while the other refolding buffers (1,4-7,12) produce soluble protein.

in buffer j with “high” pH, then the level means $\overline{x}^-$ and $\overline{x}^+$ are calculated as follows:

$$\overline{x}^- = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i^-$$
$$\overline{x}^+ = \frac{1}{n} \cdot \sum_{j=1}^{n} x_j^+$$

In the presented experiment “low” pH buffers were buffers 2, 4, 6, 8, 10 and 12, whereas “high” pH buffers were buffers 1, 3, 5, 7, 9 and 11, thus the level means were calculated as such:

$$\overline{x}^+ = \frac{x_1^+ + x_3^+ + x_5^+ + x_7^+ + x_9^+ + x_{11}^+}{6} = 24.63$$
$$\overline{x}^- = \frac{x_2^- + x_4^- + x_6^- + x_8^- + x_{10}^- + x_{12}^-}{6} = 11.34$$

The ME-plots for protein concentration suggest, that arginine was likely an important factor for the refolding of SOMT-2 (Figure 7.7b). Furthermore, the addition of SAH or glycerin and the pH seemed to have an influence, whereas the other factors seemed to play only minor roles to achieve high concentrations of soluble protein after refolding. However, the *An*alysis *o*f *Va*riance (ANOVA) test, which gave a p-value of 0.0158 for factor *arginine*, suggests that only arginine addition had a significant influence on refolding, when the significance level is set to 5 % (appendix, Table C.1). The other p-values are all higher than 0.05, which suggests the other factors had no influence on the yield of soluble protein. Only the p-value for factor *SAH* (0.0897) would suggest significance, if the significance level was raised to 10 %.

However, soluble protein is not necessarily active. Therefore activity tests were conducted with the refolded protein samples to check for naringenin conversion (Figure 7.7c). The conversion of naringenin to ponciretin was calculated from the area under the curve (AUC) of the susbtrate and product peaks as follows:

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

Although the substrate naringenin was already contaminated with about 2.5 % ponciretin, this value was not subtracted from the measured conversions to avoid introduction of unnecessary errors. 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 % (6.2 %) was observed for the protein sample refolded in buffer 7. The activity of the protein samples did not correlate well with the amount of soluble protein (Figure 7.7). This becomes clear from the samples refolded in buffers 4 and 5, where the amount of soluble protein was high but the observed activity was at a baseline level.

The ME-plot suggests that the main effects for obtaining high amounts of soluble protein and obtaining active protein after refolding are different (Figure 7.7d). Most notably, the redox state of the refolding reaction seemed to have a big influence on the protein sample's activity. The redox state, however, had almost no influence on the yield of soluble protein (Figure 7.7b). 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 judged by the p-value 0.0218 of the factor *redox* (appendix, Table C.2). However, there is the possibility for SOMT-2 to form intramolecular disulfide bridges, as the modelled structure suggests (Figure C.4). 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 [216–218]. Nevertheless, most MTs are only active under reducing conditions and literature suggests that sometimes assays of MTs are explicitly conducted under reducing conditions [64, 78]. Reducing environments decrease the chance of disulfide cross-linked protein mono- and oligomers and allow the enzyme to be more flexible, which might be important for catalysis. When using a significance level of 10 %, the *p*-value for *arginine* is 0.0824, which also suggests a significant influence of this factor on the chance to obtain active protein after refolding. This is plausible, since there cannot be any activity when no soluble protein is present, and refolding reactions without added arginine did not afford any soluble protein. Judging from the ANOVA test, the remaining factors like *glycerin*, *ionic strength* or *divalent cations* had no significant impact on the protein activity after refolding.

Due to the promising results obtained from the refolding trials, which implicated that the best overall refolding performance (i.e. soluble protein and activity) was achieved in buffer 7 (50 mM borate/NaOH, 0.5 M arginine, 2 mM $CaCl_2$, 2 mM $MgCl_2$, 10 mM NaCl, 0.5 mM KCl, 30 mM α-cyclodextrin, 5 mM DTT pH 8.5), this buffer was used to scale up the refolding reaction from a total volume of 1.05 ml to a volume of 50 ml. After concentration of soluble protein from the scaled-up refolding reaction, activity tests were conducted. Unfortunately, the refolded SOMT-2 showed no activity for naringenin methylation, which was evidence that the scaled-up refolding was unsuccessful. Nonetheless, gel filtration chromatography and circulary dichroism (CD) spectroscopy were used as a tool to study the three-dimensional structure of the refolded SOMT-2. The retention volume for the major peak eluted during the gel filtration run was 12.47 ml, but in the chromatogram a shoulder at 14.26 ml was clearly distinguishable (Figure C.5). Molecular masses of the proteins corresponding to the peaks were estimated from commercial gel filtration protein standards. The first peak corresponds to a molecular weight of approximately 165 kDa, whereas the shoulder at 14.26 ml corresponds to a globular protein of approximately 65.5 kDa. 65.5 kDa is roughly the weight of one SOMT-2 monomer (40 kDa) or a dimer, whereas a mass of 165 kDa would indicate a tetramer. These results were further indication, that the majority of the refolded protein was not in the expected native dimeric state. However, the refolded SOMT-2 had adopted some kind of fold that allowed for it to be in soluble form. Furthermore, the CD spectrum (Figure 7.8) suggested, that the refolded SOMT-2 possessed a secondary structure and was not present as an unfolded random coil. The secondary structure was estimated from the measured CD-spectrum by the K2D3 web service [219]. According to the K2D3 calculations, the secondary structure elements consisted of 12.39 % α-helix and 32.51 % β-sheet. However, the calculated protein model (Figure C.4) suggests the α-helix content might be much higher (52.3 %), whereas the β-sheet content might accordingly be lower (15.4 %). These findings further indicate, that the refolded protein was not in a native state, which might be the cause of the lack of enzymatic activity. 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.

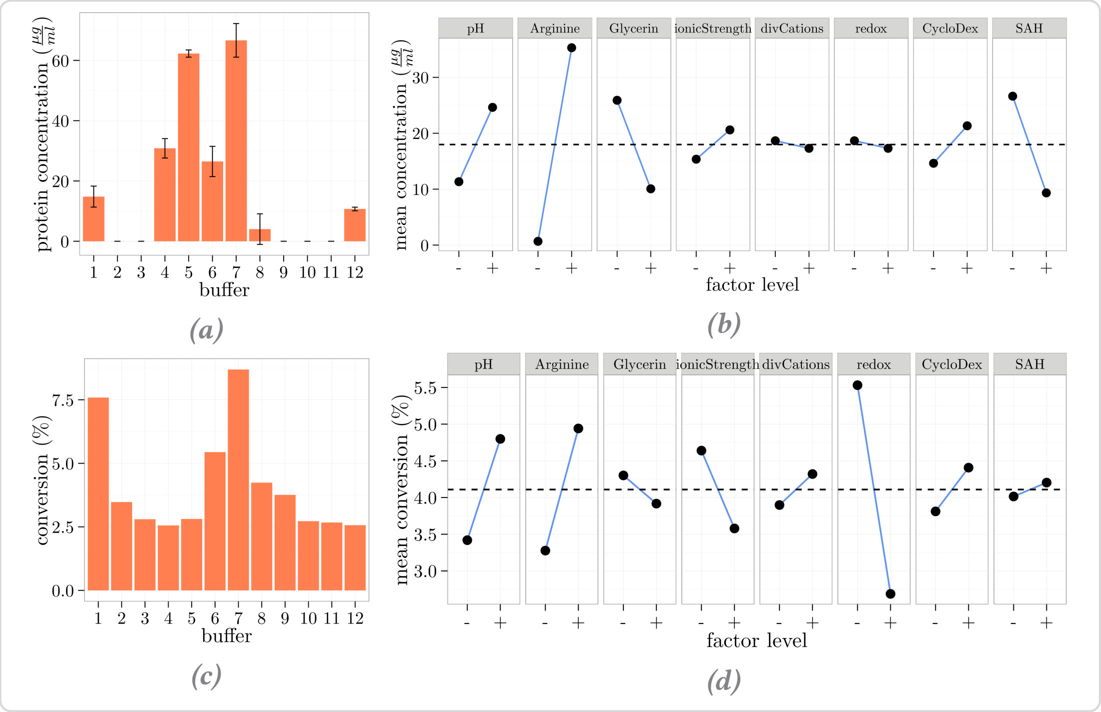


Figure 7.7.: Results of in vitro protein refolding trials. Measured data (left) is presented alongside the main effects plots (ME-plots) (right). The dashed line through the ME-plots illustrates the overall mean. ***a*** *– Protein concentration after refolding and rebuffering into a universal buffer. 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, suggesting that high pH and arginine concentration might have been 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 afforded 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 was important to achieve active protein.*

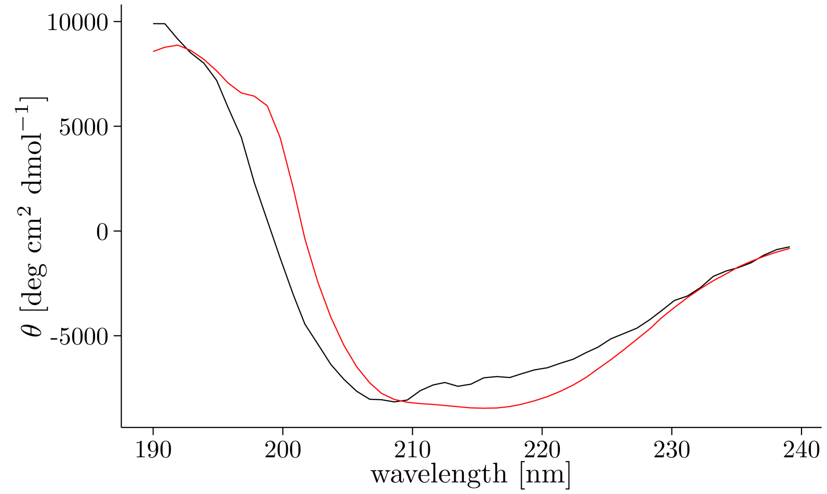


Figure 7.8.: CD-spectrum of refolded SOMT-2 (black) compared to the spectrum that was calculated (red) by the K2D3 web service (http://cbdm-01.zdv.uni-mainz.de/~andrade/k2d3//index.html) from the SOMT-2 sequence. Secondary structure estimates from the measured spectrum are 12.39 % *α-helix and* 32.51 % *β-sheet.*

7.3 PFOMT

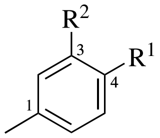
PFOMT is a class I plant *O*-MT, that was throuroughly characterized in previous studies [33, 65, 78, 79, 162]. It is easily obtained by heterologous expression in *E. coli* after which it is fairly active and stable, which is important for *in vitro* experiments. The activity of PFOMT in relation to the pH of the buffer and magnesium addition was the main aim of this part of the study.

7.3.1 Phenols

Phenols have pK_a-values of around 10 as demonstrated by four *p*-cresol derivatives (Table 7.3). Catecholic systems have two pK_a-values, one for each hydroxyl group. The 3-hydroxyl (R^2) of the displayed example has a much smaller pK_a than the 4-hydroxyl (R^1). This is in part due to the mesomeric (+M) and inductive (+I/-I) properties the substituents exhibit. The M and I–effects let the 3-OH to be deprotonated preferentially, which in turn significantly lowers the acidity and thus increases the pK_a of the 4-OH. (4-Hydroxy-3-methoxy)- and (3-hydroxy-4-methoxy)-derivatives have a similar pK_a, with the *meta*-position being 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 group with the lower pK_a always reacts first with an electrophile.

However, different enzymes are able to regioselectivly methylate the 3- or the 4-OH of such catecholic systems. Enzyme's active sites create a micro-environment, 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 7.3.: pK_a-values of phenolic groups examplified by p-cresol derivatives. Substituent numberings on the aromatic ring do not reflect conventions of the International Union of Pure and Applied Chemistry (IUPAC).

	R^1	R^2	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 single phenolic (e.g. naringenin), (3'-hydroxy-4-methoxy)- (e.g. hesperetin) or (4'-hydroxy-3-methoxy)-moieties (e.g. homoeriodictyol) [78]. In these previous studies, the reactions were all run under the same "standard" conditions. However, the reaction buffer 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 [220, 221].

Using PFOMT, reaction conditions were screened to assess if any would promote the methylation of non-catecholic substrates. Although enzymes create a specific environment for catalysis, changes in the pH of the medium can still affect said environment and therefore enzymatic activity, especially if charged groups are part of the catalytic mechanism. For PFOMT a catalytic triad of Lys-Asn-Asp, two of which are charged, is proposed to play a major role [65]. Furthermore, PFOMT is a magnesium dependent enzyme and the activity is affected by altering the concentration of Mg^{2+} or substitution of Mg^{2+} by other divalent cations [78]. 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.

7.3.2 PFOMT pH-profiles are influenced by Mg^{2+}

PFOMT was dialyzed against 50 mM succinate/sodium phosphate/glycine (SSG)-buffer pH 7.5 containing 5 mM ethylenediaminetetraacetic acid (EDTA) and again against the same buffer with the EDTA omitted, 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 in the same unified buffer system (succinate/sodium phosphate/glycine, 2:7:7 molar ratio). The pH-profiles were measured without and with the addition of 10 mM $MgCl_2$. Maximum methylation activity (≈1500 pkat/mg) towards the catecholic substrates (i.e. caffeic acid, eriodictyol) was observed when magnesium was added and the pH was about 6.5 (Figure 7.9 and Table 7.4). However, the observed maximum activity shifted towards basic pH values (pH 9.45), when magnesium was omitted from the reaction. The maximum activity for methylation of *iso*-ferulic acid was measured at pH 9.45, regardless of whether magnesium was added. The catecholic substrates caffeic acid and eriodictyol were

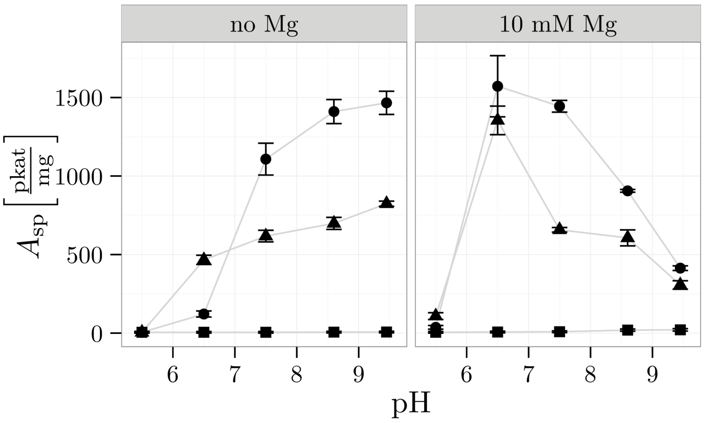


Figure 7.9.: Specific activity/pH-profiles for the conversion of three different substrates (● caffeic acid [catecholic], ▲ eriodictyol [catecholic], ■ iso-ferulic acid [3-hydroxy-4-methoxy]) by PFOMT. The specific activity for the non-catecholic substrate iso-ferulic acid was much lower than the specific activity for the catecholic substrates. When magnesium is omitted, the activity is increased by increasing the pH

converted by PFOMT much more quickly than *iso*-ferulic acid, which is a (3-hydroxy-4-methoxy)-substituted cinnamic acid (Figure 7.9).

The highest specific activity A_{sp} for *iso*-ferulic acid conversion was two orders of magnitude lower than the highest activity for the conversion of the other two substrates. Nonetheless, conversion was observed at low rates even for *iso*-ferulic acid with increasing pH and an influence of magnesium was also observed (Figure 7.9 and Table 7.4). *Iso*-ferulic acid was converted rather slowly without the addition of Mg^{2+} ($A_{sp} = 7$ pkat/mg). However, addition of 10 mM Mg^{2+} increased the rate by 3-fold, from 7 pkat/mg to 21 pkat/mg at pH 9.45 (Table 7.4).

The specific activities observed for the conversion of caffeic acid are comparable to published data [133]. For the two catecholic substrates, the pH-optimum shifted from neutral to alkaline pHs, when Mg^{2+} was omitted. Seemingly, a basic pH can substitute for Mg-complexation for activation towards a phenolate. The maximum activity towards catechols was roughly the same with and without magnesium, even though magnesium addtion seemed to have a slight rate increasing effect. Rate enhancements of up to 3-fold were observed towards catechols at pH 9.45, when Mg^{2+} was omitted compared to when it was present. The maximum activities without magnesium were observed at pH 9.45 with 1466 pkat/mg and 824 pkat/mg for caffeic acid and eriodictyol respectively, while with 10 mM Mg^{2+} the maximum activities were recorded at pH 6.5 and increased to 1572 pkat/mg and 1354 pkat/mg respectively.

Table 7.4.: 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 by the column titled "pH".

substrate	Mg^{2+}	pH	$A_{sp}\left[\frac{mU}{mg}\right]$	$A_{sp}\left[\frac{pkat}{mg}\right]$
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

During the catalysis of a methyl transferase reaction, the acceptor moiety of the substrate is activated by abstraction of a proton. At high pH-values, the substrates might already be deprotonated, thus increasing the rates of reaction, while the enzyme just acts as a scaffold. High pH-values would also mean that the mainly the 4'-hydroxyl of the catechols would be deprotonated (Table 7.3). However, an increasing amount of 4'-methylation was not observed contradicting the notion, that an already deprotonated substrate entered the active site. It is more likely, that the external milieu influences the enzymes active site and makes the active side chains more basic. However, why the addition of magnesium would shift the pH-optimum back to neutral pH-values is not clear.

To statistically support the presented findings, the collected data were also studied from a statisticians point of view, which is described in the following paragraph in a bit more detail. The shown results are solely included for purposes of making statistics based inferences in the context of domain knowledge. Relationships between independent variables $X_{1,\cdots,p}$ influencing a system and the outcome Y of such a system can be mathematically described. The simplest relationship between one independent variable X_1 and the dependent variable Y is a linear one and is defined mathematically by a linear equation $Y = \beta_1 X_1 + \beta_0$. The coefficient β_1 describes how much Y is altered by one unit X, while β_0 is the offset. Even seemingly non-linear outcomes of Y might be sufficiently described by multi-term linear equations such as

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p + \epsilon,$$

when the independent variables $X_{1,\cdots,p}$ are known. Linear regression models of this form can be used to make statistically sound inferences about a studied system and were thus used to assess the relationship between PFOMT's methylation activity and pH-modulation as well as Mg^{2+} addition. Two subsets of the activity data were prepared first. The subsets split the data into substrates with catecholic (i.e. caffeic acid, eriodictyol) and substrates without catecholic (i.e. *iso*-ferulic acid) motifs. This was done to simplify the interpretability of the results, since the activities of the catecholic and non-catecholic substrates differed greatly. The *iso*-ferulic acid data was fit to the linear model

$$activity = \beta_0 + \beta_1 \times \mathrm{Mg} + \beta_2 \times \mathrm{pH} + \beta_3 \times (\mathrm{Mg} \times \mathrm{pH}), \tag{7.1}$$

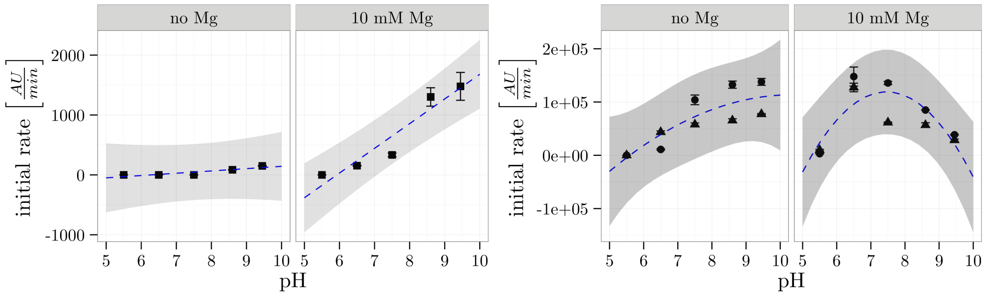
which contains one term, $\beta_3 \times \mathrm{pH} \times \mathrm{Mg}$, to account for an interaction effect between magnesium and pH besides the main effects terms. This model explains about 93.6 % of the variance ($R^2 = 0.9355$) of the measured data (Table 7.5). Fitting the data in the R software using the *stats::lm()* function also calculates p-values associated with each parameter estimate. The smallest p-value 0.0030 was calculated for the interaction factor $\mathrm{pH} \times \mathrm{Mg}$, and strongly suggests that there is a significant interaction between methylation activity on *iso*-ferulic acid and Mg^{2+} addition combined with pH-modulation. The parameter estimate β_3 for this term is almost 374. Since in this case Mg is categorical and can only be 0 (no magnesium added) or 1 (magnesium added) the interaction term $\beta_3 \times \mathrm{pH} \times \mathrm{Mg}$ resolves to 0, if no magnesium was added. This means that when Mg^{2+} is

Table 7.5.: Coefficients of the model (Equation 7.1) 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	**

significance codes: '*' 5 % level; '**' 1 % level added, the activity (in AU min^{-1}) is increased by 374 for each unit the pH is raised. The p-value for pH as a main effect is rather high, suggesting pH alone has no significant impact on the activity. The parameter estimate β_1 for factor Mg, however, has a low p-value of 0.010, which suggests that its impact on activity is significant. Due to the categorical nature of Mg^{2+}, this means, that the activity is decreased by 2201 AU min^{-1} when magnesium is added. However, this information is only of importance when predicting the activity outside of the measured pH range, which was not the aim here. Together, these results illustrate, that neither magnesium addition, nor the pH alone would have such a strong effect on the activity as both factors combined had.

The data of the second subset containing of the catecholic substrates was also modelled in a similar fashion (see Appendix C.2.1). It was found, that the pH has a much stronger influence on the methylation activity of PFOMT towards catechols than *iso*-ferulic acid. Interaction between magnesium addition and pH-modulation were also suggested for the catecholic substrates from these data. For a more thorough discussion see Appendix C.2.1. In addition to pure inference based modelling, models can also be used to make predictions based on (new) independent variables (Figure 7.10). However, it cannot be stressed enough that models do not reflect the truth, but are rather another tool to gain insight into a system.



(a) *iso-ferulic acid ($R^2 = 0.9355$)* (b) *Catecholic substrates ($R^2 = 0.6855$)*

Figure 7.10.: pH-profiles of substrate conversion along with predicted data. Predicted data from 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.

To the knowledge of the author, this is the first time the effects of Mg^{2+} and pH on methyl transferase activity were systematically analyzed. It was shown, that 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. It would be of interest to analyze this behavior with further systematic studies and multiple levels of Mg^{2+} concentrations.

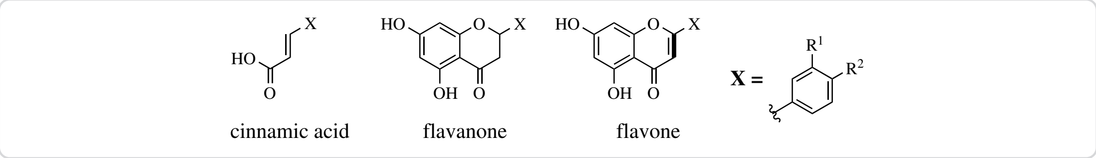
7.3.3 Methylation of different chemical motifs

The previous section showed the conversion of non-catecholic *iso*-ferulic acid by PFOMT and prompted additional experiments with other non-catechols from multiple flavonoid subgroups (Table 7.6). The tested substrates were selected from three different compound groups (cinnamic acids, flavones, flavanones) and each group contained each of four substitution patterns – single phenol, catechol, $3^{(\text{'})}$-hydroxy-$4^{(\text{'})}$-methoxy (3O4M) and $4^{(\text{'})}$-hydroxy-$3^{(\text{'})}$-methoxy (4O3M). Each substrate was assessed for conversion with two enzymes (PFOMT wild-type and 4'-specific variant Y51R N202W) at four different conditions. Magnesium addition and pH were varied for the 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 4.6.3).

Conversion of all substrates, catecholic and non-catecholic, by the wild-type could be demonstrated. The highest conversion of non-catecholic substrates was shown for 3'-hydroxy-4'-methoxy substituted compounds, especially cinnamic acids and flavanones, where conversions of up to 25 % were observed. Conversions of substances with free 4'-hydroxyl groups did not extend beyond 7 % (chrysoeriol). This was to be expected, due to the fact that PFOMT is a 3'-specific MT at physiological conditions. However, unexpectedly the 4'specific variant hardly showed any conversion of non-catecholic substrates. High pH favoured the conversion of non-catechols, as expected.

There was almost complete conversion of the catecholic substrates (eriodictyol, luteolin and caffeic acid) after 16 h of incubation, regardless of the reaction conditions, at least when the wild-type enzyme was used (Figure 7.11 and Table 7.8). This suggests, that the reaction period was chosen too long for this group of substrates and effects

Table 7.6.: 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, single phenolic, catecholic, $3^{(\prime)}$-hydroxy-$4^{(\prime)}$-methoxy (3O4M) and $4^{(\prime)}$-hydroxy-$3^{(\prime)}$-methoxy (4O3M). $R^{1/2}$ denote possible methylation positions.



	substrate	group	motif	R^1	R^2
A.1	*p*-coumaric acid	cinnamic acid	phenolic	H	OH
A.2	caffeic acid	cinnamic acid	catecholic	OH	OH
A.3	*iso*-ferulic acid	cinnamic acid	3O4M	OH	OMe
A.4	ferulic acid	cinnamic acid	4O3M	OMe	OH
B.1	naringenin	flavanon	phenolic	H	OH
B.2	eriodictyol	flavanon	catecholic	OH	OH
B.3	hesperetin	flavanon	3O4M	OH	OMe
B.4	homoeriodictyol	flavanon	4O3M	OMe	OH
C.1	apigenin	flavone	phenolic	H	OH
C.2	luteolin	flavone	catecholic	OH	OH
C.2	diosmetin	flavone	3O4M	OH	OMe
C.4	chrysoeriol	flavone	4O3M	OMe	OH

of pH or magnesium addition on this group cannot be distinguished. Conversion was observed for all tested substrates at least under high pH-conditions, when the wild-type was employed (Figure 7.11). The 4'-variant, however, hardly showed any conversion of non-catecholic substrates (Figure 7.11b). Generally, the highest conversions of non-catecholic substrates were observed at high pH and high Mg^{2+} conditions (Figure 7.11). For example, conversion of *iso*-ferulic acid was up to 25 % at pH 8.6 with 10 mM Mg^{2+} in the reaction. This is a 10-fold increase over the conversion of 2.5 %, which was observed at pH 7.5 with no magnesium added. These findings support the claims from the previous subsection 7.3.2, that PFOMT activity might be modulated enough by pH and magnesium to achieve methylation of non-catecholic phenyl propanoid substrates. Again, the trend in the data suggests, that methylation efficiency of non-catecholic moieties increases with pH, but especially in combination with the addition of magnesium .

Overall, methylation of 3O4M motifs was highest apart from the catecholic substrates.

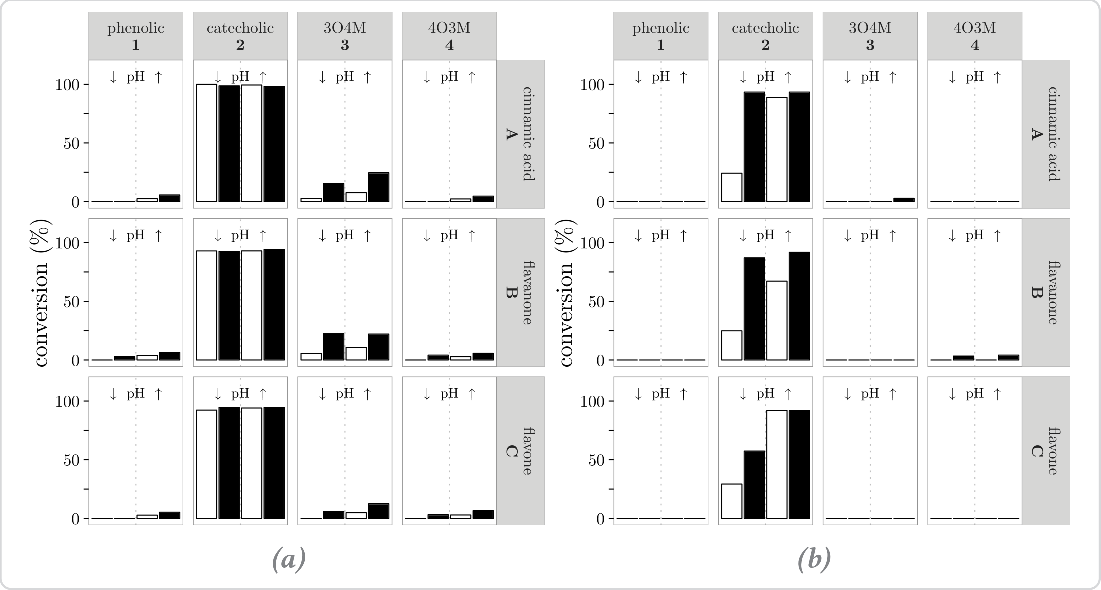


*Figure 7.11.: Conversion of multiple different substrates, catecholic and non-catecholic, by PFOMT wild-type (**a**) and the 4'-specific variant Y51R N202W (**b**). Every individual box represents one substrate p-coumaric acid (A.1), . . ., chrysoeriol (C.4). Low pH conditions are displayed on the left side of each panel, high pH conditions on the right. Mg^{2+}-conditions are color-coded as follows: no Mg^{2+} – white bars,* 10 mM *Mg^{2+} – black bars.*

Observed conversions of close to 25 % for the cinnamic acid and flavanone substrates (*iso*-ferulic acid and hesperetin) were observed. 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 (subsection 7.3.2). Similar results have been shown for SaOMT5, an *O*-MT from *Streptomyces avermitilis*, where the enzymatic activity towards quercetin increased by about 5-fold from metal-free conditions to magnesium addition [222]. Conversion of the somewhat more rigid flavone diosmetin (max. conversion 12 %) was lower by at least factor two compared to hesprereting and *iso*-ferulic acid (Table 7.8, Figure 7.11). 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 *para*-phenolic and 4O3M substrates could be due to the fact that the wild-type of PFOMT by and large methylates 3'-hydroxyls at physiological conditions.

The 4'-specific variant for the most part only showed conversion of the catecholic substrates. Only some conversion of homoeriodictyol (≈4 %) and *iso*-ferulic acid (≈3 %)

Table 7.8.: Conversion of substrates after 16 hours incubation. Only the maximum conversion is displayed, along with the conditions it was achieved under.

	substrate	wild-type (3'-specific)			4'-specific variant Y51R N202W		
		conversion %	pH	Mg^{2+}	conversion %	pH	Mg^{2+}
A.1	*p*-coumaric acid	6	↗	✓			
A.2	caffeic acid[†]	100[‡]	(all)		93	(all)	
A.3	*iso*-ferulic acid	25	↗	✓	3	↗	✓
A.4	ferulic acid	5	↗	✓			
B.1	naringenin	6	↗	✓			
B.2	eriodictyol[†]	94	(all)		92	↗	✓
B.2	hesperetin	22	↘	✓			
B.4	homoeriodictyol	6	↗	✓	4	↗	✓
C.1	apigenin	5	↗	✓			
C.2	luteolin[†]	95	(all)		92	↗	
C.3	diosmetin	12	↗	✓			
C.4	chrysoeriol	7	↗	✓			

[†] wild-type: substrate conversion was maximal for all pH/Mg^{2+} combinations.
[‡] conversion of caffeic acid by the wild-type was set to 100 %.

was observed under high pH/Mg^{2+} conditions (Table 7.8, Figure 7.11). However, for the catecholic substrates the same trend – increasing pH/Mg^{2+} increases activity – as before holds true. Control experiments without enzyme at high pH and 10 mM Mg^{2+} revealed, that no substrate conversion took place under these conditions, meaning the enzyme must be involved in the conversion.

Products and methylation sites were identified by comparison to authentic standards, or by LC-MS/MS (Table 7.9). As previous studies demonstrated, the products for the conversion of the catecholic substrates by the wild-type and variant were the 3'-methylated and 4'-methylated substrates respectively [33, 78, 79, 162]. As expected, methylation took place on the B-ring of the flavonoids. Ponciretin and acacetin were produced, when naringenin and apigenin were converted by PFOMT respectively. However, conversion of the 3O4M and 4O3M flavonoids (hesperetin/chrysoeriol and homoeriodictyol/disometin respectively), afforded the 3',4'-dimethylated compounds. This demonstrates, that even the PFOMT wild-type is able to methylate the 4'-position of flavonoids, given the right conditions (prolonged incubation, high pH, no free 3'-

hydroxyl). Furthermore, another type of product, eluting earlier than the corresponding substrates, was observed for the flavones apigenin, chrysoeriol and diosmetin after conversion with the wild-type in the LC/MS runs. Unfortunately, these products could not be identified. Nonetheless, the generation of these products seemed favoured over the 3' or 4'-*O*-methylated ones, when a free 3'-hydroxyl was absent (Appendix, Figure C.9).

Enzymatic methylation of the non-catecholic cinnamic acids also afforded two different types of product, methyl esters and methyl ethers. Methylation of *p*-coumaric acid only gave rise to the corresponding *methyl ester* (Appendix C.3). Two different products were observed for the enzymatic methylation of ferulic acid and *iso*-ferulic acid. One product was the methyl ester of the corresponding cinnamic acid, whereas the other product was the di-ether, 3,4-dimethyl caffeic acid.

Table 7.9.: Products of the enzymatic methylation of the studied substrates. The products were confirmed by authentic standards or via LCMS.

substrate	product	
	wild-type	4'variant
p-coumaric acid	*p*-hydroxy methylcinnamate† [223]	
caffeic acid	ferulic acid	ferulic acid
iso-ferulic acid	caffeic acid dimethyl ether, *iso*-ferulic acid methyl ester†	n/d
ferulic acid	caffeic acid dimethyl ether, ferulic acid methyl ester†	
naringenin	ponciretin	
eriodictyol	homoeriodictyol	hesperetin
hesperetin	3',4'-dimethyl eriodictyol†	
homoeriodictyol	3',4'-dimethyl eriodictyol†	3',4'-dimethyl eriodictyol†
apigenin	acacetin	
luteolin	chrysoeriol	n/d
diosmetin	3',4'-dimethyl luteolin†	
chrysoeriol	3',4'-dimethyl luteolin†	

†determined via LCMS; n/d – not determined

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 methylate the 4’-position, when the substrates B-ring possessed a 4O3M substitution [83]. 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 [224]. However, the authors state, that methylation of dihydroxy-derivatives such as luteolin and eriodictyol by TaOMT2 only afforded 3’-mono-methylated products, which is similar to PFOMT. Nonetheless, TaOMT2 can methylate tamarixetin, the 4O3M derivative of quercetin, albeit with low activity.

Of the 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 7.3.2). 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. Furthermore, methylation of the acid functionalities of cinnamic acids was demonstrated using PFOMT.

The N-terminus of PFOMT is important for the function of the enzyme, as was demonstrated in previous studies. However, its role *in vivo* is still not fully understood [33, 78, 225]. It cannot be ruled out that it 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 millieu can be quite different in different cell compartments in plants [226].

7.4 Conclusion

Enzymatic methylation of non-catecholic moieties, was studied using the two methyl transferases PFOMT and SOMT-2, of classes I and II, respectively. Therefore different flavonoid and phenylproanoid substrates, displaying either single phenolic, catecholic, $3^{(')}$-hydroxy-$4^{(')}$-methoxy or $4^{(')}$-hydroxy-$3^{(')}$-methoxy moieties, were compared. 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 of the exact moiety (phenolic, catecholic, $4^{(')}$-hydroxy-$3^{(')}$-methoxy). Although overall the conversions were very low, the conversion of the stilbene resveratrol was superior over all other tested substrates (≥86 % vs. ≥55 %). SOMT-2 showed methylating activity exclusively when a free 4'-OH was present, suggesting it only acts on 4'-hydroxyl groups. 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 (DoE) and fractional factorial design (FrFD) 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 single 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 in principle. The best conversions were obtained using the PFOMT wild-type at elevated pH and after Mg^{2+} addition. Magnesium addition and pH displayed synergistic effects, meaning the effects of both are not just additive. pH optimum of PFOMT shifted from around pH 7 to more basic conditions (pH >8), when Mg^{2+} was omitted. Although no magnesium was present under these high pH conditions, 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 without the addition of magnesium. These findings also 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.

8 Summary

Plant secondary metabolites comprise a vast collection of compounds with a myriad of biological activities and functions, that include defense against biotic and abiotic stresses, communication as well as gene regulation. These biological activities are largely determined by the compounds substitution patterns Phenyl propanoid derived phenolic compounds such as flavonoids make up a major part of the secondary metabolites in plants and tailoring enzymes such as methyl-, prenyl- and glycosyl-transferases are responsible of putting a finish on their molecular structure. This work aimed to study the ability of two plant *O*-methyl transferases (*O*-MTs), phenylpropanoid and flavonoid *O*-methyl transferase (PFOMT) and soy *O*-methyl transferase (SOMT-2), for the alkylation of common structural motifs in polyphenols.

In the first part, PFOMT was characterized by biophysical methods. It was shown, that the enzyme is able to bind the non-natural ethylated *S*-adenosyl-L-methionine (SAM) analogue *S*-adenosyl-L-ethionine (SAE). Conversion of SAE was, however, not observed. In addition, a novel crystal structure of the *apo*-form of PFOMT was solved that gave insights into the movements associated with substrate binding.

The second part comprised of a tandem mass-spectrometry (MS/MS)-study of 15 flavonoids of different classes (i.e. flavanones, flavones and flavonoles). Two activation methods, collision induced dissociation (CID) and higher-energy collisional dissociation (HCD), were shown to produce complementary information that enables a fast and reliable structural characterization of the studied substances.

Finally, a comprehensive study showed the methylation of several non-catecholic flavonoids, stilbenes and cinnamic acids by the enzymes SOMT-2 and PFOMT *in vivo* and *in vitro*. For the first time methylation of phenolic, $3^{(')}$-hydroxy-$4^{(')}$-methoxy and $4^{(')}$-hydroxy-$3^{(')}$-methoxy, which was achieved through the modulation of the reaction conditions (e.g. pH, magnesium concentration), by PFOMT could be demonstrated.

Together, the results presented herein show that both class I and II plant *O*-MTs can be used to methylate non-catecholic moities of polyphenolic compounds. Modulation of the reaction conditions is a worthwhile study and might bring to light new activities, that were not to be expected. This is of specific interest for biocatalytic reactions, where the naturally observed reaction might not be the desired behavior of an enzyme. Furthermore, these results might spark new discussions about the regulation and function of enzymes in different compartments (i.e. reaction milieus).

Affidavit / Eidesstattliche Erklärung

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. I furthermore declare that it is the first time that I have used this document to try and gain a doctoral degree.

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig und ohne fremde Hilfe verfasst habe. Andere als die von mir angegebenen Quellen und Hilfsmittel habe ich nicht benutzt und die den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen habe ich als solche kenntlich gemacht. Außerdem erkläre ich, dass ich mich mit der vorliegenden wissenschaftlichen Arbeit erstmals um die Erlangung des Doktorgrades bewerbe.

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

Curriculum vitæ

Personal Information

Degree	Diplom-Biochemiker
Surname	Weigel
Name	Benjamin
Address	Elcknerplatz 16b, 12555 Berlin
Date of Birth	06.04.1986
Place of Birth	Bernburg (Saale)
Nationality	German

Education

2011–2015	**Graduate Studies.** Leibniz-Institute of Plant Biochemistry - Group of Bioorganic Chemistry, Halle (Saale), Advisor: Prof. Dr. Ludger A. Wessjohann, Thesis: “Test“
2010–2011	**Diploma thesis.** Leibniz-Institute of Pflant Biochemistry, Halle(Saale), Advisor: Prof. Dr. Ludger A. Wessjohann, Thesis: „Biotransformationen mit *Cannabis sativa* Monoterpensynthasen.“
2009	**LANDO/NSF research fellow.** Department of Chemistry, University of Minnesota, Minneapolis, MN (USA), Advisor: Prof. Dr. Mark D. Distefano
2006–2011	**Studies in Biochemistry.** Martin-Luther-University Halle-Wittenberg, Halle(Saale), Degree: Diplom-Biochemiker

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

III

Appendix

A Engineering of PFOMT

Table A.1.: *Crystallographic data, phasing and refinement statistics.*

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

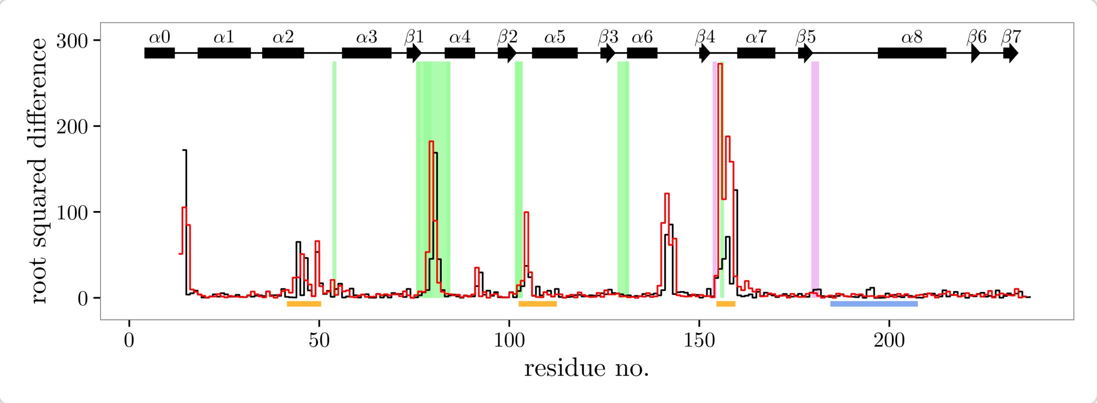


Figure A.1.: *Differences in the dihedral-angles ψ (red) and φ (black) of the solved apo-PFOMT and the structure with bound S-adenosyl-L-homocysteine (SAH) (pdb: 3C3Y). The secondary structure is displayed at the top. Helices are displayed as rectangles and sheets are shown as arrows. Graphical background 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.*

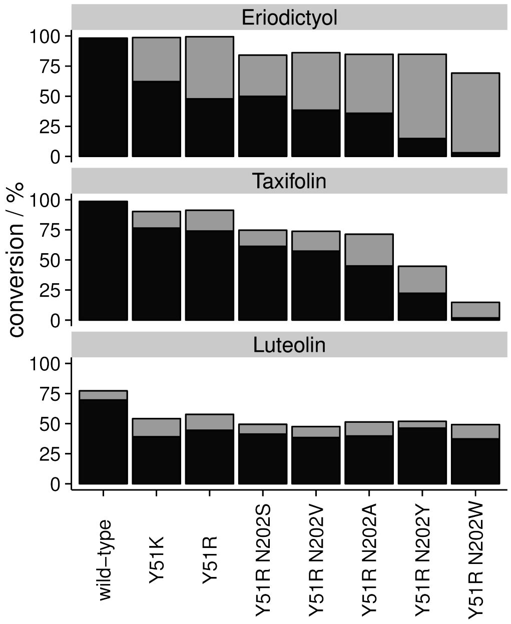


Figure A.2.: *Differences in the regioselectivity of some phenylpropanoid and flavonoid O-methyl transferase (PFOMT) variants. The products observed in high-performance liquid chromatography (HPLC) and liquid chromatography coupled mass-spectrometry (LC/MS) measurements switched from 3'-O-methylated (dark grey) to 4'-O-methylated (light grey) for the displayed variants. The height of the bars corresponds to the total conversion of substrate.*

B Tandem mass-spectrometry studies of flavonoids

Table B.1.: Key ions in the positive mode CID and HCD ESI-MS^2 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 $[M{+}H{-}H_2O]^+$	255 (1)	271 (18)	269 (1)	285 (10)	285 (4)										
4 $[M{+}H{-}C_2H_2O]^+$	231 (4)	247 (3)	245 (3)	261 (2)	261 (2)										
5 $[M{+}H{-}2C_2H_2O]^+$	189 (5)	205 (3)	203 (4)	219 (2)	219 (1)										
8 AC^+	179 (4)	179 (20)	179 (5)	179 (28)	179 (30)		179 (1)		179 (2)	179 (2)				179 (2)	179 (1)
11 $1{,}3A^+$	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 $1{,}3A^+{-}CO$											125 (1)		125 (1)	125 (1)	
13 $1{,}3A^+{-}C_2H_2O$							111 (2)	111 (1)	111 (1)		111 (2)	111 (2)	111 (4)	111 (2)	111 (2)
14 $1{,}3A^+{-}2CO$						97 (3)	97 (4)	97 (3)	97 (4)	97 (4)	97 (10)	97 (8)	97 (15)	97 (9)	97 (9)
15 $1{,}3A^+{-}2CO{-}C_2H_4$						69 (2)	69 (2)	69 (2)	69 (2)	69 (2)	69 (9)	69 (8)	69 (13)	69 (9)	69 (8)
17 $1{,}4B^+{-}2H$	147 (84)	163 (100)	161 (100)	177 (100)	177 (100)	147 (15)	163 (10)	161 (10)	177 (4)	177 (2)	147 (3)	163 (1)	161 (3)		
18 $1{,}4B^+{-}2H{-}H_2O$		145 (5)					145 (7)					145 (1)			
19 $1{,}4B^+{-}2H{-}CO$	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 $1{,}4B^+{-}2H{-}CO{-}CH_3$								118 (11)	134 (11)	134 (7)			118 (57)	134 (20)	134 (13)
22 $1{,}4B^+{-}2H{-}2CO$						91 (24)					91 (100)				
23 $1{,}4B^+{-}2H{-}2CO{-}CH_3$								90 (3)					90 (49)		
24 $1{,}4B^+{-}2H{-}C_2H_2O{-}H_2O$									117 (15)	117 (32)				117 (13)	117 (26)
25 $1{,}4B^+{-}2H{-}H_2O{-}CO$							117 (18)					117 (21)			
26 $C_7H_7^+$						91 (24)		91 (1)	91 (3)	91 (2)	91 (100)		91 (5)	91 (7)	91 (5)
27 $C_7H_5^+$							89 (23)		89 (29)	89 (24)		89 (100)	89 (22)	89 (100)	89 (100)

Table B.2.: Key ions in the positive mode CID and HCD ESI-MS^2 spectra of flavones.

fragment	CID, 45 % NCE					HCD, 75 % NCE					HCD, 100 % NCE				
	(6)	**(7)**	**(8)**	**(9)**	**(10)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(6)**	**(6)**	**(7)**	**(9)**	**(10)**
1 $[M+H]^+$	271 (2)					271 (84)	287 (66)	285 (4)			271 (2)	287 (2)			
2 $[M+H-CH_3]^{\bullet+}$			270 (100)	286 (100)	286 (100)			270 (9)	286 (20)	286 (18)					
3 $[M+H-H_2O]^+$	253 (1)	269 (9)				253 (3)	269 (6)								
4 $[M+H-CO]^+$	243 (7)	259 (9)				243 (7)					243 (2)				
5 $[M+H-CHO]^{\bullet+}$	242 (14)	258 (47)				242 (1)	258 (3)				242 (2)	258 (2)			
6 $[M+H-C_2H_2O]^+$	229 (21)	245 (13)	243 (1)			229 (4)									
7 $[M+H-CH_3-CO]^{\bullet+}$			242 (7)					242 (100)	258 (100)	258 (100)			242 (49)	258 (49)	258 (46)
8 $[M+H-CH_4-CO]^+$								241 (1)	257 (7)	257 (7)			241 (7)	257 (38)	257 (37)
9 $[M+H-H_2O-CO]^+$	225 (13)	241 (13)				225 (4)	241 (16)					241 (4)			
14 $[M+H-CH_4-2CO]^+$									229 (12)	229 (13)			213 (6)	229 (75)	229 (76)
15 $[M+H-H_2O-2CO]^+$		213 (2)				197 (4)	213 (7)				197 (3)	213 (4)			
16 $[M+H-2C_2H_2O]^+$	187 (3)	203 (4)				187 (2)	203 (2)				187 (1)	203 (3)			
17 $[M+H-CH_3OH-2CO]^{\bullet+}$							199 (1)							213 (4)	213 (4)
18 $[M+H-CH_4-2CO-C_2H_2]^+$														203 (29)	203 (31)
20 $[M+H-CH_4-3CO]^+$												187 (2)	185 (1)	201 (7)	201 (6)
23 $[M+H-CH_4-4CO]^+$											143 (1)			173 (3)	173 (2)
24 $[M+H-2CO-2C_2H_2O]^+$						131 (2)	147 (1)				131 (5)	147 (3)		161 (4)	161 (5)
27 $0,4B^+$	163 (6)	179 (7)				163 (8)	179 (3)				163 (2)				
28 $0,4B^+-H_2O$	145 (13)	161 (12)				145 (17)	161 (16)				145 (41)	161 (29)			
25 $0,4B^+-C_2H_2O$	121 (6)	137 (7)				121 (16)	137 (12)				121 (25)	137 (16)			
29 $1,3A^+$	153 (100)	153 (100)	153 (5)			153 (100)	153 (100)	153 (11)	153 (8)	153 (8)	153 (84)	153 (87)	153 (100)	153 (100)	153 (100)
30 $1,3A^+-CO$						125 (1)	125 (2)				125 (3)	125 (2)			
31 $1,3A^+-C_2H_2O$						111 (2)	111 (2)				111 (4)	111 (3)		111 (1)	111 (1)
32 $1,3A^+-2CO$						97 (2)	97 (2)				97 (9)	97 (9)	97 (1)	97 (2)	97 (2)
33 $1,3A^+-2CO-C_2H_4$						69 (4)	69 (5)				69 (24)	69 (22)	69 (1)	69 (2)	69 (2)
34 $1,3B^+$	119 (12)	135 (11)	133 (2)			119 (49)	135 (40)	133 (3)			119 (35)				
35 $1,4A^++2H$							127 (1)					127 (2)			
39 $C_7H_7^+$	91 (1)					91 (26)					91 (100)				
40 $C_7H_5^+$							89 (17)				89 (7)	89 (100)	89 (3)		89 (1)

Table B.3.: Key ions in the positive mode CID and HCD ESI-MS^2 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)**
1 $[M+H]^+$						287 (25)	303 (8)	319 (1)	301 (9)						
2 $[M+H-CH_3]^{\bullet +}$				286 (62)	302 (100)				286 (12)	302 (6)					
3 $[M+H-CH_4]^+$				285 (2)					285 (5)	301 (3)				285 (1)	
4 $[M+H-H_2O]^+$	269 (32)	285 (63)	301 (40)	283 (2)	299 (3)	269 (2)	285 (6)	301 (1)			269 (2)	285 (3)			
5 $[M+H-CO]^+$	259 (24)	275 (14)	291 (6)	273 (3)		259 (3)									
6 $[M+H-CHO]^{\bullet +}$	258 (46)	274 (20)	290 (22)	272 (20)	288 (3)	258 (10)	274 (4)		272 (2)						
7 $[M+H-CH_3OH]^{\bullet +}$				269 (11)	285 (36)				269 (11)	285 (11)				269 (3)	285 (3)
9 $[M+H-C_2H_2O]^+$	245 (4)														
10 $[M+H-CH_3-CO]^{\bullet +}$				258 (9)	274 (5)				258 (58)	274 (6)				258 (2)	
12 $[M+H-CH_4-CO]^+$				257 (3)					257 (17)	273 (13)				257 (6)	273 (2)
13 $[M+H-H_2O-CO]^+$	241 (99)	257 (100)	273 (100)	255 (3)	271 (5)	241 (5)	257 (13)	273 (6)			241 (1)	257 (2)			
14 $[M+H-2CO]^+$	231 (40)	247 (37)	263 (25)	245 (17)	261 (8)	231 (5)	247 (2)								
15 $[M+H-CH_3OH-CO]^{\bullet +}$				241 (16)	257 (8)				241 (2)	257 (7)					257 (2)
17 $[M+H-CH_3-2CO]^{\bullet +}$						216 (2)	232 (2)		230 (94)	246 (9)	216 (2)			230 (9)	
19 $[M+H-CH_4-2CO]^+$									229 (100)	245 (42)				229 (100)	245 (34)
20 $[M+H-H_2O-2CO]^+$	213 (77)	229 (70)	245 (32)	227 (2)	243 (2)	213 (20)	229 (49)	245 (22)	227 (2)		213 (12)	229 (16)	245 (6)		
21 $[M+H-3CO]^+$						203 (2)		235 (2)							
22 $[M+H-2C_2H_2O]^+$	203 (7)	219 (4)	235 (4)	217 (2)			219 (1)				203 (1)	219 (2)			
23 $[M+H-H_2O-CO-C_2H_2O]^{\bullet +}$	199 (3)	215 (2)	231 (2)	213 (4)	229 (3)			231 (1)	213 (11)	229 (33)				213 (10)	229 (15)
23 $[M+H-CH_3OH-2CO]^{\bullet +}$				213 (4)	229 (3)				213 (11)	229 (33)				213 (10)	229 (15)
24 $[M+H-CH_4-2CO-2C_2H_2]^+$									203 (3)	219 (7)				203 (11)	219 (13)
25 $[M+H-H_2O-2CO-C_2H_2]^+$							203 (3)	219 (8)	201 (25)	217 (41)	187 (2)	203 (10)	219 (21)	201 (50)	217 (56)
26 $[M+H-CH_4-3CO]^+$									201 (25)	217 (41)				201 (50)	217 (56)
27 $[M+H-CH_4-2CO-C_2H_2O]^+$									187 (4)	203 (18)				187 (13)	203 (37)
28 $[M+H-CH_4-4CO]^+$									173 (2)	189 (4)				173 (13)	189 (11)
29 $[M+H-2CO-2C_2H_2O]^+$	147 (13)	163 (7)	179 (7)	161 (13)	177 (2)	147 (9)	163 (7)	179 (9)	161 (4)		147 (8)	163 (5)	179 (9)		
31 $0,2A^+$	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 (2)	165 (1)
32 $0,2A^+-CO$	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 (8)	137 (13)
33 $0,2B^+$	121 (36)	137 (23)	153 (35)	135 (18)	151 (2)		137 (47)	153 (100)	135 (14)	151 (2)	121 (69)	137 (38)	153 (100)	135 (5)	151 (1)
35 $0,3A^++2H$				139 (100)	139 (9)			139 (3)	139 (7)	139 (1)			139 (3)	139 (4)	
37 $1,3A^+$	153 (61)	153 (20)	153 (35)	153 (28)	153 (4)		153 (100)	153 (100)	153 (65)	153 (100)	153 (100)	153 (100)	153 (100)	153 (67)	153 (100)
38 $1,3A^+-CO$								125 (3)			125 (2)	125 (1)	125 (5)		
39 $1,3A^+-C_2H_2O$	111 (19)	111 (7)	111 (4)			111 (5)	111 (6)	111 (6)	111 (1)		111 (4)	111 (5)	111 (7)	111 (3)	111 (6)
40 $1,3A^+-2CO$						97 (2)	97 (2)	97 (3)			97 (9)	97 (10)	97 (14)	97 (3)	97 (3)
41 $1,3A^+-2CO-C_2H_4$						69 (7)	69 (8)	69 (8)	69 (3)	69 (1)	69 (33)	69 (31)	69 (29)	69 (10)	69 (8)
42 $1,3B^+-2H$	133 (25)	149 (10)	165 (41)		163 (2)	133 (3)	149 (3)	165 (6)		163 (1)	133 (1)	149 (4)	165 (4)	147 (3)	163 (5)
49 $C_7H_7^+$						91 (2)	91 (2)	91 (2)	91 (4)						
50 $C_7H_5^+$											89 (6)	89 (9)	89 (7)	89 (2)	89 (2)

C Enzymatic methylation of non-catechols

C.1 SOMT expression studies

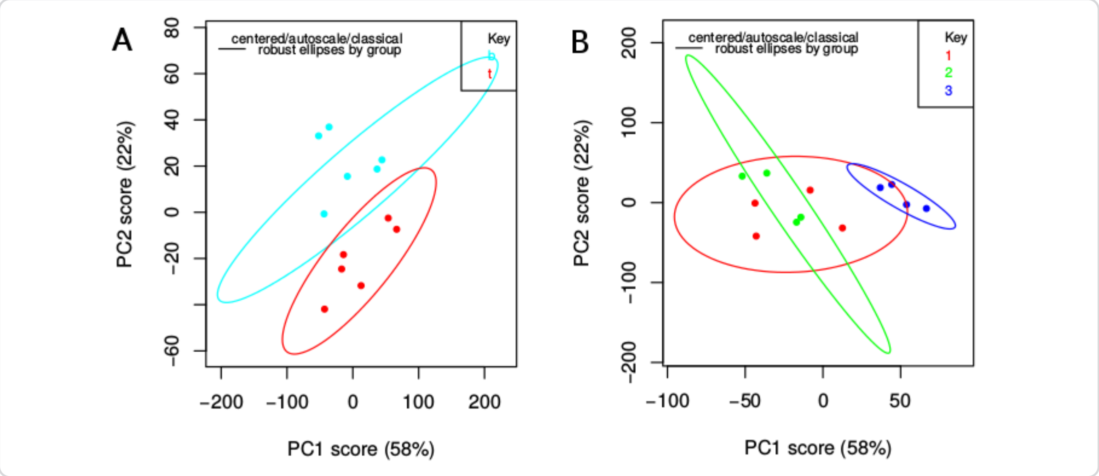


Figure C.1.: Additional scatterplots of the principal component analysis (PCA) of high-performance liquid chromatography (HPLC) data obtained from N. benthamiana leaves infiltrated by A. tumefaciens harbouring different constructs. ***A*** *– samples colored by leaf position (top: red; bottom: cyan),*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

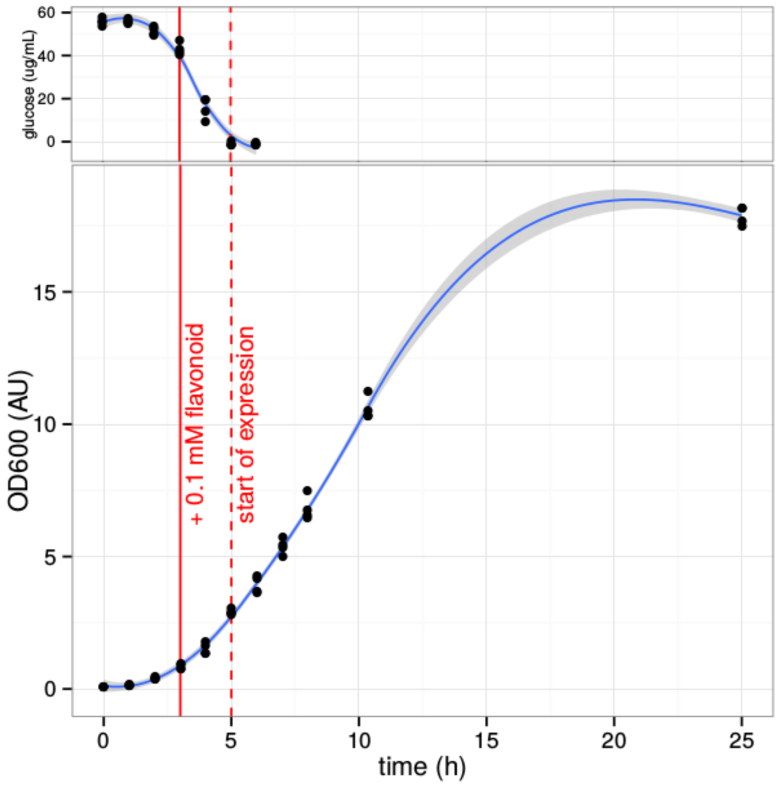


Figure C.2.: Growth curve of E. coli BL21(DE3) expressing soy O-methyl transferase (SOMT-2) at 37 °C. Glucose is depleted about 5 hours into growth, at which point the start of SOMT-2 expression is expected. The OD_{600} after inoculation was about 0.1.

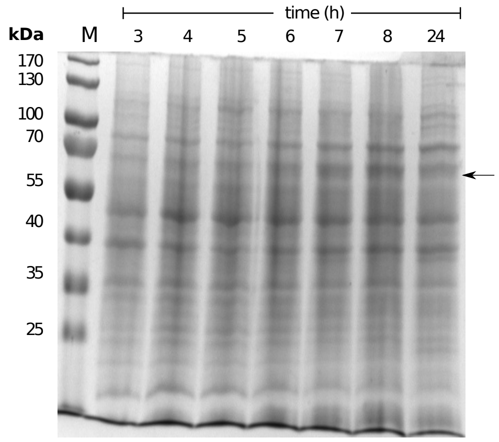


Figure C.3.: sodium dodecylsulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) gel of samples aquired during growth curve measurements. The arrow indicates the band that could correspond to the GST-tagged SOMT-2 protein.

Table C.1.: Results for the Analysis of Variance (ANOVA) of the main effects model describing soluble protein obtained after refolding.

	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			

significance codes: '•' 10 % level; '*' 5 % level

Table C.2.: Results for the ANOVA of the main effects model describing protein activity after refolding.

	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			

significance codes: '•' 10 % level; '*' 5 % level

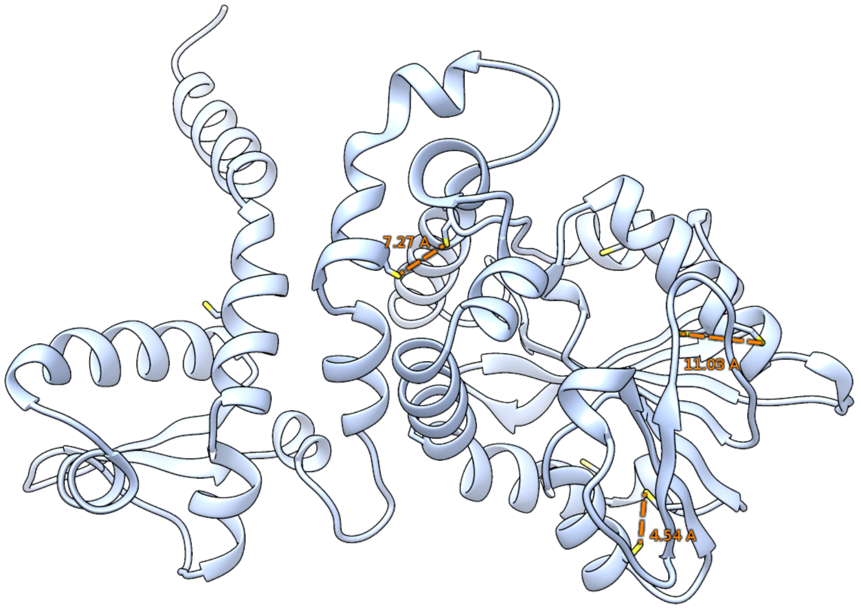


Figure C.4.: Graphical representation of a soy O-methyl transferase (SOMT-2) model obtained from the PHYRE2 web portal (http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index) [227]. Cysteines are shown as sticks. The distance between neighboring cysteines that could by oxidized to disulfide bridges is shown in orange.

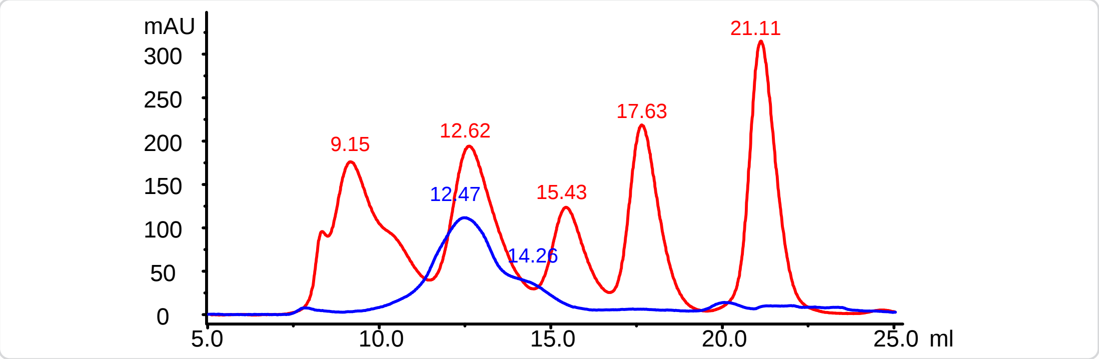


Figure C.5.: *Chromatogram of the gel filtration analysis of refolded SOMT-2 (blue). Gel-filtrations standards (red) were used to assess the size of the SOMT-2 protein. The estimated molecular weights for the eluting peaks were* 165 kDa *(*12.47 ml*) and* 65.5 kDa *(*14.26 ml*). Protein standard:* 9.15 ml – *thyroglobulin (*670 kDa*),* 12.62 ml – *γ-globulin (*158 kDa*),* 15.43 ml – *ovalbumin (*44 kDa*),* 17.63 ml – *myoglobin (*17 kDa*),* 21.11 ml – *vitamin B12 (*1.35 kDa*)*

C.2 Conversion of non-catechols by PFOMT

C.2.1 Modelling and shrinkage of catechols subset (pH profile)

The bell-shaped pH profile for the catecholic substrates showed, that there might be a quadratic relationship between pH and activity. A bell-shaped pH profile is common for most enzymatic reactions, where ionizable groups are involved [228]. A quadratic term was thus included into the linear model to capture this relationship:

$$activity = \beta_0 + \beta_1 \times \mathrm{Mg} + \beta_2 \times \mathrm{pH} + \beta_3 \times (\mathrm{Mg} \times \mathrm{pH}) + \beta_4 \times \mathrm{pH}^2 + \beta_5 \times (\mathrm{pH}^2 \times \mathrm{Mg}). \quad \text{(C.1)}$$

The model describes the actual data reasonable well, with about 68.6 % of the variance explained ($R^2 = 0.6855$)(Figure 7.10). Also, the p-values for the coefficients β_3 (0.074) and β_4 (0.0492) suggest an interdependency of Mg^{2+} and pH (Table C.3), at significant levels of 10 and 5 %, respectively. The coefficient estimate of 266 920 for β_3 illustrates that for the catecholic substrates the effect of the pH is much larger than for the methylation of *iso*-ferulic acid. In addition to the simplified linear model (Equation C.1) a more complex linear model,

Table C.3.: Coefficients of the model (C.1) for activity of catechol methylation by PFOMT. 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^2	-4977.7406	6512.6996	-0.76	0.4574	
Mg×pH	266920.7964	138224.0638	1.93	0.0740	•
Mg×pH^2	-19830.2264	9210.3481	-2.15	0.0492	*

significance codes: ’•’ 10 % level; ’*’ 5 % level

$$\begin{aligned} activity = \beta_0 &+ \beta_1 \times \mathrm{Mg} + \beta_2 \times \mathrm{pH} + \beta_3 \times \mathrm{pH}^2 + \beta_4 \times (\mathrm{Mg} \times \mathrm{pH}) + \\ &\beta_5 \times (\mathrm{pH}^2 \times \mathrm{Mg}) + \beta_6 \times (\mathrm{pH} \times \mathrm{pH}^2) + \beta_7 \times (\mathrm{Mg} \times \mathrm{pH} \times \mathrm{pH}^2), \quad \text{(C.2)} \end{aligned}$$

was prepared and shrunken via the Lasso method and 5-fold cross validation (Table C.4) [229]. The Lasso is a shrinkage method. It can shrink coefficients to exactly zero, thus making a model less complex and therefore more interpretable [229]. The shrunken model only contained the factors pH, pH×Mg, pH×pH^2 and pH×Mg×pH^2. The large coefficient estimate for parameter β_2 (pH) suggests, that in fact the pH has a large influence on the activity. This is contrary to the linear model (Equation C.1), which, judged by the *p*-value for this coefficient, suggested otherwise. However, the shrunken model also shows that the activity is dependent on the interaction of pH and magnesium, which supports the implications of the linear model (Equation C.1). The results of the the shrunken model and the results obtained by linear modelling are further statistical evidence that pH and Mg^{2+} show main effects and also interaction effects which seem to be associated with the enzyme's activity towards catecholic substrates (i.e eriodictyol, caffeic acid). Nonetheless, all of these rather simple models can not reflect the reality of such complex systems as enzymes, where many more factors play important roles.

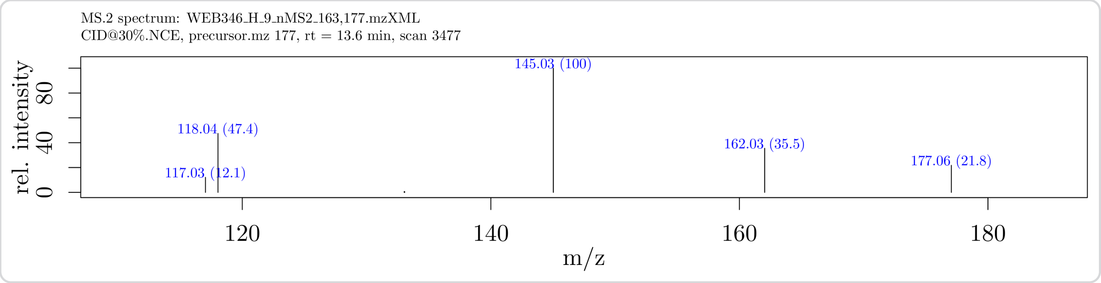
Table C.4.: Coefficients obtained for linear regression model using the catechols subset after shrinkage using the Lasso method and 5-fold cross validation. Only non-zero coefficients (variables actually do have an effect) are retained during the Lasso. Seed was set to 1336.

variable	coefficient
(Intercept)	-467632.3821
pH	94469.8366
pH×Mg	19068.9540
pH×pH^2	-381.5863
pH×Mg×pH^2	-292.3608

C.3 Identification of products from conversion of non-catechols by PFOMT

C.3.1 *p*-Coumaric acid methyl ester

The product obtained by conversion of *p*-coumaric acid by phenylpropanoid and flavonoid *O*-methyl transferase (PFOMT) was determined as 4-hydroxy cinnamic acid methyl ester. The negative mode MS^2 spectrum showed four prominent peaks *m/z* 177 (21) $[\text{M-H}]^-$, 162(35), 145(100) and 118(47). If the enzymatic product was the methyl ether, one would expect a strong *m/z* 133, corresponding to $[\text{M-H-CO}_2]^-$ [230]. However, *m/z* 133 was not observed, strongly suggesting the methyl ester. Comparison of the obtained data with literature data confirmed the methyl ester as sole product [223].

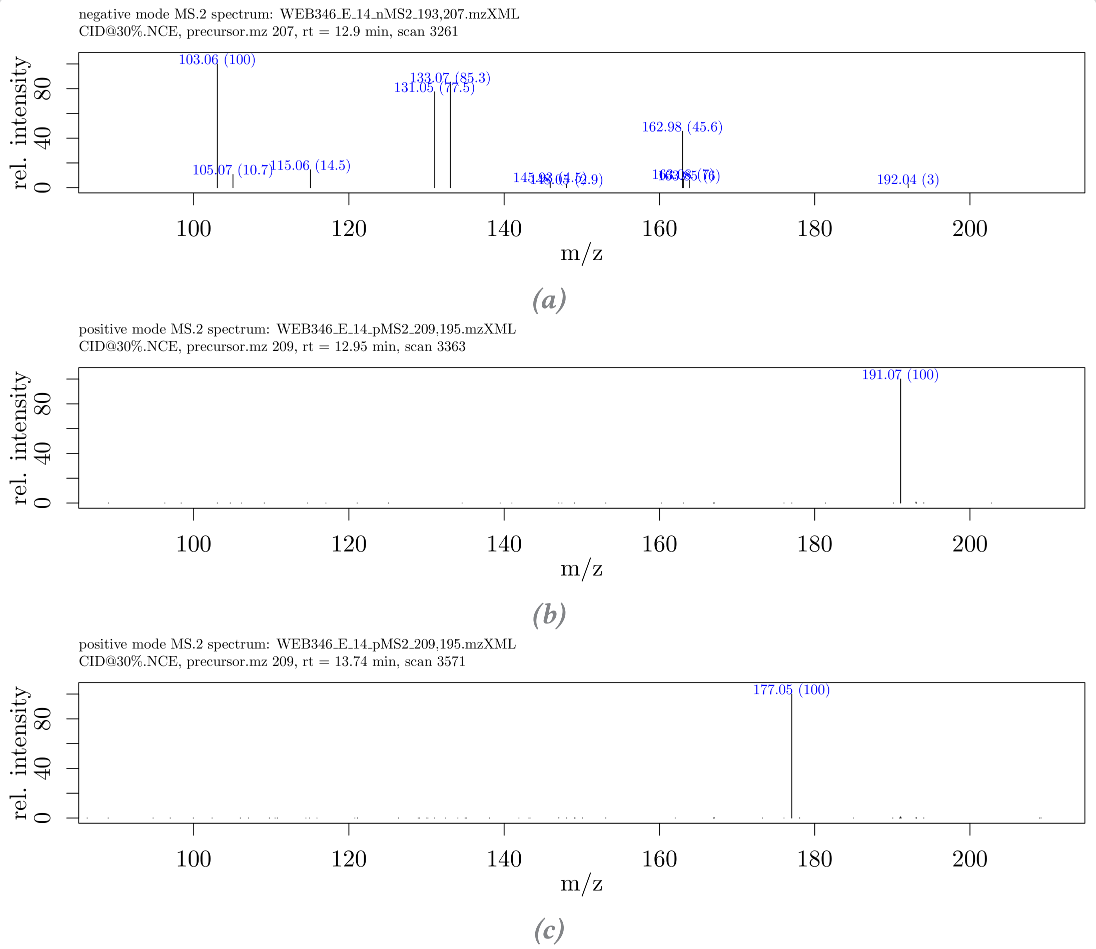


C.3.2 *iso*-Ferulic acid esters and caffeic acid dimethylether

Methylation of *iso*-ferulic acid and ferulic acid afforded two methylated products with retention times of 12.9 and 13.7 min. The compound eluting at 12.9 min was identified as caffeic acid dimethylether (**1**) through comparison to an authentic standard, whereas the compounds eluting about one minute later were identified as the ferulic (**2**) and *iso*-ferulic acid methyl esters (**3**). Since the the retention time is an indicator for the polarity of an eluting compound and the methyl ester is much more unpolar than the dimethyl ether, it comes as no surprise that the latter elutes earlier on a reversed-phase column. Ionization of the enzymatic products was difficult in negative mode, but easily achieved in positive mode. The only peaks in the positive mode MS^2 spectra of (**1**) and (**2**,**3**) were *m/z* 191 and *m/z* 177, respectively. This indicates a loss of water and methanol from the dimethylether and methyl ester, respectively.

C.3.3 3’,4’-Dimethyl eriodictyol

Conversion of homoeriodictyol or hesperetin by PFOMT afforded 3’,4’-dimethyl eriodictyol (**4**). The product was identified by liquid chromatography-tandem mass spectrometry (LC-MS/MS). Products from both conversions possessed the same retention times of 14.54 min. The collision induced dissociation (CID) spectra of these products showed five distinct signals at *m/z* 299 (14), 191(100), 179(62), 165(17) and 153(67) (Figure C.7). The *m/z* 299 corresponds to the ion $[M+H-H_2O]^+$, showing that both, homoeriodictyol as well as hesperetin were methylated. The *m/z* 153 corresponds to



negative mode MS.2 spectrum: WEB346_E_14_nMS2_193,207.mzXML
CID@30%.NCE, precursor.mz 207, rt = 12.9 min, scan 3261

(a)

positive mode MS.2 spectrum: WEB346_E_14_pMS2_209,195.mzXML
CID@30%.NCE, precursor.mz 209, rt = 12.95 min, scan 3363

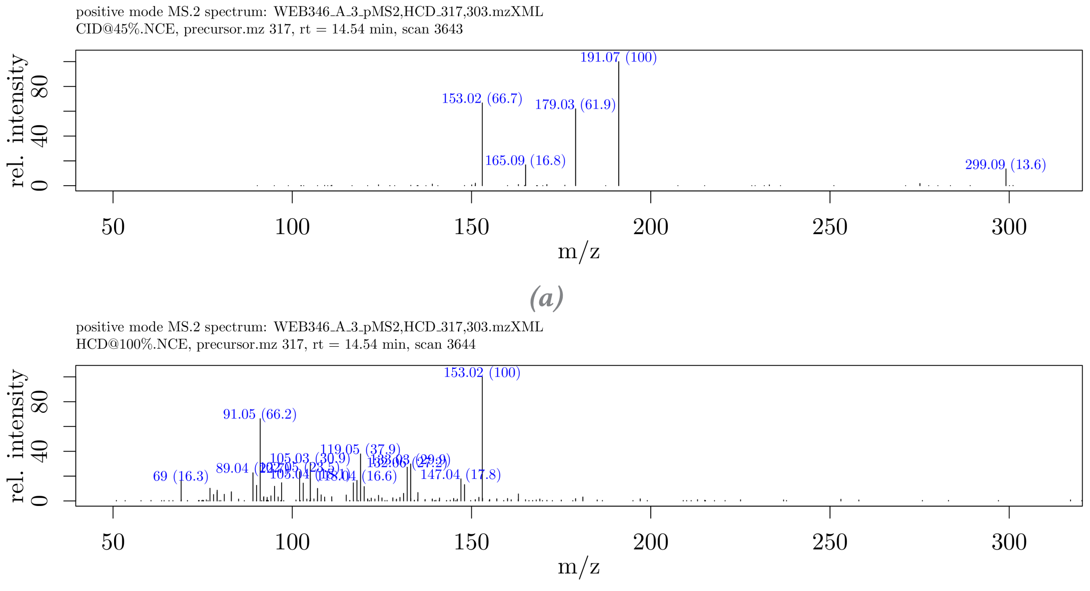
(b)

positive mode MS.2 spectrum: WEB346_E_14_pMS2_209,195.mzXML
CID@30%.NCE, precursor.mz 209, rt = 13.74 min, scan 3571

(c)

*Figure C.6.: MS^2 spectra of (**1**) and (**3**). (**a**) negative mode MS^2 of (**1**). (**b**) positive mode MS^2 of (**1**). (**c**) positive mode MS^2 of (**3**).*

the $^{1,3}A^+$ fragment, which is characteristic for 3,7-dihydroxy substituted flavonoids. This indicates a methylation of the B-ring. Further evidence of a dimethoxylated B-ring is the fragment $^{1,4}B^+$-2H with *m/z* 191. The fragmentation pattern of 3',4'-dimethyl eriodictyol agrees with the general fragmentation of flavanones described in chapter 6.



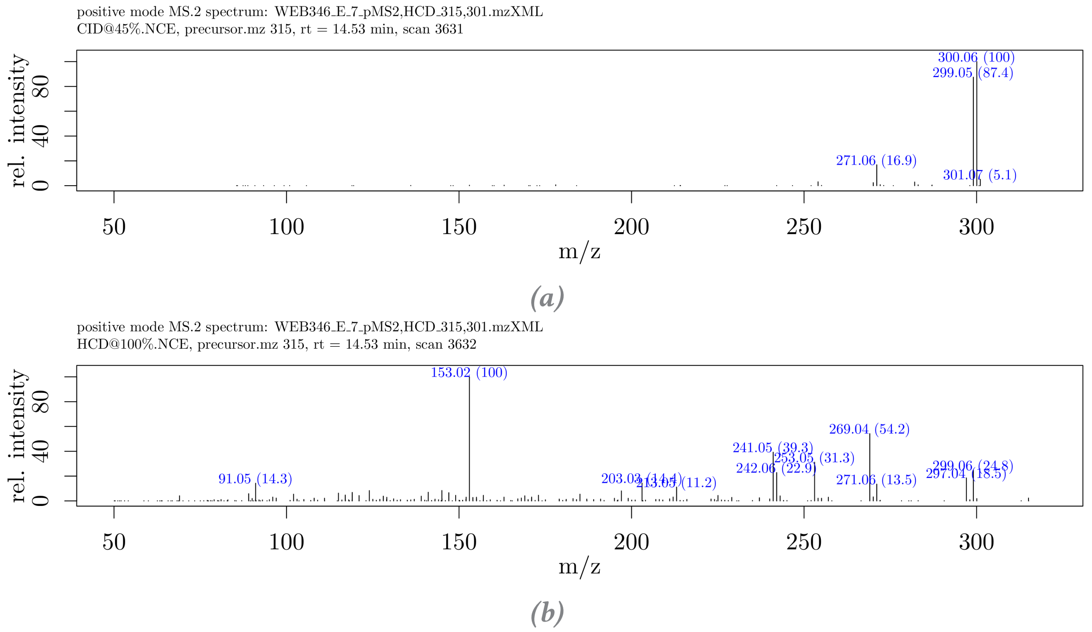
(a)

(b)

*Figure C.7.: MS^2 spectra of 3'4'-dimethyl eriodictyol (**4**). (**a**) positive mode MS^2 CID spectrum of (**4**). (**b**) positive mode MS^2 HCD spectrum of (**4**).*

C.3.4 3',4'-Dimethyl luteolin

Conversion of diosmetin and chrysoeriol by PFOMT afforded 3',4'-dimethyl luteolin (**5**) and an unidentified product. (**5**) eluted after 14.53 min. The CID spectrum of (**5**) shows three signals, *m/z* 300 (100), 299(87) and 271(17) (Figure C.8). These signals correspond to the $[M{+}H{-}CH_3]^{\bullet+}$, $[M{+}H{-}CH_4]^+$ and $[M{+}H{-}CH_4{-}CO]^+$ ions, respectively. The higher-energy collisional dissociation (HCD) spectrum of (**5**) clearly shows a peak with *m/z* 153 amongst others. Again, this is spectrometric evidence of a 3,7-dihydroxylated flavonoid (fragment $^{1,3}A^+$) demonstrating a 3',4'-dimethylation. The unidentified products of the conversions of apigenin, diosmetin and chrysoeriol eluted after 12.59, 12.77 and 12.69 min, respectively.



(a)

(b)

*Figure C.8.: MS^2 spectra of 3'4'-dimethyl luteolin (**5**). (**a**) positive mode MS^2 CID spectrum of (**5**). (**b**) positive mode MS^2 HCD spectrum of (**5**).*

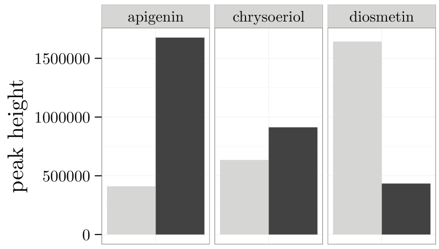


Figure C.9.: Product composition after conversion of flavones with PFOMT. Bar chart of the peak heights of the unidentified (black) and 3',4'-O-dimethylated products (gray) in the selected ion chromatograms (HCD at 100 % NCE). The conversion experiments were conducted with the wild-type PFOMT at pH 8.6 with 10 mM Mg^{2+} added.

D Additional information

Table D.2.: SAM analogues that have been used with MTs. Targets: P – *peptide/protein,* D – *DNA,* R – *RNA,* S – *small molecule.*

analogue	enzyme	target	references
SAM			
$-CH_2-CH_3$	PRMT1, M.TaqI, M.HhaI, M.BcnIB, RebM, RapM	S,P,D	[102, 112, 113, 231] [1]
$-CH_2-CH_2-CH_3$	PRMT1, M.TaqI, M.HhaI, M.BcnIB	P,D	[102, 231]
$-CH_2-CH_2-CH_2-CH_3$	PRMT1	P	[231]
$-CH_2-C_6H_5$	NovO, CouO, PRMT1	S,P	[99, 231]
$-CH_2-C(=O)-CH_3$	COMT, TPMT, CazF	S	[111, 232]

---

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

analogue	enzyme	target	references
$-CH_2-CH=CH_2$	NovO, CouO, RapM, PRMT1, M.TaqI, M.HhaI, M.BcnIB, RebM, Tgs	P,S,D	[99, 102, 104, 108, 112, 113, 231, 233, 234]
$-CH_2-CH=CH-CH_3$	NovO, CouO	S	[99]
$-CH_2-C\equiv CH$	Dim-5, *Hs*MLL, TRM1, NovO, CouO, PRMT1	P,R,S	[99, 104, 107, 233, 235]
$-CH_2-C\equiv N$	RebM	S	[113]
$-CH_2-CH_2-C\equiv CH$	PKMT	P	[235]
$-CH_2-CH_2-CH_2-C\equiv CH$	PKMT	P	[235]
$-CH_2-C\equiv C-CH_3$	NovO, CouO, M.HhaI, M.TaqI, M.BcnIB	S,D	[99, 102, 236]
$-CH_2-C\equiv C-CH_2-CH_3$	M.HhaI	D	[236]
$-CH_2-C\equiv C-CH_2-NH_2$	M.HhaI	D	[236]
$-CH_2-C\equiv C-CH_2-NH-C(=O)(-CH_2-)_3-NH_2$	M.HhaI	D	[236]
$-CH_2-C\equiv C(-CH_2-)_3-NH_2$	M.HhaI	D	[236]
$-CH_2-C\equiv C(-CH_2-)_3-NH-C(=O)(-CH_2-)_3-NH_2$	M.HhaI	D	[236]
$-CH_2-C\equiv C(-CH_2-)_3-C\equiv CH$	M.HhaI	D	[236]
$-CH_2-C\equiv C(-CH_2-)_3-N_3$	M.HhaI	D	[236]
$-CH_2-CH=CH-C\equiv CH$	Dim-5, *Hs*MLL, TRM1, PRMT1, Tgs	P,R	[104–108, 233, 235]
$-CH_2-CH=CH-CH_2-C\equiv CH$			[233, 235]
$-CH_2-CH=CH-CH_2-O-CH_2-C\equiv CH$	PRMT1	P	[104, 233]
SeAM			
$-CH_3$			

analogue	enzyme	target	references
$-CH_2-C\equiv CH$	Dim-5, *Hs*MLL, TRM1, RebM, CazF	P,R,S	[107, 109, 113, 232]
N-mustard derivatives			
$-CH_2-CH_2-I$	RebM	S	[114]

Table D.1.: Overview over the contructs 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)
pBEW103	pBEW102 with BamHI cloning site	pBEW102		
pBEW104	rhaP$_{BAD}$ promoter	pBEW4b	pBEW103	amplification (*pRha1.fw/rv*), cloning (BglII, BamHI)
pBEW106	pICH413038-somt	pET28MC-somt	pICH413038	amplification (*somt1/2/3/4*), golden gate cloning (BpiI)
pBEW107		pICH51266, pBEW106, pICH41421	pICH75044	golden gate cloning (BsaI)
pET28-pfomt	*pfomt* gene in pET-28a(+), endogenous NdeI site removed	pQE30-pfomt	pET-28a(+)	mutagenesis (*pfomt1.fw/rv*), amplification (*pfomt2.fw/rv*), cloning (NdeI, EcoRI)
pET20-somt	N-terminal pelB-tag fusion for periplasmic expression		pET20-b(+)	
pET28-somt			pET28-a(+)	
pET32-somt	N-terminal TrX-tag fusion		pET-32a(+)	
pET41-somt	N-terminal GST-tag fusion		pET-41a(+)	
pUC19*	added BglII site	pUC19	-	mutagenesis (*pUC1.fw/rv*)
pUCB1	pUC19 derivative with lsrA promoter	lsr-XX-DAS	pUC19*	cloning (NdeI, BglII)

Bibliography

(1) Croteau, R.; Kutchan, T. M.; Lewis, N. G. In *Biochemistry & Molecular Biology of Plants*, Buchanan, B., Gruissem, W., Jones, R., Eds.; 7, 2000; Vol. 7, pp 1250–1318.

(2) Barton, G. *Phytochemistry* **Jan. 1972**, *11*, 426–429.

(3) Dagne, E.; Bekele, A.; Noguchi, H.; Shibuya, M.; Sankawa, U. *Phytochemistry* **1990**, *29*, 2671–2673.

(4) Wollenweber, E. *Phytochemistry* **1985**, *24*, 1493–1494.

(5) Simonsen, H. T.; Larsen, M. D.; Nielsen, M. W.; Adsersen, A.; Olsen, C. E.; Strasberg, D.; Smitt, U. W.; Jaroszewski, J. W. *Phytochemistry* **Aug. 2002**, *60*, 817–820.

(6) Hegnauer, R., *The Flavonoids*, 1st ed.; Harborne, J. B., Ed.; 2; Chapman and Hall: 1976; Vol. 15, p 357.

(7) Agati, G.; Azzarello, E.; Pollastri, S.; Tattini, M. *Plant Science* **Nov. 2012**, *196*, 67–76.

(8) Taylor, L. P.; Grotewold, E. *Current Opinion in Plant Biology* **June 2005**, *8*, 317–23.

(9) Kumar, S.; Pandey, A. K. *The Scientific World Journal* **2013**, *2013*, 1–16.

(10) Goodrich, R. M.; Braddock, R. J. *Processing* **2006**, 1–4.

(11) Mulinacci, N.; Romani, a.; Galardi, C.; Pinelli, P.; Giaccherini, C.; Vincieri, F. F. *Journal of Agricultural and Food Chemistry* **2001**, *49*, 3509–3514.

(12) Ouallagui, Z. B.; Ouaziz, M. B.; An, J. H.; Oukhris, M. B.; Igane, G. R.; Riha, I. F.; Emai, H. J.; Ki, I. F.; Horbel, H. G.; Soda, H. I.; Ayadi, S. S. *Journal of Arid Land Studies* **2012**, *64*, 61–64.

(13) Wessjohann, L. A.; Keim, J.; Weigel, B.; Dippe, M. *Current Opinion in Chemical Biology* **Mar. 2013**, *17*, 229–235.

(14) Trantas, E. A.; Koffas, M. A. G.; Xu, P.; Ververidis, F. *Frontiers in Plant Science* **2015**, *6*, 1–16.

(15) Zhu, S.; Wu, J.; Du, G.; Zhou, J.; Chen, J. *Applied and Environmental Microbiology* **2014**, *80*, 3072–3080.

(16) Koopman, F.; Beekwilder, J.; Crimi, B.; van Houwelingen, A.; Hall, R. D.; Bosch, D.; van Maris, A. J.; Pronk, J. T.; Daran, J.-M. *Microbial Cell Factories* **2012**, *11*, 155.

(17) Louie, G. V.; Bowman, M. E.; Moffitt, M. C.; Baiga, T. J.; Moore, B. S.; Noel, J. P. *Chemistry and Biology* **2006**, *13*, 1327–1338.

(18) Grotewold, E., *The Science of Flavonoids The Science of Flavonoids*, 1st ed.; Grotewold, E., Ed.; Springer: New York, 2006, pp 1–274.

(19) Urban, P.; Werck-Reichhart, D.; Teutsch, H. G.; Durst, F.; Regnier, S.; Kazmaier, M.; Pompon, D. *European Journal of Biochemistry* **1994**, *222*, 843–850.

(20) Hu, Y.; Gai, Y.; Yin, L.; Wang, X.; Feng, C.; Feng, L.; Li, D.; Jiang, X.-N.; Wang, D.-C. *The Plant Cell* **2010**, *22*, 3093–3104.

(21) Ferrer, J.-L.; Jez, J. M.; Bowman, M. E.; Dixon, R. A.; Noel, J. P. *Nature Structural Biology* **Aug. 1999**, *6*, 775–784.

(22) Jez, J. M.; Bowman, M. E.; Dixon, R. A.; Noel, J. P. *Nature Structural Biology* **2000**, *7*, 786–791.

(23) Gholami, A.; De Geyter, N.; Pollier, J.; Goormachtig, S.; Goossens, A. *Natural Product Reports* **2014**, *31*, 356–80.

(24) Cao, G.; Sofic, E.; Prior, R. L. *Free Radical Biology and Medicine* **Jan. 1997**, *22*, 749–760.

(25) Leopoldini, M.; Russo, N.; Chiodo, S.; Toscano, M. *Journal of Agricultural and Food Chemistry* **2006**, *54*, 6343–6351.

(26) Metodiewa, D.; Jaiswal, A. K.; Cenas, N.; Dickancaité, E.; Segura-Aguilar, J. *Free Radical Biology and Medicine* **Jan. 1999**, *26*, 107–116.

(27) Cushnie, T. P.; Lamb, A. J. *International Journal of Antimicrobial Agents* **Nov. 2005**, *26*, 343–356.

(28) Friedman, M.; Henika, P. R.; Levin, C. E.; Mandrell, R. E.; Kozukue, N. *Journal of Food Protection* **2006**, *69*, 354–361.

(29) Cowan, M. M. *Clinical Microbiology Reviews* **1999**, *12*, 564–582.

(30) Thomas, D. J.; Li, J.; Waters, S. B.; Xing, W.; Adair, B. M.; Drobna, Z.; Devesa, V.; Styblo, M. *Experimental Biology and Medicine* **Jan. 2007**, *232*, 3–13.

(31) Schmidberger, J. W.; James, A. B.; Edwards, R.; Naismith, J. H.; O&apos;Hagan, D. *Angewandte Chemie - International Edition* **2010**, *49*, 3646–3648.

(32) Kim, B. G.; Sung, S. H.; Chong, Y.; Lim, Y.; Ahn, J. H. *Journal of Plant Biology* **Sept. 2010**, *53*, 321–329.

(33) Kopycki, J. G.; Rauh, D.; Chumanevich, A. a.; Neumann, P.; Vogt, T.; Stubbs, M. T. *Journal of Molecular Biology* **Apr. 2008**, *378*, 154–164.

(34) Struck, A. W.; Thompson, M. L.; Wong, L. S.; Micklefield, J. *ChemBioChem* **Nov. 2012**, *13*, 2642–2655.

(35) Copeland, R. A. *Clinical Cancer Research* **2013**, *19*, 6344–6350.

(36) Robertson, K. D. *Oncogene* **2001**, *20*, 3139–3155.

(37) Robertson, K. D. *Nature Reviews Genetics* **Aug. 2005**, *6*, 597–610.

(38) McClelland, M.; Nelson, M.; Raschke, E. *Nucleic Acids Research* **1994**, *22*, 3640–3659.

(39) Schubert, H. L.; Blumenthal, R. M.; Cheng, X. *Trends in Biochemical Sciences* **2003**, *28*, 329–335.

(40) Dixon, M. M.; Huang, S.; Matthews, R. G.; Ludwig, M. *Structure* **1996**, *4*, 1263–1275.

(41) Schubert, H. L.; Wilson, K. S.; Raux, E.; Woodcock, S. C.; Warren, M. J. *Nature Structural Biology* **1998**, *5*, 585–592.

(42) Michel, G.; Sauvé, V.; Larocque, R.; Li, Y.; Matte, A.; Cygler, M. *Structure* **2002**, *10*, 1303–1315.

(43) Xiao, B.; Jing, C.; Wilson, J.; Walker, P.; Vasisht, N.; Kelly, G.; Howell, S.; Taylor, I.; Blackburn, G.; Gamblin, S. *Nature* **2003**, *421*, 652–656.

(44) Wlodarski, T.; Kutner, J.; Towpik, J.; Knizewski, L.; Rychlewski, L.; Kudlicki, A.; Rowicka, M.; Dziembowski, A.; Ginalski, K. *PLoS ONE* **2011**, *6*, DOI: *10.1371/journal.pone.0023168*.

(45) Werner, W. J.; Allen, K. D.; Hu, K.; Helms, G. L.; Chen, B. S.; Wang, S. C. *Biochemistry* **2011**, *50*, 8986–8988.

(46) Boal, A. K.; Grove, T. L.; McLaughlin, M. I.; Yennawar, N. H.; Booker, S. J.; Rosenzweig, A. C. *Science (New York, N.Y.)* **2011**, *332*, 1089–1092.

(47) Broderick, J. B.; Duffus, B. R.; Duschene, K. S.; Shepard, E. M. *Chemical Reviews* **Apr. 2014**, *114*, 4229–4317.

(48) Zhang, Q.; van der Donk, W. A.; Liu, W. *Accounts of Chemical Research* **Apr. 2012**, *45*, 555–564.

(49) Layer, G.; Moser, J.; Heinz, D. W.; Jahn, D.; Schubert, W. D. *EMBO Journal* **2003**, *22*, 6214–6224.

(50) Cantoni, G. L. *The Journal of Biological Chemistry* **1953**, *204*, 403–416.

(51) Butte, W. *Journal of Chromatography A* **1983**, *261*, 142–145.

(52) Borchardt, R. T. *Journal of the American Chemical Society* **1979**, *101*, 458–463.

(53) Hoffman, J. L. *Biochemistry* **1986**, *25*, 4444–4449.

(54) Takusagawa, F.; Kamitori, S.; Misaki, S.; Markham, G. D. *The Journal of Biological Chemistry* **1996**, *271*, 136–147.

(55) Lozada-Ramírez, J. D.; Martínez-Martínez, I.; García-Carmona, F.; Sánchez-Ferrer, A. *Biotechnology Progress* **2008**, *24*, 120–7.

(56) Banerjee, R. V.; Matthews, R. G. *The FASEB Journal* **1990**, *4*, 1450–1459.

(57) Gray, C. H.; Coward, J. K.; Schowen, K. B.; Schowen, R. L. *Journal of the American Chemical Society* **1979**, *101*, 4351–4358.

(58) Mihel, I.; Knipe, J. O.; Coward, J. K.; Schowen, R. L. *Journal of the American Chemical Society* **July 1979**, *101*, 4349–4351.

(59) Floss, H. G.; Lee, S. *Accounts of Chemical Research* **Mar. 1993**, *26*, 116–122.

(60) Floss, H. G.; Tsai, M. D. *Advances in Enzymology and Related Areas of Molecular Biology* **Jan. 1979**, *50*, 243–302.

(61) Woodard, R. W.; Mascaro, L.; Horhammer, R.; Eisenstein, S.; Floss, H. G. *Journal of the American Chemical Society* **1980**, *102*, 6314–6318.

(62) Woodard, R. W.; Tsai, M.-D.; Floss, H. G.; Crooks, P. A.; Coward, J. K. *The Journal of Biological Chemistry* **1980**, *255*, 9124–9127.

(63) Zubieta, C.; Kota, P.; Ferrer, J.-l.; Dixon, R. a.; Noel, J. P. *The Plant Cell* **2002**, *14*, 1265–1277.

(64) Zubieta, C.; He, X.-Z.; Dixon, R. A.; Noel, J. P. *Nature Structural Biology* **Mar. 2001**, *8*, 271–279.

(65) Brandt, W.; Manke, K.; Vogt, T. *Phytochemistry* **May 2015**, *113*, 130–139.

(66) Vidgren, J.; Svensson, L. A.; Liljas, A. *Nature* **1994**, *368*, 354–358.

(67) Fersht, A., *Structure and mechanism in protein science*; Julet, M. R., Hadler, G. L., Eds.; W.H. Freeman and Company: New York, 1999.

(68) Klimasauskas, S.; Kumar, S.; Roberts, R. J.; Cheng, X. *Cell* **1994**, *76*, 357–369.

(69) Coward, J. K.; Slixz, E. P.; Wu, F. Y. *Biochemistry* **1973**, *12*, 2291–2297.

(70) Simms, S. A.; Subbaramaiah, K. *Journal of Biological Chemistry* **1991**, *266*, 12741–12746.

(71) Sawicki, A.; Willows, R. D. *The Biochemical Journal* **2007**, *406*, 469–478.

(72) Yee, W. C.; Eglsaer, S. .; Richards, W. R. *Biochemical and Biophysical Research Communications* **1989**, *162*, 483–490.

(73) Bhattacharya, S. K.; Dubey, A. K. *Journal of Biological Chemistry* **1999**, *274*, 14743–14749.

(74) Shi, Y. Q.; Rando, R. R. *Journal of Biological Chemistry* **1992**, *267*, 9547–9551.

(75) Ball, P.; Knuppen, R.; Haupt, M.; Breuer, H. *European Journal of Biochemistry* **1972**, *26*, 560–569.

(76) Joshi, C. P.; Chiang, V. L. *Plant Molecular Biology* **1998**, *37*, 663–674.

(77) Zhang, X.-H.; Chinnappa, C. C. *Journal of Biosciences* **Mar. 1997**, *22*, 161–175.

(78) Ibdah, M.; Zhang, X. H.; Schmidt, J.; Vogt, T. *Journal of Biological Chemistry* **Nov. 2003**, *278*, 43961–43972.

(79) Vogt, T. *FEBS Letters* **Mar. 2004**, *561*, 159–162.

(80) Kim, B. G.; Shin, K. H.; Lee, Y.; Hur, H. G.; Lim, Y.; Ahn, J. H. *Biotechnology Letters* **Dec. 2005**, *27*, 1861–1864.

(81) Kim, D. H.; Kim, B. G.; Lee, Y.; Ji, Y. R.; Lim, Y.; Hur, H. G.; Ahn, J. H. *Journal of Biotechnology* **Sept. 2005**, *119*, 155–162.

(82) Kim, M. J.; Kim, B. G.; Ahn, J. H. *Applied Microbiology and Biotechnology* **2013**, *97*, 7195–7204.

(83) Schröder, G.; Wehinger, E.; Lukačin, R.; Wellmann, F.; Seefelder, W.; Schwab, W.; Schröder, J. *Phytochemistry* **2004**, *65*, 1085–1094.

(84) Deavours, B. E.; Liu, C. J.; Naoumkina, M. a.; Tang, Y.; Farag, M. a.; Sumner, L. W.; Noel, J. P.; Dixon, R. a. *Plant Molecular Biology* **2006**, *62*, 715–733.

(85) Brandt, W.; Bräuer, L.; Günnewich, N.; Kufka, J.; Rausch, F.; Schulze, D.; Schulze, E.; Weber, R.; Zakharova, S.; Wessjohann, L. *Phytochemistry* **Oct. 2009**, *70*, 1758–1775.

(86) Lairson, L. L.; Henrissat, B.; Davies, G. J.; Withers, S. G. *Annual Review of Biochemistry* **2008**, *77*, 521–555.

(87) Li, K.; Frost, J. W. *Journal of the American Chemical Society* **1998**, *7863*, 10545–10546.

(88) Bong, G. K.; Kim, H.; Jeong, H. K.; Lim, Y.; Ahn, J. H. *Bulletin of the Korean Chemical Society* **2006**, *27*, 357–358.

(89) Joe, E. J.; Kim, B.-g.; An, B.-c.; Chong, Y.; Ahn, J.-h. *Molecules and Cells* **Aug. 2010**, *30*, 137–41.

(90) Kim, B.-G.; Joe, E. J.; Ahn, J.-H. *Biotechnology Letters* **Apr. 2010**, *32*, 579–84.

(91) Leonard, E.; Chemler, J.; Lim, K. H.; Koffas, M. a. G. *Applied Microbiology and Biotechnology* **Mar. 2006**, *70*, 85–91.

(92) Malla, S.; Koffas, M. A. G.; Kazlauskas, R. J.; Kim, B.-G. *Applied and Environmental Microbiology* **Feb. 2012**, *78*, 684–94.

(93) Morishige, T.; Choi, K.-B.; Sato, F. *Bioscience, Biotechnology and Biochemistry* **May 2014**, *68*, 939–941.

(94) Markoglou, N.; Wainer, I. W. *Journal of chromatography. A* **Mar. 2002**, *948*, 249–56.

(95) Palazón, J.; Navarro-Ocaña, A.; Hernandez-Vazquez, L.; Mirjalili, M. H. *Molecules* **Jan. 2008**, *13*, 1722–42.

(96) Rigbers, O.; Li, S.-M. *Journal of Biological Chemistry* **Oct. 2008**, *283*, 26859–26868.

(97) Zhang, Q.; van der Donk, W. A. *FEBS Letters* **Sept. 2012**, *586*, 3391–7.

(98) Crnovčić, I.; Süssmuth, R.; Keller, U. *Biochemistry* **Nov. 2010**, *49*, 9698–9705.

(99) Stecher, H.; Tengg, M.; Ueberbacher, B. J.; Remler, P.; Schwab, H.; Griengl, H.; Gruber-Khadjawi, M. *Angewandte Chemie - International Edition* **Jan. 2009**, *48*, 9546–9548.

(100) Hunter, S. C.; Cahoon, E. B. *Lipids* **Mar. 2007**, *42*, 97–108.

(101) Coiner, H.; Schröder, G.; Wehinger, E.; Liu, C.-J.; Noel, J. P.; Schwab, W.; Schröder, J. *The Plant Journal* **Apr. 2006**, *46*, 193–205.

(102) Dalhoff, C.; Lukinavicius, G.; Klimasăuskas, S.; Weinhold, E. *Nature Chemical Biology* **2006**, *2*, 31–32.

(103) Dalhoff, C.; Lukinavicius, G.; Klimasauskas, S.; Weinhold, E. *Nature Protocols* **2006**, *1*, 1879–1886.

(104) Wang, R.; Zheng, W.; Yu, H.; Deng, H.; Luo, M. *Journal of the American Chemical Society* **2011**, *133*, 7648–7651.

(105) Peters, W.; Willnow, S.; Duisken, M.; Kleine, H.; Macherey, T.; Duncan, K. E.; Litchfield, D. W.; Lüscher, B.; Weinhold, E. *Angewandte Chemie - International Edition* **2010**, *49*, 5170–5173.

(106) Motorin, Y.; Burhenne, J.; Teimer, R.; Koynov, K.; Willnow, S.; Weinhold, E.; Helm, M. *Nucleic Acids Research* **2011**, *39*, 1943–1952.

(107) Willnow, S.; Martin, M.; Lüscher, B.; Weinhold, E. *ChemBioChem* **2012**, *13*, 1167–1173.

(108) Schulz, D.; Holstein, J. M.; Rentmeister, A. *Angewandte Chemie - International Edition* **2013**, *52*, 7874–7878.

(109) Bothwell, I. R.; Islam, K.; Chen, Y.; Zheng, W.; Blum, G.; Deng, H.; Luo, M. *Journal of the American Chemical Society* **2012**, *134*, 14905–14912.

(110) Bothwell, I. R.; Luo, M. *Organic Letters* **June 2014**, *16*, 3056–3059.

(111) Lee, B. W. K.; Sun, H. G.; Zang, T.; Ju-Kim, B.; Alfaro, J. F.; Zhou, Z. S. *Journal of the American Chemical Society* **2010**, *132*, 3642–3643.

(112) Law, B. J. C.; Struck, A.-W.; Bennett, M. R.; Wilkinson, B.; Micklefield, J. *Chemical Science* **2015**, *6*, 2885–2892.

(113) Singh, S.; Zhang, J.; Huber, T. D.; Sunkara, M.; Hurley, K.; Goff, R. D.; Wang, G.; Zhang, W.; Liu, C.; Rohr, J.; Van Lanen, S. G.; Morris, A. J.; Thorson, J. S. *Angewandte Chemie - International Edition* **2014**, *53*, 3965–3969.

(114) Zhang, C.; Weller, R. L.; Thorson, J. S.; Rajski, S. R. *Journal of the American Chemical Society* **2006**, *128*, 2760–2761.

(115) Computing, R Foundation for Statistical Vienna, A. R: A language and environment for statistical computing., Vienna, Austria, 2008.

(116) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. *Journal of Computational Chemistry* **2004**, *25*, 1605–1612.

(117) Sambrook, J.; Russell, D. W., *Molecular Cloning: A Laboratory Manual*, 3rd ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor (NY, USA), 2001.

(118) Agilent Technologies QuikChange II Site-Directed Mutagenesis Kit: Instruction Manual., 2011.

(119) Kondou, Y.; Higuchi, M.; Ichikawa, T.; Matsui, M. *Methods* **2011**, *729*, 183–197.

(120) Engler, C.; Kandzia, R.; Marillonnet, S. *PLoS ONE* **Jan. 2008**, *3*, e3647.

(121) Novagen, *pET System Manual*, 11th ed.; 2 Pt 1; EMD Chemicals: Darmstadt, 2014; Vol. 123, p 369.

(122) Gill, S. C.; von Hippel, P. H. *Analytical Biochemistry* **Nov. 1989**, *182*, 319–326.

(123) Gasteiger E.; Hoogland C.; Gattiker A.; Duvaud S.; Wilkins M.R.; Appel R.D.; Bairoch, A. In *The Proteomics Protocols Handbook*, Walker, J. M., Ed.; Humana Press: 2005, pp 571–607.

(124) Makrides, S. C. *Microbiological Reviews* **1996**, *60*, 512–538.

(125) Ke, N.; Berkmen, M. In *Current Protocols in Molecular Biology*; John Wiley & Sons: New York, 2014, 16.1B.1–16.1B.21.

(126) Laemmli, U. K. *Nature* **1970**, *227*, 680–685.

(127) Rudolph, R.; Lilie, H. *The FASEB Journal* **1996**, *10*, 49–56.

(128) Palmer, I.; Wingfield, P. T. *Current Protocols in Protein Science* **Nov. 2012**, *1*, Unit6.3.

(129) Willis, M. S.; Hogan, J. K.; Prabhakar, P.; Liu, X.; Tsai, K.; Wei, Y.; Fox, T. *Protein Science* **2005**, *14*, 1818–1826.

(130) Akbari, N.; Khajeh, K.; Ghaemi, N.; Salemi, Z. *Protein Expression and Purification* **Apr. 2010**, *70*, 254–259.

(131) Benoit, I.; Coutard, B.; Oubelaid, R.; Asther, M.; Bignon, C. *Protein Expression and Purification* **Sept. 2007**, *55*, 166–174.

(132) Box, G. E. P.; Hunter, J. S.; Hunter, W. G., *Statistics for Experimenters: Design, Innovation, and Discovery*, 2nd ed.; Wiley-Interscience: New York, 2005.

(133) Dippe, M.; Brandt, W.; Rost, H.; Porzel, A.; Schmidt, J.; Wessjohann, L. A. *Chemical Communications* **Feb. 2015**, *51*, 3637–3640.

(134) Shapiro, S. K.; Ehninger, D. J. *Analytical Biochemistry* **May 1966**, *15*, 323–333.

(135) Dai, Y.; van Spronsen, J.; Witkamp, G. J.; Verpoorte, R.; Choi, Y. H. *Analytica Chimica Acta* **Mar. 2013**, *766*, 61–68.

(136) Huang, Z. L.; Wu, B. P.; Wen, Q.; Yang, T. X.; Yang, Z. *Journal of Chemical Technology and Biotechnology* **2014**, DOI: *10.1002/jctb.4285*.

(137) Kabsch, W. *Acta Crystallographica Section D: Biological Crystallography* **Feb. 2010**, *66*, 133–144.

(138) Kabsch, W. *Journal of Applied Crystallography* **Dec. 1993**, *26*, 795–800.

(139) Kabsch, W. *Acta Crystallographica Section D: Biological Crystallography* **Feb. 2010**, *66*, 125–132.

(140) Powell, H. R. *Acta Crystallographica Section D: Biological Crystallography* **1999**, *55*, 1690–1695.

(141) Evans, P. *Acta Crystallographica Section D: Biological Crystallography* **Jan. 2006**, *62*, 72–82.

(142) Rupp, B., *Biomolecular Crystallography: Principles, Practice, and Application to Structural Biology*, 1st ed.; Garland Science: New York, 2009.

(143) Rossmann, M. G.; Blow, D. M. *Acta Crystallographica* **Jan. 1962**, *15*, 24–31.

(144) Rossmann, M. G. *Acta Crystallographica - Section D Biological Crystallography* **Sept. 2001**, *57*, 1360–1366.

(145) McCoy, A. J. *Acta Crystallographica Section D: Biological Crystallography* **Jan. 2006**, *63*, 32–41.

(146) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. *Journal of Applied Crystallography* **Aug. 2007**, *40*, 658–674.

(147) Winn, M. D. et al. *Acta Crystallographica Section D: Biological Crystallography* **Apr. 2011**, *67*, 235–242.

(148) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. *Acta Crystallographica Section D Biological Crystallography* **Apr. 2010**, *66*, 486–501.

(149) Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. *Acta Crystallographica Section D: Biological Crystallography* **May 1997**, *53*, 240–255.

(150) Vagin, A. A.; Steiner, R. A.; Lebedev, A. A.; Potterton, L.; McNicholas, S.; Long, F.; Murshudov, G. N. en *Acta Crystallographica Section D: Biological Crystallography* **Nov. 2004**, *60*, 2184–2195.

(151) Adams, P. D. et al. en *Acta Crystallographica Section D: Biological Crystallography* **Feb. 2010**, *66*, 213–221.

(152) Chen, V. B.; Arendall, W. B.; Headd, J. J.; Keedy, D. a.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. *Acta Crystallographica Section D: Biological Crystallography* **Jan. 2010**, *66*, 12–21.

(153) Huey, R.; Morris, G. M.; Olson, A. J.; Goodsell, D. S. *Journal of Computational Chemistry* **Apr. 2007**, *28*, 1145–1152.

(154) Morris, G. M.; Ruth, H.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. *Journal of Computational Chemistry* **Dec. 2009**, *30*, 2785–2791.

(155) Trott, O.; Olson, A. J. *Journal of Computational Chemistry* **Jan. 2010**, *31*, 455–461.

(156) Sigma-Aldrich Technical Bulletin no. 2003-03: freezing of microbial samples prior to testing. Parenteral Drug Association., 2003.

(157) Claiborne, A.; Fridovich, I. *Biochemistry* **1979**, *18*, 2324–2329.

(158) Josephy, P. D.; Eling, T.; Mason, R. P. *Journal of Biological Chemistry* **1982**, *257*, 3669–3675.

(159) Newman, J. *Acta Crystallographica Section D: Biological Crystallography* **2004**, *60*, 610–612.

(160) Schweigert, N.; Zehnder, a. J. B.; Eggen, R. I. L. *Environmental Microbiology* **2001**, *3*, 81–91.

(161) Mentasti, E.; Pelizzetti, E.; Saini, G. en *Journal of the Chemical Society, Dalton Transactions* **Jan. 1973**, 2609.

(162) Dippe, M.; Weigel, B.; Heinke, R.; Vogt, T.; Wessjohann, L. A. *manuscript submitted* **2015**.

(163) Freyer, M. W.; Lewis, E. a. *Methods in Cell Biology* **Jan. 2008**, *84*, 79–113.

(164) Mabry, T. J.; Markham, K. R.; Thomas, M. B., *The Systematic Identification of Flavonoids*; Springer Berlin Heidelberg: Berlin, Heidelberg, 1970.

(165) Shulgin, A.; Shulgin, A., *Tihkal: The Continuation*, 1st ed.; Joy, D., Ed.; Transform Press: Berkeley, 1997, p 804.

(166) Hofmann, A., *Die Mutterkornalkaloide*; Nachtschatten Verlag: Solothurn, 2000.

(167) Santos, C. B. R.; Vieira, J. B.; Lobato, C. C.; Hage-Melim, L. I. S.; Souto, R. N. P.; Lima, C. S.; Costa, E. V. M.; Brasil, D. S. B.; Macêdo, W. J. C.; Carvalho, J. C. T. *Molecules* **2014**, *19*, 367–399.

(168) Alexander, M.; Lustigman, B. K. *Journal of Agricultural and Food Chemistry* **July 1966**, *14*, 410–413.

(169) Mani, S. V.; Connell, D. W.; Braddock, R. D. *Critical Reviews in Environmental Control* **Jan. 1991**, *21*, 217–236.

(170) Ley, J. P.; Krammer, G.; Reinders, G.; Gatfield, I. L.; Bertram, H. J. *Journal of Agricultural and Food Chemistry* **2005**, *53*, 6061–6066.

(171) Elsevier Reaxys, version 2.19790.2.

(172) Murshudov, G. N.; Dodson, E. J. *CCP4 Newsletter on Protein Crystallography* **1997**, *33*, 25–30.

(173) Searle, M. S.; Westwell, M. S.; Williams, D. H. *Journal of the Chemical Society, Perkin Transactions 2* **Jan. 1995**, 141.

(174) Gilli, P.; Ferretti, V.; Gilli, G.; Borea, P. A. *Journal of Physical Chemistry* **Feb. 1994**, *98*, 1515–1518.

(175) Dunitz, J. D. *Chemistry & Biology* **Nov. 1995**, *2*, 709–712.

(176) Li, L.; Dantzer, J. J.; Nowacki, J.; O'Callaghan, B. J.; Meroueh, S. O. *Chemical Biology and Drug Design* **June 2008**, *71*, 529–532.

(177) Dunitz, J. D. *Science* **Apr. 1994**, *264*, 670.

(178) Sawada, Y.; Hirai, M. Y. Integrated LC-MS / MS system for plant metabolomics., en, May 2013.

(179) Lu, L.; Wang, J.; Xu, Y.; Wang, K.; Hu, Y.; Tian, R.; Yang, B.; Lai, Q.; Li, Y.; Zhang, W.; Shao, Z.; Lam, H.; Qian, P.-Y. *Scientific Reports* **Jan. 2014**, *4*, 6537.

(180) Ernst, M.; Silva, D. B.; Silva, R. R.; Vêncio, R. Z. N.; Lopes, N. P. *Natural Product Reports* **June 2014**, *31*, 784–806.

(181) Liebeke, M.; Dörries, K.; Meyer, H.; Lalk, M. *Methods in Molecular Biology* **Jan. 2012**, *815*, 377–398.

(182) Whitehouse, C. M.; Dreyer, R. N.; Yamashita, M.; Fenn, J. B. *Analytical Chemistry* **Mar. 1985**, *57*, 675–679.

(183) Horning, E.; Carroll, D.; Dzidic, I.; Haegele, K.; Horning, M.; Stillwell, R. *Journal of Chromatography A* **Jan. 1974**, *99*, 13–21.

(184) Venzie, J. L.; Castro, J.; Balarama Krishna, M. V.; Nelson, D. M.; Marcus, R. K. *Analytical and Bioanalytical Chemistry* **2007**, *387*, 321–333.

(185) Sleno, L.; Volmer, D. A. *Journal of Mass Spectrometry* **2004**, *39*, 1091–1112.

(186) Williams, C. A.; Grayer, R. J. *Natural Product Reports* **2004**, *21*, 539–573.

(187) Chen, H. J.; Inbaraj, B. S.; Chen, B. H. *International Journal of Molecular Sciences* **2012**, *13*, 260–285.

(188) Cuyckens, F.; Claeys, M. *Journal of Mass Spectrometry* **2004**, *39*, 1–15.

(189) Fabre, N.; Rustan, I.; De Hoffmann, E.; Quetin-Leclercq, J. *Journal of the American Society for Mass Spectrometry* **2001**, *12*, 707–715.

(190) Gates, P. J.; Lopes, N. P. *International Journal of Analytical Chemistry* **2012**, *2012*, 1–7.

(191) Hughes, R. J.; Croley, T. R.; Metcalfe, C. D.; March, R. E. *International Journal of Mass Spectrometry* **2001**, *210-211*, 371–385.

(192) Lee, J. S.; Kim, D. H.; Liu, K. H.; Oh, T. K.; Lee, C. H. *Rapid Communications in Mass Spectrometry* **2005**, *19*, 3539–3548.

(193) Li, C.; Huang, C.; Lu, T.; Wu, L.; Deng, S.; Yang, R.; Li, J. *Rapid Communications in Mass Spectrometry* **2014**, *28*, 2363–2370.

(194) Ma, Y. L.; Li, Q. M.; Van Den Heuvel, H.; Claeys, M. *Rapid Communications in Mass Spectrometry* **1997**, *11*, 1357–1364.

(195) March, R. E.; Miao, X. S. *International Journal of Mass Spectrometry* **2004**, *231*, 157–167.

(196) March, R. E.; Brodbelt, J. *Journal of Mass Spectrometry* **Dec. 2008**, *43*, 1581–1617.

(197) Kuhn, F.; Oehme, M.; Romero, F.; Abou-Mansour, E.; Tabacchi, R. *Rapid Communications in Mass Spectrometry* **2003**, *17*, 1941–1949.

(198) Ma, Y. L.; Heuvel, H. V.; Claeys, M. *Rapid Communications in Mass Spectrometry* **1999**, *13*, 1932–42.

(199) Kang, J.; Hick, L. a.; Price, W. E. *Rapid Communications in Mass Spectrometry* **2007**, *21*, 857–868.

(200) Wolfender, J.-L.; Waridel, P.; Ndjoko, K.; Hobby, K. R.; Major, H. J.; Hostettmann, K. *Analusis* **Dec. 2000**, *28*, 895–906.

(201) Böcker, S.; Rasche, F. *Bioinformatics* **2008**, *24*, i49–i55.

(202) Wolf, S.; Schmidt, S.; Müller-Hannemann, M.; Neumann, S. *BMC Bioinformatics* **2010**, *11*, 148.

(203) Yordi, E. G.; Pérez, E. M.; Matos, M. J.; Villares, E. U. *Nutrition, Well-Being and Health* **2012**, 23–48.

(204) König, K. Engineering of the Anthocyanin Biosynthetic Pathway in Nicotiana benthamiana., Master Thesis, Martin-Luther-Universität Halle-Wittenberg, 2014.

(205) Studier, F. W. *Protein Expression and Purification* **May 2005**, *41*, 207–234.

(206) Kang, S.-Y.; Lee, J. K.; Choi, O.; Kim, C. Y.; Jang, J.-H.; Hwang, B. Y.; Hong, Y.-S. *BMC Biotechnology* **2014**, *14*, 67.

(207) Yamaguchi, H.; Miyazaki, M. *Biomolecules* **2014**, *4*, 235–51.

(208) Montgomery, D., *Design and Analysis of Experiments*, 5th ed.; John Wiley & Sons: New York, 2008, p 680.

(209) Vincentelli, R.; Canaan, S.; Campanacci, V.; Valencia, C.; Maurin, D.; Frassinetti, F.; Scappucini-Calvo, L.; Bourne, Y.; Cambillau, C.; Bignon, C. *Protein Science* **2004**, *13*, 2782–2792.

(210) Tobbell, D. A.; Middleton, B. J.; Raines, S.; Needham, M. R. C.; Taylor, I. W. F.; Beveridge, J. Y.; Abbott, W. M. *Protein Expression and Purification* **Mar. 2002**, *24*, 242–254.

(211) Tsumoto, K.; Umetsu, M.; Kumagai, I.; Ejima, D.; Philo, J. S.; Arakawa, T. *Biotechnology Progress* **2004**, *20*, 1301–1308.

(212) Anselment, B.; Baerend, D.; Mey, E.; Buchner, J.; Weuster-Botz, D.; Haslbeck, M. *Protein Science* **2010**, *19*, 2085–2095.

(213) Fujimoto, A.; Hirano, A.; Shiraki, K. *Protein Journal* **2010**, *29*, 161–166.

(214) Chen, J.; Liu, Y.; Wang, Y.; Ding, H.; Su, Z. *Biotechnology Progress* **2008**, *24*, 1365–1372.

(215) Bradford, M. M. *Analytical Biochemistry* **May 1976**, *72*, 248–254.

(216) Guelorget, A.; Roovers, M.; Guérineau, V.; Barbey, C.; Li, X.; Golinelli-Pimpaneau, B. *Nucleic Acids Research* **2010**, *38*, 6206–6218.

(217) Jorda, J.; Yeates, T. O. *Archaea* **2011**, *2011*, DOI: *10.1155/2011/409156.*

(218) Gee, C. L.; Nourse, A.; Hsin, a. Y.; Wu, Q.; Tyndall, J. D.; Grunewald, G. L.; McLeish, M. J.; Martin, J. L. *Biochimica et Biophysica Acta - Proteins and Proteomics* **June 2005**, *1750*, 82–92.

(219) Louis-Jeune, C.; Andrade-Navarro, M. a.; Perez-Iratxeta, C. *Proteins: Structure, Function and Bioinformatics* **Feb. 2012**, *80*, 374–381.

(220) Bornscheuer, U. T.; Huisman, G. W.; Kazlauskas, R. J.; Lutz, S.; Moore, J. C.; Robins, K. *Nature* **May 2012**, *485*, 185–194.

(221) Koeller, K. M.; Wong, C. H. *Nature* **Jan. 2001**, *409*, 232–240.

(222) Youngdae, Y.; Park, Y.; Yi, Y. S.; Lee, Y.; Jo, G.; Park, J. C.; Ahn, J. H.; Lim, Y. *Journal of Microbiology and Biotechnology* **2010**, *20*, 1359–1366.

(223) Jarrell, T. M.; Marcum, C. L.; Sheng, H.; Owen, B. C.; O'Lenick, C. J.; Maraun, H.; Bozell, J. J.; Kenttämaa, H. I. *Green Chemistry* **2014**, *16*, 2713.

(224) Zhou, J. M.; Gold, N. D.; Martin, V. J. J.; Wollenweber, E.; Ibrahim, R. K. *Biochimica et Biophysica Acta* **2006**, *1760*, 1115–1124.

(225) Vogt, T. *FEBS Letters* **Mar. 2004**, *561*, 159–162.

(226) Martinière, A.; Desbrosses, G.; Sentenac, H.; Paris, N. *Frontiers in Plant Science* **2013**, *4*, 523.

(227) Kelley, L. a.; Mezulis, S.; Yates, C. M.; Wass, M. N.; Sternberg, M. J. E. *Nature Protocols* **June 2015**, *10*, 845–858.

(228) Cornish-Bowden, A., *Fundamentals of Enzyme Kinetics*, 3rd ed.; Portland Press Ltd: London, 2004.

(229) Tibshirani, R. Regression Selection and Shrinkage via the Lasso., 1994.

(230) Song, Q.; Song, Y.; Zhang, N.; Li, J.; Jiang, Y.; Zhang, K.; Zhang, Q.; Tu, P. *RSC Advances* **2015**, *5*, 57372–57382.

(231) Thomsen, M.; Vogensen, S. B.; Buchardt, J.; Burkart, M. D.; Clausen, R. P. *Organic & Biomolecular Chemistry* **2013**, *11*, 7606–10.

(232) Winter, J. M.; Chiou, G.; Bothwell, I. R.; Xu, W.; Garg, N. K.; Luo, M.; Tang, Y. *Organic Letters* **2013**, *15*, 3774–3777.

(233) Wang, R.; Ibáñez, G.; Islam, K.; Zheng, W.; Blum, G.; Sengelaub, C.; Luo, M. *Molecular BioSystems* **2011**, *7*, 2970.

(234) Tengg, M.; Stecher, H.; Remler, P.; Eiteljörg, I.; Schwab, H.; Gruber-Khadjawi, M. *Journal of Molecular Catalysis B: Enzymatic* **Dec. 2012**, *84*, 2–8.

(235) Islam, K.; Zheng, W.; Yu, H.; Deng, H.; Luo, M. *ACS Chemical Biology* **2011**, *6*, 679–684.

(236) Lukinavičius, G.; Tomkuvienè, M.; Masevičius, V.; Klimašauskas, S. *ACS Chemical Biology* **2013**, *8*, 1134–1139.

Acronyms

Å Ångström, 0.1 nm

3O4M $3^{(\prime)}$-hydroxy-$4^{(\prime)}$-methoxy
4CL 4-coumarate:CoA ligase
4O3M $4^{(\prime)}$-hydroxy-$3^{(\prime)}$-methoxy

ABPP activity based protein profiling
AC-9 anthracene-9-carboxylic acid
AI auto-induction 98, *see* ZYP-5052
ANOVA *An*alysis *o*f *Va*riance
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 collision 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-enthropy compensation
EI electron ionization
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, γ-L-glutamyl-L-cysteinylglycine
GSSG glutathione disulfide
GST glutathione *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

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 49, 50
PFOMT phenylpropanoid and flavonoid *O*-methyl transferase
PHENIX Phyton-based Hierarchial Environment for Integrated Xtallography
PKS polyketide synthase
PMSF phenylmethylsulfonylfluoride
P-MT protein methyl transferase

QSAR quantitative structure activity relationship

rmsd root mean squared deviation
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-[[(2*S*,3*S*,4*R*,5*R*)-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
***Se*AM** *Se*-adenosyl selenomethionine
SET *s*uvar3-9, *e*nhancer-of-zeste, *t*rithorax
SID surface-induced dissociation
***sm*MT** small molecule methyl transferase
***S*-MT** *S*-methyl transferase
SOMT-2 soy *O*-methyl transferase
SPOUT *SpoU*-*T*rmD
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, *y*east extract, *p*hosphate 42, 171, 173, *see* ZYP-5052

Glossary

His_6-tag Hexa-histidine tag commonly used for recombinant protein production.

PFOMT Phenylpropanoid and flavonoid *O*-methyl transferase from *Mesembryanthemum crystallinum*, which was first described by Ibdah et al. in 2003 [78]

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 [205]. The naming stems from the components N-*Z*-amine, *y*east extract and *p*hosphate. The numbering designates the composition; e.g. 5052 refers to 0.5 % glycerol, 0.05 % glucose and 0.2 % lactose. 171