

# Dissertation

## Biotransformations from and to methylated flavo- noids

### *Subtitle*

Benjamin Weigel  
Leibniz-Institute of Plant Biochemistry  
Department of Bioorganic Chemistry  
Weinberg 3  
06120 Halle(Saale)  
August 20, 2015

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

noch nicht bekannt

---

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

– *unknown* –

# Contents

<b>I</b>	<b>Preface</b>	1
<hr/>		
<b>1</b>	<b>Abstracts</b>	2
1.1	English Abstract . . . . .	3
1.2	Deutsche Zusammenfassung . . . . .	3
<hr/>		
<b>II</b>	<b>Thesis</b>	4
<hr/>		
<b>2</b>	<b>Introduction</b>	5
2.1	Natural products and secondary metabolites . . . . .	6
2.1.1	General . . . . .	6
2.1.2	Classes of natural products . . . . .	6
2.2	Alkylating reactions in nature . . . . .	6
2.2.1	Methylation . . . . .	6
2.2.2	Prenylation . . . . .	6
2.2.3	Glycosylation . . . . .	6
2.3	Usage and expansion of natures reaction toolbox . . . . .	7
2.3.1	Terpene synthases and elongases . . . . .	7
2.3.2	Methyl transferases . . . . .	7
2.3.3	Glycosyl transferases . . . . .	7
2.3.4	Other important enzymes in biotech research . . . . .	7
2.4	Conclusion . . . . .	7
<b>3</b>	<b>Material And Methods</b>	8
3.1	Materials . . . . .	9
3.1.1	Chemicals . . . . .	9
3.1.2	Commonly used solutions and buffers . . . . .	9
3.1.3	Culture media used to grow bacteria . . . . .	10
3.1.4	Bacterial strains . . . . .	11
3.1.5	Plasmids . . . . .	12
3.1.6	Oligonucleotides and synthetic genes . . . . .	12
3.1.7	Instruments . . . . .	13
3.1.8	Software . . . . .	14
3.2	Molecular Biology . . . . .	14
3.2.1	Golden Gate Cloning . . . . .	15
3.2.2	Subcloning of genes . . . . .	15

3.2.3 Transformation of electrocompetent <i>Agrobacterium tumefaciens</i> cells . . . . .	16
3.3 Treatment of plant material . . . . .	16
3.3.1 Infiltration of <i>Nicotiana benthamiana</i> . . . . .	16
3.3.2 Plant material harvest . . . . .	16
3.3.3 Extraction of flavonoids from <i>N. benthamiana</i> leaves . . . . .	17
3.4 Protein biochemistry . . . . .	17
3.4.1 Determination of protein concentration . . . . .	17
3.4.2 Protein production test (expression test) . . . . .	18
3.4.3 Protein subfractionation . . . . .	19
3.4.4 Protein sample concentration by TCA precipitation . . . . .	19
3.4.5 Preparation of periplasmic protein . . . . .	20
3.4.6 Discontinuous SDS-polyacrylamide gel electrophoresis (SDS-PAGE) . . . . .	20
3.4.7 Buffer change of protein samples . . . . .	21
3.4.8 Production of recombinant protein . . . . .	21
3.4.9 Preparation of inclusion bodies (IBs) . . . . .	22
3.4.10 Purification of His-tagged proteins using immobilized metal affinity chromatography (IMAC) . . . . .	23
3.4.11 Refolding of SOMT-2 on a micro scale using design of experiments (DoE) . . . . .	24
3.4.12 Enzymatic production of SAM and SAE . . . . .	26
3.5 Crystallographic Procedures . . . . .	27
3.5.1 Crystallization of proteins . . . . .	27
3.5.2 Data collection and processing . . . . .	28
3.5.3 Structure solution . . . . .	29
3.5.4 Model building, refinement and validation . . . . .	29
3.5.5 <i>In silico</i> substrate docking . . . . .	30
3.6 Analytics . . . . .	30
3.6.1 Recording of growth curves . . . . .	30
3.6.2 <i>In vitro</i> determination of glucose . . . . .	31
3.6.3 <i>In vitro</i> O-methyl transferase (O-MT) assay . . . . .	31
3.6.4 Photospectrometric assay for the methylation of catecholic moieties . . . . .	34
3.6.5 Concentration of SOMT-2 using hydrophobic interaction chromatography (HIC) . . . . .	35
3.6.6 Analytical gel filtration . . . . .	36
3.6.7 Binding experiments using Isothermal Titration Calorimetry (ITC) . . . . .	36
3.6.8 High-performance liquid chromatography (HPLC) analytics .	37

3.6.9 liquid chromatography coupled mass-spectrometry (LC/MS) measurements . . . . .	37
<b>4 Engineering of phenylpropanoid and flavonoid O-methyl transferase (PFOMT)</b>	38
4.1 Introduction . . . . .	39
4.2 Crystallization of PFOMT . . . . .	42
4.3 Substrate binding studies using ITC . . . . .	44
4.4 Study of variants for long-chain alkylations . . . . .	47
4.5 Conclusion/Discussion . . . . .	50
4.6 Contributions . . . . .	52
<b>5 Enzymatic methylation of Non-catechols</b>	53
5.1 Introduction . . . . .	55
5.2 SOMT-2 . . . . .	55
5.2.1 <i>In vivo</i> biotransformation in <i>N. benthamiana</i> . . . . .	55
5.2.2 <i>In vivo</i> biotransformation in <i>E. coli</i> . . . . .	59
5.2.3 <i>In vitro</i> studies using recombinantly produced SOMT-2 . . . . .	63
5.3 PFOMT . . . . .	70
5.3.1 Phenolic hydroxyls . . . . .	70
5.3.2 PFOMT pH-profiles are influenced by Mg <sup>2+</sup> . . . . .	72
5.3.3 Methylation of different chemical motifs . . . . .	75
5.4 Conclusion/Discussion . . . . .	80
<b>6 Structural studies of flavonoids using tandem mass-spectrometry (MS/MS)</b>	82
6.1 Introduction . . . . .	83
6.2 Fragmentation of flavanones . . . . .	83
6.3 Fragmentation of flavones . . . . .	83
6.4 Fragmentation of flavonols . . . . .	83
6.5 Fragmentation of cinnamic acids . . . . .	83
6.6 Conclusion/Discussion . . . . .	83
6.7 Contributions . . . . .	87
<b>7 DES in protein crystallography</b>	88
7.1 Introduction . . . . .	89
7.2 Solubility enhancement of hydrophobic substances by addition of DES . . . . .	89
7.3 Enzymatic O-methylation in DES . . . . .	89
7.4 DES as precipitants in protein crystallization . . . . .	89
7.5 Conclusion/Discussion . . . . .	89

<b>8 Acknowledgements</b>	90
<b>III Appendix</b>	92
<hr/>	
<b>A Figures</b>	93
A.1 Chapter 3 . . . . .	94
A.2 Chapter 4 . . . . .	97
<b>B Tables</b>	101
<b>C Affidavit</b>	108
<b>Acronyms</b>	121
<b>Glossary</b>	126

# List of Figures

3.1 Pruned <i>N. benthamiana</i> plant, with two bottom and one top leaf, ready to be infiltrated. . . . .	17
3.2 Oxidation of the reporter substrate <i>o</i> -dianisidine. Consecutive one-electron transfers lead to the fully oxidized diimine form of <i>o</i> -dianisidine. The first electron transfer is believed to produce a charge transfer complex intermediate. [16, 44] . . . . .	32
3.3 Calibration curves of different relative compositions of ferulic acid to caffeic acid, that were taken as described in 3.6.4. The total concentration was always 0.4 mM. At lower pH values around 4, the method seems to overestimate the concentration of caffeic acid. However, the slope of the curves stays the same. . . . .	35
4.1 Labelling of macromolecules by using a combination of novel alkene-derivatized SAM analogues and Cu <sup>I</sup> -catalyzed azide-alkyne 1,3-dipolar cycloaddition (CuAAC). Depending on the type of label used, it can be employed for detection (e.g. through fluorophores, coupled assays) or affinity purification (e.g. biotin). This technique is also feasable for use in activity based protein profiling (ABPP) approaches. . . . .	41
4.2 Some crystal and pseudo-crystal shapes that were observed during the crytsallization screen. a – high (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> , b-c – CaCl <sub>2</sub> , PEG-4000, e – LiCl, PEG-6000 . . . . .	43
4.3 An overview of the features in the <i>apo</i> -PFOMT structure. <b>a</b> – The assymetric unit of <i>apo</i> -PFOMT consists of two homodimers (4 monomers). Individual monomers are rainbow colored from N- (blue) to C-terminus (red). <b>b</b> – Comparison of 3C3Y (steelblue) and <i>apo</i> -PFOMT (green). The N-terminus of <i>apo</i> -PFOMT was resolved up to the N-terminus (red) and even the His-tag (red, transparent) was partly resolved. The N-terminus fits into a cleft on the surface of the 3C3Y structure, shown as a surface model on the right. SAH (white ball-and-sticks) and Ca <sup>2+</sup> (green sphere) are featured in the published structure, whereas a sulphate ion (red/yellow spheres) was bound in the newly solved structure. . . . .	45

4.4 Positional differences between the individual residues of the solved <i>apo</i> -PFOMT and the structure with bound SAH (pdb: 3C3Y). The diffraction precision indicator [20] (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 [55]. . . . .	46
4.5 The binding of different SAM analogues was measured via ITC. . . . .	46
4.6 ITC measurements of PFOMT:effector binding. <b>a</b> – Binding of SAH, SAM and SAE to PFOMT. <b>b</b> – SAH is injected into a PFOMT solution, with (red) or without (black) addition of Mg <sup>2+</sup> and caffeic acid. When Mg <sup>2+</sup> and caffeic acid were already present, the binding process seems to happen quicker, but is less enthalpic. <b>c</b> – Upon addition of caffeic acid to the protein heat is produced, however no sensible binding curve could be obtained. . . . .	48
4.7 The active site of PFOMT (pdb: 3C3Y). The outline of the protein backbone is displayed, with active site residues portrayed as colored sticks (cyan – F103, red – F80, turquoise – M52, yellow – Y51, white – F198, blue – W184, orange – N202, grey – as labelled). The co-substrate SAM (ball-and-stick model) was docked into the structure. . . . .	49
4.8 Activities of different PFOMT variants towards caffeic acid methylation. Colorations correspond to the ones used in Figure 4.7. . . . .	50
4.9 Comparison of the active sites of <b>a</b> – the solved <i>apo</i> -structure (green) and <b>b</b> – the ligand-bound structure (steelblue; pdb: 3C3Y). Waters are represented as small red spheres, calcium as a green sphere (complexing bonds are dashed) and SAH is displayed as a white ball-and-stick model. A possible hydrogen bond network (blue lines) for the ligand-bound state is displayed. . . . .	52

5.1 Semi-synthetic pathway to naringenin and 4'-O-methyl naringenin in <i>N. benthamiana</i> . Enzymes not endogenous to <i>N. benthamiana</i> are in gray. PAL - phenylalanine ammonia lyase, C4H - cinnamic acid 4-hydroxylase, 4CL - 4-coumaric acid:CoA ligase, CHS - chalcone synthase, CHI - chalcone isomerase, SOMT2 - soy O-methyl transferase	56
2 . . . . .	56
5.2 Cloning of SOMT-2 . . . . .	57
5.3 PCA of leaf material. The samples are colored by leaf side. . . . .	58
5.4 Biotransformation methods as described by Kim <i>et al.</i> (left) and developed in this work (right). . . . .	60
5.5 Growth curve of <i>E. coli</i> BL21(DE3) expressing SOMT-2 at 37 °C. Glucose is depleted about 5 hours into growth, at which point the start of SOMT-2 expression is expected. The OD <sub>600</sub> after inoculation was about 0.1. . . . .	61
5.6 The masses resulting from the fragmentation into A- and B-ring along the C-ring (dashed line, <b>b</b> ) are evidence, that the 4'-hydroxyl on the B-ring is methylated by SOMT-2 . . . . .	62
5.7 SDS-PAGE of pET28a(+) SOMT-2 expressed in <i>E. coli</i> BL21(DE3) in autoinduction medium at different temperatures (shown above). The insoluble fractions show a protein band the same height as the 40 kDa marker band, which corresponds to the SOMT-2 protein (40 425 Da). M – protein size marker, S – soluble fraction, I – insoluble fraction . . . . .	65
5.8 SDS-PAGE of the insoluble and soluble fractions of the refolding reactions. . . . .	67
5.9 Results of <i>in vitro</i> protein refolding trials. . . . .	69
5.10 Initial rate/pH-profiles for the conversion of three different substrates (red – caffeic acid, green – eriodictyol, blue – <i>iso</i> -ferulic acid) by PFOMT. The non-catecholic substrate <i>iso</i> -ferulic acid is converted much less quickly than the catecholic substrates. . . . .	72
5.11 pH-profiles of substrate conversion. The linear regression models (blue, dashed lines) grasp the general trend of the data reasonably well to draw inferences. 95 % prediction intervals are displayed as shaded areas.	74
5.12 Substrates from four different groups were screened for methylation by PFOMT under different conditions (pH/Mg <sup>2+</sup> ). Figure refers to Table 5.10.	77

5.13 Conversion of multiple different substrates, catecholic and non catecholic, by PFOMT. pH/Mg <sup>2+</sup> color coded from light to dark: ✕/✗, ✕/✓, ✕/✗, ✕/✓ . . . . .	78
6.1 Major fragmentation pathways of hesperetin upon HCD at 75 eV . . . . .	84
6.2 Proposed mechanisms for the formation of <b>a</b> – <sup>1,3</sup> B <sup>+</sup> and <b>b</b> – <sup>1,4</sup> B <sup>+</sup> ions from hesperetin upon HCD. . . . .	85
6.3 Major fragmentation pathways of chrysoeriol upon HCD at 100 eV . . . . .	86
6.4 Proposed mechanism for the formation of <sup>0,2</sup> B <sup>+</sup> from the molecular ion of Chrysoeriol under HCD conditions. . . . .	87
A.1 Graphical representation of the work that has been done on MTs in combination with SAM analogues. The grey areas represent individual groups of SAM analogues (aliphatic, allylic, propargylic, aromatic, SeAM analogues, nitrogen analogues and miscelaneous others). The height of the grey areas represents the number of times a member of the corresponding group has been described as tested in the MT literature. The height of the colored bars represents the times that individual substrate has been tested. The colors represent the different types of MT (red – DNA MT, green – P-MT, lilac – small molecule MT, blue – RNA MT). The black dash across the bar shows the number of times this substrate was actually converted by either enzyme. . . . .	94
A.2 Differences in the dihedrals $\psi$ (red) and $\varphi$ (black) of the solved <i>apo</i> -PFOMT and the structure with bound SAH (pdb: 3C3Y). The secondary structure is displayed at the top. Helices are displayed as rectangles and sheets are shown as arrows. Graphical background annotations are used to display the binding sites of SAH (green) and the metal ion (plum). The orange bars indicate regions, where much movement seems to happen upon binding or release of the co-substrate. The blue bar shows the region that was annotated as "insertion loop" in previous studies. . . . .	95
A.3 Differences in the regioselectivity of some PFOMT variants. The products observed in HPLC and LC/MS measurements switched from 3'-methylated (dark grey) to 4'-methylated (light grey) for the displayed variants. . . . .	96
A.4 PCA of <i>N. benthamiana</i> leaves infiltrated by <i>A. tumefaciens</i> carrying different genes. . . . .	97

A.5 SDS-PAGE gel of samples aquired during growth curve measurements. The arrow indicated the band that could correspond to the GST-tagged SOMT-2 protein. . . . .	98
A.6 Graphical representation of a SOMT-2 model obtained from the PHYRE2 web portal ( <a href="http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index">http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index</a> ) [49] . . . . .	98
A.7 Chromatogram of the gel filtration analysis of refolded SOMT-2. . . . .	99
A.8 CD-spectrum of refolded SOMT-2 (black) compared to the spectrum that was calculated by the K2D3 web service ( <a href="http://cbdm-01.zdv.uni-mainz.de/~andrade/k2d3//index.html">http://cbdm-01.zdv.uni-mainz.de/~andrade/k2d3//index.html</a> ). Secondary structure estimates from the spectrum are 12.39 % $\alpha$ -helix and 32.51 % $\beta$ -sheet. . . . .	99
A.9 . . . . .	100

## List of Tables

3.1 NADES-mixtures used within this work. . . . .	10
3.3 Plasmids used in this work. . . . .	12
3.4 Primers used in this work. Recognition sites for endonucleases are underlined. Positions used for site directed mutagenesis are in lower case font. . . . .	13
3.6 Calculated extinction coefficients of proteins used in this work.. . .	18
3.7 Factors used in the construction of the FrFD. . . . .	24
3.8 Experimental design matrix for the FrFD. . . . .	24
4.1 Results of fitting a simple one-site binding model to the data obtained from ITC experiments. . . . .	47
5.2 Naringenin and 4'-methylated derivatives. . . . .	58
5.3 <i>In vivo</i> biotransformation of different flavonoids, phenylpropanoids and anthraquinones by SOMT-2 in <i>E. coli</i> . . . . .	64
5.5 Results for the ANOVA of the main effects model describing soluble protein. Significance codes: 0.05 ( * ), 0.1 ( .). . . . .	68
5.6 Results for the ANOVA of the main effects model describing protein activity. Significance codes: 0.05 ( * ), 0.1 ( .) . . . . .	69
5.7 pK <sub>a</sub> -values of phenolic hydroxyl groups exemplified by <i>p</i> -cresole derivatives. Substituent positions on the aromatic ring are arbitrary and do not reflect conventions of the International Union of Pure and Applied Chemistry (IUPAC). . . . .	71
5.8 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. . . . .	73
5.9 Coefficients of the model (5.2) for activity of <i>iso</i> -ferulic acid methylation. The factor Mg is a categorical variable (addition/no addition) and can therefore only be 0 or 1. . . . .	74

5.10 Substrate grid that was tested for methylation with PFOMT. Four different groups of compounds were screened. The groups of flavones, flavanones and cinnamic acids each contained one representative of each motif (phenolic, catecholic, 3'-hydroxy-4'-methoxy (3O4M), 4'-hydroxy-3'-methoxy (4O3M). The substitution patterns refer to Figure 5.12. . . . .	76
5.12 Conversion of substrates after 16 hours incubation. Only the maximum conversion is displayed, along with the conditions it was achieved under. <sup>†</sup> – wild-type: substrate conversion was maximal for all pH/Mg <sup>2+</sup> combinations. <sup>‡</sup> – conversion of caffeic acid by the wild-type was set to 100 %. . . . .	79
B.3 SAM analogues that have been used with MTs. Targets: <i>P</i> – peptide/protein, <i>D</i> – DNA, <i>R</i> – RNA, <i>S</i> – small molecule. . . . .	102
B.1 Overview over the constructs produced for the present thesis. Each step during the production of the construct is given in the workflow steps column. Primers (italic font) or restriction sites used during each step are displayed in parenthesis. . . . .	105
B.4 Crystallographic data, phasing and refinement statistics. . . . .	106
B.5 Coefficients of the model (5.3) for activity of catecholic methylation. The factor Mg is a categorical variable (addition/no addition) and can therefore only be 0 or 1. . . . .	107
B.7 Coefficients obtained for linear regression model suing the iso-ferulic acid subset after shrinkage using the Lasso method and 5-fold cross validation. Only non-zero coefficients (variables actually do have an effect) are retained during the Lasso. Seed was set to 1337. . . . .	107
B.9 Coefficients obtained for linear regression model using the catechols subset after shrinkage using the Lasso method and 5-fold cross validation. Only non-zero coefficients (variables actually do have an effect) are retained during the Lasso. Seed was set to 1336. . . . .	107

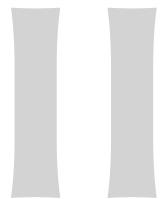


# Preface

# 1 Abstracts

## **1.1 English Abstract**

## **1.2 Deutsche Zusammenfassung**



# Thesis

## 2 Introduction

Some introductory text

## 2.1 Natural products and secondary metabolites

### 2.1.1 General

### 2.1.2 Classes of natural products

#### Terpenoids and Steroids

... here is some text

#### Polyketides and non-ribosomal peptides

... here is some text

#### Alkaloids

... here is some text

#### Phenylpropanoids

... here is some text Flavonoids and phenyl propanoids have important functions in nature and can function as protection against high UV-exposure, signaling molecules or transcriptional regulators [5, 36].

## 2.2 Alkylating reactions in nature

### 2.2.1 Methylation

### 2.2.2 Prenylation

### 2.2.3 Glycosylation

## **2.3 Usage and expansion of natures reaction toolbox**

### **2.3.1 Terpene synthases and elongases**

### **2.3.2 Methyl transferases**

### **2.3.3 Glycosyl transferases**

### **2.3.4 Other important enzymes in biotech research**

**BMVOs**

**Esterases/Lipases**

**Oxidases**

**Lyases**

**Transaminases**

## **2.4 Conclusion**

c

### **3 Material And Methods**

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

## 3.1 Materials

### 3.1.1 Chemicals

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

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

### 3.1.2 Commonly used solutions and buffers

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

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

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

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

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

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

### 3.1.3 Culture media used to grow bacteria

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

### 3.1.4 Bacterial strains

#### *E.coli*

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

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

### ***Agrobacterium tumefaciens***

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

### **3.1.5 Plasmids**

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

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

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

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

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

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

### 3.1.7 Instruments

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

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

### 3.1.8 Software

All mathematical and statistical computations and graphics were done with the R software (versions 3.1.X, <http://cran.r-project.org/>) [80]. Visualizations of macromolecules were arranged using the PyMol Molecular Graphics System, version 1.7.0.0 (Schrödinger, New York, USA).

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

## 3.2 Molecular Biology

Basic molecular biology methods like polymerase chain reaction (PCR), DNA restriction/ligation, DNA gel electrophoresis, preparation of competent cells and transformation were performed based on the protocols summarized by Sambrook and Russell [85].

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 [2] offered by Agilent Technologies (Santa Clara, USA).

Nucleotide fragments obtained by PCR, restriction/ligation procedures or excision from electrophoresis gels were purified and concentrated using the *Nucleospin Gel and PCR Clean-up* kit provided by Machery-Nagel (Düren, Germany) according to the instructions provided by the manufacturer.

### 3.2.1 Golden Gate Cloning

The Golden Gate cloning procedure is a one-pot method, meaning the restriction digestion and ligation are carried out in the same reaction vessel at the same time [27, 53]. Consequently PCR-fragments, destination vector, restriction endonuclease and ligase are added together in this reaction. The methodology employs type II restriction enzymes, which together with proper design of the fragments allow for a ligation product lacking the original restriction sites.

For digestion/ligation reactions of fragments containing BpiI sites, 20 fmol of each fragment or vector, together with 5 U of BpiI and 5 U of T4 ligase were combined in a total volume of 15 µl 1× ligase buffer. For fragments to be cloned via BsaI sites, BpiI in the above reaction was substituted by 5 U BsaI.

The reaction mixture was placed in a thermocycler and incubated at 37 °C for 2 min and 16 °C for 5 min. These two first steps were repeated 50 times over. Finally, the temperature was raised to 50 °C (5 min) and 80 °C (10 min) to inactivate the enzymes.

### 3.2.2 Subcloning of genes

All subcloning procedures were performed according to section 3.2 and specifically subsection 3.2.1. Specific steps for the subcloning of any genes discussed can be found in the appendix (p.105). The *pfromt* gene was subcloned from the pQE-30 vector kindly provided by Thomas Vogt (Leibniz-Institute of Plant Biochemistry (IPB), Halle, Germany) into the pET-28a(+) vector. The *somt-2* gene was subcloned from the pQE-30 vector kindly provided by Martin Dippe (IPB, Halle, Germany) into the pET-28-MC vector.

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

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

## 3.3 Treatment of plant material

### 3.3.1 Infiltration of *Nicotiana benthamiana*

Before infiltration *N. benthamiana* plants were pruned, such that only leaves to be infiltrated remained with the plant (Figure 3.1). 5 ml cultures of transformed *A. tumefaciens* in LB-medium (with 40 µg/ml rifampicin and 50 µg/ml carbencillin) were grown over night at 28 °C and 220 rpm. OD<sub>600</sub> of the culture was measured and adjusted to 0.2 by dilution with infiltration buffer (10 mM MES/NaOH, 10 mM MgSO<sub>4</sub> pH 5.5). When multiple *A. tumefaciens* transformed with different constructs/plasmids were used for infiltration, the cultures were mixed and diluted using infiltration buffer, such that OD<sub>600</sub> of each culture in the mix was 0.2. The solution was infiltrated into the abaxial side of *N. benthamiana* leaves using a plastic syringe. The leaf material was harvested after 7 days.

### 3.3.2 Plant material harvest

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



**Figure 3.1:** Pruned *N. benthamiana* plant, with two bottom and one top leaf, ready to be infiltrated.

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

Two tips of a small spatula of freeze-dried material ( $\approx 6$  mg), were weighed exactly and extracted with 500  $\mu$ l 75 % aqueous methanol containing 1 mM ascorbic acid, 0.2 % formic acid and 0.1 mM flavone (internal standard). Therefore the suspension was vortexed for 30 s, rotated on an orbital shaker for 10 min and vortexed again for 30 s. The suspension was centrifuged ( $20\,000 \times g$ , 4 °C, 10 min) and the supernatant transferred to a new tube, to remove the insoluble plant material. The supernatant was centrifuged again ( $20\,000 \times g$ , 4 °C, 10 min) and the resulting supernatant was transferred to a HPLC-vial and stored at –20 °C until analysis.

## 3.4 Protein biochemistry

Stock solutions of antibiotics, IPTG or sugars were prepared according to the pET System Manual by Novagen [76], unless otherwise noted.

### 3.4.1 Determination of protein concentration

Protein concentrations were estimated using the absorption of protein solutions at 280 nm, which is mainly dependent on the amino acid composition of the protein

studied [34]. Extinction coefficients of proteins were calculated from the amino acid sequence using the ExpPASy servers's ProtParam tool [32].

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

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

### 3.4.2 Protein production test (expression test)

The heterologous production of proteins in *E. coli* was assessed in a small scale protein production test, henceforth called expression test. Single colonies of *E. coli* transformed with the constructs to be studied were used to inoculate a 2 ml starter culture in LB-medium containing the appropriate antibiotics. The working concentrations of antibiotics used was as follows: 200 µg/ml ampicillin, 150 µg/ml kanamycin, 50 µg/ml chloramphenicol, 20 µg/ml tetracycline.

The starter culture was allowed to grow at 37 °C and 200 rpm over night. A 5 ml sampling culture of the medium to be studied containing the appropriate antibiotics was prepared. The media tested included LB, terrific broth (TB) and auto-induction media like ZYP-5052. The sampling culture was inoculated to an OD<sub>600</sub> of 0.075 using the starter culture and incubated at different temperatures and 200 rpm in a shaking incubator. 1 mM isopropyl-D-thiogalactopyranosid (IPTG) was added when the OD<sub>600</sub> reached 0.6-0.8, if appropriate for the studied construct. 1 ml samples were removed after different times of incubation (e.g. 4, 8, 12 hours), subfractionated (3.4.3) and analyzed via SDS-polyacrylamide gel electrophoresis (PAGE) (3.4.6). Exact specifications of growth conditions (e.g. temperature, time, constructs) are discussed in the individual sections.

### 3.4.3 Protein subfractionation

The protein subfractionation procedure described herein was adapted from the protocol described in the pET Manual [76]. 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<sub>600</sub> of the culture sample was measured and the cells harvested by centrifugation at 10 000 × *g*, 4 °C for 5 minutes. The protein in the supernatant medium was concentrated by precipitation with trichloro acetic acid (TCA) (3.4.4) for SDS-PAGE analysis. The periplasmic protein was prepared (3.4.5) and also concentrated by TCA precipitation for SDS-PAGE. Cells were lysed by resuspending the cell pellet in (OD<sub>600</sub> × V × 50) µl of bacterial protein extraction reagent (B-PER) and vortexing vigorously for 30 s. The suspension was incubated at room temperature (RT) for 30 min to assure complete lysis. To separate insoluble protein and cell debris from the soluble cytosolic protein, the suspension was centrifuged at 10 000 × *g* and 4 °C for 10 min. Soluble cytoplasmic protein was contained in the supernatant, whereas insoluble protein remained in the pellet. For SDS-PAGE analysis of the insoluble protein, the pellet was resuspended in the same volume of B-PER. To obtain only the total cell protein fraction, the preparation of periplasmic and soluble cytosolic protein was omitted. Sample volumes of 10 µl of each fraction were used for SDS-PAGE analysis.

### 3.4.4 Protein sample concentration by TCA precipitation

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

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

### 3.4.5 Preparation of periplasmic protein

Target proteins may be directed to the periplasmic space by N-terminal signal sequences like *pelB* or *DsbA/C* [65]. The periplasma is, other than the cytosol, an oxidizing environment and often used for the production of proteins containing disulfide linkages. The preparation of periplasmic protein was accomplished by an osmotic shock protocol modified from Current Protocols in Molecular Biology [6]. The cell pellet was resuspended in the same volume as the culture sample of 30 mM tris-HCl, 20 % (w/v) sucrose, pH 8 and 1 mM ethylenediaminetetraacetic acid (EDTA) was added. The suspension was stirred for 10 min at RT and the cells were collected by centrifugation at 10 000 × *g*, 4 °C for 10 min. The supernatant was discarded and the cell pellet was resuspended in the same volume of ice-cold 5 mM MgSO<sub>4</sub>. The suspension was stirred for 10 min on ice, while the periplasmic proteins were released into the solution. The cells were collected by centrifugation as before. Periplasmic proteins were contained in the supernatant.

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

The analysis of samples via SDS-PAGE was realized via the discontinuous system first described by Laemmli, which allows separation of proteins based on their electrophoretic mobility, which in turn depends on their size [56].

The SDS-PAGE procedure was carried out according to standard protocols described by Sambrook and Russell [85]. Very dilute and/or samples with high ionic strength were concentrated and/or desalting by the TCA precipitation procedure described in subsection 3.4.4. Generally a 10 % (acrylamide/bisacrylamide) running gel combined with a 4 % stacking gel was used. Reducing SDS-PAGE sample buffer was added to the protein sample to be analyzed, whereafter the sample was heated to 95 °C for 5 min, to allow for total unfolding of the protein. After cooling to RT the samples were transferred into the gel pockets for analysis. The *PageRuler™ Prestained*

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

### 3.4.7 Buffer change of protein samples

The buffer in protein samples was exchanged either by dialysis, or by centrifugal filter concentrators (Amicon® Ultra Centrifugal Filter; Merck, Darmstadt, Germany). Large volumes of highly concentrated protein solutions were preferably dialyzed. Respectively, very dilute samples were concentrated and rebuffered using centrifugal concentrators.

Dialysis was carried out at least twice against a minimum of 100 times the sample volume. Dialysis steps were carried out at RT for 2 hours, or over-night at 4 °C. Centrifugal concentrators were used according to the manufacturers instructions.

### 3.4.8 Production of recombinant protein

#### Heterologous production of PFOMT

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

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

### Heterologous production of SOMT-2

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

### 3.4.9 Preparation of inlusion bodies (IBs)

Often, when recombinant protein is produced in high levels in *E. coli* it is accumulated in so-called inlusion bodies (IBs) [83]. 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 [77].

The cells were resuspended in 5 ml/g<sub>cells</sub> IB lysis buffer (100 mM Tris/HCl, 1 mM EDTA pH 7), 0.5 mM phenylmethylsulfonylfluoride (PMSF) was added as protease inhibitor. The solution was homogenized using a tissue grinder homogenizer (Ultra Turrax®; IKA®-Werke GmbH & Co. KG, Staufen, Germany). 200 µg/ml lysozyme was added to aid in the breakage of cells and the cells were lysed by sonicating thrice at 70 % amplitude (1 s on-off-cycle) for 30 seconds. DNase I (10 µg/ml) was added and the solution was incubated on ice for 10 min. The lysate was clarified by

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

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

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

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

Design of experiments (DoE) and fractional factorial design (FrFD) have been successfully used to optimize the refolding conditions of several proteins [3, 7, 111]. Thus, an approach using fractional factorial design (FrFD) was used to find optimal refolding conditions for SOMT-2.

Factors studied were pH (buffer), arginine, glycerol, divalent cations, ionic strength, redox system, cyclodextrin and co-factor addition. The experimental matrix was constructed using the FrF2 package (<http://cran.r-project.org/web/packages/FrF2/index.html>) in the R software.

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

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

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

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

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

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

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

Experiment	A	B	C	D	E	F	G	H
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 (3.4.4) and analyzed by SDS-PAGE. The remaining supernatant was rebuffered into 50 mM 2-[Bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)propane-1,3-diol (BisTris) pH 7.5 using Amicon® Ultra 0.5 ml centrifugal filters (Merck, Darmstadt, Germany) according to the manufacturers instructions. The pre-weighed collection tubes were re-weighed after recovery and the volume of recovered liquid calculated ( $\rho \approx 1 \text{ g/cm}^3$ ). The sample was filled up to 100 µl using 50 mM BisTris pH 7.5 and

the protein concentration was assessed using the Roti<sup>®</sup>-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 (3.6.3). The reactions were incubated over night and stopped by the extraction method. However, before the actual extraction 1 µl of anthracene-9-carboxylic acid (AC-9) was added as internal standard. The samples were analyzed by high-performance liquid chromatography (HPLC).

### Assessment of refolding performance

The performance of each buffer on the refolding of SOMT-2 was examined by comparing the SDS-PAGE results, as well as the amount of soluble protein and the conversion of substrate. Main effects were analyzed qualitatively using main effects plots [11].

### Upscaling of refolding reactions

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

#### 3.4.12 Enzymatic production of SAM and SAE

*S*-adenosyl-L-methionine (SAM) and *S*-adenosyl-L-ethionine (SAE) were prepared according to the method described by Dippe, et. al [22].

Preparative reactions (20 ml) were performed in 0.1 M Tris/HCl, 20 mM MgCl<sub>2</sub>, 200 mM KCl pH 8.0 and contained 7.5 mM adenosine triphosphate (ATP), 10 mM D,L-methionine or D,L-ethionine, for the production of SAM or SAE respectively, and 0.2 U *S*-adenosylmethionine synthase (SAMS) variant I317V. The reaction was stopped by lowering the pH to 4 using 10 M acetic acid after 18 h of incubation at 30 °C, 60 rpm. After 10 min incubation on ice the solution was centrifuged (15 000 × g, 10 min) to remove insoluble matter. The supernatant was transferred to a round bottom flask, frozen in liquid nitrogen and lyophilized.

Crude products were extracted from the pellet using 73 % ethanol and purified using ion exchange chromatography (IEX). IEX was performed on a sulfopropyl sepharose matrix (25 ml) via isocratic elution (500 mM HCl). Before injection, the crude extract was acidified to 0.5 M HCl using concentrated hydrochloric acid. After elution, the product containing fractions were dried via lyophilization. The amount of product was determined by UV/VIS-spectroscopy at 260 nm using the published extinction coefficient of SAM ( $\epsilon_0 = 15\,400\text{ M}^{-1}\text{ cm}^{-1}$ ) after resuspension in water [91].

## 3.5 Crystallographic Procedures

### 3.5.1 Crystallization of proteins

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

#### PFOMT

PFOMT protein was concentrated to (6 to 8) mg/ml and rebuffered to 10 mM Tris/HCl pH 7.5 using Amicon® Ultracel centrifugal concentrators (10 kDa MWCO).

The concentrated protein solution was centrifuged at  $14\,000 \times g$ , 4 °C for 10 min to remove any insoluble material or aggregates. Flavonoids and phenylpropanoid substrates were added to the protein solution from 10 mM stock solution in dimethyl sulfoxide (DMSO). Crystallization screens were set up as described above. *apo*-PFOMT was crystallized using the following conditions – 2 M  $(\text{NH}_4)_2\text{SO}_4$ , 20 %glycerol. The protein solution contained 0.25 mM SAE, 0.25 mM  $\text{MgCl}_2$ , 0.25 mM eriodictyol and 7.53 mg/ml (0.262 mM) PFOMT .

### Crystallization of proteins using NADES

NADES have the potential to be excellent solvents for hydrophobic compounds such as flavonoids or cinnamic acids [17] and in addition they are able to stabilize and activate enzymes [39].

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

### 3.5.2 Data collection and processing

Crystallographic data were collected at the beamline of the group of Professor Stubbs (MLU, Halle, Germany). The beamline was equipped with a rotating anode X-ray source MicroMax007 (Rigaku/MSC, Tokio, Japan), which had a maximum

power of 0.8 kW (40 kV, 20 mA) and supplied monochromatic Cu-K $\alpha$ -radiation with a wavelength of 1.5418 Å. Diffraction patterns were detected with a Saturn 944+ detector (CCD++, Rigaku/MSC, Tokio, Japan).

Indexing and integration of the reflexes via Fourier transformation (FT) was accomplished using *XDS* [45, 46, 47] or *MOSFLM* [79]. *Scala* [28], which is integrated in the Collaborative Computational Project No. 4 (CCP4)-Suite, was used for scaling of the intensities.

### 3.5.3 Structure solution

For the determination of the electron density  $\rho(\mathbf{r})$ , where  $\mathbf{r}$  is the positional vector, from the diffraction images by FT two terms are necessary as coefficients; the *structure factor amplitudes*,  $F_{\text{obs}}(\mathbf{h})$  and the *phase angles* or *phases*,  $\alpha(\mathbf{h})$ , where  $\mathbf{h}$  is the reciprocal index vector. The structure factor amplitudes can be directly determined from the measured and corrected diffraction intensities of each spot. However, the phase information is lost during the detection of the diffracted photons and there is no direct way to determine the phases. This constitutes the so-called *phase problem*. Thus, additional phasing experiments are necessary in order to obtain the phases. A variety of phasing experiments are available, which include *marker atom substructure methods*, *density modification* and *molecular replacement* (MR) techniques [84]. Phases of the structures herein were exclusively determined by MR [81, 82].

MR was performed using the software *Phaser* [68, 69], which is included in the CCP4-Suite [113]. A previously published PFOMT structure (PDB-code: 3C3Y [55]) was used as a template during MR procedure for the PFOMT structure solution. For the MR of the lysozyme structure the PDB-entry 4NHI was used.

### 3.5.4 Model building, refinement and validation

Macromolecular model building and manipulation, as well as real space refinement and Ramachandran idealization were performed using the Crystallographic Object-Oriented Toolkit (*Coot*) software [26]. Structure refinement was done using the software REFMAC5 [74, 105] as part of the CCP4-suite or the Phyton-

based Hierarchical Environment for Integrated Xtallography (PHENIX) [1]. Validation of the structures was carried out using the web service MolProbity (<http://molprobity.biochem.duke.edu/>) [15]. Structure visualization and the preparation of figures was performed using PyMOL (Schrödinger, New York, USA).

### 3.5.5 *In silico* substrate docking

*In silico* molecular docking studies were performed using the AutoDock Vina 1.1.2 or AutoDock 4.2.6 software in combination with the AutoDockTools-Suite (<http://autodock.scripps.edu/>) [40, 72, 103]. Substrates were docked into the PFOMT structure with the PDB-code 3C3Y. The grid box, which determines the search space, was manually assigned to center at 1.581, 5.196 and 25.718 (x, y, z) and had size of (22, 20 and 25) Å (x, y, z). The exhaustiveness of the global search for AutoDock Vina was set to 25, whereas the rest of the input parameters were kept at their defaults.

## 3.6 Analytics

### 3.6.1 Recording of growth curves

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

MTP used was opaque. Excitation ( $\lambda^{\text{ex}}$ ) and emission ( $\lambda^{\text{em}}$ ) wavelengths were (470 and 510) nm respectively.

### 3.6.2 *In vitro* determination of glucose

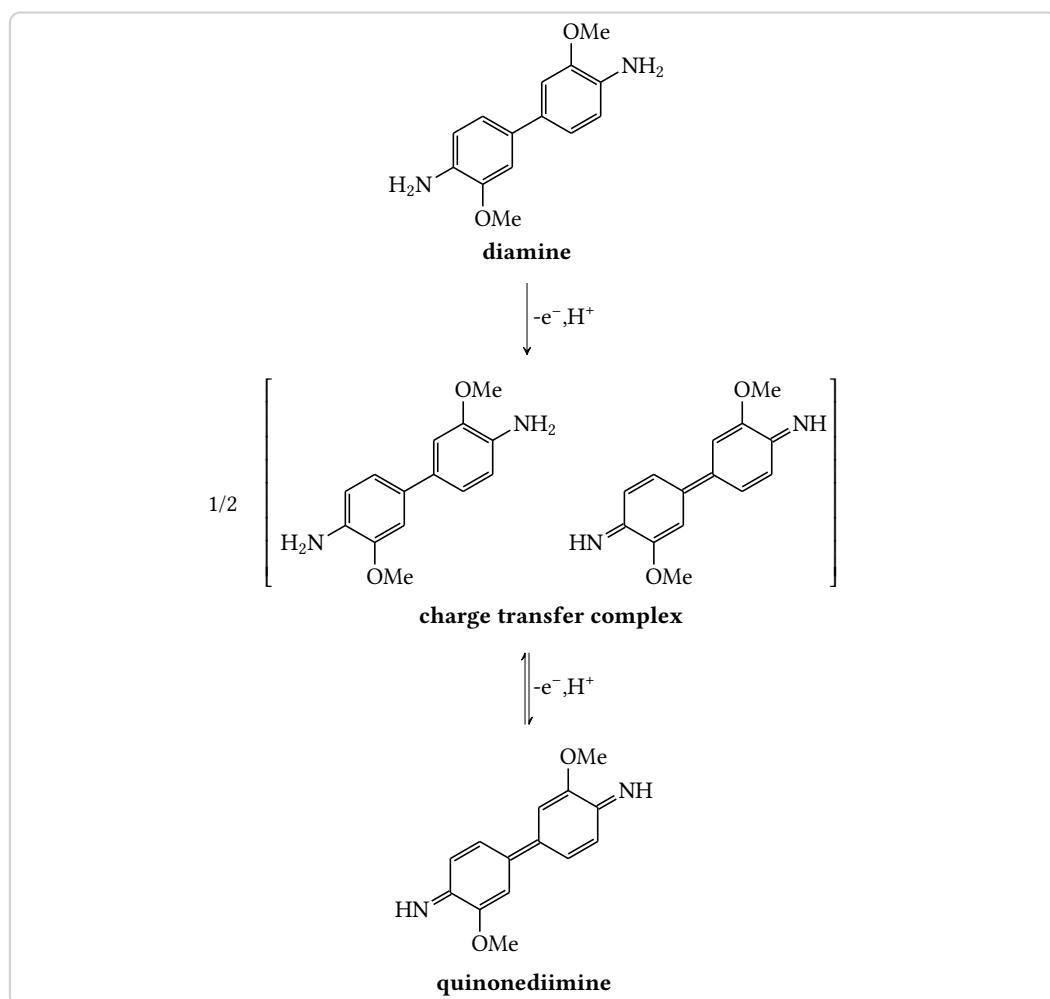
The glucose concentration in clarified, aqueous samples was determined by a modified version of the glucose assay kit procedure provided by Sigma-Aldrich [93]. Glucose oxidase (GOD) oxidizes D-glucose to gluconic acid, whereby hydrogen peroxide is produced. The hydrogen peroxide can be detected and quantified by horseradish peroxidase (HRP), which reduces the produced H<sub>2</sub>O<sub>2</sub> and thereby oxidizes its chromogenic substrate *o*-dianisidine via consecutive one-electron transfers. The oxidized diimine form of *o*-dianisidine can then be measured photospectrometrically [16].

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

Sample solutions from culture supernatants were typically diluted in 9 volumes of water. The reaction was started, by adding 50 µl reaction solution to 25 µl of sample and was incubated at 37 °C and 200 rpm for 30 min in a shaking incubator. 50 µl 6 M sulfuric acid was added to stop the reaction and achieve maximum color development (full oxidation of any *o*-dianisidine charge transfer complexes) (Figure 3.2). The developed pink color was measured at 540 nm in a MTP-reader. A calibration curve of a standard D-glucose solutions (0 to 100 µg/ml), that was always part of the experiments, was used to quantify the sample measurements.

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

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



**Figure 3.2.:** 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. [16, 44]

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

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

### Measurement of activity/pH profiles

Assays to measure activity over larger pH ranges were set up in 50 mM L-malic acid/MES/Tris (MMT)- (pH 4 to 9) or succinate/sodium phosphate/glycine (SSG)-buffer (pH 4 to 10) to keep the concentrations of buffer salts constant for each pH [75].

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

### Estimation of product concentration

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

Under the assumption, that

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

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

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

and

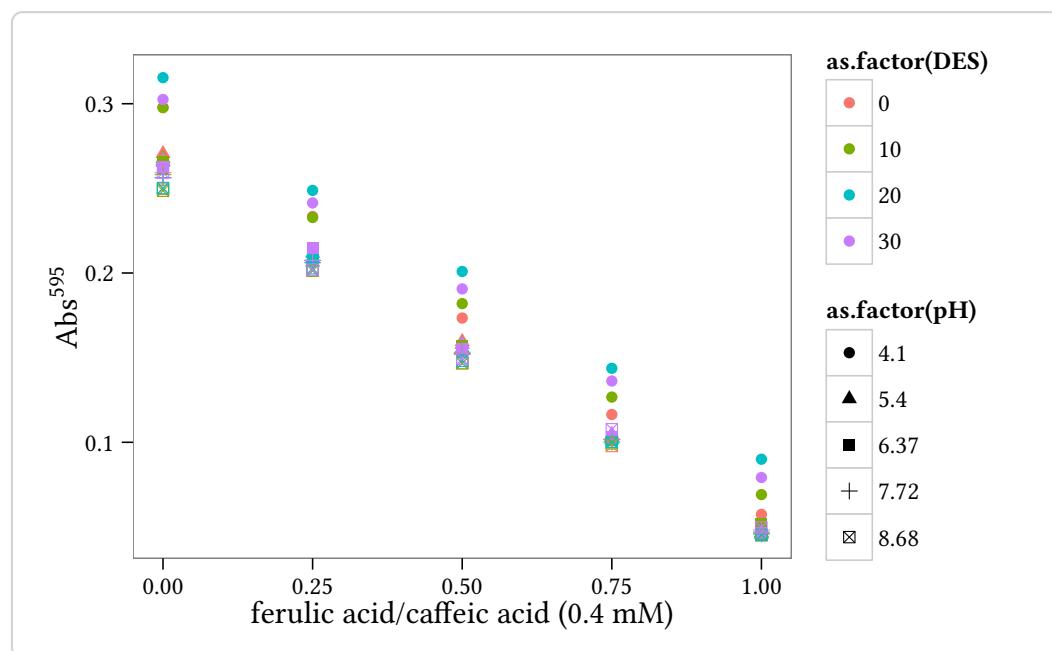
$$c^{\text{SAH}} = x^{\text{SAH}} \times c_0^{\text{SAM}}.$$

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

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

Catecholic moieties can form stable complexes in the presence of heavy metals such as copper or iron [70, 89]. Hence, caffeic acid can complex ferric ( $\text{Fe}^{3+}$ ) ions and form a colored complex with  $\lambda_{\text{max}} = 595 \text{ nm}$  [21]. Since the complex formation is specific for caffeic acid and methylated derivatives (i.e. ferulic and iso-ferulic acid) cannot complex  $\text{Fe}^{3+}$ , this can be used as a measure for methylation reactions. O-MT assays were prepared as before (3.6.3). However, the reactions were stopped by addition of 0.1 volumes 1 M Tris/HCl pH 8, immediately followed by 0.5 volumes

catechol reagent (2 mM FeCl<sub>3</sub> 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.



**Figure 3.3.:** Calibration curves of different relative compositions of ferulic acid to caffeic acid, that were taken as described in 3.6.4. The total concentration was always 0.4 mM. At lower pH values around 4, the method seems to overestimate the concentration of caffeic acid. However, the slope of the curves stays the same.

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

After refolding using rapid dilution protein samples are very dilute and a concentration step is required. Refolded SOMT-2 was concentrated directly from the 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<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> solution and the pH was adjusted to 7 using 5 M NaOH. The sample was centrifuged (20 000 × 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 ( $1\text{ M }(\text{NH}_4)_2\text{SO}_4$ ,  $50\text{ mM HEPES pH 7}$ ). The target protein was eluted using a stepwise gradient (( $1, 0.8, 0.6, 0.4, 0.2$  and  $0\text{ M }(\text{NH}_4)_2\text{SO}_4$ ,  $50\text{ mM HEPES pH 7}$ ;  $5\text{ CV each}$ ) to remove the ammonium sulfate. The column was washed using  $20\%$  ethanol. Before SDS-PAGE analysis the eluted high salt fractions were desalting using TCA precipitation (3.4.4).

### 3.6.6 Analytical gel filtration

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

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

Isothermal Titration Calorimetry (ITC) can be used to directly characterize the thermodynamics of an observed process, be this a binding interaction or an enzymatic reaction [30].

ITC measurements to describe the interaction between PFOMT and its substrates/-effector were performed using a MicroCal iTC200 device (Malvern, Worcestershire, UK). PFOMT protein was extensively dialyzed against  $50\text{ mM MMT-buffer pH 7}$  prior to ITC experiments. The solution was subsequently centrifuged ( $14\,000 \times g$ ,  $4^\circ\text{C}$ ,  $10\text{ min}$ ), to remove insoluble matter and aggregates. The dialysate was stored at  $4^\circ\text{C}$  and used to prepare substrate and effector solutions. Generally  $50\text{ }\mu\text{M}$  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  $20^\circ\text{C}$  unless otherwise stated. The stirring speed was set to  $500\text{ rpm}$ . The injection volume was set to (2 to 4)  $\mu\text{l}$ , amounting to a total of 10 to 19 injections.

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

Due to their aromaticity, methanolic extracts of flavonoids exhibit two major absorption peaks in the UV/VIS region of the light spectrum in the range of (240 to 400) nm [64]. However, even the more simple phenyl propanoids (e.g. cinnamic acids) show absorption of light in the UV/VIS-region.

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

### 3.6.9 liquid chromatography coupled mass-spectrometry (LC/MS) measurements

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

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

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

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

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

Keywords: methyl transferase, pfomt, SAM

### Abstract

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

---

### 4.1 Introduction

Small changes to molecules can have profound influences on their chemical, physical and biological properties. For example, butyric acid esters differing only by a few methylene groups already exhibit quite divergent smells. However, not only the macroscopically qualitative properties can differ. The quantifiable psychotomimetic effect of methylated and ethylated lysergic acid amids differ by at least an order of magnitude [38, 92]. There are many more of these so-called structure activity relationship (SAR) and quantitative structure activity relationship (QSAR) studies on any number of compounds [4, 66, 86].

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 [60, 97]. Methyl transferases (MTs) catalyze the transfer of a methyl group from the co-substrate SAM to an activated atom of the acceptor molecule [97].

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

It has recently been shown however, that a wide array of SAM analogues are used as co-substrates by a variety of MTs [97]. The majority of the work so far has been done on protein methyl transferases (P-MTs) and DNA methyl transferases (DNA MTs) (Figure A.1), since epi-genetics and finding regions of gene-reulation is of great interest. There have been a great many of SAM analogues synthesized, both chemically and with the help of enzymes [18, 94].

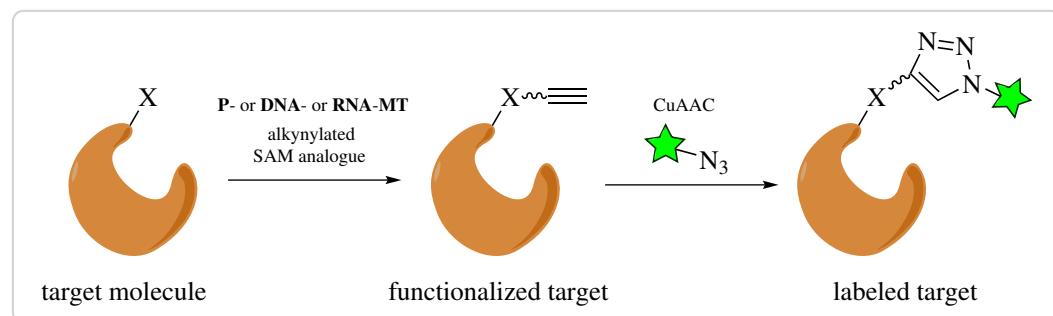
However, the first description of novel synthetic SAM analogues with extended carbon chains, including SAE, allyl and propargyl derivatives, that were also shown to be useful in modifying DNA via the action of several DNA MTs was provided by Dalhoff, *et al.* [18, 19]. A whole variety of allyl derivatives 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 [78, 110]. However, the larger substrate analogues were not neccesarily accomodated by the native P-MTs making engineering efforts for the accomodation of larger substrates inevitable [110]. The specific introduction of alkine functionalized groups made it then possible to use click chemistry for

---

<sup>1</sup>Reaxys searches for natural product isolates with a molecular mass between (150 and 1500) containing the substructures methyl, ethyl or propyl connected to a heteroatom return 66759, 2797 and 52 results respectively. However, it stands to note that 70 % of the propyl results were either esters or otherwise activated moieties. [25]

further functionalization and/or detection of the labelled proteins, DNA or RNA (Figure 4.1) [73, 78, 88, 110, 112].

In 2012 Bothwell and Luo even described the exchange of the sulfonium with a selenonium center, which afforded Se-adenosyl selenomethionine (SeAM) analogues that have since then been described as substrates for several P-MTs [9, 10]. SeAM analogues have the advantage of being more resistant to chemical decomposition than their sulfur counterparts, but also show enhanced transmethylation reactivity [9].



**Figure 4.1:** Labelling of macromolecules by using a combination of novel alkyne-derivatized SAM analogues and Cu<sup>I</sup>-catalyzed azide-alkyne 1,3-dipolar cycloaddition (CuAAC). Depending on the type of label used, it can be employed for detection (e.g. through fluorophores, coupled assays) or affinity purification (e.g. biotin). This technique is also feasible for use in activity based protein profiling (ABPP) approaches.

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-methyl transferases (C-MTs) NovO and CouO along with synthetic SAM analogues to accomplish biocatalytic Friedel-Crafts alkylations of some aminocoumarine antibiotics [96]. Lee *et al.* were the first ones to describe the transfer of a keto-group from an SAM derivative by means of the small molecule MTs catechol O-methyl transferase (EC 2.1.1.6) and thiopurine S-methyl transferase (EC 2.1.1.67) [58]. Furthermore the work done on the O-methyl transferases (O-MTs) RebM and RapM, which modify the antitumor active natural products rebeccamycin and rapamycin respectively, shows the general feasibility of using SAM analogues in combination with MTs to modify small molecules [57, 94, 117]. However, no bioactivity data has been reported that shows the biological activity of the newly produced compounds.

The *O*-MT PFOMT is a highly promiscuous enzyme with regards to its flavonoid substrates and has extensively been characterized [13, 41, 55, 107]. 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.

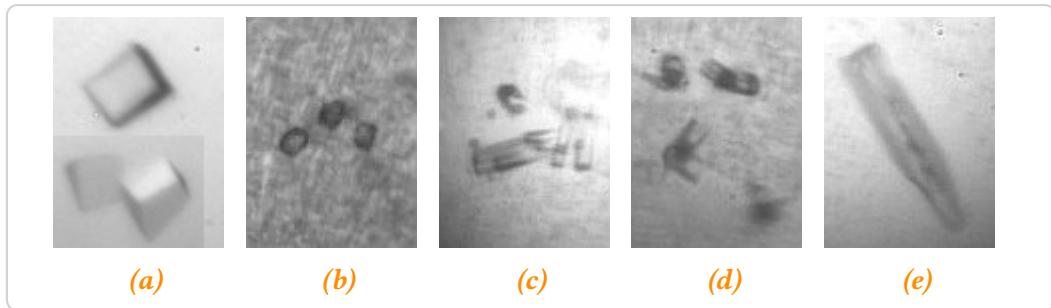
## 4.2 Crystallization of PFOMT

The crystal structure of PFOMT was published in 2008, however binding of substrates could not be accomplished [55]. 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 [55]. However, reproduction of these results could not be accomplished and new crystallization conditions had to be found.

Several commercially available buffer solutions (see section 3.5) were screened in combination with different protein solutions (e.g. solutions containing co-substrates and acceptor substrates or not) to obtain protein crystals co-crystallized with substrates or of the *apo*-form. Crystals were obtained in various wells after a few days. The crystal shape varied from very smooth and almost cubic (high

ammonium sulfate) over sphreulites and intergrown crystals ( $\text{CaCl}_2$ , PEG-4000) to brittle and ragged needles ( $\text{LiCl}$ , PEG-6000) (Figure 4.2).



**Figure 4.2.**: Some crystal and pseudo-crystal shapes that were observed during the crystallization screen. a – high  $(\text{NH}_4)_2\text{SO}_4$ , b-c –  $\text{CaCl}_2$ , PEG-4000, e –  $\text{LiCl}$ , PEG-6000

Crystals that were large enough ( $\geq 50 \mu\text{m}$ ), where screened for diffraction at the home-source after cryoprotection. A rough estimate of the resolution, cell parameters and the space group was acquired, if the diffraction images could be indexed. The screened crystals all had similar cell parameters and belonged to the same space group,  $P2_12_12_1$ , as the previously published structure (pdb: 3C3Y)[55]. However, the unit cell of crystals that grew out of high ammonium sulfate concentrations ( $\geq 1.8 \text{ M}$ ) was approximately four times as large as that of the published structure. Several datasets were collected of crystals from high  $(\text{NH}_4)_2\text{SO}_4$ , since these seemed to be promising candidates to find differences in the bound substrates. Datasets of crystals that grew from other conditions were insufficient for structure solution.

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

PFOMT crystallized without any bound substrates under conditions of high  $(\text{NH}_4)_2\text{SO}_4$ . One dataset was solved to completion to obtain a complete structure of this novel *apo*-PFOMT at a resolution of  $1.95 \text{ \AA}$  (Table B.4). The assymetric unit of *apo*-PFOMT contained two homodimers (4 monomers) (Figure 4.3a), rather than just one homodimer (3C3Y). The active site of each monomer was found to be empty except for a sole sulfate ion, which was positioned where the amino- and carboxylate groups of the SAH residue in the 3C3Y structure (Figure 4.3b). Shifts in

the structure of some loops were observed and contrary to the previously published structure the entire N-terminus was resolved up to and including the His-tag.

The resolved N-terminus contained another N-terminal  $\alpha$ -helix, which was positioned in a cleft on the surface, where substrates may be bound [55]. This interaction extends up to the His-tag. Considerable movement was observed in different parts of the protein, when no substrate was bound, some of which can be attributed to SAM and metal ion binding residues (Figure 4.4 and Figure A.2) as is obvious for the loop region between  $\beta$ -sheet 1 and  $\alpha$ -helix 4. Nonetheless, most of the movement seemed to be restricted to areas, which are not directly involved in the binding of either SAM or metal ions. However, all of the regions that moved are located at or near the active site.

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

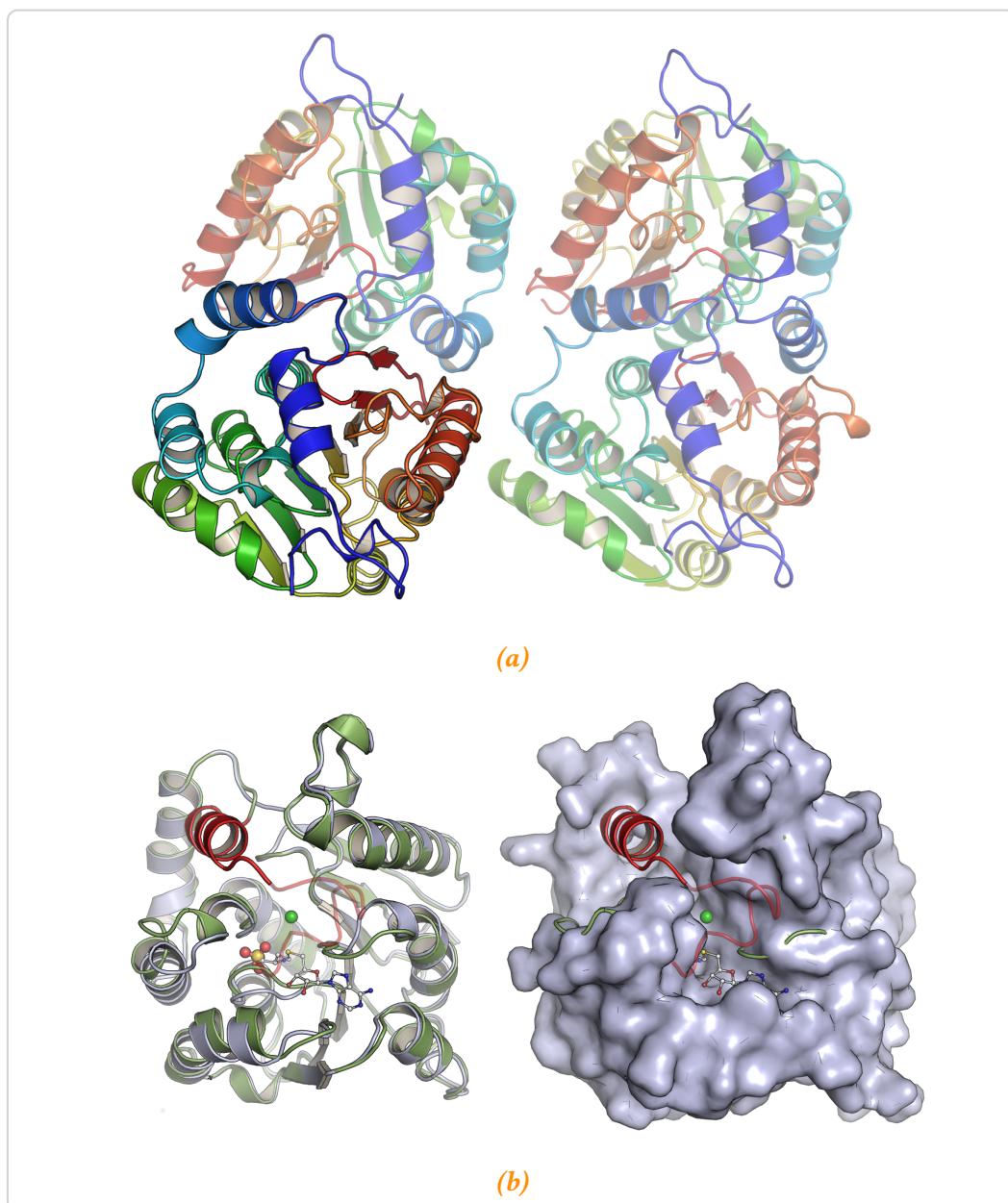
### 4.3 Substrate binding studies using ITC

The binding of different substrates to PFOMT was examined by Isothermal Titration Calorimetry (ITC), to determine whether the enzyme can bind non-natural SAM analogues. The homologues SAH, SAM and SAE were selected to also study the influence of the alkyl chain length on binding (Figure 4.5). Furthermore the binding of the substrate caffeic acid and the influence of  $Mg^{2+}$  addition on substrate binding was investigated.

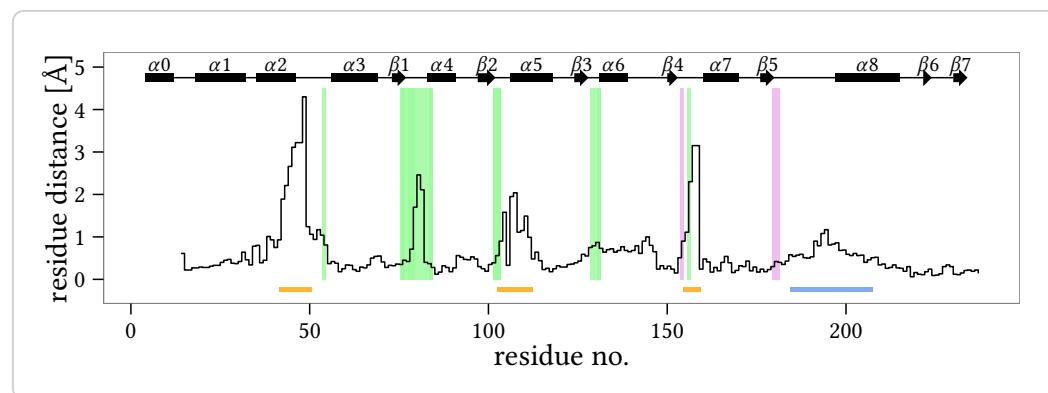
The  $K_D$  values of SAH, SAM and SAE were all in the low micromolar range, around 2  $\mu$ M. However, the binding enthalpy clearly decreased with the length of the aliphatic chain connected to the sulfur atom (Figure 4.6a). The binding of SAH, gave off more heat than the binding of SAM, which in turn gave off more heat than the binding of SAE (Table 4.1). Thus, the entropic influence must get larger with increasing chain length in order for equations (4.1) and (4.2) to still hold true.

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

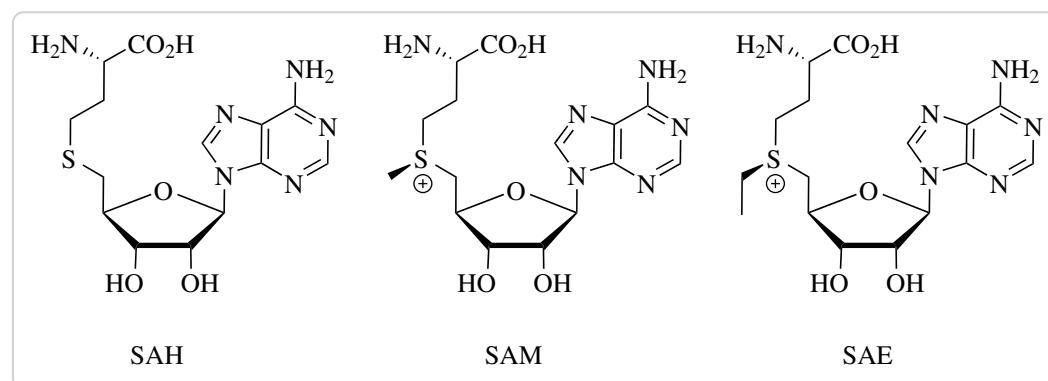
$$\Delta G = \Delta G^0 - RT \ln K \quad (4.2)$$



**Figure 4.3.**: An overview of the features in the apo-PFOMT structure. **a** – The assymmetric unit of apo-PFOMT consists of two homodimers (4 monomers). Individual monomers are rainbow colored from N-terminus (blue) to C-terminus (red). **b** – Comparison of 3C3Y (steelblue) and apo-PFOMT (green). The N-terminus of apo-PFOMT was resolved up to the N-terminus (red) and even the His-tag (red, transparent) was partly resolved. The N-terminus fits into a cleft on the surface of the 3C3Y structure, shown as a surface model on the right. SAH (white ball-and-sticks) and  $\text{Ca}^{2+}$  (green sphere) are featured in the published structure, whereas a sulphate ion (red/yellow spheres) was bound in the newly solved structure.



**Figure 4.4:** Positional differences between the individual residues of the solved apo-PFOMT and the structure with bound SAH (pdb: 3C3Y). The diffraction precision indicator [20] (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 [55].



**Figure 4.5:** The binding of different SAM analogues was measured via ITC.

Indeed, the value for  $\Delta S$  was negative for binding of SAH, but positive for the binding of SAM and SAE (Table 4.1). This relationship between the change of entropy and the change of enthalpy has been found for many biological systems and is called enthalpy-entropy compensation (EEC) [24, 35, 90]. The stoichiometry for the binding process is given by the parameter  $N$ . For all the ligands SAH, SAM and SAE this value was found to be about 0.5, which corresponds to one bound molecule ligand per dimer of PFOMT (Table 4.1).

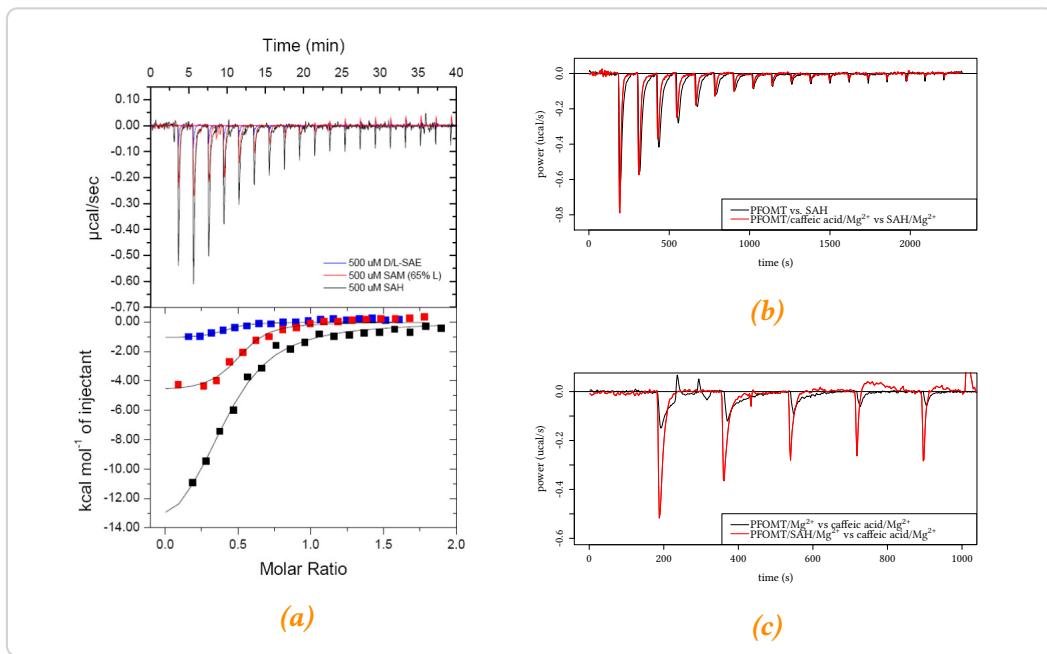
Upon titration of caffeic acid to PFOMT small amounts of released heat were detected for the system (Figure 4.6c). When the enzyme was incubated with SAH prior to addition of caffeic acid the released heat was slightly increased. The slope of the ITC profile also got steeper. However, the data obtained could not be fitted to afford a sensible solution. When caffeic acid and  $Mg^{2+}$  were incubated with PFOMT prior to addition of SAH, the process of heat production as observed by ITC had a steeper slope (Figure 4.6b). 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.

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

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

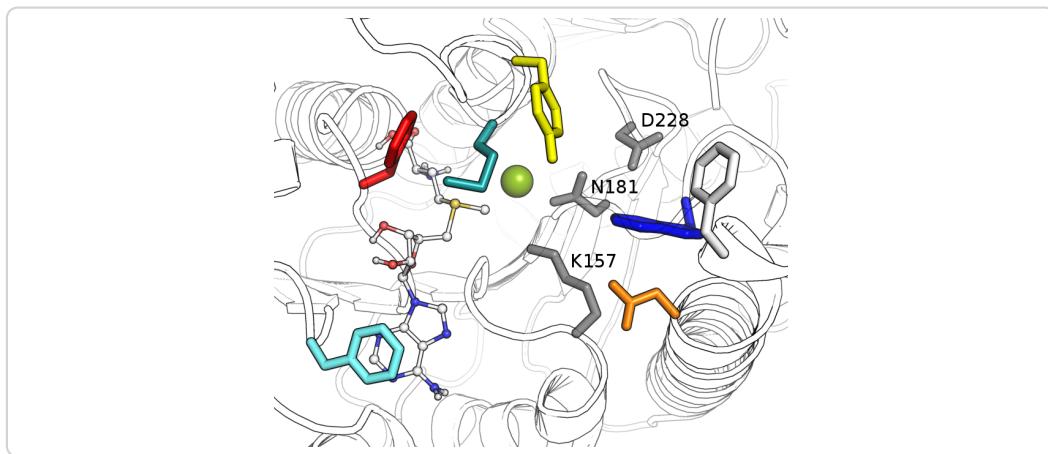
## 4.4 Study of variants for long-chain alkylations

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



**Figure 4.6:** 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  $\text{Mg}^{2+}$  and caffeic acid. When  $\text{Mg}^{2+}$  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.

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 [110]. The available crystal structures of PFOMT were consulted to select suitable residues. Residues that were exchanged were selected based upon their position in the active site and in relation to the substrate(s) (Figure 4.7). 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.

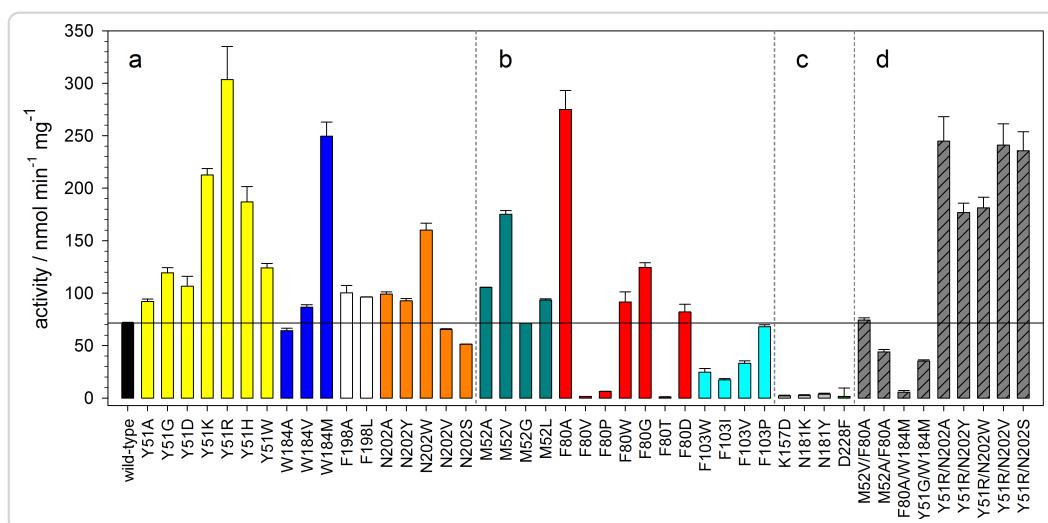


**Figure 4.7.:** The active site of PFOMT (pdb: 3C3Y). The outline of the protein backbone is displayed, with active site residues portrayed as colored sticks (cyan – F103, red – F80, turquoise – M52, yellow – Y51, white – F198, blue – W184, orange – N202, grey – as labelled). The co-substrate SAM (ball-and-stick model) was docked into the structure.

Over 20 enzyme variants were prepared to assess, whether PFOMT ethylation activity would improve over the wild-type. However, no ethylation activity was observed for either variant. Some of the new variants however displayed an increased methylation activity with the substrates caffeic acid and SAM (Figure 4.8). The methylation activity of some of the variants increased by over 4-fold. Interestingly most amino acid substitutions proved as beneficial.

Methylation activity benifited greatly from the replacement of bulky hydrophobic residues by smaller and/or charged residues in the vicinity of the acceptor substrates (Tyr51, Trp184 and Phe198). However, this was not a general trend since the substitutions N202W and Y51W also improved methylation activity. Looking more

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



**Figure 4.8:** Activities of different PFOMT variants towards caffeic acid methylation. Colorations correspond to the ones used in Figure 4.7.

## 4.5 Conclusion/Discussion

Whereas the binding of SAH was solely dependent on the large negative enthalpy, the binding of SAE was almost entirely driven by entropy, since  $\Delta H$  was close to 0 (Table 4.1). Entropy gain can be a major driving force for ligand-protein interactions and in some cases ligand binding can be entirely attributed this gain in entropy [61]. Displacement of protein-bound water molecules contributes strongly to the

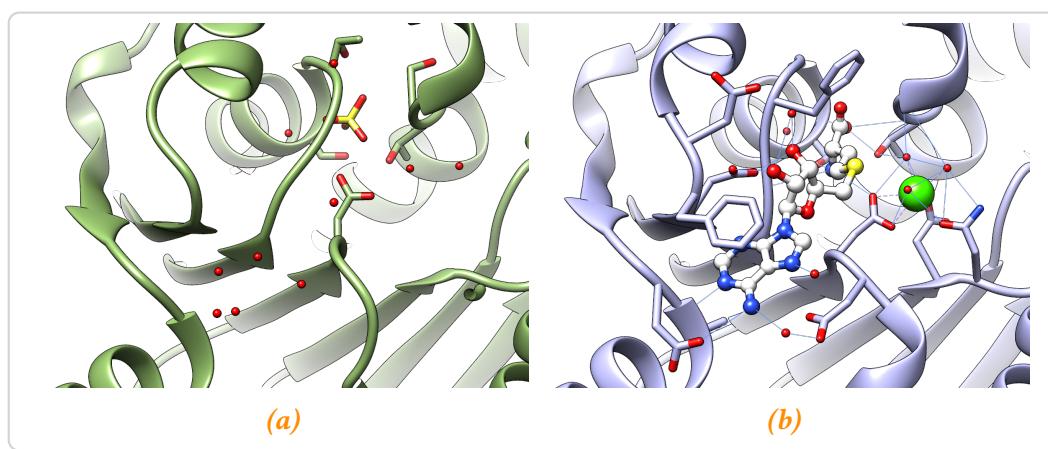
entropic gain. There were some waters present in the active site of PFOMT in the crystal structure developed herein. However, no metal ion was present in the active site in the *apo*-PFOMT structure. Furthermore Mg<sup>2+</sup> titration via ITC did not afford significant signals, suggesting the notion, that the metal is only bound along with the co-substrate (Figure 4.9). It has been suggested, that the entropy cost to transfer one water molecule from bulk to the protein-bound state can be up to 7 cal mol<sup>-1</sup> K<sup>-1</sup> [23]. The replacement of ordered waters from the active site or from a hydrated metal ion by a growing aliphatic chain could therefore explain the gain in entropy, and SAH is positioned in a way to warrant exactly that (Figure 4.9). Also, the hydrogen and metal complexing bonds consequently lost could explain the less negative enthalpy. However, this is purely hypothetical since more evident data is missing. Additional insight might be gained by expanding the ITC experiments to even longer SAM analogues. The limited space in the active site, which forces the growing side chain to expel water and possibly the metal ion might also be the reason for the inactivity of PFOMT towards SAE. If the metal ion is blocked from its complexing moieties, activation of the substrate hydroxyl would be hindered.

Comparison of the novel *apo*-PFOMT and the published structure (pdb: 3C3Y) suggests that the movement (upon ligand binding) along multiple parts of the backbone proximal to the active site pocket is a main contributor to the overall rmsd of 0.9 Å (Figure 4.4).

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 [55, 107]. 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



**Figure 4.9.:** 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.

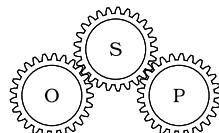
indirect proximal regions in the PFOMT structure that might be studied using site-directed mutagenesis, gene-shuffling or similar approaches in order to work towards a variant that can in fact employ SAE for transalkylation reactions. Furthermore variation of these regions might provide variants with altered substrate specificities which are of high interest.

## 4.6 Contributions

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

## 5 Enzymatic methylation of Non-catechols

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



**Benjamin Weigel<sup>1,a</sup>, Martin Dippe, Annegret Laub<sup>1,b</sup>,**

**Ludger A. Wessjohann<sup>1,c</sup>**

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

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

**Keywords:** methyl transferase, SAM, biocatalysis

## Abstract

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

---

## 5.1 Introduction

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

The 4'-hydroxyl of naringenin is non-catecholic in nature. However it is much more acidic than the 4'-hydroxyl of eriodictyol ( $pK_a$  9.8 vs 12.7) and thus about equally as acidic as the 3'-hydroxyl of eriodictyol ( $pK_a$  9.7)<sup>1</sup> [Schweigert2001a].

## 5.2 SOMT-2

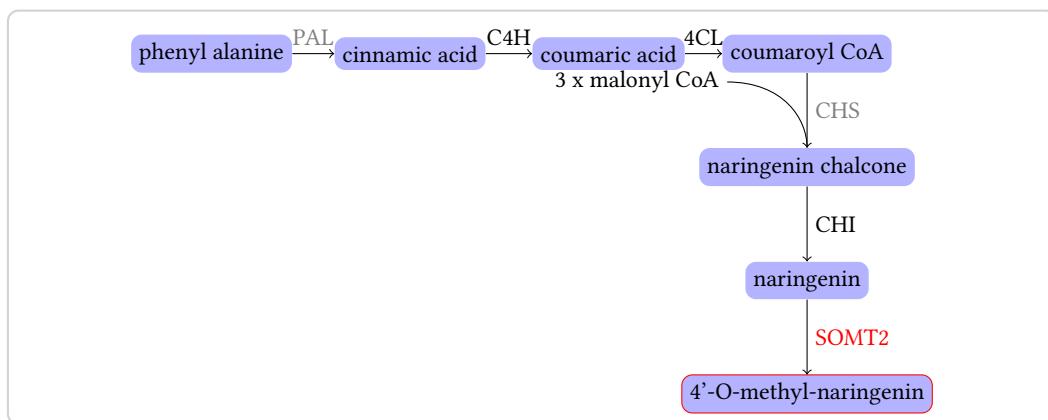
SOMT-2 has been described in the literature to methylate multiple flavonoids at the 4'-position of the B-ring [50, 51]. There it also showed the highest activity towards naringenin, to produce ponciretin (also known as isosakuranetin). Furthermore it has been the only characterized enzyme of this nature described in literature. Thus, SOMT-2 was selected as a model candidate for enzymes that can methylate 4'-hydroxyls of non-catecholic flavonoids.

### 5.2.1 *In vivo* biotransformation in *N. benthamiana*

The group of Sylvestre Marillonet (IPB) established an efficient system to clone and assemble multi enzyme pathways in *N. benthamiana*, using a modular cloning toolbox, which has already been used to produce flavonoids [54]. The group had already cloned all the enzymes required to establish the pathway up to naringenin in *N. benthamiana* (Figure 5.1). However, the *SOMT2* gene needed to be cloned into suitable vectors to be transiently expressed by *N. benthamiana*. The gene was first cloned into a level 0 module using *BpiI* and consequently subcloned into a level 1 module flanked by a promoter and a terminator using *BsaI* (Figure 5.2). The 35S-promoter was used alongside the nopaline synthase (*nos*)-terminator to achieve the highest possible transcription rates.

Both sides of *N. benthamiana* leaves were infiltrated with different samples. The left side was infiltrated with *A. tumefaciens* cultures transformed with pAGM10733

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



**Figure 5.1.**: 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

(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. After 7 days, the plant material was harvested. The average weight loss after freeze drying was 87.5 %.

The dried material was extracted and analyzed via high-performance liquid chromatography (HPLC) to determine whether ponciretin or related compounds were produced (Table 5.2). 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 resulting 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. Numerous more experiments were required to address the issue of non-detection, which however was outside the scope of this work.

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 5.3 and A.4). The principal component analysis (PCA)-plot shows that the samples of the different

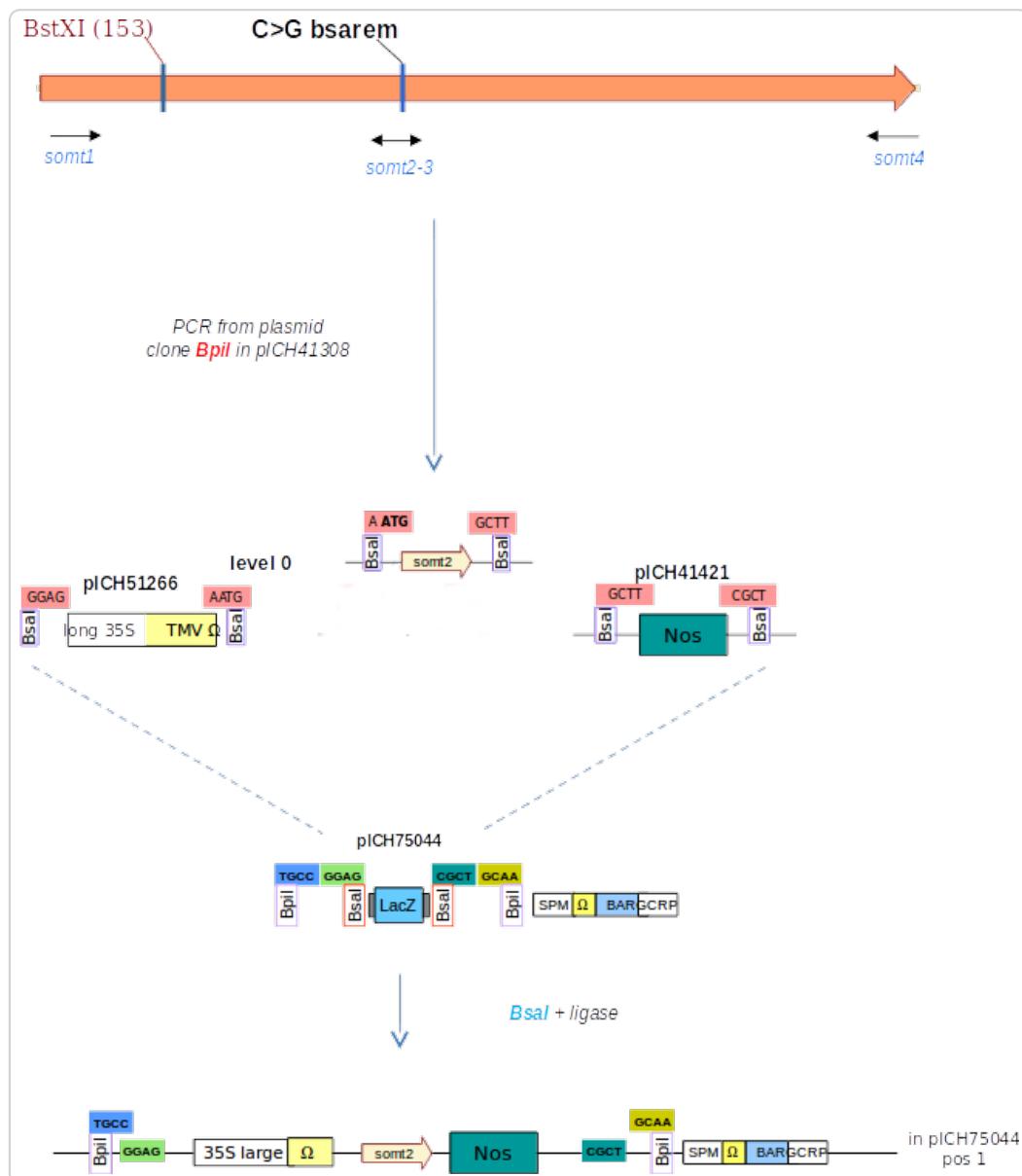
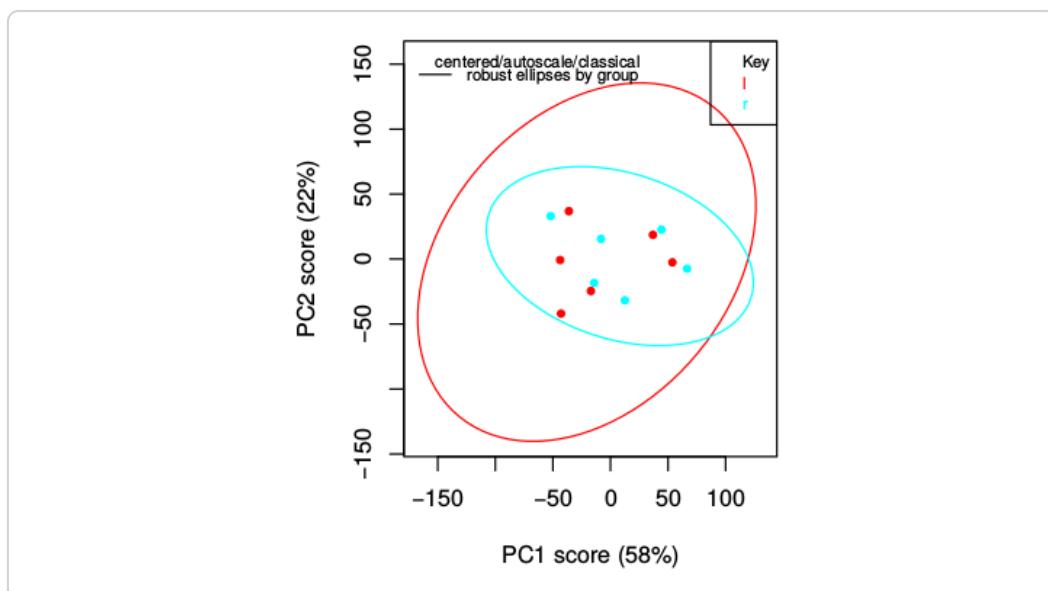


Figure 5.2.: Cloning of SOMT-2

**Table 5.2.:** Naringenin and 4'-methylated derivatives.

<chem>O=C1OC(Oc2ccc(Oc3ccccc3)cc2)C=C1O</chem>	R <sup>1</sup>	R <sup>2</sup>	name
	H	H	naringenin
	CH <sub>3</sub>	H	ponciretin
	CH <sub>3</sub>	rutinose <sup>1</sup>	poncirin
	CH <sub>3</sub>	neohesperidose <sup>2</sup>	didymine

leaf sides do not separate, indicating no difference between infiltration with the *SOMT* gene and vector control between the first two principal components, which account for 80 % of the variance. However, there is a slight separation between top and bottom leaves in the second principal component and between plant 3 and plants 1/2 in the first principal component. This suggest, that the chemical composition as detected by HPLC is slightly different in the top and bottom leaves, as well as between plants.

**Figure 5.3.:** PCA of leaf material. The samples are colored by leaf side.

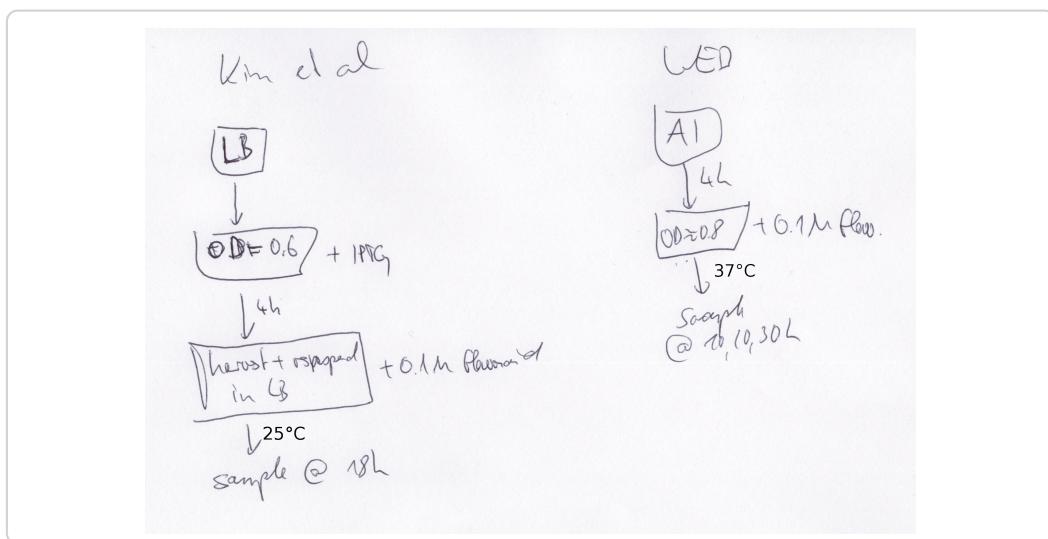
### 5.2.2 *In vivo* biotransformation in *E. coli*

Kim *et al.* already showed, that SOMT-2 could be used for the biotransformation of different flavonoids in *E. coli* live cultures [50, 51]. Their technique was to be studied and expanded upon to gain additional insight into the enzyme. The SOMT-2 gene was cloned into the pET28a(+) and pET41a(+) vectors, to obtain constructs for the production of SOMT-2 without and with a N-terminal Glutathion S-transferase (GST)-tag, respectively since both have been used successfully by Kim *et al.*. However, methylated flavonoids were not detected when biotransformations were prepared according to the methods of the aforementioned authors (Figure 5.4).

Thus, the biotransformation medium was changed to auto-induction medium (N-Z-amino, yeast extract, phosphate (ZYP-5052)) [98]. 0.05 % glucose were present in the medium. Growth curves were measured, alongside glucose depletion, to establish the time at which substrate addition would be most beneficial. Furthermore, sodium dodecylsulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) samples were prepared to observe protein accumulation. The glucose present in the medium was depleted after about 5 hours into growth. At that time expression of the SOMT-2 gene is expected to begin, because the catabolite repression on the *lac* promoter is relieved.

Thus, 0.1 mM flavonoid were added at 4 hours to minimize its influence on growth and possible degradation. Although SDS-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 (Figure A.5). Nonetheless, methylation of some of the tested substrates was observed over a course of 30 hours (Table 5.3).

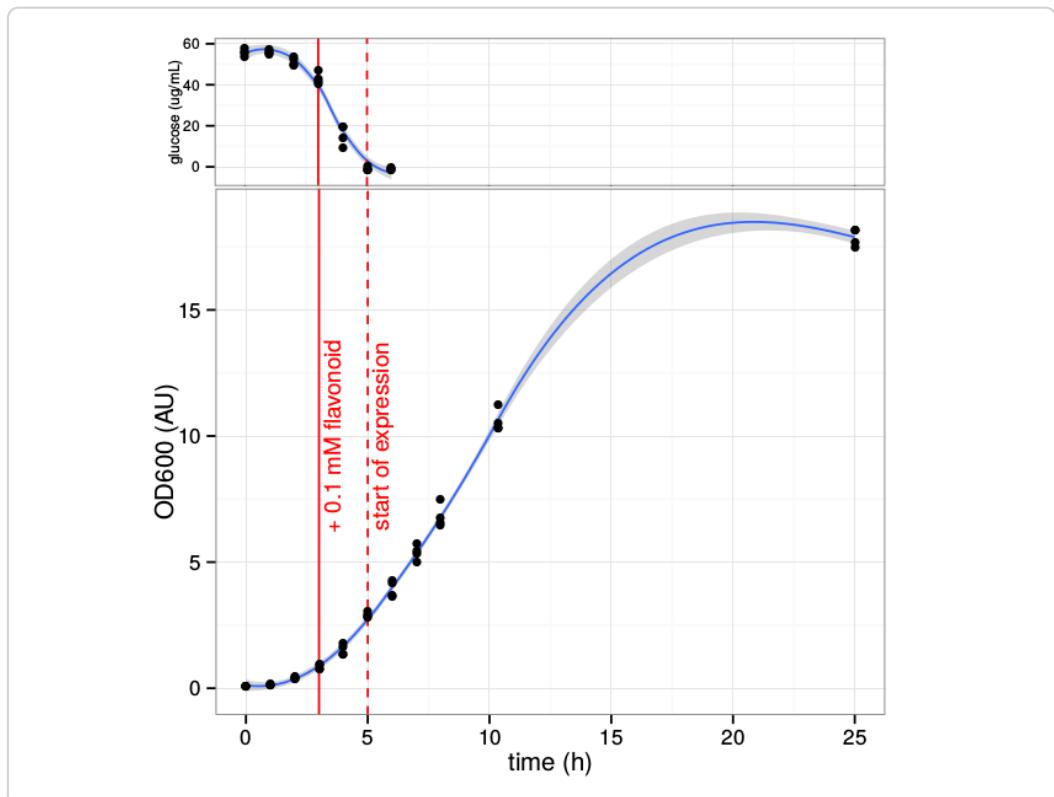
Liquid chromatography coupled mass-spectrometry (LC/MS) was employed to determine the site of methylation, since this method is highly sensitive and numerous structural studies on flavonoids using tandem-mass spectrometry experiments have highlighted the feasibility of this approach [29, 59]. Collision 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



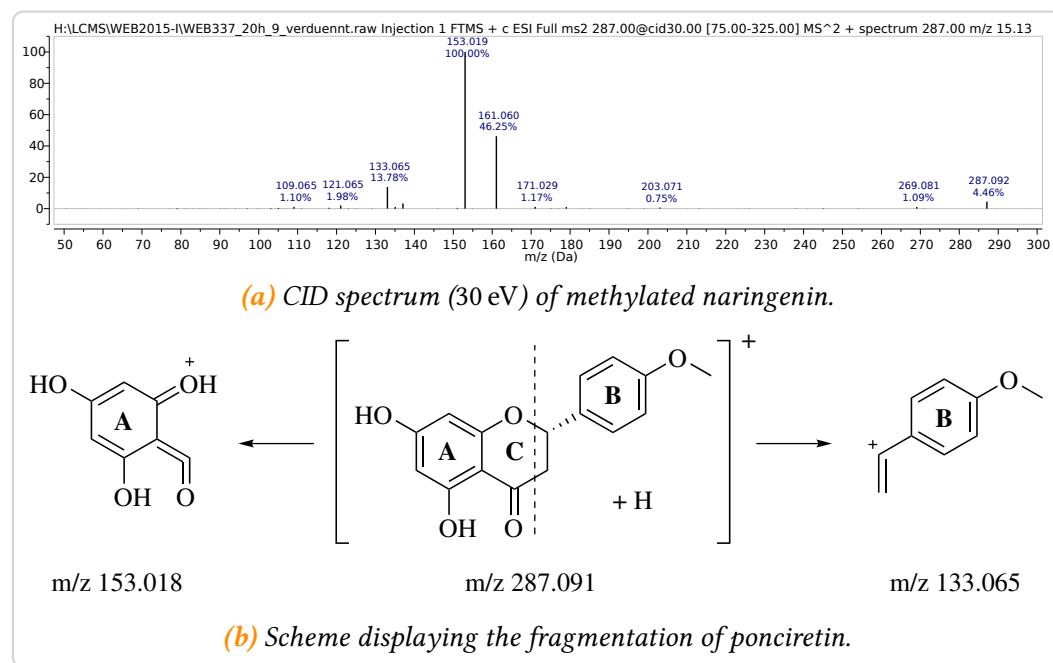
**Figure 5.4.:** Biotransformation methods as described by Kim *et al.* (left) and developed in this work (right).

yield fragments [95]. 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 [29, 59]. The fragmentation of interest in this work, was the one along the C-ring, which produces two fragments (A- and B-ring) (Figure 5.6b). The mass of the A- and B-ring fragments gives strong evidence for the position (ring) at which methylation occurred. Using the CID technique, an energy of 30 eV proved sufficient to fragment most flavonoids along the C-ring as is shown here for the methylated naringenin (Figure 5.6). 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 5.6b). If the A-ring was methylated, the expected fragment ions of A and B-ring would have  $m/z$ -values of 167 and 119 respectively.

The LC/MS results suggest, that methylation occurred exclusively at the 4'-hydroxyl. There was no conversion detected, when a 4'-hydroxyl was absent (Table 5.3). A free 4'-hydroxyl seems therefore necessary for a substance to be a substrate for SOMT-2, which confirms the previous results by Kim *et al.* [51].



**Figure 5.5:** Growth curve of *E. coli* BL21(DE3) expressing SOMT-2 at 37 °C. Glucose is depleted about 5 hours into growth, at which point the start of SOMT-2 expression is expected. The OD<sub>600</sub> after inoculation was about 0.1.



**Figure 5.6.:** The masses resulting from the fragmentation into A- and B-ring along the C-ring (dashed line, **b**) are evidence, that the 4'-hydroxyl on the B-ring is methylated by SOMT-2

Conversion was observed for flavonoids and the stilbene resveratrol, although 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. SOMT-2 acts on phenolic, catecholic as well as (4-hydroxy-3-methoxy-phenyl)-moeities, as is suggested by the assay results that showed methylation of naringenin, eriodictyol and homoeriodictyol respectively. The methylation of (4-hydroxy-3-methoxy-phenyl)-moieties and of stilbenes are properties of SOMT-2 that have not been described before.

The conversion ratios were assessed, but are beset with large errors due to the nature of *E. coli* rich medium extracts. The highest conversions were observed for flavanones and flavones (up to  $\geq 55\%$ ). The tested isoflavones and flavonols showed much lesser conversion ratios (less than 10%). The conversion ratios of apigenin ( $\geq 54\%$ ) and naringenin ( $\geq 55\%$ ) are comparable to the ones published by Kim *et al.*. However, genistein only showed minute conversions, which is converse to the data published by their group [51]. Conversion of eriodictyol, homoeriodictyol and kaempferol were not reported before.

The biotransformation of resveratrol to 3,5-dihydroxy-4'-methoxy-stilbene showed a conversion ratio of  $\geq 86\%$  in 30 hours. This is roughly double the conversion which was recently reported for *in vivo* biotransformations using the specific resveratrol O-methyl transferase (O-MT) sbCOM1, which only achieved 42 % conversion in 36 hours [48].

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

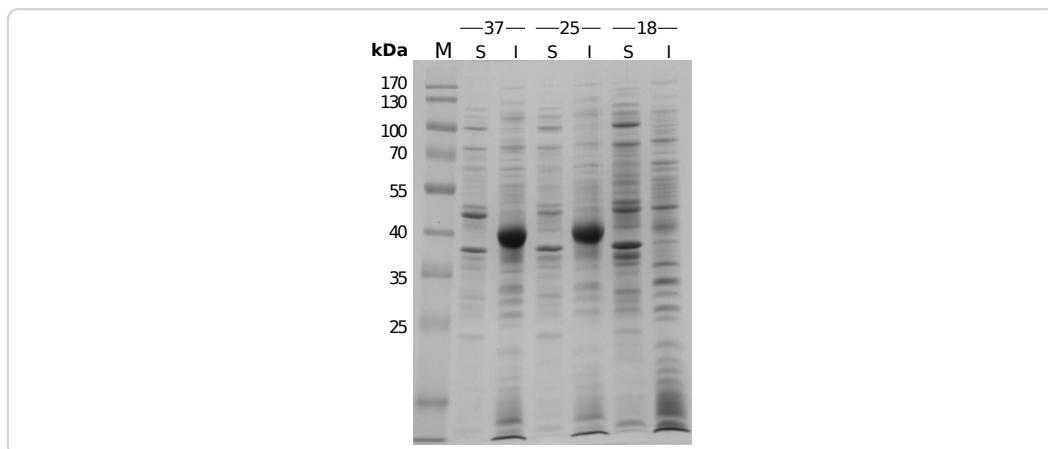
*In vivo* biotransformations are an important tool for the primary characterization of enzymes. However, because live organisms are used and lots of variables are unknown, these systems can cause large errors and are not fit to thoroughly characterize an enzyme. Initially SOMT-2 was to be purified to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

**Table 5.3.: In vivo biotransformation of different flavonoids, phenylpropanoids and anthraquinones by SOMT-2 in *E. coli***

substrate	class	4'-OH	conversion <sup>3</sup>	product <sup>4</sup>
alizarin	anthraquinone	✗	✗	–
purpurin	anthraquinone	✗	✗	–
apigenin	flavone	✓	✓(≥54 %)	4'-O-methyl apigenin
chrysin	flavone	✗	✗	–
genistein	isoflavone	✓	✓(<1 %)	Biochanin A <sup>5</sup>
galangin	flavonol	✗	✗	–
kaempferol	flavonol	✓	✓(≥6 %)	kaempferide <sup>6</sup>
naringenin	flavanone	✓	✓(≥55 %)	ponciretin
eriodictyol	flavanone	✓	✓(≥40 %)	hesperetin
homoeriodictyol	flavanone	✓	✓(>6 %)	3',4'-( <i>O,O</i> )-dimethyl eriodictyol
hesperetin	flavanone	✗	✗	–
phloretin	chalcone	✓	✗	–
resveratrol	stilbene	✓	✓(≥86 %)	4'-O-methyl resveratrol
<i>p</i> -coumaric acid	cinnamic acid	✓	✗	–
caffein acid	cinnamic acid	✓	✗	–
reosmin	cinnamic acid <sup>7</sup>	✓	✗	–

### Protein production test

Initial protein production tests were carried out using *SOMT-2* cloned into pET28a(+) with an N-terminal His<sub>6</sub>-tag. However, *SOMT-2* was not produced in soluble form (Figure 5.7). Numerous systems were tested for the expression of *SOMT-2*. *E. coli* strains used for the trials included BL21(DE3), Rosetta(DE3), Origami(DE3), C41(DE3), C43(DE3), C41(DE3) pLys, C43(DE3) pLys and DH5 $\alpha$ . The *SOMT-2* gene was cloned into multiple other vectors, including pET20b for periplasmic protein production, pET32 for expression with an Trx-tag and vectors that carry promoters for induction by rhamnose. Multiple media, including terrific broth (TB), lysogeny broth (LB) and autoinduction media were used along with different inducers (e.g. lactose, rhamnose, isopropyl-D-thiogalactopyranosid (IPTG)) at different temperatures. Nonetheless *SOMT-2* could not be produced in a soluble form and expression only resulted in inclusion bodies (IBs).



**Figure 5.7.:** 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

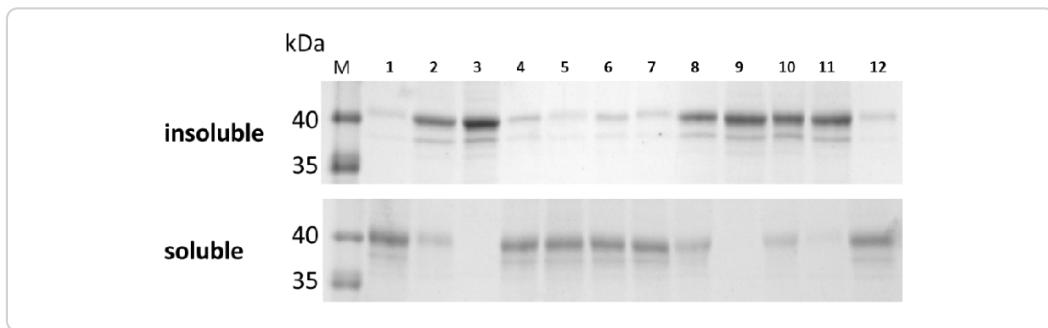
### In vitro protein refolding

Since the SOMT-2 protein could not be obtained in soluble formen, when recombinantly expressed in *E. coli*, the IBs were prepared and used for *in vitro* refolding studies. Protein refolding is no trivial task. 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 [83, 111, 115]. Refolding efficiency is best measured via biological activity, but even with adequate assays refolding studies are a time-consuming process of trial-and-error. 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 [11, 71]. 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 [102, 106, 111].

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 (see tables 3.7 and 3.8). 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.

The SDS-PAGE gels that were prepared of the soluble and insoluble fraction already show big differences between folding buffers. Refolding buffers 2,3 and 8–11 mainly produced insoluble protein, whereas the majority of the protein in refolding buffers 1, 4–7 and 12 was in soluble form in these buffers after an overnight refolding reaction (Figure 5.8). After rebuffering the cleared refolding reactions into a unified buffer the protein concentrations were estimated by Bradford [12]. The protein concentrations are consistent with the SDS-PAGE gels. 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.



**Figure 5.8.:** SDS-PAGE of the insoluble and soluble fractions of the refolding reactions.

The common denominator of those buffers is that all of them contained arginine, whose addition has proven beneficial for many refolding applications [14, 31, 104]. Main effects plots (ME-plots) illustrate the difference between level means for each factor. The ME-plot for protein concentration shows, that arginine seems to be an important factor for the refolding of SOMT-2 (Figure 5.9b). Furthermore, the addition of SAH or glycerin and the pH seem to have an influence, whereas the other factors seem to play only a minor role for high protein concentrations after refolding. However, the *Analysis of Variance* (ANOVA) test suggests that only arginine addition has a significant influence on refolding, when the significance level is set to 5 % (*p*-value 0.0158). The other *p*-values are all higher, which suggests no influence. Only the *p*-value for SAH addition (0.0897) would show significance, if the significance level was raised to 10 %.

Activity tests were conducted with the refolded protein samples to check for naringenin conversion. The amount of produced ponciretin was determined relative to the internal standard anthracene-9-carboxylic acid (AC-9). The protein activity in the refolded samples was generally very low. The maximum conversion of about 8.7 % was observed for the refolded protein sample 7. The activity of the samples did not correlate well with the amount of soluble protein. This becomes clear from the samples 4 and 5, where the amount of soluble protein was high but the observed activity was at a baseline level. The ME-plot also suggests that the main effects have changed. Most notably the redox state of the refolding reaction seemed to have a

**Table 5.5.:** Results for the ANOVA of the main effects model describing soluble protein.  
Significance codes: 0.05 ( \* ), 0.1 ( . )

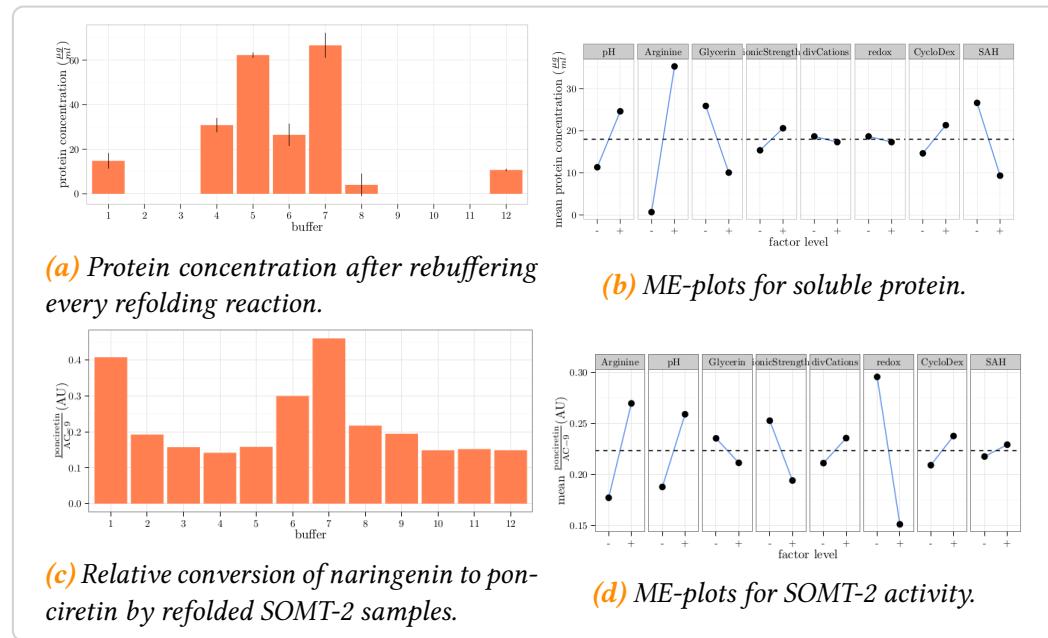
	df	Sum Sq	Mean Sq	F value	Pr(>F)	
Arginine	1	3595.63	3595.63	24.56	0.0158	*
pH	1	529.87	529.87	3.62	0.1533	
Glycerin	1	752.08	752.08	5.14	0.1083	
ionicStrength	1	82.37	82.37	0.56	0.5077	
divCations	1	5.49	5.49	0.04	0.8588	
redox	1	5.52	5.52	0.04	0.8584	
CycloDex	1	134.67	134.67	0.92	0.4083	
SAH	1	896.83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

big influence on the protein activity. Indeed, the ANOVA test suggests that using reducing refolding conditions (DTT) over a redox-shuffling system (GSH:GSSG, oxidizing) has a significant influence on methylation activity (*p* – value = 0.0210). However, there is the possibility for SOMT-2 to form intramolecular disulfide bridges, as the modelled structure suggests (Figure A.6). 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 [33, 37, 43]. Nevertheless most MTs are only active under reducing conditions and literature suggests, that sometimes assays of MTs are explicitly conducted under reducing conditions [41, 119].

The *p*-value for arginine addition is 0.0649, which also suggests a significant contribution when using a significance level of 10 %. This is plausible, since there cannot be any activity when no soluble protein is present. Judging from the ANOVA test, the remaining factors have no significant impact on the protein activity after refolding.

### Scaleup of *in vitro* refolding reactions

Buffer 7 (50 mM borate/NaOH, 0.5 M arginine, 2 mM CaCl<sub>2</sub>, 2 mM MgCl<sub>2</sub>, 10 mM NaCl, 0.5 mM KCl, 30 mM  $\alpha$ -cyclodextrin, 5 mM DTT pH 8.5) was used to scale up the refolding reaction from a total volume of 1.05 ml to a volume of 50 ml. After

**Figure 5.9:** Results of *in vitro* protein refolding trials.**Table 5.6:** Results for the ANOVA of the main effects model describing protein activity. Significance codes: 0.05 (\*), 0.1 (.)

	df	Sum Sq	Mean Sq	F value	Pr(>F)	
Arginine	1	0.03	0.03	8.14	0.0649	.
pH	1	0.02	0.02	4.83	0.1153	
Glycerin	1	0.00	0.00	0.55	0.5122	
ionicStrength	1	0.01	0.01	3.27	0.1682	
divCations	1	0.00	0.00	0.57	0.5047	
redox	1	0.06	0.06	19.88	0.0210	*
CycloDex	1	0.00	0.00	0.78	0.4428	
SAH	1	0.00	0.00	0.13	0.7439	
Residuals	3	0.01	0.00			

refolding and concentration the protein solution was analyzed by gel filtration, circular dichroism (CD) spectrometry and activity tests were conducted.

The retention time for SOMT-2 during the gel filtration run was 14.5 min (Figure A.7). From the gel filtration standard measurements, this corresponds to a molecular weight of approximately 57 853 Da, which is roughly the weight of one SOMT-2 monomer (40 kDa). The CD spectrum (Figure A.8) suggest, that the refolded SOMT-2 possesses a secondary structure and is not present as an unfolded random coil. The secondary structure was estimated from the measured CD-spectrum by the K2D3 web service [62]. According to the calculations, the secondary structure elements consist of 12.39 %  $\alpha$ -helix and 32.51 %  $\beta$ -sheet. However, the calculated protein model (A.6) suggests the helix content is much higher (52.3 %), whereas the sheet content is accordingly lower (15.4 %). This finding suggests, that the refolded protein is not in a native state. Also, the enzyme showed no activity for naringenin methylation, which was further evidence that the scaled-up refolding was unsuccessful. Even over the course of many trials a successful large scale refolding of SOMT-2 yielding active protein could not be achieved.

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

## 5.3 PFOMT

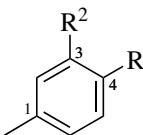
### 5.3.1 Phenolic hydroxyls

Phenolic hydroxyl groups have  $pK_a$ -values of around 10 as demonstrated by four *p*-cresole derivatives (Table 5.7). Catecholic systems have two  $pK_a$ s. The 3-hydroxyl in this example has a much smaller  $pK_a$  than the 4-hydroxyl. This is in part due to the mesomeric (+M) and inductive (+I/-I) properties the substituents display. The M and I-effects let the 3-OH be deprotonated first, which in turn really lowers the acidity and thus increases the  $pK_a$  of the 4-OH. (4-hydroxy-3-methoxy)- and (3-hydroxy-4-methoxy)-derivatives again have a similar  $pK_a$ , with the *meta*-position slightly more

acidic due to the +I-effect of the methyl substituent. The nucleophilicity of these phenolic groups happens to coincide with their BRØNSTED acidity. Chemically speaking the hydroxyl with the lower  $pK_a$  always reacts first with an electrophile.

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

**Table 5.7.:**  $pK_a$ -values of phenolic hydroxyl groups exemplified by *p*-cresole derivatives. Substituent positions on the aromatic ring are arbitrary and do not reflect conventions of the International Union of Pure and Applied Chemistry (IUPAC).

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

Previous studies have established that PFOMT is a 3'-*O*-methyl transferase, which is not able to methylate substrates that bear either phenolic (e.g. naringenin), (3'-hydroxy-4-methoxy)- (e.g. hesperetin) or (4'-hydroxy-3-methoxy)-moieties (e.g. homo-eriodictyol) [41]. The reactions were all run under the same "standard" conditions. However, the reaction buffer/medium can have a tremendous impact on enzymes and their reactions. Therefore reaction conditions require optimization, just as the enzymes themselves, to augment an enzymatic process [8, 52].

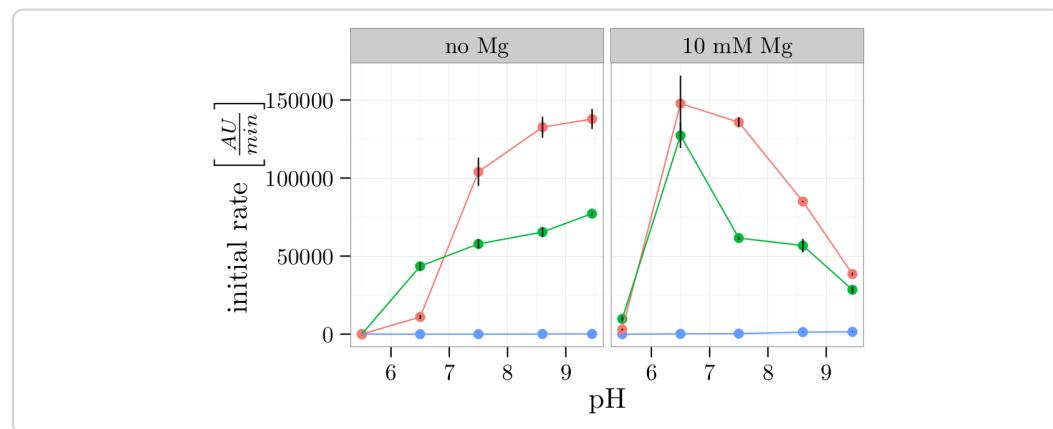
Using PFOMT the reaction conditions were screened, to assess if any would promote the methylation of non-catecholic substrates. Changes in the pH of the medium effect enzymatic activity, especially if charged groups are part of the catalytic mechanism, which was proposed for PFOMT [13]. Furthermore, PFOMT is a magnesium dependent enzyme and the activity is effected by altering the concentration of Mg<sup>2+</sup>. Thus, the pH was chosen to be varied along with Mg<sup>2+</sup> concentration in order to study the influence of those two factors on the methylation reaction.

### 5.3.2 PFOMT pH-profiles are influenced by Mg<sup>2+</sup>

PFOMT was dialyzed extensively against buffer containing 5 mM ethylenediaminetetraacetic acid (EDTA) to obtain enzyme that was virtually free of bound divalent cations. pH-profiles (pH 5.5 – 9.5) of three different substrates (caffeic acid, *iso*-ferulic acid, eriodictyol) were obtained. The pH-profiles were measured without and with the addition of 10 mM MgCl<sub>2</sub>.

The catecholic substrates caffeic acid and eriodictyol were converted by PFOMT much more quickly than *iso*-ferulic acid, which is a (3-hydroxy-4-methoxy)-substituted cinnamic acid (Figure 5.10). The highest rate of *iso*-ferulic acid conversion was two orders of magnitude lower than the highest rate for conversion of the other two substrates. Nonetheless conversion was observed for *iso*-ferulic acid with increasing pH and even an influence of magnesium was observed. Addition of 10 mM Mg<sup>2+</sup> increased the rate of *iso*-ferulic acid conversion by 3-fold from 7 pkat/mg to 21 pkat/mg at pH 9.45 (Table 5.8).

The specific activities observed for the conversion of caffeic acid are comparable to published data [22]. For the two catecholic substrates, the pH-optimum shifted from neutral to alkaline pHs, when Mg<sup>2+</sup> was omitted. However, the maximum activity remained roughly the same. This was not true for *iso*-ferulic acid, which was hardly converted at all without the addition of Mg<sup>2+</sup>.



**Figure 5.10:** Initial rate/pH-profiles for the conversion of three different substrates (red – caffeic acid, green – eriodictyol, blue – iso-ferulic acid) by PFOMT. The non-catecholic substrate iso-ferulic acid is converted much less quickly than the catecholic substrates.

**Table 5.8.:** Maximum specific activity for the conversion of three different substrates with and without addition of magnesium. The pH at which the maximal activity was reached is indicated.

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

Linear regression models of the form

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p + \epsilon \quad (5.1)$$

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

$$\text{activity} = \beta_0 + \beta_1 \times \text{Mg} + \beta_2 \times \text{pH} + \beta_3 \times (\text{Mg} \times \text{pH}), \quad (5.2)$$

which explains about 93.5 % of the variance (Table 5.9). The small p-values of the coefficients for Mg and pH×Mg suggest that there is a strong interaction between Mg<sup>2+</sup> addition and pH. However, the pH as a main effect has little significance.

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

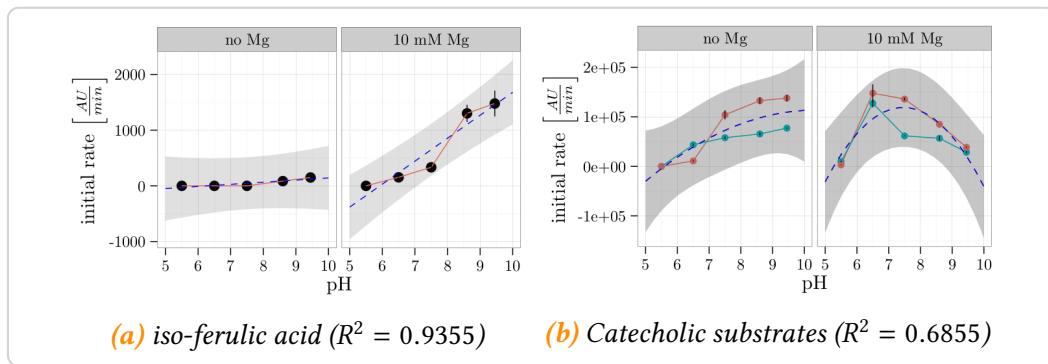
$$\text{activity} = \beta_0 + \beta_1 \times \text{Mg} + \beta_2 \times \text{pH} + \beta_3 \times (\text{Mg} \times \text{pH}) + \beta_4 \times \text{pH}^2 + \beta_5 \times (\text{pH}^2 \times \text{Mg}), \quad (5.3)$$

This model describes the actual data reasonably well, with about 68.6 % of the variance explained (Figure 5.11). Also here the p-values suggest an interaction

between  $Mg^{2+}$  and pH (Table B.5). In addition more complex linear models were prepared and shrunken via the Lasso method and 5-fold cross validation (see Table B.7 and B.9) [101]. The Lasso is a shrinkage method that can shrink coefficients to exactly zero and thus make a model more interpretable. The results of the shrunken models match well with the results obtained by linear modelling and are further statistical evidence that pH and  $Mg^{2+}$  show main effects and also interaction effects which seem to be associated with the activity towards iso-ferulic acid and catechol (eriodictyol, caffeic acid) methylation by PFOMT. Nonetheless, all of these rather simple models can not reflect the reality of such complex systems as enzymes where lots of factors play important roles. The shown results are solely included for purposes of making statistics based inferences in the context of domain knowledge.

**Table 5.9.:** Coefficients of the model (5.2) for activity of iso-ferulic acid methylation. The factor Mg is a categorical variable (addition/no addition) and can therefore only be 0 or 1.

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



**Figure 5.11.:** pH-profiles of substrate conversion. The linear regression models (blue, dashed lines) grasp the general trend of the data reasonably well to draw inferences. 95 % prediction intervals are displayed as shaded areas.

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. Addition of  $Mg^{2+}$ , seems to remedy the effect of high pHs on catalysis at least for caffeic acid, since the initial rates observed were not as high. It would be of interest to analyze this behavior with further systematic studies and multiple levels of  $Mg^{2+}$  concentrations.

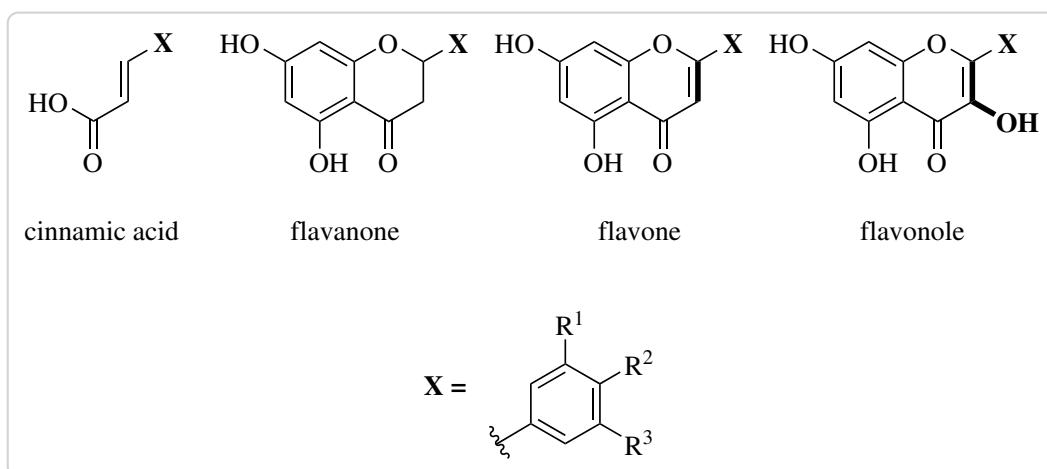
### 5.3.3 Methylation of different chemical motifs

The initial results, that showed the conversion of non-catecholic *iso*-ferulic acid by PFOMT prompted additional experiments with other different non-catechols from multiple flavonoid subgroups (Table 5.10, Figure 5.12). The tested substrates were selected from four different compound groups (cinnamic acids, flavones, flavanones and flavonols) and each group contained each of the structural motifs tested (phenol, catechol, 3O4M, 4O3M), if possible. Each substrate was assessed for conversion with two enzymes (wild-type and 4'-specific variant; Y51R N202W) at four different conditions (pH/ $Mg^{2+}$ : low/no, low/yes, high/no, high/yes). The low and high pH-values were 7.5 and 8.6, respectively. When  $Mg^{2+}$  was added the concentration was 10 mM. The reactions were incubated at 30 °C for 16 h (see section 3.6.3).

There was almost complete conversion of the catecholic substrates under any condition after 16 h of incubation, at least when the wild-type enzyme was used (Table 5.12). This suggests, that the reaction period was chosen too long. Effects on the conversion of catecholic substrates will therefore be disregarded or only discussed superficially. When the wild-type enzyme was employed conversion was observed for all substrates under at least one condition. The highest conversion of non-catecholic substrates was observed at high pH and high  $Mg^{2+}$  conditions

**Table 5.10.**: Substrate grid that was tested for methylation with PFOMT. Four different groups of compounds were screened. The groups of flavones, flavanones and cinnamic acids each contained one representative of each motif (phenolic, catecholic, 3'-hydroxy-4'-methoxy (3O4M), 4'-hydroxy-3'-methoxy (4O3M). The substitution patterns refer to Figure 5.12.

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



**Figure 5.12.:** Substrates from four different groups were screened for methylation by PFOMT under different conditions ( $\text{pH}/\text{Mg}^{2+}$ ). Figure refers to Table 5.10.

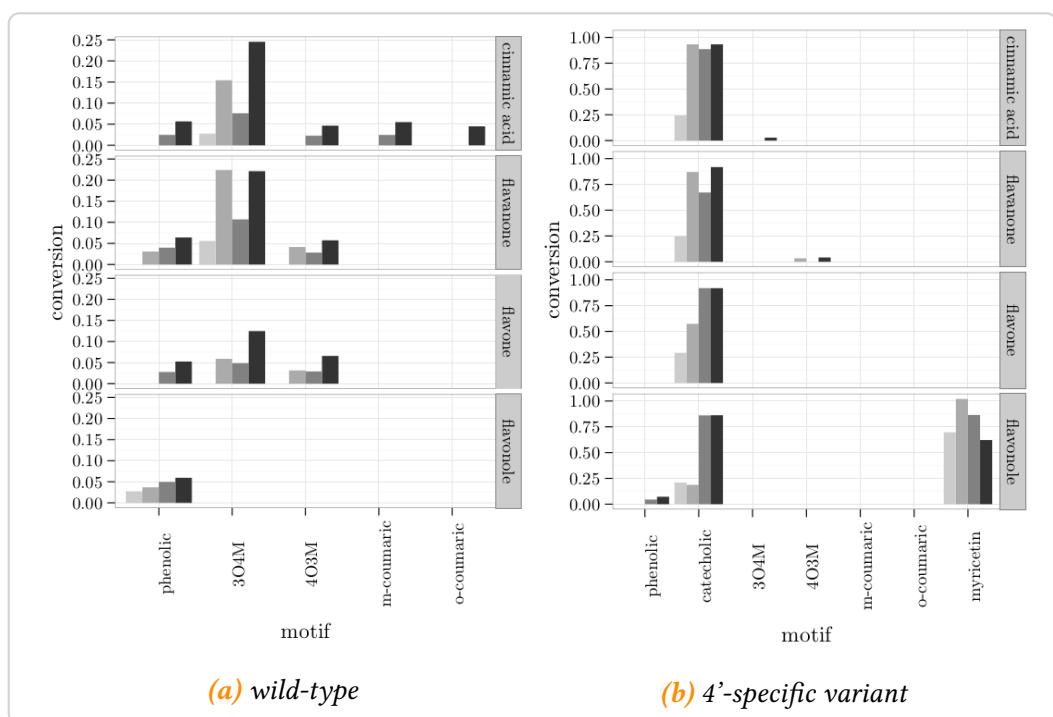
(Figure 5.13). The trend in the data suggests, that methylation efficiency increases with pH, but especially when magnesium is also added. The ME-plot and interaction effects plot (IAE-plot) suggest the same, however there is no statistical evidence to support this notion (see Figure A.9a, A.9b ). Although the statistical tests performed on this data do not back the hypothesis, this does not rule out an influence of pH or  $\text{Mg}^{2+}$ . It should be kept in mind, that enzymatic systems are quite complex and simple linear regression models that allow for inferences often cannot describe what is going on accurately and therefore fail. Furthermore this very complex dataset is limited to only about 150 observations.

Overall, methylation of 3O4M motifs was highest, with observed conversions of close to 25 % for the cinnamic acid and flavanone substrates (*iso*-ferulic acid and hesperetin). For these substrates the conversion increased by almost 5-fold upon  $\text{Mg}^{2+}$  addition, which is close to the observed increase of the initial rate of *iso*-ferulic acid methylation (5.3.2). Similar results have been shown for SaOMT5, an O-MT from *Streptomyces avermitilis* [116]. Conversion of the somewhat more rigid flavone diosmetin was lower by at least factor two compared to hesperetin and *iso*-ferulic acid. At low pH-values and without addition of  $\text{Mg}^{2+}$  barely any conversion of the non-catecholic substrates was observed. The fact, that conversion of 3O4M-moiety bearing substrates is greater than that of the 4'-phenolic and 4O3M

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

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

The site of methylation was studied by LC-tandem mass-spectrometry (MS/MS). As expected, methylation only took place on the B-ring of the flavonoids. For a throughout discussion refer to chapter 6.



**Figure 5.13.: Conversion of multiple different substrates, catecholic and non catecholic, by PFOMT. pH/Mg<sup>2+</sup> color coded from light to dark: ✕/✗, ✕/✓, ✘/✗, ✘/✓.**

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

**Table 5.12.:** Conversion of substrates after 16 hours incubation. Only the maximum conversion is displayed, along with the conditions it was achieved under.

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

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

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

methylate the 4'-position, when the substrates B-ring possessed a 4O3M substitution [87]. 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 [118]. Methylation of dihydroxy-derivatives such as luteolin and eriodictyol by TaOMT2 only gave rise to 3'-mono-methylated products, which is similar to PFOMT. However, there it was also demonstrated that TaOMT2 could methylate tamarixetin, the 4O3M derivative of quercetin, albeit at low activities.

Of the two PFOMT enzymes, 4'-specific variant and wild-type, only the wild-type showed significant methylation of non-catecholic moieties. These findings support the previous results, that could show a pH and magnesium-dependent rate of methylation of *iso*-ferulic acid (subsection 5.3.2). Although methylation of 3'-hydroxyl groups was preferred by the wild-type, a tendency to methylate 4'-hydroxyls, when these were the only ones present, could be demonstrated.

Whereas it is clear, that the N-terminus of PFOMT is important for the function of the enzyme, the role of it *in vivo* is still not fully understood [41, 55, 108]. However, it cannot be ruled out that acts as a signal sequence that can direct the enzyme to different compartments. The findings presented here might give some implications as to the regulation of *O*-MTs, such as PFOMT, since the pH can be quite different in different cell compartments in plants [67].

## 5.4 Conclusion/Discussion

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

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

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

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

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

## **6 Structural studies of flavonoids using tandem mass-spectrometry (MS/MS)**

# Characterization of flavanoid aglycones and cinnamic acids by liquid chromatography coupled tandem mass spectrometry

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

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

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

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

## Abstract

---

### 6.1 Introduction

### 6.2 Fragmentation of flavanones

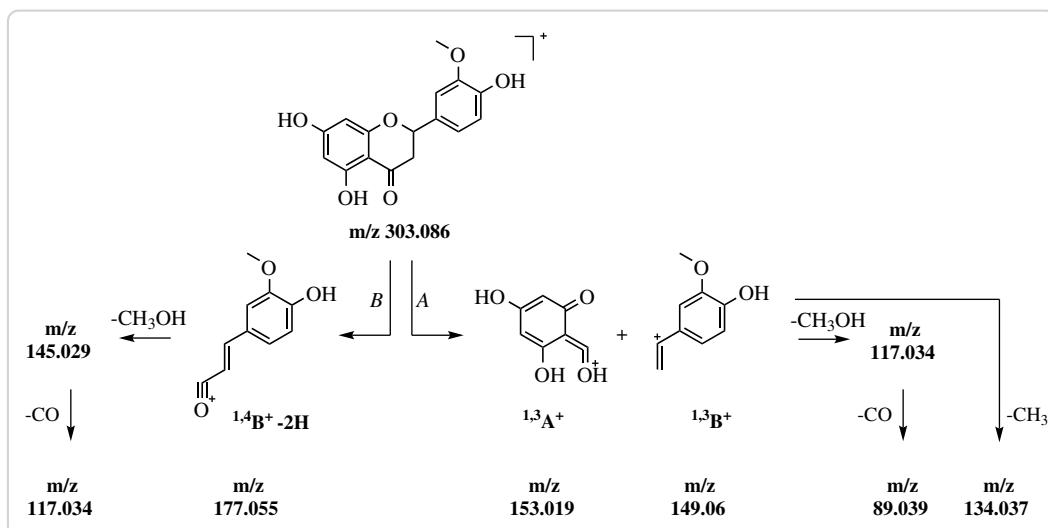
### 6.3 Fragmentation of flavones

### 6.4 Fragmentation of flavonols

### 6.5 Fragmentation of cinnamic acids

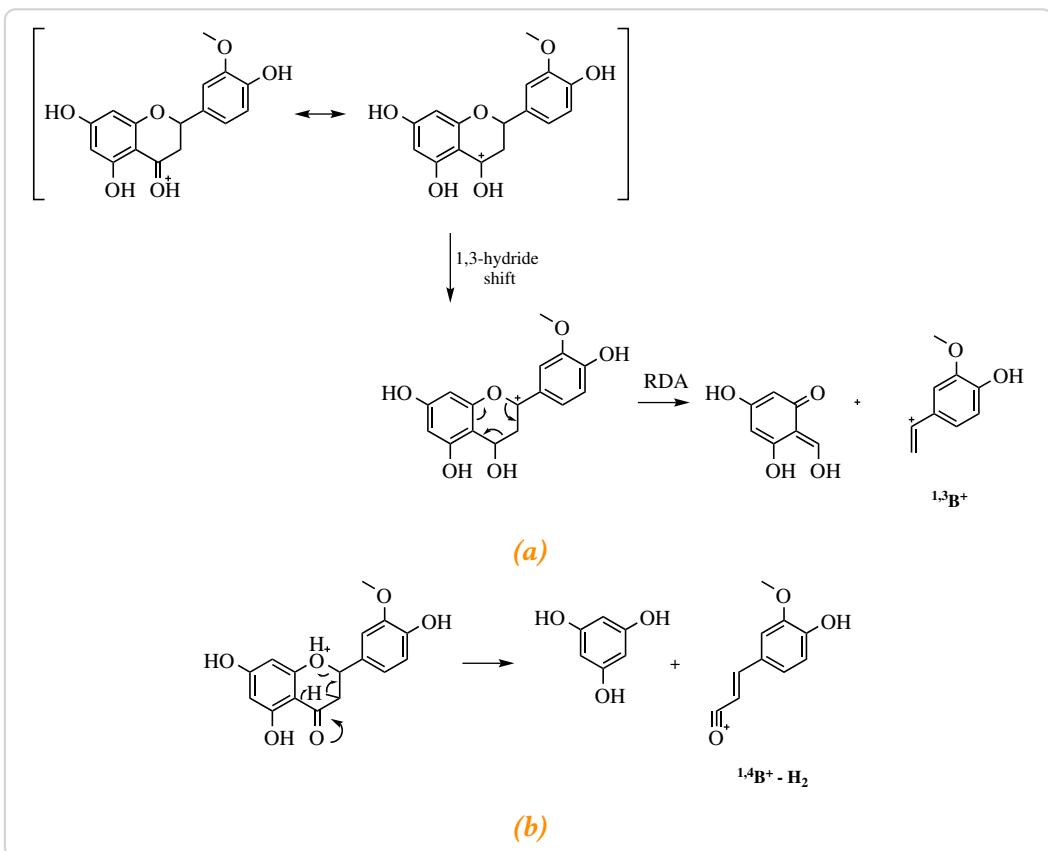
### 6.6 Conclusion/Discussion

- unmethylated compounds give better fragmentation (CID)
- HCD better for methylated flavonoids
- cinnamic acids: eV makes no difference (30 and 40 eV) give the same spectra
- cinnamic acids: dimethylcaffeic acid and isoferulic acid methyl ester differ in retention time (methyl ester is less polar and elutes later), fragmentation: meoh cleavage of ifa-mester, h2o cleavage of dimethylcaffeic acid

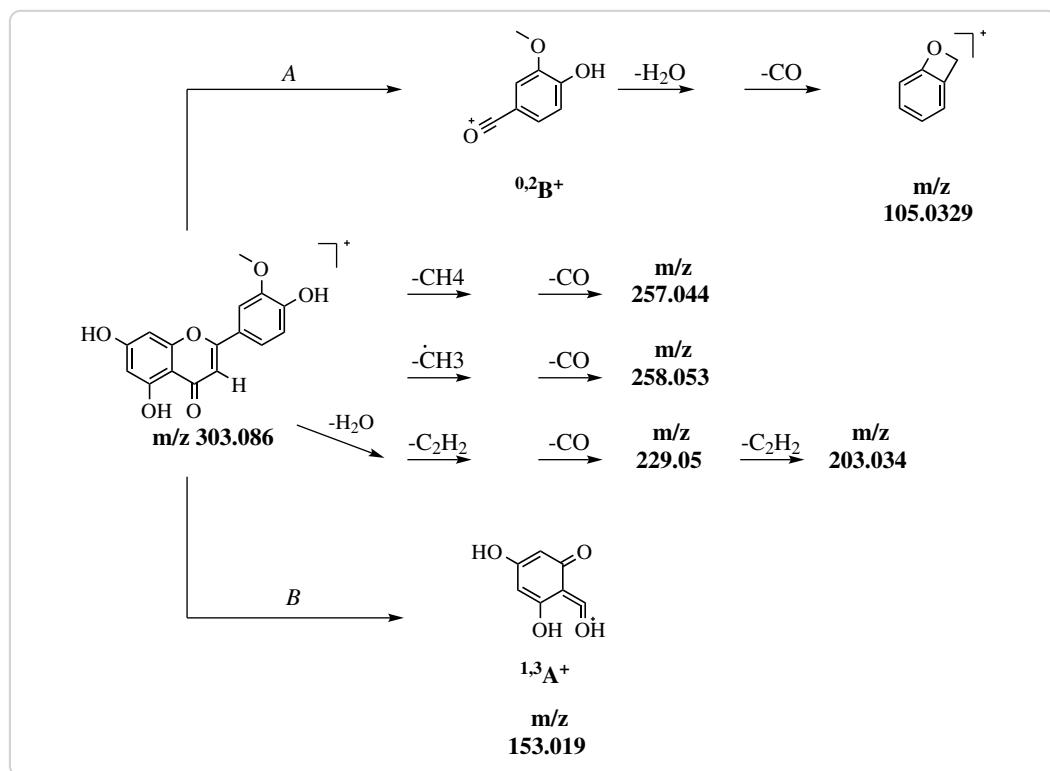


**Figure 6.1:** Major fragmentation pathways of hesperetin upon higher-energy collisional dissociation (HCD) at 75 eV

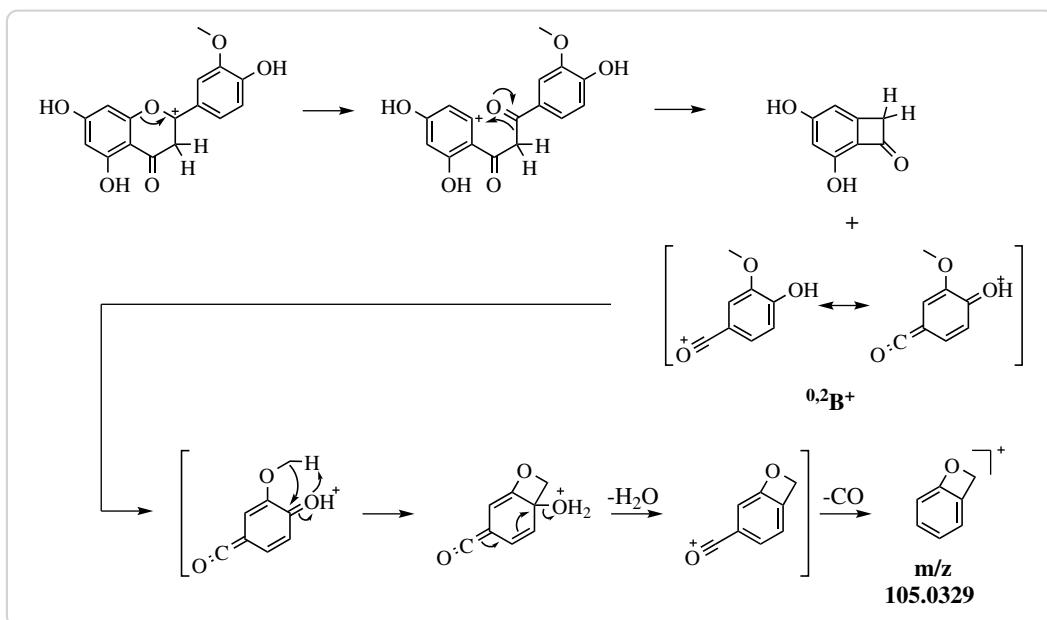
- cinnamic acid: methyl esters do not ionize well in neg. mode
- HCD: flavanes → mainly fragmentation first (e.g.  $^{1,2}\text{A}^+$ ,  $^{1,4}\text{B}^+$ ), then cleavages off the fragments ( $(^{1,3}\text{B}^+ \text{-CO-CH}_3\text{OH})$ )
- HCD: flavones → higher energy (100 eV) increases ring fragmentation (main fragment  $^{1,3}\text{A}^+$ ), lower energy tends to show losses off  $[\text{M+H}]^+$  (e.g.  $[\text{M+H-CH}_3]^{++}$ ,  $[\text{M+H-CH}_3\text{-CO}]^{++}$ )
- HCD: flavonols → mainly losses off the  $[\text{M+H}]^+$  (e.g.  $[\text{M+H-CH}_4]^+$ ,  $[\text{M+H-CH}_3\text{OH}]^+$ , etc.)
- most informative collision energy 30 eV [Cuyckens2004 ]
- $^{1,3}\text{A}^+$  is usually most abundant and most readily formed [Cuyckens2004 ] → originates from acetate–malonate pathway, therefore m/z 153; when 4-keto is missing e.g. flavanones flavanols m/z 139
- losses of small molecules and radicals common from  $[\text{M+H}]^+$ ; methyl ethers usually by loss of 15 u ( $\text{CH}_3^*$ ) [Cuyckens2004 ] → not flavones!, there loss of water (18 u) or oxa-cyclopropene ( $\text{C}_2\text{H}_2\text{O}$ )
- when methoxy substituent, the dominant ion is  $[\text{M+H-CH}_3]^{++}$  [Wolfender2000, Ma1997 ] (also in neg. mode [Cuyckens2004 ]) → not for flavones



**Figure 6.2.:** Proposed mechanisms for the formation of **a** –  $^{1,3}\mathbf{B}^+$  and **b** –  $^{1,4}\mathbf{B}^+$  ions from hesperetin upon HCD.



**Figure 6.3.:** Major fragmentation pathways of chrysoeriol upon HCD at 100 eV



**Figure 6.4.:** Proposed mechanism for the formation of  $^{0.2}\mathbf{B}^+$  from the molecular ion of Chrysoeriol under HCD conditions.

## 6.7 Contributions

Benjamin Weigel prepared substances, analyzed mass spectral data and prepared manuscript. Annelgret Laub and Jürgen Schmidt conducted liquid chromatography coupled mass-spectrometry (LC/MS) measurement runs.

## 7 DES in protein crystallography

**7.1 Introduction**

**7.2 Solubility enhancement of hydrophobic substances by addition of DES**

**7.3 Enzymatic *O*-methylation in DES**

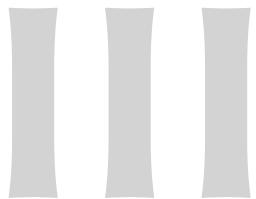
**7.4 DES as precipitants in protein crystallization**

**7.5 Conclusion/Discussion**

## 8 Acknowledgements

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

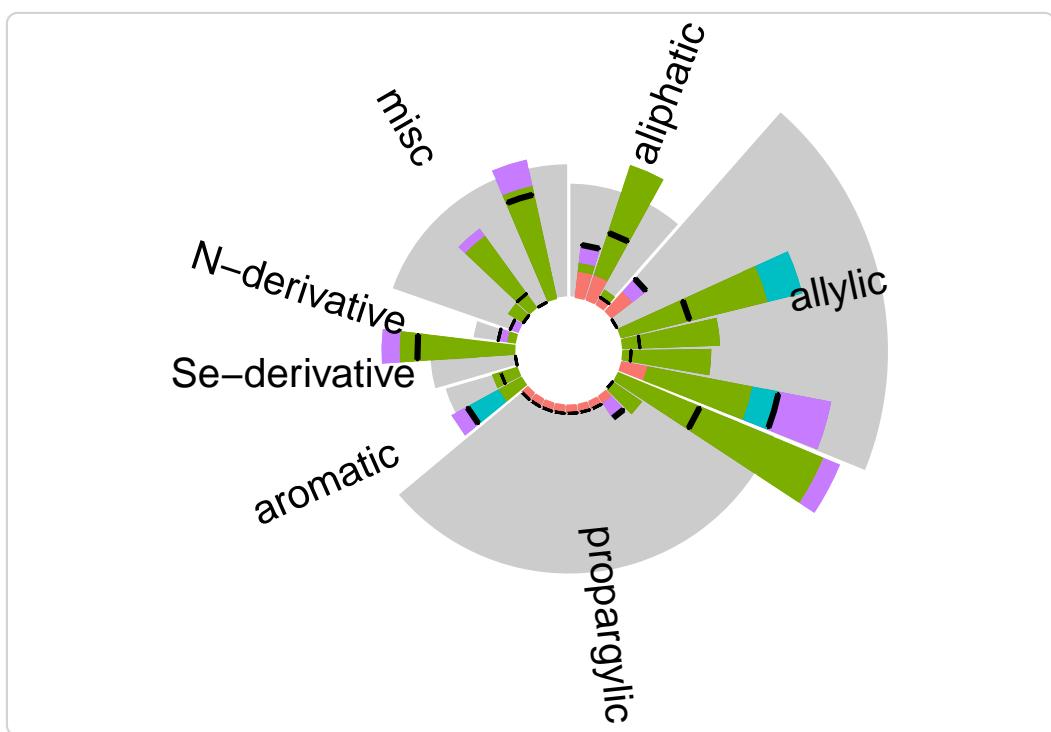
Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien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



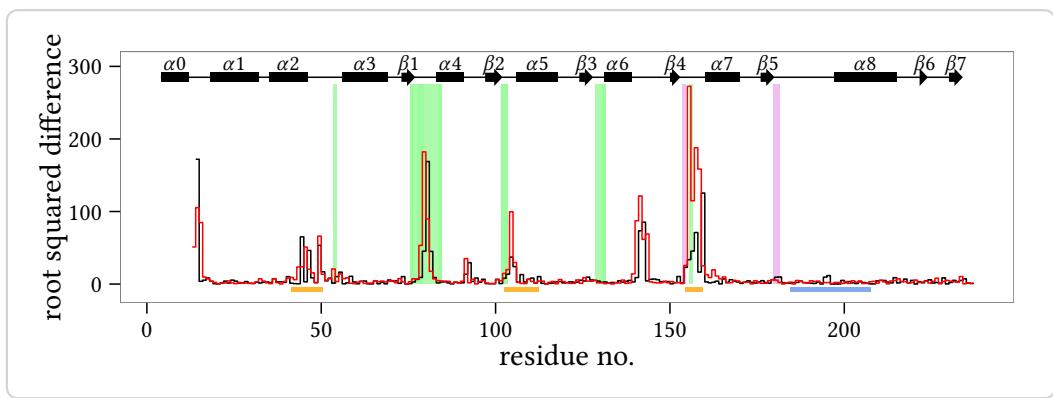
# Appendix

## A Figures

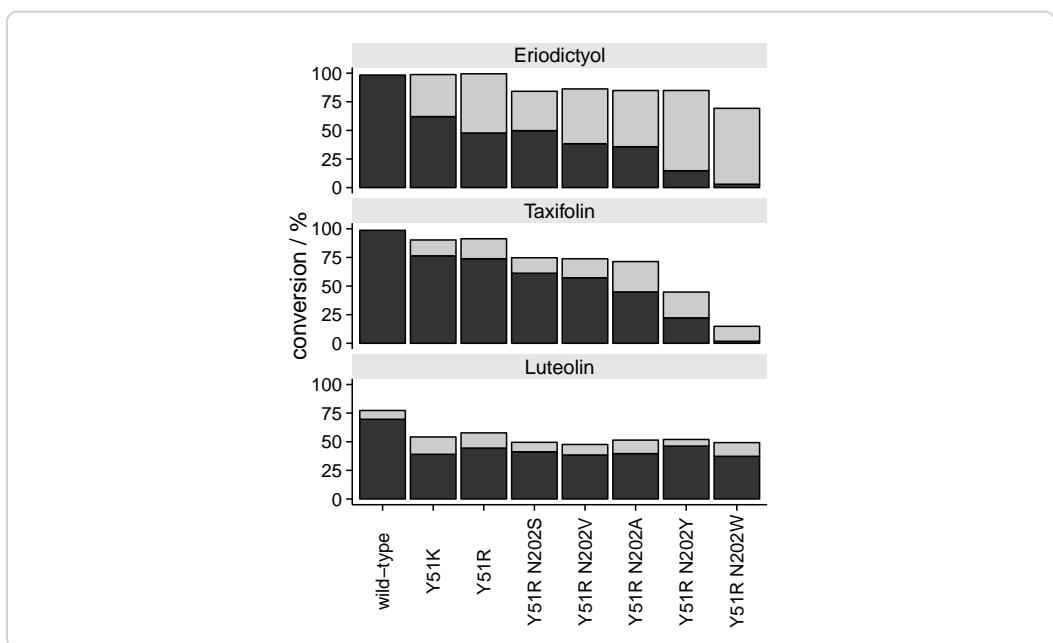
## A.1 Chapter 3



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

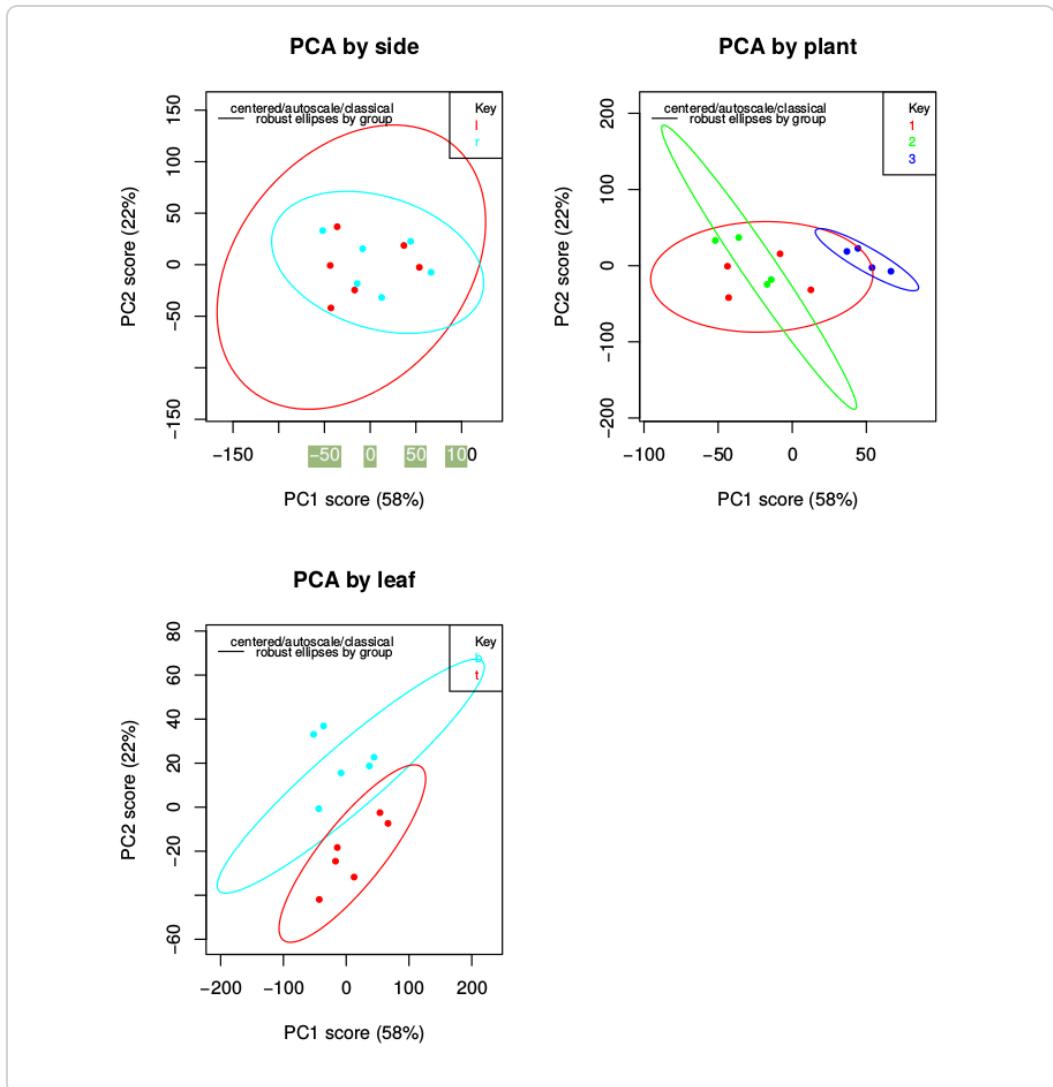


**Figure A.2.:** Differences in the dihedrals  $\psi$  (red) and  $\varphi$  (black) of the solved apo-PFOMT and the structure with bound S-adenosyl-L-homocysteine (SAH) (pdb: 3C3Y). The secondary structure is displayed at the top. Helices are displayed as rectangles and sheets are shown as arrows. Graphical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

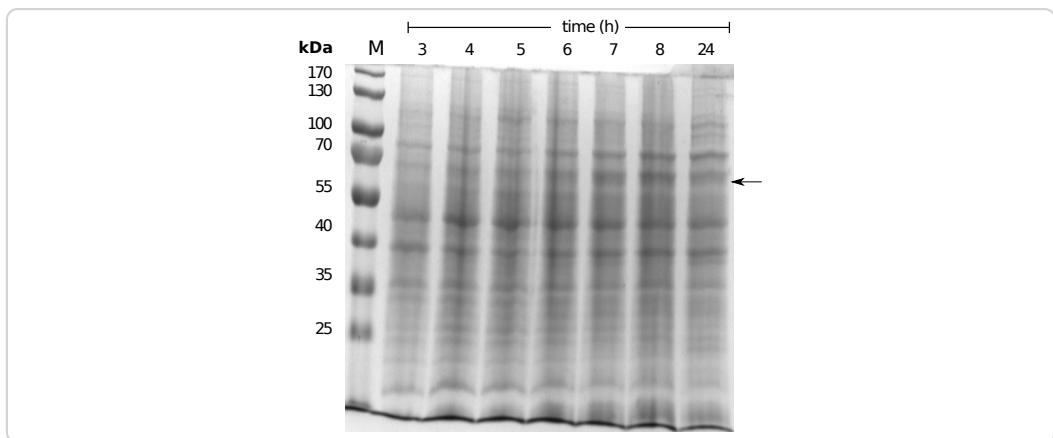


**Figure A.3.:** Differences in the regioselectivity of some phenylpropanoid and flavonoid O-methyl transferase (PFOMT) variants. The products observed in high-performance liquid chromatography (HPLC) and liquid chromatography coupled mass-spectrometry (LC/MS) measurements switched from 3'-methylated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

## A.2 Chapter 4



**Figure A.4.:** principal component analysis (PCA) of *N. benthamiana* leaves infiltrated by *A. tumefaciens* carrying different genes.



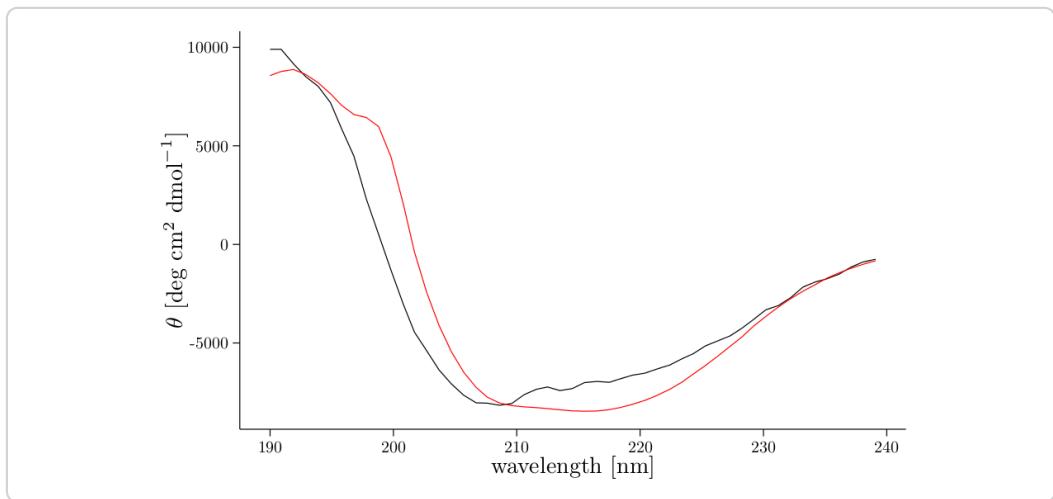
**Figure A.5.:** sodium dodecylsulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) gel of samples aquired during growth curve measurements. The arrow indicated the band that could correspond to the GST-tagged soy O-methyl transferase (SOMT-2) protein.



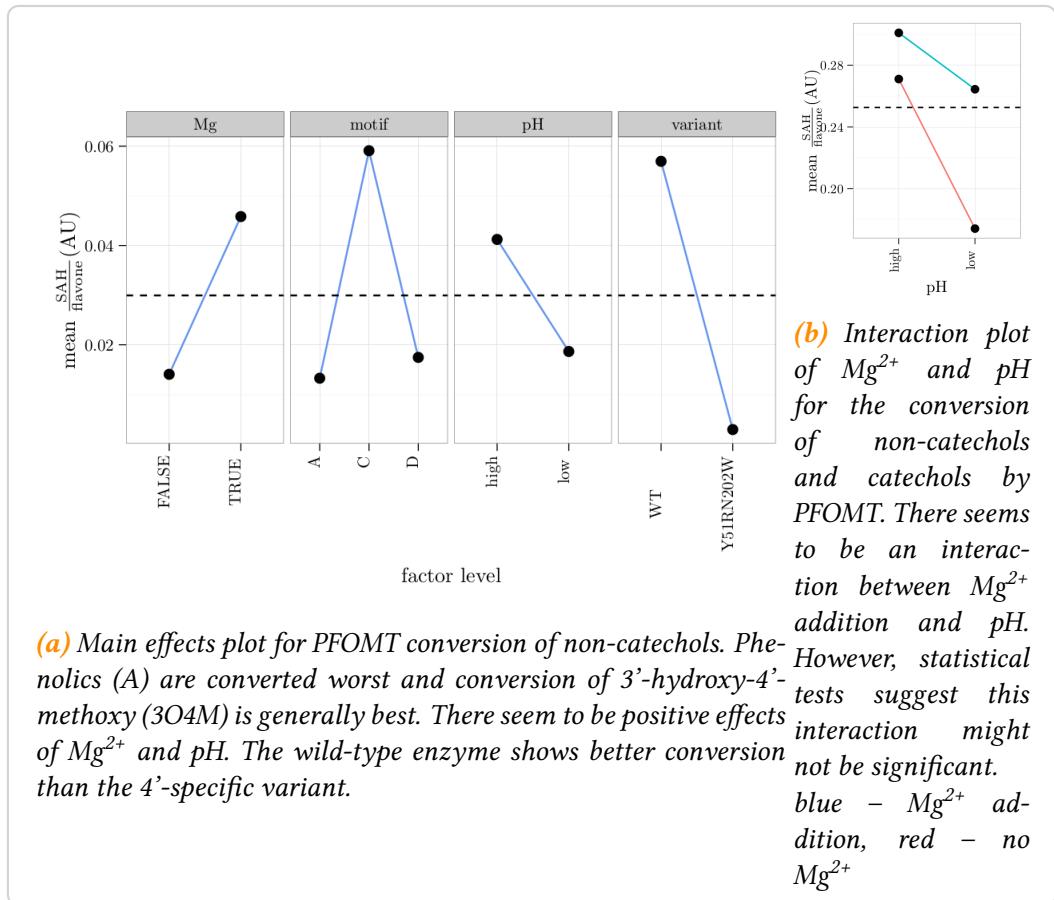
**Figure A.6.:** Graphical representation of a SOMT-2 model obtained from the PHYRE2 web portal (<http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index>) [49]



**Figure A.7.:** Chromatogram of the gel filtration analysis of refolded SOMT-2.



**Figure A.8.:** CD-spectrum of refolded SOMT-2 (black) compared to the spectrum that was calculated by the K2D3 web service (<http://cbdm-01.zdv.uni-mainz.de/~andrade/k2d3/index.html>). Secondary structure estimates from the spectrum are 12.39 %  $\alpha$ -helix and 32.51 %  $\beta$ -sheet.

**Figure A.9.**

## B Tables

## Appendix B. Tables

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

analogue	enzyme	target	references
<i>SAM</i>			
-CH <sub>2</sub> -CH <sub>3</sub>	PRMT1, M.TaqI, M.HhaI, M.BcnIB, RebM, RapM	S,P,D	[18, 57, 94, 100] <sup>1</sup>
-CH <sub>2</sub> -CH <sub>2</sub> -CH <sub>3</sub>	PRMT1, M.TaqI, M.HhaI, M.BcnIB	P,D	[18, 100]
-CH <sub>2</sub> -CH <sub>2</sub> -CH <sub>2</sub> -CH <sub>3</sub>	PRMT1	P	[100]
-CH <sub>2</sub> -C <sub>6</sub> H <sub>5</sub>	NovO, CouO,	S,P	[96, 100]
-CH <sub>2</sub> -C(=O)-CH <sub>3</sub>	PRMT1 COMT, TPMT, CazF	S	[58, 114]
-CH <sub>2</sub> -CH=CH <sub>2</sub>	NovO, CouO, RapM, PRMT1, M.TaqI, M.HhaI, M.BcnIB, RebM, Tgs	P,S,D	[18, 57, 88, 94, 96, 99, 100, 109, 110]
-CH <sub>2</sub> -CH=CH-CH <sub>3</sub>	NovO, CouO	S	[96]

<sup>1</sup>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.

*Appendix B. Tables*

analogue	enzyme	target	references
$-\text{CH}_2-\text{C}\equiv\text{CH}$	Dim-5, <i>HsMLL</i> ,	P,R,S	[42, 96, 109, 110, 112]
	TRM1,		
	NovO, CouO,		
	PRMT1		
$-\text{CH}_2-\text{C}\equiv\text{N}$	RebM	S	[94]
$-\text{CH}_2-\text{CH}_2-\text{C}\equiv\text{CH}$	PKMT	P	[42]
$-\text{CH}_2-\text{CH}_2-\text{CH}_2-\text{C}\equiv\text{CH}$	PKMT	P	[42]
$-\text{CH}_2-\text{C}\equiv\text{C}-\text{CH}_3$	NovO, CouO, M.HhaI, M.TaqI, M.BcnIB	S,D	[18, 63, 96]
$-\text{CH}_2-\text{C}\equiv\text{C}-\text{CH}_2-\text{CH}_3$	M.HhaI	D	[63]
$-\text{CH}_2-\text{C}\equiv\text{C}-\text{CH}_2-\text{NH}_2$	M.HhaI	D	[63]
$-\text{CH}_2-\text{C}\equiv\text{C}-\text{CH}_2-\text{NH}-\text{C}(=\text{O})(-\text{CH}_2-)_3-\text{NH}_2$	M.HhaI	D	[63]
$-\text{CH}_2-\text{C}\equiv\text{C}(-\text{CH}_2-)_3-\text{NH}_2$	M.HhaI	D	[63]
$-\text{CH}_2-\text{C}\equiv\text{C}(-\text{CH}_2-)_3-\text{NH}-\text{C}(=\text{O})(-\text{CH}_2-)_3-\text{NH}_2$	M.HhaI	D	[63]
$-\text{CH}_2-\text{C}\equiv\text{C}(-\text{CH}_2-)_3-\text{C}\equiv\text{CH}$	M.HhaI	D	[63]
$-\text{CH}_2-\text{C}\equiv\text{C}(-\text{CH}_2-)_3-\text{N}_3$	M.HhaI	D	[63]
$-\text{CH}_2-\text{CH}=\text{CH}-\text{C}\equiv\text{CH}$	Dim-5, <i>HsMLL</i> ,	P,R	[42, 73, 78, 88, 109, 110, 112]
	TRM1,		
	PRMT1, Tgs		
$-\text{CH}_2-\text{CH}=\text{CH}-\text{CH}_2-\text{C}\equiv\text{CH}$			[42, 109]
$-\text{CH}_2-\text{CH}=\text{CH}-\text{CH}_2-\text{O}-\text{CH}_2-\text{C}\equiv\text{CH}$	PRMT1	P	[109, 110]
<i>SeAM</i>			
$-\text{CH}_3$			
$-\text{CH}_2-\text{C}\equiv\text{CH}$	Dim-5, <i>HsMLL</i> ,	P,R,S	[10, 94, 112, 114]
	TRM1, RebM,		
	CazF		

*Appendix B. Tables*

---

analogue	enzyme	target	references
<i>N</i> -mustard derivatives			
–CH <sub>2</sub> –CH <sub>2</sub> –I	RebM	S	[117]

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

construct name	description	entry constructs	destination	workflow steps (primers/cloning sites)
pBEW101				
pBEW102	lsrA promoter			
pBEW103	pBEW102 with BamHI cloning site	pBEW102 pBEW4b	pBEW103	amplification ( <i>pRhal1fw/rv</i> ), cloning (BglII, BamHI)
pBEW104	rhaP_BAD promoter			
pBEW106	pICH413038-somt	pET28MC-somt	pICH413038	amplification ( <i>somt1/2/3/4</i> ), cloning (BpiI) golden gate cloning (BsAl)
pBEW107		pICH51266, pBEW106, pICH41421	pICH75044	golden gate cloning (BsAl)
pBEW1a				
pBEW1b				
pBEW2a				
pBEW2b				
pBEW3a				
pBEW3b				
pBEW4a				
pBEW4b				
pET28-pfomt	<i>pfomt</i> gene in pET-28a(+), endogenous NdeI site removed	pQE30-pfomt	pET-28a(+)	mutagenesis ( <i>pfomt1fw/rv</i> ), amplification ( <i>pfom2fw/rv</i> ), cloning (NdeI, EcoRI)
pET20-somt	N-terminal pelB-tag fusion for periplasmic expression			pET20-b(+) pET28-a(+)
pET28-somt				
pET28MC-somt				
pET32-somt				
pET41-somt	N-terminal TrX-tag fusion			pET-32a(+) pET-41a(+)
pUC19*	N-terminal GST-tag fusion added BglII site	pUC19 lsr-XX-DAS	-	mutagenesis ( <i>pUC1.fw/rv</i> ) cloning (NdeI, BglII)
pUCB1	pUC19 derivative with lsrA promoter			
pUCB1-sfGFP-DAS+4				

	140519_PFOMT	MC001413-G10.1
<i>data collection</i>		
wavelength (Å)		
resolution (Å)	1.95	
total reflections	392 368	
unique reflections	125 822	
completeness (%)	99.12	
$I/\sigma(I)$	9.9	
$R_{\text{sym}}^a$		
redundancy		
space group	$P2_12_12_1$	
cell dimensions (Å)		
<i>a</i>	86.16	48.88
<i>b</i>	128	71.36
<i>c</i>	129.3	127.80
<i>refinement</i>		
$R_{\text{work}}/R_{\text{free}}$	0.21369 / 0.24700	
rmsd bond lengths (Å)	0.0199	
rmsd bond angles (°)	2.0568	
B-values (Å <sup>2</sup> )	21.593	
water		
ions		
<i>Ramachandran plot (%)</i>		
favoured	96.82	
allowed	2.38	
outliers	0.8	

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

## Appendix B. Tables

---

**Table B.5.:** Coefficients of the model (5.3) for activity of catecholic methylation. The factor Mg is a categorical variable (addition/no addition) and can therefore only be 0 or 1.

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

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

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

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

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

## C Affidavit

---

*Appendix C. Affidavit*

---

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

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

# Bibliography

- [1] Paul D. Adams et al. “PHENIX: A comprehensive Python-based system for macromolecular structure solution”. en. In: *Acta Crystallographica Section D: Biological Crystallography* 66.2 (Feb. 2010), pp. 213–221.
- [2] Agilent Technologies. *QuikChange II Site-Directed Mutagenesis Kit: Instruction Manual*. 2011.
- [3] Neda Akbari et al. “Efficient refolding of recombinant lipase from Escherichia coli inclusion bodies by response surface methodology.” In: *Protein expression and purification* 70.2 (Apr. 2010), pp. 254–9.
- [4] Martin Alexander and B. K. Lustigman. “Effect of Chemical Structure on Microbial Degradation of Substituted Benzenes”. In: *Journal of Agricultural and Food Chemistry* 14.4 (July 1966), pp. 410–413.
- [5] Oyvind M. Andersen and Kenneth R. Markham, eds. *Flavonoids: Chemistry, Biochemistry and Applications*. 1st ed. Boca Raton (FL): Taylor & Francis Group, 2006.
- [6] Frederick M Ausubel et al. “Current Protocols in Molecular Biology”. In: (2008), p. 23.
- [7] Isabelle Benoit et al. “Expression in Escherichia coli, refolding and crystallization of Aspergillus niger feruloyl esterase A using a serial factorial approach.” In: *Protein expression and purification* 55.1 (Sept. 2007), pp. 166–74.
- [8] U T Bornscheuer et al. “Engineering the third wave of biocatalysis.” In: *Nature* 485.7397 (May 2012), pp. 185–94.
- [9] Ian R Bothwell and Minkui Luo. “Large-scale, protection-free synthesis of Se-adenosyl-l-selenomethionine analogues and their application as cofactor surrogates of methyltransferases”. In: *Organic Letters* 16.11 (2014), pp. 3056–3059.
- [10] Ian R. Bothwell et al. “Se-adenosyl-L-selenomethionine cofactor analogue as a reporter of protein methylation”. In: *Journal of the American Chemical Society* 134.36 (2012), pp. 14905–14912.
- [11] George E. P. Box, J. Stuart Hunter, and William G. Hunter. *Statistics for Experimenters: Design, Innovation, and Discovery*. 2nd ed. New York: Wiley-Interscience, 2005.
- [12] Marion M. Bradford. “A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding”. In: *Analytical Biochemistry* 72.1-2 (May 1976), pp. 248–254.

- [13] Wolfgang Brandt, Kerstin Manke, and Thomas Vogt. "A catalytic triad – Lys-Asn-Asp – Is essential for the catalysis of the methyl transfer in plant cation-dependent O-methyltransferases". In: *Phytochemistry* 113 (May 2015), pp. 130–139.
- [14] Jing Chen et al. "Different Effects of L -Arginine on Protein Refolding : Suppressing Aggregates of Hydrophobic Interaction, Not Covalent Binding". In: (2008), pp. 1365–1372.
- [15] Vincent B Chen et al. "MolProbity: all-atom structure validation for macromolecular crystallography." In: *Acta crystallographica. Section D, Biological crystallography* 66.Pt 1 (Jan. 2010), pp. 12–21.
- [16] Al Claiborne and Irwin Fridovich. "Chemical and Enzymatic Intermediates in the Peroxidation of o-Dianisidine by Horseradish Peroxidase. 1. Spectral Properties of the Products of Dianisidine Oxidation". In: *Biochemistry* 18 (1979), pp. 2324–2329.
- [17] Yuntao Dai et al. "Natural deep eutectic solvents as new potential media for green technology." In: *Analytica chimica acta* 766 (Mar. 2013), pp. 61–8.
- [18] Christian Dalhoff et al. "Direct transfer of extended groups from synthetic cofactors by DNA methyltransferases." In: *Nature chemical biology* 2.1 (2006), pp. 31–32.
- [19] Christian Dalhoff et al. "Synthesis of S-adenosyl-L-methionine analogs and their use for sequence-specific transalkylation of DNA by methyltransferases." In: *Nature protocols* 1.4 (2006), pp. 1879–1886.
- [20] Daresbury Laboratory. "No Title". In: *Newsletter on protein crystallography* 33 (1997), pp. 25–30.
- [21] Martin Dippe et al. "Engineering of a Mg<sup>2+</sup>-dependent O-methyltransferase towards novel regiospecificity". In: *manuscript submitted* (2015).
- [22] Martin Dippe et al. "Rationally engineered variants of S-adenosylmethionine (SAM) synthase: reduced product inhibition and synthesis of artificial co-factor homologues." en. In: *Chemical communications (Cambridge, England)* 51.17 (Feb. 2015), pp. 3637–40.
- [23] J D Dunitz. "The entropic cost of bound water in crystals and biomolecules." In: *Science (New York, N.Y.)* 264.5159 (Apr. 1994), p. 670.
- [24] Jack D. Dunitz. "Win some, lose some: enthalpy-entropy compensation in weak intermolecular interactions". In: *Chemistry & Biology* 2.11 (Nov. 1995), pp. 709–712.
- [25] Elsevier. *Reaxys, version 2.19790.2*.

- [26] P. Emsley et al. “Features and development of Coot”. In: *Acta Crystallographica Section D: Biological Crystallography* 66.4 (2010), pp. 486–501.
- [27] Carola Engler, Romy Kandzia, and Sylvestre Marillonnet. “A one pot, one step, precision cloning method with high throughput capability”. In: *PLoS ONE* 3.11 (Jan. 2008), e3647.
- [28] Philip Evans. “Scaling and assessment of data quality”. In: *Acta Crystallographica Section D: Biological Crystallography* 62.1 (Jan. 2006), pp. 72–82. arXiv: S0907444905036693 [doi: 10.1107].
- [29] Nicolas Fabre et al. “Determination of flavone, flavonol, and flavanone aglycones by negative ion liquid chromatography electrospray ion trap mass spectrometry”. In: *Journal of the American Society for Mass Spectrometry* 12.6 (2001), pp. 707–715.
- [30] Matthew W Freyer and Edwin a Lewis. “Isothermal titration calorimetry: experimental design, data analysis, and probing macromolecule/ligand binding and kinetic interactions.” In: *Methods in cell biology* 84.07 (Jan. 2008), pp. 79–113.
- [31] Akiko Fujimoto, Atsushi Hirano, and Kentaro Shiraki. “Ternary system of solution additives with Arginine and salt for refolding of beta-galactosidase”. In: *Protein Journal* 29.3 (2010), pp. 161–166.
- [32] E. Gasteiger et al. “Protein Identification and Analysis Tools on the ExPASy Server”. In: *The Proteomics Protocols Handbook*. Ed. by John M. Walker. Humana Press, 2005, pp. 571–607.
- [33] Christine L Gee et al. “Disulfide-linked dimers of human adrenaline synthesizing enzyme PNMT are catalytically active.” In: *Biochimica et biophysica acta* 1750.1 (June 2005), pp. 82–92.
- [34] S C Gill and P H von Hippel. “Calculation of protein extinction coefficients from amino acid sequence data.” In: *Analytical biochemistry* 182.2 (Nov. 1989), pp. 319–26.
- [35] Paola Gilli et al. “Enthalpy-entropy compensation in drug-receptor binding”. In: *The Journal of Physical Chemistry* 98.5 (Feb. 1994), pp. 1515–1518.
- [36] Erich Grotewold, ed. *The Science of Flavonoids*. 1st ed. New York: Springer, 2006.
- [37] Amandine Guelorget et al. “Insights into the hyperthermostability and unusual region-specificity of archaeal Pyrococcus abyssi tRNA m1A57/58 methyltransferase”. In: *Nucleic Acids Research* 38.18 (2010), pp. 6206–6218.

- [38] Albert Hofmann. *Die Mutterkornalkaloide*. Solothurn: Nachtschatten Verlag, 2000.
- [39] Ze Lin Huang et al. “Deep eutectic solvents can be viable enzyme activators and stabilizers”. In: *Journal of Chemical Technology and Biotechnology* October 2013 (2014).
- [40] Ruth Huey et al. “A semiempirical free energy force field with charge-based desolvation.” In: *Journal of computational chemistry* 28.6 (Apr. 2007), pp. 1145–52.
- [41] Mwafaq Ibdah et al. “A Novel Mg<sup>2+</sup>-dependent O-Methyltransferase in the Phenylpropanoid Metabolism of *Mesembryanthemum crystallinum*”. In: *Journal of Biological Chemistry* 278.45 (Nov. 2003), pp. 43961–43972.
- [42] Kabirul Islam et al. “Expanding cofactor repertoire of protein lysine methyltransferase for substrate labeling”. In: *ACS Chemical Biology* 6.7 (2011), pp. 679–684.
- [43] Julien Jorda and Todd O. Yeates. “Widespread disulfide bonding in proteins from thermophilic archaea”. In: *Archaea* 2011 (2011).
- [44] PD Josephy, T Eling, and RP Mason. “The horseradish peroxidase-catalyzed oxidation of 3, 5, 3’, 5’-tetramethylbenzidine. Free radical and charge-transfer complex intermediates.” In: *Journal of Biological Chemistry* 257 (1982), pp. 3669–3675.
- [45] Wolfgang Kabsch. “Automatic processing of rotation diffraction data from crystals of initially unknown symmetry land cell constants”. In: *Journal of Applied Crystallography* 26.pt 6 (Dec. 1993), pp. 795–800.
- [46] Wolfgang Kabsch. “Integration, scaling, space-group assignment and post-refinement”. In: *Acta Crystallographica Section D: Biological Crystallography* 66.2 (Feb. 2010), pp. 133–144.
- [47] Wolfgang Kabsch. “Xds”. In: *Acta Crystallographica Section D: Biological Crystallography* 66.2 (Feb. 2010), pp. 125–132.
- [48] Sun-Young Kang et al. “Biosynthesis of methylated resveratrol analogs through the construction of an artificial biosynthetic pathway in *E. coli*.” In: *BMC biotechnology* 14.1 (2014), p. 67.
- [49] Lawrence A Kelley et al. “The Phyre2 web portal for protein modeling, prediction and analysis.” In: *Nature protocols* 10.6 (June 2015), pp. 845–858.
- [50] Bong Gyu Kim et al. “Multiple regiospecific methylations of a flavonoid by plant O-methyltransferases expressed in *E. coli*.” In: *Biotechnology letters* 27.23-24 (Dec. 2005), pp. 1861–4.

- [51] Dae Hwan Kim et al. “Regiospecific methylation of naringenin to ponciretin by soybean O-methyltransferase expressed in *Escherichia coli*.” In: *Journal of biotechnology* 119.2 (Sept. 2005), pp. 155–62.
- [52] K M Koeller and C H Wong. “Enzymes for chemical synthesis.” In: *Nature* 409.6817 (Jan. 2001), pp. 232–40.
- [53] Youichi Kondou et al. “cDNA Libraries”. In: *Methods in Molecular Biology* 729 (2011), pp. 183–197.
- [54] Kristin König. “Engineering of the Anthocyanin Biosynthetic Pathway in *Nicotiana benthamiana*”. Master. Martin-Luther-Universität Halle-Wittenberg, 2014.
- [55] Jakub G. Kopycki et al. “Biochemical and Structural Analysis of Substrate Promiscuity in Plant Mg<sup>2+</sup>-Dependent O-Methyltransferases”. In: *Journal of Molecular Biology* 378.1 (Apr. 2008), pp. 154–164.
- [56] Ulrich K Laemmli. “Cleavage of structural proteins during the assembly of the head of bacteriophage T4.” In: *Nature* 227.5259 (1970), pp. 680–685.
- [57] Brian J. C. Law et al. “Site-specific bioalkylation of rapamycin by the RapM 16-O-methyltransferase”. In: *Chem. Sci.* (2015), pp. 2885–2892.
- [58] Bobby W K Lee et al. “Enzyme-catalyzed transfer of a ketone group from an S-adenosylmethionine analogue: A tool for the functional analysis of methyltransferases”. In: *Journal of the American Chemical Society* 132.11 (2010), pp. 3642–3643.
- [59] Jong Suk Lee et al. “Identification of flavonoids using liquid chromatography with electrospray ionization and ion trap tandem mass spectrometry with an MS/MS library”. In: *Rapid Communications in Mass Spectrometry* 19.23 (2005), pp. 3539–3548.
- [60] Jakob P. Ley et al. “Evaluation of bitter masking flavanones from Herba Santa (*Eriodictyon californicum* (H. & A.) Torr., Hydrophyllaceae)”. In: *Journal of Agricultural and Food Chemistry* 53.15 (2005), pp. 6061–6066.
- [61] Liwei Li et al. “PDBcal: a comprehensive dataset for receptor-ligand interactions with three-dimensional structures and binding thermodynamics from isothermal titration calorimetry.” In: *Chemical biology & drug design* 71.6 (June 2008), pp. 529–32.
- [62] Caroline Louis-Jeune, Miguel A Andrade-Navarro, and Carol Perez-Iratxeta. “Prediction of protein secondary structure from circular dichroism using theoretically derived spectra.” In: *Proteins* 80.2 (Feb. 2012), pp. 374–81.

- [63] Gražvydas Lukinavičius et al. “Enhanced chemical stability of AdoMet analogues for improved methyltransferase-directed labeling of DNA”. In: *ACS Chemical Biology* 8.6 (2013), pp. 1134–1139.
- [64] Tom J. Mabry, K. R. Markham, and M. B. Thomas. *The Systematic Identification of Flavonoids*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1970.
- [65] Savvas C Makrides and Savvas C Makrides. “Strategies for Achieving High-Level Expression of Genes in *Escherichia coli*”. In: *Microbiological reviews* 60.3 (1996), pp. 512–538.
- [66] S. V. Mani, D. W. Connell, and R. D. Braddock. “Structure activity relationships for the prediction of biodegradability of environmental pollutants”. en. In: *Critical Reviews in Environmental Control* 21.3-4 (Jan. 1991), pp. 217–236.
- [67] Alexandre Martinière et al. “Development and properties of genetically encoded pH sensors in plants.” In: *Frontiers in plant science* 4.December (2013), p. 523.
- [68] Airlie J. McCoy. “Solving structures of protein complexes by molecular replacement with Phaser”. In: *Acta Crystallographica Section D: Biological Crystallography* 63.1 (Jan. 2006), pp. 32–41.
- [69] Airlie J. McCoy et al. “Phaser crystallographic software”. In: *Journal of Applied Crystallography* 40.4 (Aug. 2007), pp. 658–674.
- [70] Edoardo Mentasti and Ezio Pelizzetti. “Reactions between iron(III) and catechol (o-dihydroxybenzene). part I. Equilibria and kinetics of complex formation in aqueous acid solution”. en. In: *Journal of the Chemical Society, Dalton Transactions* 23 (Jan. 1973), p. 2605.
- [71] Douglas C. Montgomery. *Design and Analysis of Experiments*. 5th ed. New York: John Wiley & Sons, 2001.
- [72] Garrett M Morris et al. “AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility.” In: *Journal of computational chemistry* 30.16 (Dec. 2009), pp. 2785–91.
- [73] Yuri Motorin et al. “Expanding the chemical scope of RNA:methyltransferases to site-specific alkynylation of RNA for click labeling”. In: *Nucleic Acids Research* 39.5 (2011), pp. 1943–1952.
- [74] Garib N. Murshudov, Alexei a. Vagin, and Eleanor J. Dodson. “Refinement of macromolecular structures by the maximum-likelihood method”. In: *Acta Crystallographica Section D: Biological Crystallography* 53.3 (May 1997), pp. 240–255.

- [75] Janet Newman. "Novel buffer systems for macromolecular crystallization". In: *Acta Crystallographica Section D: Biological Crystallography* 60.3 (2004), pp. 610–612.
- [76] Novagen. *pET System Manual*. 11th ed. Darmstadt: EMD Chemicals, 2010.
- [77] Ira Palmer and Paul T. Wingfield. "Preparation and extraction of insoluble (Inclusion-body) proteins from Escherichia coli". In: *Current Protocols in Protein Science* 1.SUPPL.70 (Nov. 2012), Unit6.3.
- [78] Wibke Peters et al. "Enzymatic site-specific functionalization of protein methyltransferase substrates with alkynes for click labeling". In: *Angewandte Chemie - International Edition* 49.30 (2010), pp. 5170–5173.
- [79] Harold R. Powell. "The Rossmann Fourier autoindexing algorithm in MOSFLM". In: *Acta Crystallographica Section D: Biological Crystallography* 55.10 (1999), pp. 1690–1695.
- [80] R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2015.
- [81] M. G. Rossmann and D. M. Blow. "The detection of sub-units within the crystallographic asymmetric unit". In: *Acta Crystallographica* 15.1 (Jan. 1962), pp. 24–31.
- [82] Michael G. Rossmann. "Molecular replacement - Historical background". In: *Acta Crystallographica - Section D Biological Crystallography* 57.10 (Sept. 2001), pp. 1360–1366.
- [83] Rainer Rudolph and Hauke Lilie. "In vitro folding of inclusion body proteins". In: *FASEB Journal* 10 (1996), pp. 49–56.
- [84] Bernhard Rupp. *Biomolecular Crystallography: Principles, Practice, and Application to Structural Biology*. 1st ed. New York: Garland Science, 2009, p. 800.
- [85] J Sambrook and D W Russell. *Molecular Cloning: A Laboratory Manual*. 3rd ed. Cold Spring Harbor (NY, USA): Cold Spring Harbor Laboratory Press, 2001.
- [86] Cleydson Breno R Santos et al. "A SAR and QSAR study of new artemisinin compounds with antimalarial activity". In: *Molecules* 19.1 (2014), pp. 367–399.
- [87] Gudrun Schröder et al. "Flavonoid methylation: A novel 4'-O-methyltransferase from Catharanthus roseus, and evidence that partially methylated flavanones are substrates of four different flavonoid dioxygenases". In: *Phytochemistry* 65.8 (2004), pp. 1085–1094.

- [88] Daniela Schulz, Josephin Marie Holstein, and Andrea Rentmeister. “A chemo-enzymatic approach for site-specific modification of the RNA cap”. In: *Angewandte Chemie - International Edition* 52.30 (2013), pp. 7874–7878.
- [89] N. Schweigert, a. J B Zehnder, and R. I L Eggen. “Chemical properties of catechols and their molecular modes of toxic action in cells, from microorganisms to mammals”. In: *Environmental Microbiology* 3.2 (2001), pp. 81–91.
- [90] Mark S. Searle, Martin S. Westwell, and Dudley H. Williams. “Application of a generalised enthalpy?entropy relationship to binding co-operativity and weak associations in solution”. en. In: *Journal of the Chemical Society, Perkin Transactions 2* 1 (Jan. 1995), p. 141.
- [91] Stanley K. Shapiro and Dimis J. Ehninger. “Methods for the analysis and preparation of adenosylmethionine and adenosylhomocysteine”. In: *Analytical Biochemistry* 15.2 (May 1966), pp. 323–333.
- [92] Alexander Shulgin and Ann Shulgin. *TiHKAL - The Continuation*. Ed. by Dan Joy. 1st ed. Berkeley: Transform Press, 1997.
- [93] Sigma-Aldrich. *Technical Bulletin no. 2003-03: freezing of microbial samples prior to testing*. Parenteral Drug Association. 2003.
- [94] Shanteri Singh et al. “Facile chemoenzymatic strategies for the synthesis and utilization of S-adenosyl-L-methionine analogues”. In: *Angewandte Chemie - International Edition* 53.15 (2014), pp. 3965–3969.
- [95] Lekha Sleno and Dietrich a. Volmer. “Ion activation methods for tandem mass spectrometry”. In: *Journal of Mass Spectrometry* 39.10 (2004), pp. 1091–1112.
- [96] Harald Stecher et al. “Biocatalytic Friedel-Crafts alkylation using non-natural cofactors.” In: *Angewandte Chemie (International ed. in English)* 48.50 (Jan. 2009), pp. 9546–8.
- [97] Anna-Winona Struck et al. “S-Adenosyl-Methionine-Dependent Methyltransferases: Highly Versatile Enzymes in Biocatalysis, Biosynthesis and Other Biotechnological Applications.” In: *Chembiochem : a European journal of chemical biology* (Nov. 2012), pp. 1–15.
- [98] F William Studier. “Protein production by auto-induction in high density shaking cultures.” In: *Protein expression and purification* 41.1 (May 2005), pp. 207–234. arXiv: *NIHMS150003*.

- [99] Martin Tengg et al. “Molecular characterization of the C-methyltransferase NovO of *Streptomyces sphaeroides*, a valuable enzyme for performing Friedel–Crafts alkylation”. In: *Journal of Molecular Catalysis B: Enzymatic* 84 (Dec. 2012), pp. 2–8.
- [100] Marie Thomsen et al. “Chemoenzymatic synthesis and in situ application of S-adenosyl-L-methionine analogs.” In: *Organic & biomolecular chemistry* 11.43 (2013), pp. 7606–10.
- [101] Robert Tibshirani. *Regression Selection and Shrinkage via the Lasso*. 1994.
- [102] Dominique A Tobbell et al. “Identification of in Vitro Folding Conditions for Procathepsin S and Cathepsin S Using Fractional Factorial Screens”. In: *Protein Expression and Purification* 24.2 (Mar. 2002), pp. 242–254.
- [103] Oleg Trott and Arthur J Olson. “AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading.” In: *Journal of computational chemistry* 31.2 (Jan. 2010), pp. 455–61.
- [104] Kouhei Tsumoto et al. “Role of arginine in protein refolding, solubilization, and purification.” In: *Biotechnology progress* 20.5 (2004), pp. 1301–1308.
- [105] Alexei a. Vagin et al. “REFMAC5 dictionary: Organization of prior chemical knowledge and guidelines for its use”. en. In: *Acta Crystallographica Section D: Biological Crystallography* 60.12 I (Nov. 2004), pp. 2184–2195.
- [106] Renaud Vincentelli et al. “High-throughput automated refolding screening of inclusion bodies.” In: *Protein science : a publication of the Protein Society* 13.10 (2004), pp. 2782–2792.
- [107] Thomas Vogt. “Regiospecificity and kinetic properties of a plant natural product O-methyltransferase are determined by its N-terminal domain”. In: *FEBS Letters* 561.1-3 (Mar. 2004), pp. 159–162.
- [108] Thomas Vogt. “Regiospecificity and kinetic properties of a plant natural product O-methyltransferase are determined by its N-terminal domain.” In: *FEBS letters* 561.1-3 (Mar. 2004), pp. 159–62.
- [109] Rui Wang et al. “Formulating a fluorogenic assay to evaluate S-adenosyl-L-methionine analogues as protein methyltransferase cofactors”. In: *Molecular BioSystems* 7.11 (2011), p. 2970.
- [110] Rui Wang et al. “Labeling substrates of protein arginine methyltransferase with engineered enzymes and matched S-adenosyl-l-methionine analogues”. In: *Journal of the American Chemical Society* 133.20 (2011), pp. 7648–7651.

- [111] Melissa Swope Willis et al. “Investigation of protein refolding using a fractional factorial screen: a study of reagent effects and interactions.” In: *Protein Science* 14.7 (2005), pp. 1818–1826.
- [112] Sophie Willnow et al. “A Selenium-Based Click AdoMet Analogue for Versatile Substrate Labeling with Wild-Type Protein Methyltransferases”. In: *ChemBioChem* 13.8 (2012), pp. 1167–1173.
- [113] Martyn D. Winn et al. “Overview of the CCP4 suite and current developments”. In: *Acta Crystallographica Section D: Biological Crystallography* 67.4 (Apr. 2011), pp. 235–242.
- [114] Jaclyn M. Winter et al. “Expanding the structural diversity of polyketides by exploring the cofactor tolerance of an inline methyltransferase domain”. In: *Organic Letters* 15.14 (2013), pp. 3774–3777.
- [115] Hiroshi Yamaguchi and Masaya Miyazaki. “Refolding techniques for recovering biologically active recombinant proteins from inclusion bodies.” In: *Biomolecules* 4.1 (2014), pp. 235–51.
- [116] Yoon Youngdae et al. “Characterization of an O-methyltransferase from *Streptomyces avermitilis* MA-4680”. In: *Journal of Microbiology and Biotechnology* 20.9 (2010), pp. 1359–1366.
- [117] Changsheng Zhang et al. “Natural product diversification using a non-natural cofactor analogue of S-adenosyl-L-methionine”. In: *Journal of the American Chemical Society* 128.9 (2006), pp. 2760–2761.
- [118] Jian Min Zhou et al. “Sequential O-methylation of tricetin by a single gene product in wheat”. In: *Biochimica et Biophysica Acta - General Subjects* 1760.7 (2006), pp. 1115–1124.
- [119] C Zubieta et al. “Structures of two natural product methyltransferases reveal the basis for substrate specificity in plant O-methyltransferases.” In: *Nature structural biology* 8.3 (Mar. 2001), pp. 271–9.

## **Acronyms**

- Å** Ångström, 0.1 nm
- 3O4M** 3'-hydroxy-4'-methoxy 54, 75, 77, 80, 100
- 4O3M** 4'-hydroxy-3'-methoxy 54, 75, 77, 80
- ABPP** activity based protein profiling 42
- AC-9** anthracene-9-carboxylic acid 26
- ANOVA** Analysis of Variance xv, 67–69
- ATP** adenosine triphosphate 26
- AUC** area under the curve 34
- BisTris** 2-[Bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)propane-1,3-diol
- B-PER** bacterial protein extraction reagent
- CCP4** Collaborative Computational Project No. 4 29
- CD** circular dichroism 13, 70
- CHS** chalcone synthase 56
- CID** collision induced dissociation 59, 60
- C-MT** C-methyl transferase 41
- COMT** catechol O-methyl transferase 18
- Coot** Crystallographic Object-Oriented Toolkit 29
- CV** column volumes
- DMSO** dimethyl sulfoxide 28
- DNA** desoxyribonucleic acid
- DNA MT** DNA methyl transferase xii, 40, 94
- DoE** design of experiments 24, 54, 70
- DTT** dithiothreitol; (2*S*,3*S*)-1,4-bis(sulfanyl)butane-2,3-diol
- EDTA** ethylenediaminetetraacetic acid 20, 22–24, 33, 72
- EEC** enthalpy-entropy compensation 47
- ESI** electrospray ionization 59
- FPLC** fast protein liquid chromatography 23, 36
- FrFD** fractional factorial design xv, 24, 54, 66, 70

**FT** Fourier transformation 29

**GdmCl** guanidinium hydrochloride

**GFP** green fluorescent protein 30

**GOD** glucose oxidase 31, 127

**GSH** glutathione,  $\gamma$ -L-glutamyl-L-cysteinylglycine 24, 31, 33

**GSSG** glutathione disulfide 24

**GST** Glutathion S-transferase 59

**HCD** higher-energy collisional dissociation xii, 84–87

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

**HIC** hydrophobic interaction chromatography 35

**HPLC** high-performance liquid chromatography xii, 17, 26, 33, 34, 37, 56, 58, 96

**HRP** horseradish peroxidase 31

**IAE-plot** interaction effects plot 77

**IB** inclusion body 22, 23, 25, 65, 66

**IEX** ion exchange chromatography 27

**IMAC** immobilized metal affinity chromatography

**IPB** Leibniz-Institute of Plant Biochemistry

**IPTG** isopropyl-D-thiogalactopyranosid 18, 21, 22, 65

**ITC** Isothermal Titration Calorimetry x, xv, 36, 44, 46, 47, 51, 52, 127

**LB** lysogeny broth 16, 18, 21, 22, 65

**LC/MS** liquid chromatography coupled mass-spectrometry xii, 59, 60, 87, 96

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

**ME-plot** main effects plot 67, 69, 77

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

**MLU** Martin-Luther-Universität

**MMT** L-malic acid/MES/Tris 9, 36

**MR** molecular replacement

**MS/MS** tandem mass-spectrometry 78

**MT** methyl transferase xii, xvi, 40–42, 54, 68, 94, 102

**MTP** micro-titer plate 27, 28, 30, 31, 127

- MW** molecular weight 21
- MWCO** molecular weight cut-off
- NADES** natural deep eutectic solvent xv, 10, 28, 124
- nos** nopaline synthase 55
- NPS** nitrogen, phosphate, sulfate buffer
- NRPS** non-ribosomal peptide synthase 40
- NTA** nitrilo triacetic acid 23
- O-MT** O-methyl transferase 31, 34, 37, 41, 42, 49, 63, 77, 78, 80, 81
- PAGE** polyacrylamide gel electrophoresis xi, xiii, 18–20, 25, 26, 36, 59, 65–67, 98
- PAL** phenylalanine ammonia-lyase 56
- PBS** phosphate buffered saline 19, 25, 30
- PCA** principal component analysis xi, xii, 56, 58, 97
- PCH** propane-1,2-diol/choline chloride,natural deep eutectic solvent (NADES)-mixture 28
- PCR** polymerase chain reaction 14, 15
- PDA** photo diode array 37
- PDB** Protein Data Base 29, 30
- PFOMT** phenylpropanoid and flavonoid O-methyl transferase ix–xii, xvi, 18, 21, 22, 27–31, 36, 39, 42–52, 54, 71, 72, 74–78, 80, 81, 96, 100, 127
- PHENIX** Phyton-based Hierachial Environment for Integrated Xtallography 29
- PKS** poly ketide synthase 40
- PMSF** phenylmethylsulfonylfluoride
- P-MT** protein methyl transferase xii, 40, 41, 94
- QSAR** quantitative structure activity relationship 39
- rmsd** root mean squared deviation x, 46, 51, 106
- RNA** ribonucleic acid xii, 94
- RT** room temperature
- SAE** S-adenosyl-L-ethionine, (2*S*)-2-amino-4-[[*(2S,3S,4R,5R)*-5-(6-aminopurin-9-yl)-3,4-dihydroxyoxolan-2-yl]methyl-ethylsulfonio]butanoat x, 26, 28, 39, 40, 42, 44, 47, 48, 50–52

- SAH** *S*-adenosyl-L-homocysteine ix, x, xii, 24, 33, 34, 39, 42–48, 50–52, 66, 67, 95  
**SAM** *S*-adenosyl-L-methionine ix, x, xii, xvi, 26, 27, 31, 33, 34, 39–42, 44, 46–49, 51, 54, 94, 102  
**SAMS** *S*-adenosylmethionine synthase 26  
**SAR** structure activity relationship 39  
**SDS** sodium dodecylsulfate xi, xiii, 10, 18–20, 25, 26, 36, 59, 65–67, 98  
**SeAM** *Se*-adenosyl selenomethionine xii, 41, 94, 102  
**SOMT-2** soy O-methyl transferase xi, xiii, xv, 18, 22, 24–26, 35, 54–57, 59–70, 80, 81, 98, 99  
**SSG** succinate/sodium phosphate/glycine 9, 10  
**TB** terrific broth 18, 65  
**TCA** trichloro acetic acid 19, 20, 25, 33, 36  
**Ti-plasmid** tumor inducing plasmid 12, 127  
**Tris** tris(hydroxymethyl)-aminomethane  
**U** enzyme unit; measure for enzymatic activity (1 U = 1 µmole/min = 1/60 µkat)  
**UV/VIS** ultra violet/visible (light spectrum) 27, 37  
**V** volume  
**ZYP** N-Z-amino, yeast extract, phosphate 21, 59, 127

# Glossary

**GOD** Glucose oxidase is an enzyme.... 123

**His<sub>6</sub>-tag** Hexa-histidine tag commonly used for recombinant protein production. 65

**Isothermal Titration Calorimetry (ITC)** Fill in description here 123

**MTP** Micro-titer plate. Small format rectangular plastic plate containing wells to allow for storage of multiple small samples or the containment multiple simultaneous reactions. Typical sizes include 24, 96 and 384-wells 123

**PFOMT** Phenylpropanoid and flavonoid O-methyl transferase from *Mesembryanthemum crystallinum*, which was first described by Ibdah et al. in 2003 [41] 124

**T7-tag** Initial 11 amino acids of the T7 gene 10 protein. 63

**Ti-plasmid** Commonly found plasmids in *A. tumefaciens* and *A. rhizogenes* that confer virulence 125

**Trx-tag** Thioredoxin tag used to increase solubility and stability of recombinantly expressed proteins. 65

**ZYP-5052** Autoinduction medium developed by Studier [98]. The naming stems from the components N-Z-amine, yeast extract and phosphate. The numbering designates the composition; e.g. 5052 refers to 0.5 % glycerol, 0.05 % glucose and 0.2 % lactose. 125