Engineering of Methyltransferases: Exploring Their Catalytic Promiscuity

Inauguraldissertation

zur

Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Universität Greifswald

vorgelegt von

Qingyun Tang

Greifswald, Dezember 2020

Dekan: Prof. Dr. Gerald Kerth

1. Gutachter: Prof. Dr. Uwe T. Bornscheuer

2. Gutachter: Prof. Dr. Stephan Hammer

Tag der Promotion: 26. Februar 2021

Table of Contents

List of abbreviations

µM	micromolar
Å	Angstrom
ADK	adenosine kinase
Allyl-I	allyl iodide
AMP	adenosine monophosphate
*At*HMT	*Arabidopsis thaliana* halide methyltransferase
ATP	adenosine triphosphate
BuI	butyl iodide
CaOMT	Caffeic acid/5-hydroxyferulic acid 3/5-*O*-methyltransferase
CH_3Cl	chloromethane
*Ci*VCPO	*Curvularia inaequalis* vanadium-dependent chloroperoxidase
CMT	*C*-methyltransferase
COMT	catechol *O*-methyltransferase
*Ct*HMT	*Chloracidobacterium thermophilum* halide methyltransferase
cxSAM	carboxy-*S*-adenosyl-L-methionine
E. coli	*Escherichia coli*
EAE	experimental autoimmune encephalomyelitis
et al.	*et alia*
EtI	ethyl iodide
FOMT	flavonoid *O*-methyltransferase
h	hour
Hcys	L-homocysteine
HMT	halide methyltransferase
HOI	hypoiodous acid
HPLC	high-performance liquid chromatography
IeOMT	isoeugenol 4-*O*-methyltransferase
IOMT	isoflavonoid *O*-methyltransferase
LC-MS	liquid chromatography-mass spectrometry
MAT	methionine adenosyltransferase
MB-COMT	membrane-bound catechol *O*-methyltransferase

MeI	methyl iodide
meta-OH/3-OH	*meta*-hydroxyl group
mg	milligram
min	minute
mM	millimolar
MT	methyltransferase
NCBI	National Center for Biotechnology Information
nm	nanometre
NMR	nuclear magnetic resonance
NMT	*N*-methyltransferase
OMT	*O*-methyltransferase
para-OH/4-OH	*para*-hydroxyl group
PDB	Protein Data Bank
PolyP	polyphosphate
POMT	phenylpropanoid *O*-methyltransferase
PPK2	polyphosphate kinase 2
PrI	propyl iodide
RMSD	root-mean-square deviation
*Rs*HMT	*Raphanus sativus* halide methyltransferase
SAA	*S*-adenosyl-L-allylionine
SAE	*S*-adenosyl-L-ethionine
SAH	*S*-adenosyl-L-homocysteine
SAHH	*S*-adenosyl-L-homocysteine hydrolase
SAM	*S*-adenosyl-L-methionine
SAP	*S*-adenosyl-L-propionine
S-COMT	soluble catechol *O*-methyltransferase
TMB	3,3′,5,5′-tetramethylbenzidine

Additionally, the one and three letter codes for proteogenic amino acids were used.

Scope and outline

S-adenosyl-L-methionine- (SAM) dependent methyltransferases (MTs) catalyse methylation of halide ions and the *C, O, N, S, Se*, and *As* atoms of biomolecules ranging from biopolymers to small molecules. They display different chemo-, regio- and stereoselectivity according to their specific functions. This thesis focuses on the engineering of *O*-methyltransferases (OMTs) and halide methyltransferases (HMTs) through rational design and directed evolution to study their structure-function relationship and to explore their catalytic promiscuity. The influence of substrate binding residues on the substrate scope and regioselectivity of a plant OMT against various phenolic substrates (**Article I**) and flavonoids (**Article II**) has been investigated. **Article III** describes the directed evolution of an HMT for the biocatalytic synthesis of diverse SAM analogues. With the evolved HMT, regioselective alkylation of phenolic compounds and flavonoids, as well as the SAM analogue regeneration, were achieved through an HMT-MT cascade reaction.

Article I   **Specific residues expand the substrate scope and enhance the regioselectivity of a plant *O*-methyltransferase**

> Q. Tang, U. T. Bornscheuer, I. V. Pavlidis, *ChemCatChem* **2019**, *11*, 3227.

It was reported in literature that an isoeugenol 4-OMT (IeOMT) can be engineered to a caffeic acid 3-OMT (CaOMT) by replacing three consecutive residues. In this article, we investigated the effect of these residues on substrate preference and regioselectivity of IeOMT. The triple mutant T133M/A134N/T135Q and the respective single mutants were constructed and tested against a series of phenolic compounds. The variant T133M had a universal effect to improve enzymatic activities against all tested substrates while the mutant A134N had enhanced regioselectivity. The triple mutant T133M/A134N/T135Q benefits from these two mutations, which not only expanded the substrate scope, but also enhanced the regioselectivity of IeOMT. On the basis of this work, regiospecific methylated phenolics can be produced in high purity by different IeOMT variants.

Article II   **Influence of substrate binding residues on the substrate scope and regioselectivity of a plant *O*-methyltransferase against flavonoids**

> Q. Tang, Y. M. Vianney, K. Weisz, C. W. Grathwol, A. Link, U. T. Bornscheuer, I. V. Pavlidis, *ChemCatChem* **2020**, *12*, 3721.

Flavonoid OMTs (FOMTs), isoflavonoid OMTs (IOMTs) and phenylpropanoid OMTs (POMTs) display different substrate preferences. Sequence comparison showed that the substrate binding residues at positions 322 and 326 are different between these OMT groups and might be critical for the substrate discrimination. Residues at positions 322 and 326 in IeOMT (a POMT) were mutated to the commonly presented residues in FOMT and IOMT. The introduced mutants, in cooperation with the variant T133M, have improved or brought novel activities and regioselectivity against the tested flavonoids eriodictyol, naringenin, luteolin, quercetin, and also the isoflavonoid genistein compared to the wild-type IeOMT. On the basis of this work, methylated flavonoids that are rare in nature were produced in high purity.

Article III   Directed evolution of a halide methyltransferase enables biocatalytic synthesis of diverse SAM analogs

Q. Tang, C. W. Grathwol, A. S. Aslan-Üzel, S. Wu, A. Link, I. V. Pavlidis, C. P. S. Badenhorst, U. T. Bornscheuer, *Angew. Chem., Int. Ed.* **2021**, *60*, 1524; *Angew. Chem.* **2021**, *133*, 1547.

Biocatalytic alkylations to obtain chemo-, regio- and stereoselectively alkylated compounds can be achieved by MTs with the supply of SAM analogues. It was recently discovered that SAM can be directly synthesized from *S*-adenosyl-ʟ-homocysteine (SAH) and methyl iodide, catalysed by an HMT. To explore the promiscuity of HMT in the synthesis of SAM analogues, we performed directed evolution of the *Arabidopsis thaliana* HMT based on a sensitive, colorimetric iodide assay. The identified variant V140T displayed activities against ethyl-, propyl-, and allyl iodides to produce the corresponding SAM analogues. With this HMT variant, regioselective ethylation of luteolin and allylation of 3,4-dihydroxybenzaldehyde, as well as the SAM analogue regeneration, were achieved through this HMT-MT one-pot cascade reaction.

1. Introduction

1.1. The effects of methylation

Methylation is an important and ubiquitous process both in biological activities and chemical synthesis. The physical properties of a molecule such as the hydrophobicity, solubility and pK value can be modified; the odour and taste, and the bioactivity can be changed significantly. For example, vanillin (3-methoxy-4-hydroxybenzaldehyde) and veratraldehyde (3,4-dimethoxybenzaldehyde) represent vanilla aroma and woody fragrance, respectively (Figure 1a).[1-2] Methylation on specific carbon atoms of the drug molecule against the experimental autoimmune encephalomyelitis (EAE) greatly increase its binding potency and improve the IC$_{50}$ value more than 2000-fold (Figure 1b).[3-5] Various biological activities and diseases are regulated through the methylation of nucleotides, proteins and metabolites.[6-9] Methylation of DNA, RNA and proteins are involved in regulation of gene expression, DNA labelling, RNA stability, also post-translational modification for proteins (Figure 1c and 1d).[10-12] Metabolism of neurotransmitters and hormones are tightly connected to methylation (Figure 1e).[13-15] Other small molecules or natural products such as signalling and host defence compounds, cell wall components and cell membrane, pigments, prosthetic groups and cofactors are also methylated.[16-19]

1.2. SAM-dependent methyltransferases

Methylation can be achieved by a specific group of enzymes called methyltransferases (MTs, E.C. 2.1.1.x), using the biomolecule S-adenosyl-L-methionine (SAM) as methyl donor. The methyl group at the sulfonium of SAM is transferred to the substrate, producing the methylated product and the by-product S-adenosyl-L-homocysteine (SAH) (Figure 2). Enzymatic methylation processes are environmentally friendly alternatives to chemical methylation methods and are superior due to their chemo-, regio- and stereoselectivity. Depending on the atoms to be methylated, MTs can be classified into C-, O-, N-, S-, Se-, As- and also halide MTs, among which O-methyltransferases (OMTs) are the most abundant class (54% of their EC subclass).[20]

a)

b)

Vanillin   Veratraldehyde   NIBR-0213

c)

d)

5-Methylcytosine   6-Methyladenine N_ε-Trimethyllysine N_ω-Dimethylarginine

e)

3-Methoxytyramine   Melatonin   Adrenaline
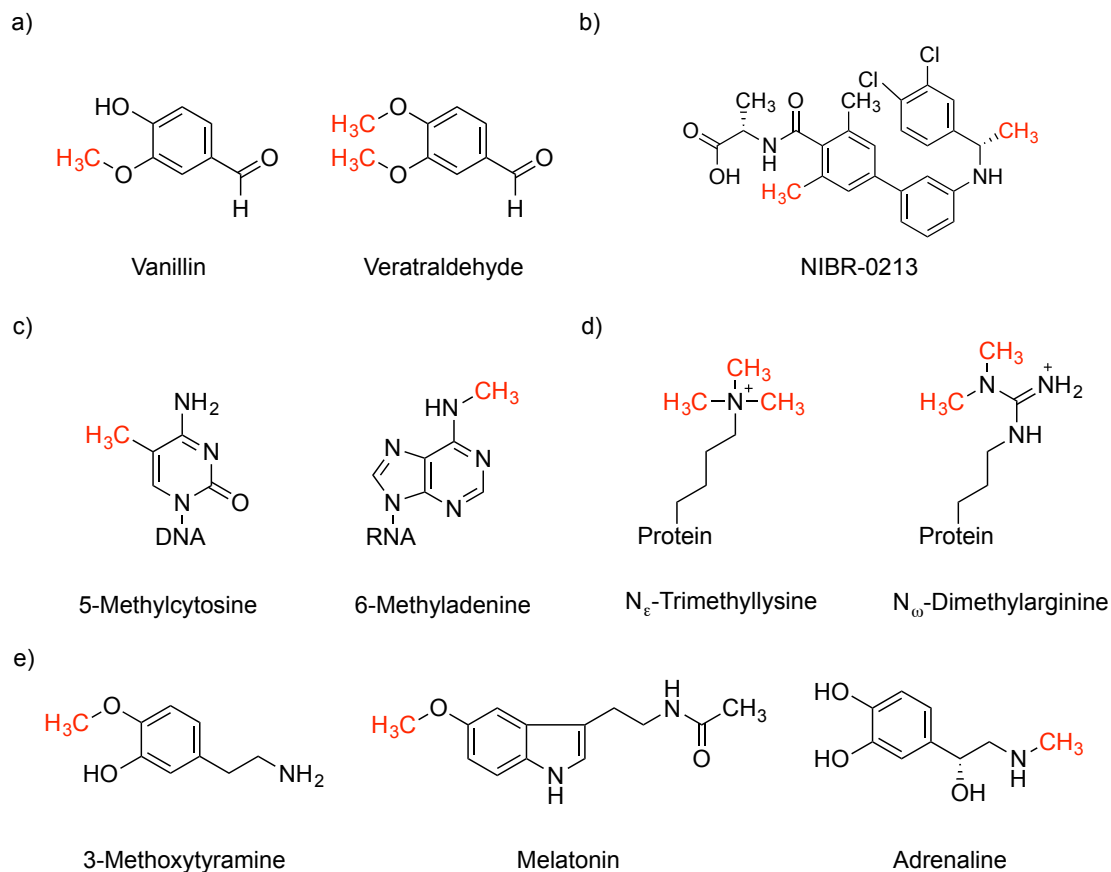
Figure 1. Methylation of different molecules with the added methyl groups coloured in red. a) Structures of vanillin and veratraldehyde which present vanilla aroma and woody fragrance, respectively. b) The structure of NIBR-0213, a drug for EAE which shows more than 2000-fold boosted potency compared to the unmethylated precursor. c) Methylation of cytosine and adenine in DNA and RNA. d) Methylation of lysine and arginine residues within proteins. e) Methylation of neurotransmitters and hormones.
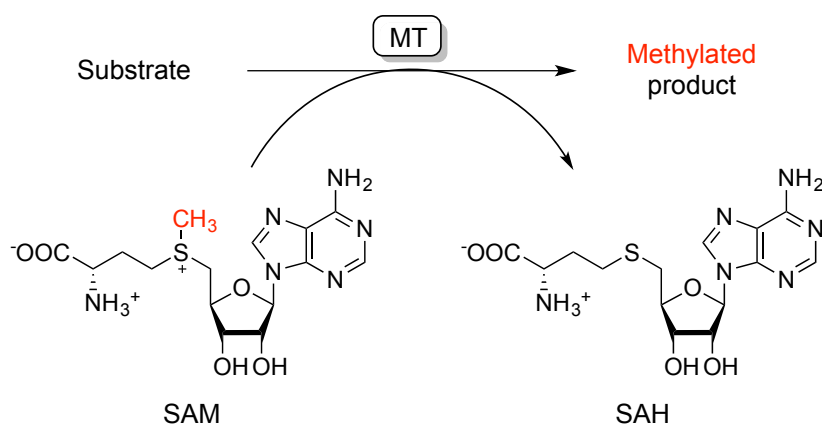
Figure 2. MT-catalysed methylation of the substrate using SAM as the methyl group donor. The methylated product and SAH are thus produced.

The MT-catalysed methyl group transfer reaction is a S_N2 like nucleophilic substitution.[16, 21-23] Due to the strong electron-withdrawing effect of the positively charged sulfur atom, the adjacent methyl group is activated as an electrophile. The nucleophilic, methyl group-accepting atom is activated through deprotonation from a general base in the enzyme and attacks the electrophilic SAM methyl group, yielding the methylated product (Figure 3).
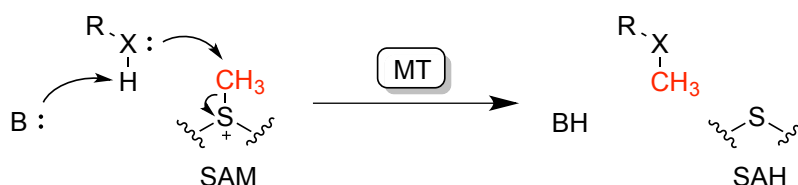

Figure 3. Reaction mechanism of the transfer of the methyl group from SAM to the substrate, catalysed by MTs. X and B represent the methyl group-accepting atom and the general base in the enzyme, respectively.

1.3. *O*-Methyltransferases classified by substrate preference

OMTs are the most abundant methyltransferases and catalyse the methylation of hydroxyl groups. The common substrates are phenolics, flavonoids, chalcones, alkaloids, carbohydrates, CoASH esters, and also carboxylic acids (Figure 4).[24-28]
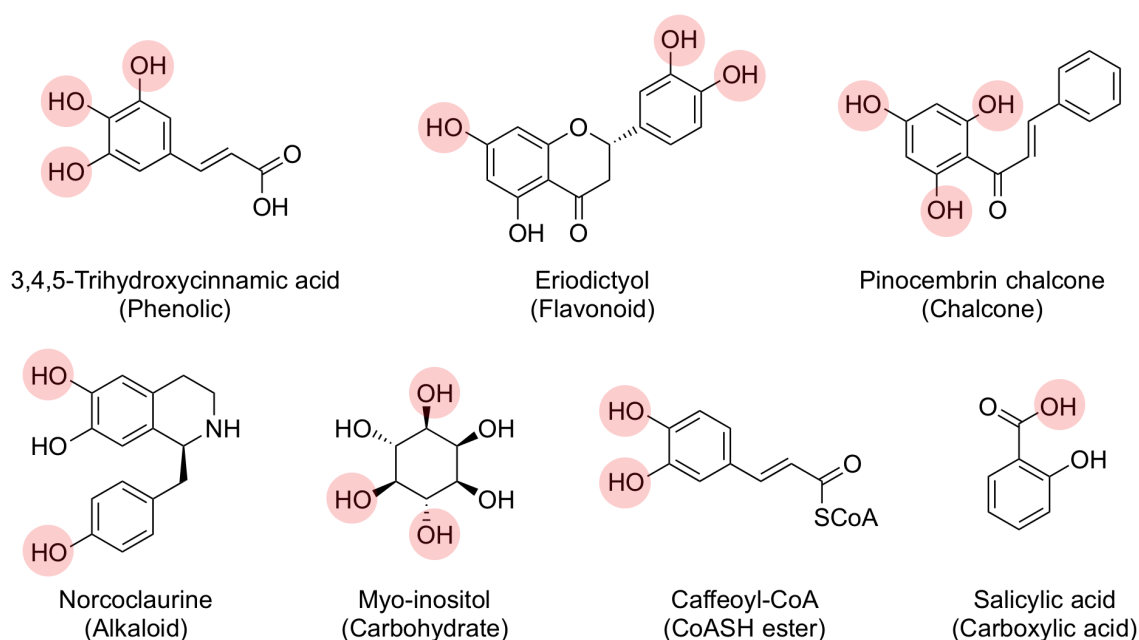

Figure 4. Examples of common substrates of OMTs. The methylation target sites are highlighted.

Plant OMTs are responsible for methylation of a broad spectrum of plant metabolites which are involved in a variety of processes such as plant growth, signalling, and development.[29-30] They are categorized into three classes depending on their substrate preferences.[31-32] Type I OMTs accept phenylpropanoids, flavonoids, isoflavonoids or chalcones as substrates with molecular weights ranging from 40 to 43 kDa and are not bivalent cation-dependent. Type II and type III OMTs methylate the phenylpropanoid-CoA esters and the carboxylic acids, respectively.[28]

OMTs are also widely found in animals. Mammalian catechol OMTs (COMTs) are well-studied due to their roles in the methylation of neurotransmitters and hormones, such as dopamine and adrenaline, as well as their roles as drug targets.[14, 33] COMTs are bivalent cation-dependent and there are two isoforms encoded by the same gene found in different parts of human body.[34] The membrane-bound COMT (MB-COMT, 30 kDa) contains an additional 50 amino acids at the N-terminus and displays higher substrate affinity but lower catalytic activity compared to the soluble COMT (S-COMT, 26 kDa).[35]

1.3.1. Phenylpropanoid OMTs

Caffeic acid/5-hydroxyferulic acid 3/5-OMTs (CaOMTs) are key enzymes in the biosynthesis of lignin precursors found in all lignin-producing plants.[36] They catalyse methylation of the lignin precursors caffeic alcohol/aldehyde/acid and 5-hydroxyconiferyl alcohol/aldehyde/acid exclusively at the *meta*-hydroxyl groups (*meta*-OH or 3-OH), leaving the *para*-hydroxyl groups (*para*-OH or 4-OH) free to form lignin polymers and to avoid oxidative radical coupling (Figure 5a).[37-38] These compounds are also the building blocks of chalcones and flavonoids.[24] On the other hand, isoeugenol 4-OMTs (IeOMTs) catalyse methylation of eugenol and isoeugenol at the only available *para*-OH to produce methyl eugenol and methyl isoeugenol, respectively (Figure 5b).[39] Both the substrates and products are volatiles providing floral scent in plants and are applied in perfumes and food seasonings as flavour additives, or as building blocks in organic synthesis.[1, 40-41]
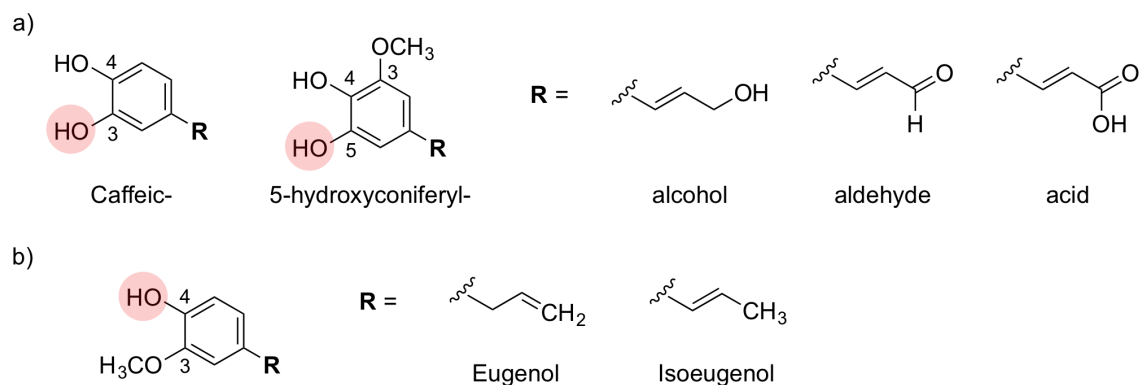
Figure 5. Substrates of phenylpropanoid OMTs. a) Structures of caffeic alcohol/aldehyde/acid and 5-hydroxyconiferyl alcohol/aldehyde/acid. b) Structures of eugenol and isoeugenol. The target hydroxyl groups for the corresponding OMTs are highlighted.

Wang *et al*. have isolated an IeOMT and a CaOMT from *Clarkia breweri*.[42] Although the two OMTs share 83% amino acid sequence identity, they display distinct substrate preference and regioselectivity. While CaOMT serves to produce *meta*-methylated lignin precursors, IeOMT only accepts isoeugenol or eugenol as substrates and methylates the solely free *para*-OH. Through mutagenesis of the substrate binding residues, Wang *et al*. discovered that substitution of the amino acid cluster T133-A134-T135 in IeOMT with the corresponding residues in CaOMT (M135-N136-Q137) have brought IeOMT the CaOMT characteristics, regarding the substrate specificity as well as the regiospecificity.[43] However, the mechanism of the substrate discrimination had not been elucidated yet.

1.3.2. Flavonoid OMTs

Flavonoids are a large group of natural polyphenols and secondary metabolites from plants. They attract a lot of attention due to their nutritional, health-beneficial and pharmacological properties including free radical-scavenging antioxidative activities, anti-inflammatory, antimicrobial and anticancer activities.[44-50] They are classified into flavonoids (2-phenylbenzopyrans), isoflavonoids (3-phenylbenzopyrans) and neoflavonoids (4-phenylbenzopyrans), based on the positions of the phenyl rings (Figure 6a).[51] Flavonoids are further divided into flavanes, flavanones, flavonols, flavones and anthocyanins (Figure 6b).[52] Various modification reactions such as oxidation, hydroxylation, glycosylation and methylation lead to a huge variety of flavonoids.[53] *O*-methylation of free hydroxyl groups on dietary flavonoids greatly increases their absorption and oral bioavailability through the improvement of their

metabolic stability and better membrane transportation.[54-56] Methylation of flavonoids can also bring new biological activities. For example, chrysoeriol (4′,5,7-trihydroxy-3′-methoxyflavone, 3′-methylluteolin) and isohamnetin (3,4′,5,7-tetrahydroxy-3′-methoxyflavone, 3′-methylquercetin) show strong and selective inhibition of the formation of a carcinogenic oestrogen metabolite related to breast cancer (Figure 6c).[57-58] Besides, eriodictyol (3′,4′,5,7-tetrahydroxyflavanone) and homoeriodictyol (4′,5,7-trihydroxy-3′-methoxyflavanone, 3′-methyleriodictyol) are known for their remarkable bitter-masking effect (Figure 6d).[59-60]
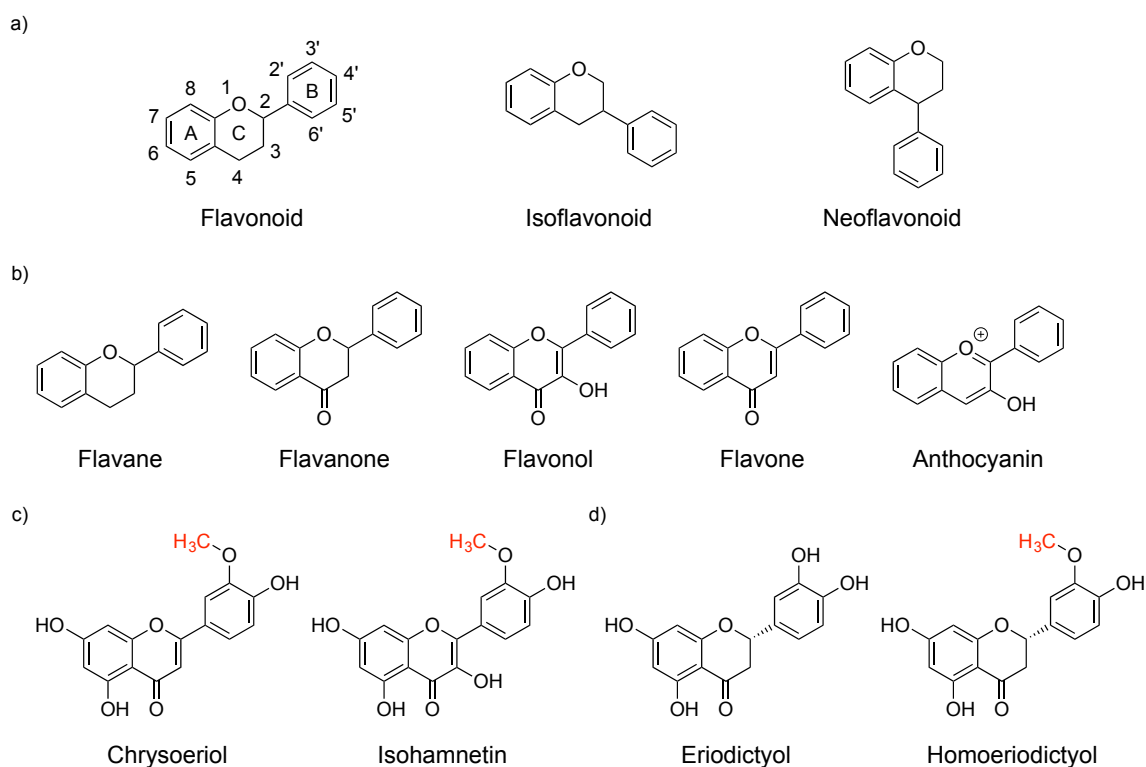

Figure 6. a) Structures of flavonoids, isoflavonoids and neoflavonoids. Ring assignment and the backbone atom numbering are shown. b) Structures of flavanes, flavanones, flavonols, flavones and anthocyanins. c) Structures of chrysoeriol and isohamnetin; d) Structures of eriodictyol and homoeriodictyol. The added methyl groups in c) and d) are coloured in red.

Although phenylpropanoid OMTs (POMTs), flavonoid OMTs (FOMTs) and isoflavonoid OMTs (IOMTs) all belong to the type I OMTs due to highly similar protein sequences and structures, they are divided into different OMT groups based on their substrate preferences. POMTs catalyse methylation of phenylpropanoids such as caffeic alcohol/aldehyde/acid which serve for lignin and flavonoid biosynthesis, and the flower scent compounds eugenol and isoeugenol.[61] FOMTs accept a wide range of flavonoid

substrates and show strict regioselectivity. Isoflavonoids are exclusively produced in leguminous plants and function as phytoalexins. Only SOMT2 from soybean (*Glycine max*) and SaOMT2 from *Streptomyces avermitilis* were discovered to show both flavonoid and isoflavonoid methylation activities.[62-63] Methylation reactions mostly take place at the 7-, 3′-, and 4′-hydroxyl groups of flavonoids, the 7- and 4′-hydroxyl groups of isoflavonoids and the 3- and 4-hydroxyl groups of phenylpropanoids.

1.3.3. Mammalian catechol OMTs

COMTs are involved in the regulation of neurotransmitters and hormones (L-DOPA, dopamine, noradrenaline, adrenaline etc., Figure 7).[14, 33] They also accept a wide range of catechol derivatives as substrates and the majority of them prefer to methylate the *meta*-OH.[15, 33] Owing to the regioselectivity of COMTs, they are also applied in the biocatalytic synthesis of the flavor compound vanillin (3-methoxy-4-hydroxybenzaldehyde) from glucose (Figure 1a).[64] Few COMTs are characterized to have *para*-preference, which serves as a regio-complementary catalyst to the common COMTs and enables the synthesis of the uncommon methylated products.[65-66] Structural and mechanistic studies showed that the two catechol hydroxyl groups of the substrate are stabilized by the active site residues and the bound Mg^{2+}; the general base assists one of the hydroxyl groups to attack SAM (Figure 8). Substrates with polar and ionizable substituents (R groups) tend to orientate their side chains into solvent (aqueous buffer) and expose their *meta*-OH to SAM. Conversely, neutral and hydrophobic side chains are more likely to interact with the hydrophobic wall of COMT and thus present their *para*-OH to SAM for methylation. Based on this reaction mode, regioselectivity of COMT can be altered though engineering of active site residues.[67]
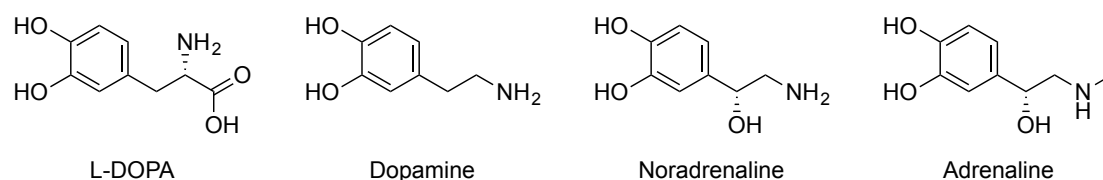

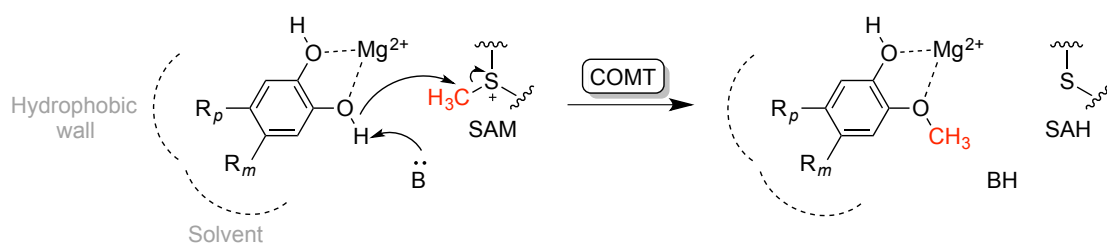Figure 7. Chemical structures of the neurotransmitter and hormone substrates of COMT.

Figure 8. Reaction mechanism of COMT. Catechol hydroxyl groups of the substrate are stabilized by the active site residues and the bound Mg^{2+}. One of the hydroxyl group is activated by the general base (B:) and attacks the methyl group of SAM. *Meta*-methylated product is produced when the substrate side chain (R_m group) orientates into solvent, while *para*-methylated product is produced when the substrate side chain (R_p group) interacts with the enzyme hydrophobic wall.

1.4. Other methyltransferases

1.4.1. *N*-Methyltransferases

Methylation of the nucleic acid amino groups, for example the N6 position of adenine and the N4 position of cytosine, are catalysed by the corresponding DNA and RNA *N*-methyltransferases (NMTs) (Figure 1c).[68-69] Protein NMTs install one or more methyl groups at the terminal amino and imine groups of lysine and arginine, respectively (Figure 1d).[70] Apart from that, NMTs also play a prominent role in the synthesis of many natural products. For example, several NMTs are involved in the production of caffeine in tobacco plants (Figure 9a).[71-72] The coclaurine NMT in cooperation with other OMTs are responsible for the synthesis of the bioactive benzylisoquinoline (*S*)-reticuline which is the precursor of morphinans, aporphines and protoberberines (Figure 9b).[73-75] Besides, the activity of signalling molecules such as noradrenalin, phenylethanolamine, histamine and tryptamine are modulated by the corresponding NMTs (Figure 9c).[13, 25]

1.4.2. *C*-Methyltransferases

Carbon-carbon bond formations are ubiquitous in natural systems and are important reactions in organic synthesis.[3, 76-77] Apart from the classical Friedel–Crafts alkylation, *C*-methylation can also be achieved by SAM-dependent *C*-methyltransferases (CMTs). However, methylation of carbon atoms requires more energy to generate the active carbanion intermediate comparing to the methylation of oxygen and nitrogen. Therefore, *C*-methylation usually takes place at the more nucleophilic carbon atoms in the enolizable

ketones and phenolic substrates. For example, CMTs catalyse the stereoselective β-methylation of α-keto acids in amino acids and polyketides (Figures 10a and 10b).[78-81] An aromatic CMT displays substrate promiscuity on a group of phenolic amino acids such as 3-hydroxykynurenine (Figure 10c), *p*-tyrosine, *m*-tyrosine and L-DOPA.[82-84] The CMTs CouO and NovO discovered from the coumermycin A1 and novobiocin antibiotic biosynthesis pathways not only accept a wide range of coumarin compounds as substrates but are also able to transfer an extensive repertoire of alkyl groups (Figure 10d).[85-86]

a)

b)

Caffeine

(*S*)-Reticuline

c)

Noradrenaline

Phenylethanolamine

Histamine

Tryptamine
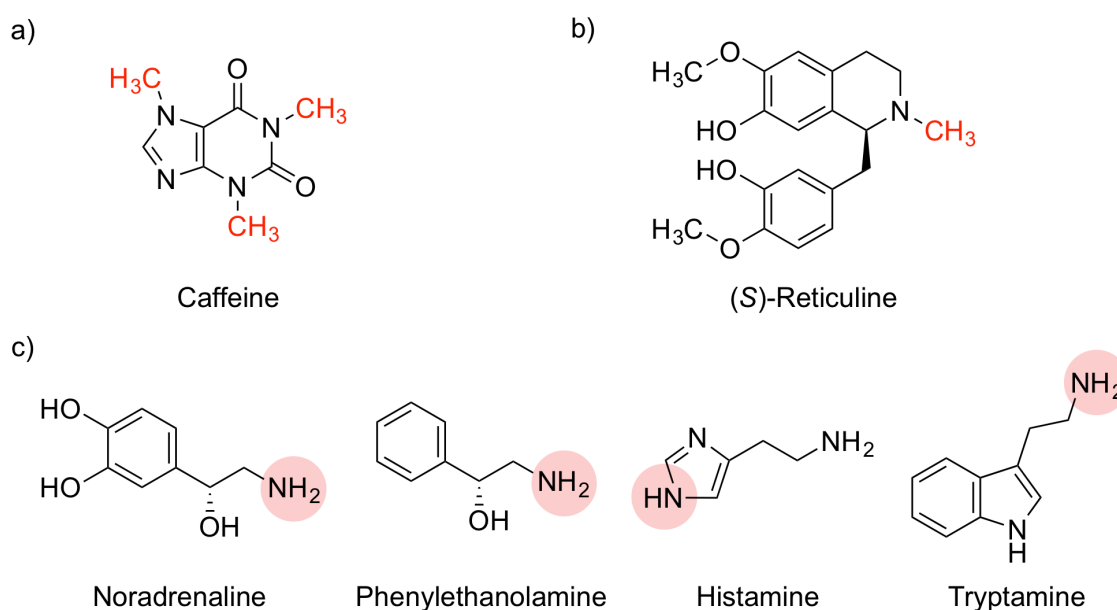
Figure 9. Structures of NMT products a) caffeine and b) (*S*)-reticuline. The added methyl groups are coloured in red. c) Signalling molecules as NMT substrates. The methylation target sites are highlighted.

a)

b)

c)

d)

(3*R*)-methyl-2-oxoglutaric acid

Polypeptide

3-hydroxy-4-methylkynurenine

Coumarin derivatives

Figure 10. Structures of CMT products. a) (3*R*)-methyl-2-oxoglutaric acid; b) β-methylated polyketide; ACP, acyl carrier protein; c) 3-hydroxy-4-methylkynurenine; d) alkylated coumarin derivatives; R_1 represents the transferred alkyl group which can be the allyl, propargyl and benzyl groups and are coloured in red.

1.5. Regeneration of SAM

So far, MTs are not yet commonly used in industrial processes mainly due to the lack of bulk availability of the methyl donating cofactor SAM. It is a complex, unstable and expensive molecule and needs to be provided in stoichiometric amount to the substrate. Besides, the by-product SAH produced after transferring the methyl g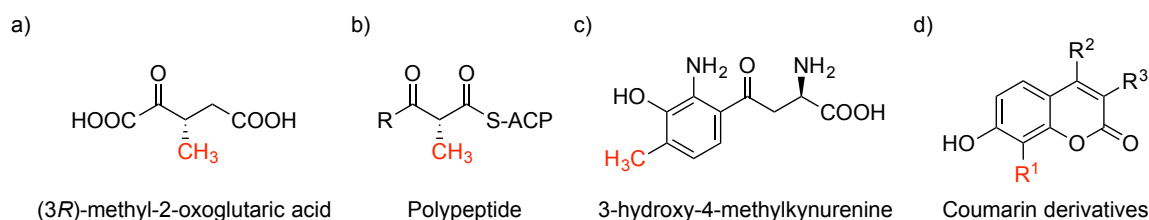roup is a potent inhibitor of MTs.[87-88] Regeneration of SAM can be a powerful approach to address these problems. It produces SAM *in situ* and degrades the produced SAH, which makes full use of the molecule and eliminates the inhibition effect of SAH.

1.5.1. A multistep, cyclic regeneration system

The very first attempt was a multistep, cyclic, *in vitro* regeneration system, based on the physiological cycle of the metabolites in cells.[89] The SAH produced from SAM was first hydrolysed to adenosine and L-homocysteine (Hcys) by the SAH hydrolase (SAHH), after which adenosine was phosphorylated to adenosine monophosphate (AMP) and then to adenosine triphosphate (ATP) by the adenosine kinase (ADK) and polyphosphate kinases 2 (PPK2) I & II (Figure 11).[90] Finally, a new SAM molecule is produced from ATP and L-methionine (L-Met) catalysed by the methionine adenosyltransferase (MAT). Overall, this is a complex regeneration system requiring an additional of five enzymes and has shown rather low turnover, up to only 11 cycles.

1.5.2. HMT-MT cascade

A more versatile and much simpler SAM recycling system that requires only a halide methyltransferase for SAM production was developed by Liao *et al.*[91] HMTs were discovered in chloromethane (CH_3Cl) producing organisms and were reported to catalyse methylation of halide ions using SAM as the methyl donor.[92-94] By noticing that the reaction is actually thermodynamically unfavourable but most likely driven by the removal of the highly volatile products, Liao *et al.* in turn used HMT to produce SAM from SAH and methyl iodide. The SAM regeneration system is comprised of two enzymes, the MT and HMT, and requires only a catalytic amount of SAH and stoichiometric concentrations of methyl iodide as the methyl donor (Figure 12). The enzyme cascade is compatible with OMTs, NMTs and CMTs and can reach up to 580 cycles.
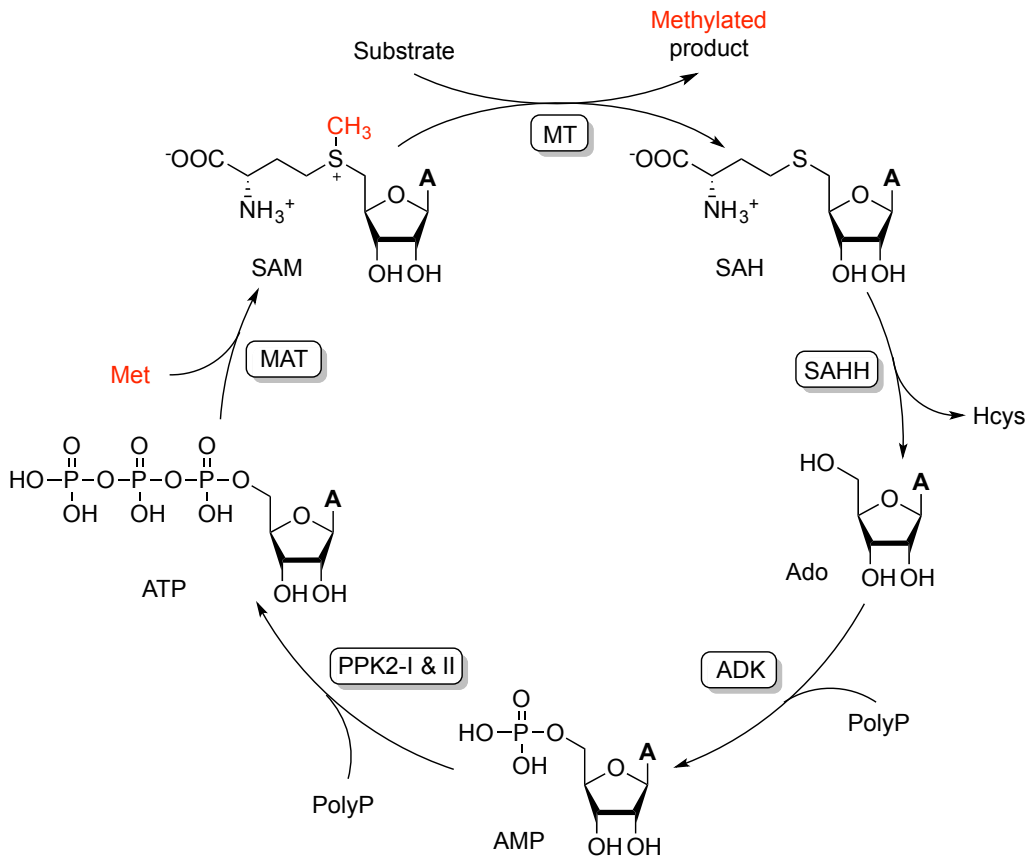
Figure 11. A multistep, cyclic, *in vitro* SAM regeneration system. The "**A**" represents the adenine moiety. SAHH, SAH hydrolase; ADK, adenosine kinase; PPK2-I & II, polyphosphate kinases 2, I & II; MAT, methionine adenosyltransferase; Hcys, L-homocysteine; PolyP, polyphosphate; Met, L-methionine.
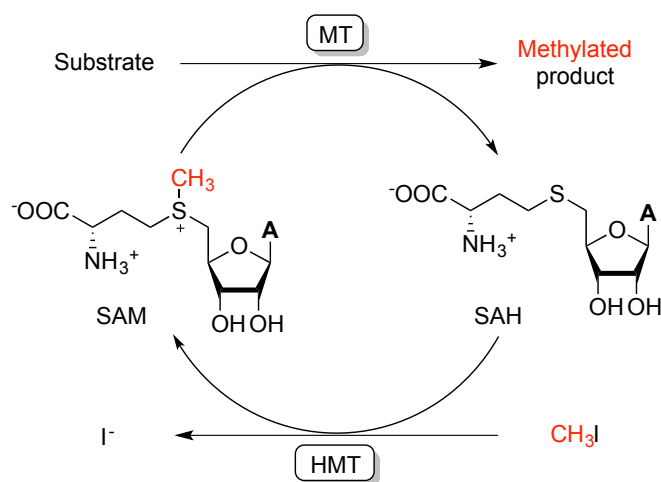

Figure 12. Regeneration of SAM through the HMT-MT cascade.

1.6. Biocatalytic alkylation

Apart from methylation, selective alkylation of molecules is highly desirable.[95] It can expand the structural and functional diversity of chemicals and enable various applications, such as site-selective modification or labelling of molecules with fluorescent or "clickable groups".[20, 96-98] Biocatalytic alkylations are important because of the exquisite chemo-, regio-, and stereospecificity achievable using enzymes.[95, 99-100] Some methyltransferases are insensitive to the size of the alkyl substituent of SAM and have been reported to catalyse alkylation reactions with the corresponding SAM analogues provided.[67, 96, 101-102] SAM analogues enabling these diverse alkylation reactions are crucial not only for expanding the industrial relevance of biocatalytic alkylation but also for discovering more promiscuous MTs.[103-104] However, very few naturally occurring SAM analogues are known, making limited access to SAM analogues one of the most serious impediments to the progress in the field.[96, 104-106]

1.6.1. Chemical synthesis of SAM analogues

SAM analogues can be synthesized chemically from SAH and an excess of alkyl triflates or alkyl halides in the formic acid-acetic acid mixture. The SAM analogues are obtained as diastereomeric mixtures at the sulfonium centre (Figure 13). Only the (*S*,*S*)-epimers can be utilized by MTs. The biologically inactive (*R*,*S*)-diastereomers are potent inhibitors of MTs and are not easily separated from the desired SAM analogues.[101, 107-108]
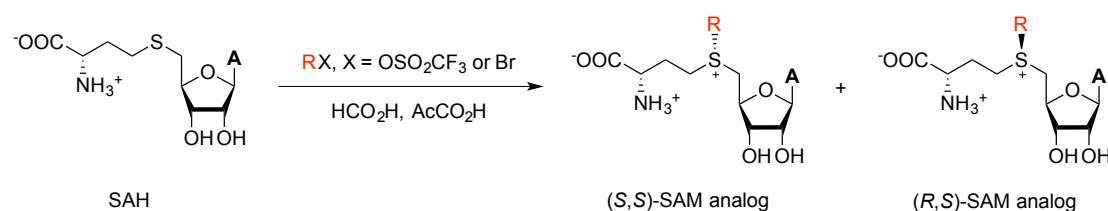
Figure 13. Chemical synthesis of SAM analogues, resulting in diastereomeric mixtures at the sulfonium centre.

1.6.2. Chemo-enzymatic synthesis of SAM analogues

Enzymatic synthesis of SAM analogues is more specific which yields only the desired isomers. There are two enzyme classes, MATs and halogenases, that have been reported to synthesize SAM analogues from L-Met analogues and the adenosyl moiety providing molecules ATP or 5′-chloro-5′-deoxyadenosine (Figure 14a).[109-113] Both enzymes are

highly promiscuous which accept diverse alkyl groups up to the benzyl group, as well as the polar amino groups, amino acids and nitriles. However, the major drawback of these approaches is that they require methionine analogues which are not commercially available (except L-ethionine) and needed to be synthesized chemically (Figure 14b).
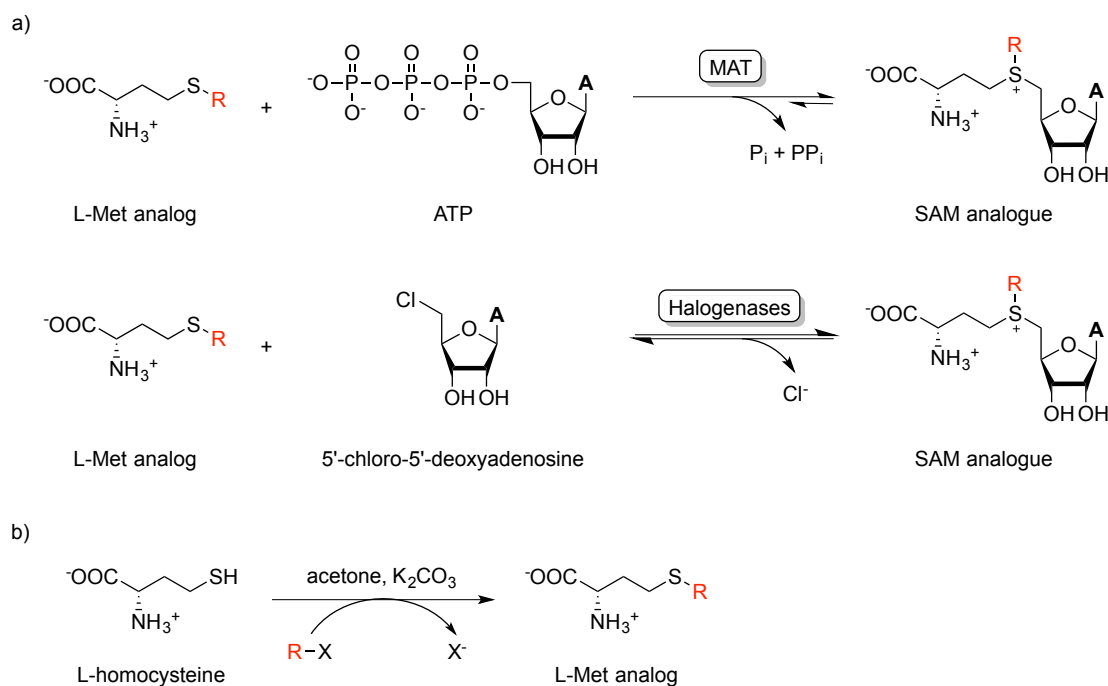
Figure 14. Chemo-enzymatic synthesis of SAM analogues. a) MATs or halogenases can then use ATP or 5'-chloro-5'-deoxyadenosine, respectively, to convert the L-methionine analogues to the corresponding SAM analogues. b) The L-Met analogues need be synthesized chemically from L-homocysteine and alkyl halides.

1.6.3. HMT-catalysed enzymatic synthesis of SAM analogues

Since the SAM synthesis function of HMT was discovered, it can be an efficient alternative to synthesize SAM analogues because the alkyl iodides are cheap and readily available (Figure 15). Most importantly, SAM analogues can be regenerated through the HMT-MT cascade reaction while achieving bioalkylation. Most of the HMTs have been reported to transfer only the methyl group. Until recently, Bengel *et al.* discovered that few natural HMTs were able to transfer the ethyl, propyl and the butyl groups.[114] Therefore, it is of great potential to discover more promiscuous HMTs, either from nature or via protein engineering, for the synthesis of different SAM analogues.
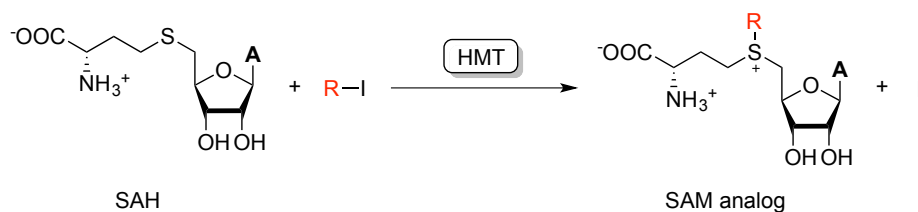
Figure 15. Production of SAM analogues from SAH and alkyl iodides, catalysed by HMT.

1.6.4. MTs that accept SAM analogues

With the supply of SAM analogues, biocatalytic alkylation can be achieved by the MTs which accommodate alkyl groups that are bulkier than the methyl group. The wild-type COMT accepts the ethyl, allyl and benzyl groups in the alkylation of 3,4-dihydroxybenzaldehyde.[67] The RapM 16-OMT is promiscuous in transferring the ethyl and allyl groups to the macrolide immunosuppressive agent rapamycin, generating analogues with altered physicochemical and pharmacological properties.[102] Besides, CMTs CouO and NovO enable the regiospecific installation of a series of unsaturated alkenyl and alkynyl groups on different coumarin compounds.[86] Interestingly, COMT and the coclaurine NMT were engineered to change their preference from methylation to carboxymethylation, with the cofactor carboxy-SAM (cxSAM) generated from SAM by a cxSAM synthase.[104, 115-116] Apart from small-molecule MTs, many promiscuous wild-type or engineered DNA/RNA MTs and protein MTs have been utilized to install various clickable terminal alkyne-containing groups for the labelling of biomolecules.[97, 117-119]

2. Results

2.1. Expanding the substrate scope and enhancing the regioselectivity of IeOMT against phenolic compounds (Article I)

As reported by Wang *et al.* which was discussed in Chapter 1.3.1, the substrate and regioselectivity discrimination between IeOMT and CaOMT is determined by the specific amino acid cluster T133-A134-T135 in IeOMT and the corresponding residues M135-N136-Q137 in CaOMT. We investigated in detail the effects of these residues. Genes encoding IeOMT from *Clarkia breweri* (NCBI accession number: O04385) and CaOMT from *Catharanthus roseus* (NCBI accession number: Q8W013) were synthesized and expressed in *E. coli*. The triple mutant T133M/A134N/T135Q, double mutant T133M/A134N and the single mutants at each position were constructed using IeOMT as scaffold. All enzymes and variants were purified and tested against a series of phenolic compounds, isoeugenol (**1b**), eugenol (**2b**), caffeic acid (**3a**), 3,4-dihydroxybenzaldehyde (**4a**), and phenolic acids 3,4-dihydroxybenzoic acid (**5a**), 3,4-dihydroxyphenylacetic acid (**6a**) and 3,4-dihydroxyphenylpropanoic acid (**7a**) (Figure 16).
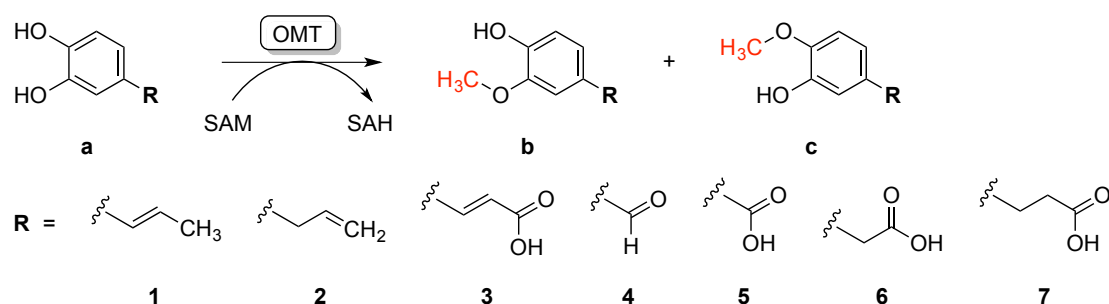
Figure 16. OMTs catalyse methylation of phenolic substrates (**a**) to give *meta*- (**b**) or *para*- (**c**) methylated products. Structures of different substrates are shown, based on the R groups and the phenolic moieties **a**, **b** or **c**.

As shown in Figure 17, the wild-type IeOMT displayed *para*-methylated activities against isoeugenol and eugenol, while CaOMT showed preference to the *meta*-OH of caffeic acid. The triple mutant T133M/A134N/T135Q of IeOMT has increased methylation activities against both substrates. Specifically, the single mutant T133M has a significant effect on the enhancement of the activities, while the T135Q largely impairs enzymatic activity. The triple mutant T133M/A134N/T135Q has conferred *meta*-specificity to IeOMT against caffeic acid, as reported by Wang et al.[43] Interestingly, the mutants T133M and A134N at adjacent positions gave different regioselectivities against caffeic acid. T133M

increases the preference for the *para*-OH, leading to the major production of isoferulic acid (**3c**), while A134N is strictly *meta*-specific.

Figure 17. Methylation efficiency and regioselectivity of the wild-type IeOMT, mutants and CaOMT against different substrates. Negative controls performed without enzyme were subtracted. Data reported are the means and standard deviations calculated from three independent measurements. Figure from Tang *et al.*, 2019.[120]

Both T133M/A134N/T135Q, T133M/A134N and the T133M variant show improved activities against the other phenolic substrates compared to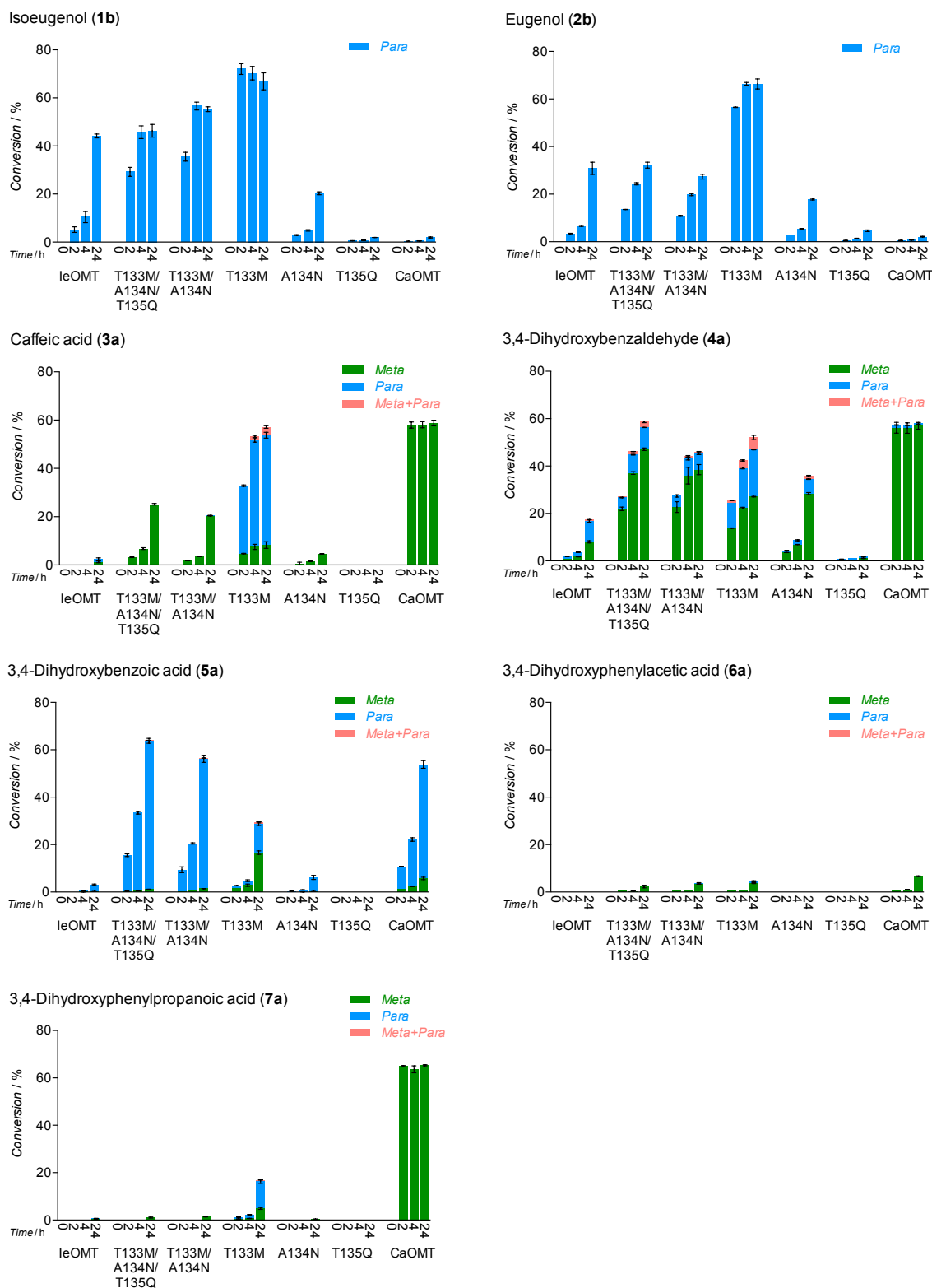 the wild-type IeOMT. The improvement is more remarkable for 3,4-dihydroxybenzaldehyde and 3,4-dihydroxybenzoic acid, which have one-carbon aldehyde and carboxylate groups attached to the phenyl ring, respectively. Interestingly, the T133M/A134N/T135Q, T133M/A134N and A134N variants are *meta*-selective against 3,4-dihydroxybenzaldehyde and produce mainly vanillin (**4b**), they show *para*-specificity towards 3,4-dihydroxybenzoic acid. CaOMT also shows totally different regiospecificity against these two substrates.

Crystal structures of the *Clarkia breweri* IeOMT (with mutations T133L/E165I/F175I) and a CaOMT from *Lolium perenne* (perennial ryegrass) were determined with phenolic substrates and SAH bound in the active site.[121-122] Both structures were obtained in closed conformations, referring to the active catalytic states. The two structures are well-superimposed, with the root-mean-square deviation (RMSD) value of 0.636 Å. As shown in Figure 18, the ligands isoeugenol and sinapaldehyde are surrounded and stabilized by a series of hydrophobic residues. Among them, two Met residues (Met183 and Met323) which are highly conserved in plant OMTs and the residue Phe179 play major roles in stabilizing and orienting the substrate phenyl rings.[17] Depending on whether the *meta*- or the *para*-hydroxyl group is closer to the methyl group of SAM (within *ca.* 3 Å to the activated methyl carbon of SAM or *ca.* 4 Å to the sulfonium moiety), it gets preferably methylated and thus determines the regioselectivity.[25] Interactions between the substrate side chain and the adjacent residues affect the substrate orientation. Since the added Met residue in mutant T133M is located above the isoeugenol side chain and stabilizes it through hydrophobic interaction, it attracts the hydrophobic side chain and makes the *para*-hydroxyl group well-exposed to SAM. Similarly, T133M is also *para*-selective to caffeic acid. However, T133M does not show obvious regioselectivity against 3,4-dihydroxy-benzaldehyde (**4a**) or 3,4-dihydroxybenzoic acid (**5a**), probably because these substrates have shorter side chains and therefore are more flexible in the enzyme active site.
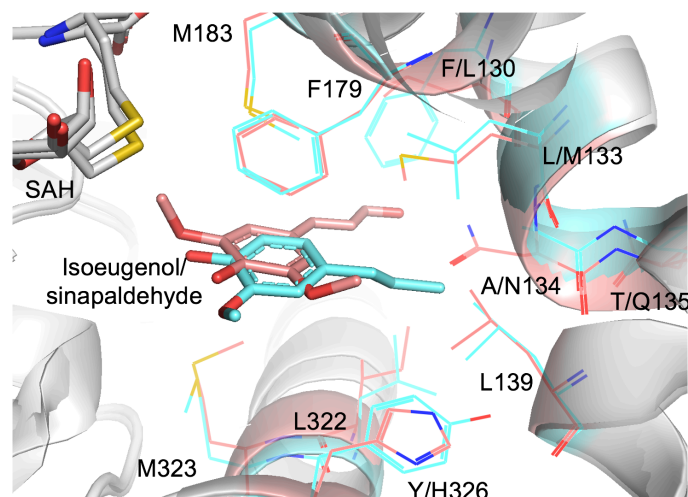
Figure 18. Superimposed active sites of the *Clarkia breweri* IeOMT (PDB ID: 3REO; with mutations T133L/E165I/F175I; cyan; with SAH and isoeugenol bound) and the *Lolium perenne* CaOMT (PDB ID: 3P9I; pink; with SAH and sinapaldehyde bound). Residues interacting with the ligands are shown in lines with elemental colouring. Interactions between the substrate side chain and residues 133 and 134 affect the orientation of the substrate. The hydroxyl group which is closer to the sulfur atom of SAH (white, with elemental colouring) will be preferably methylated.

In mutant A134N, the mutated residue has steric hindrance with the substrate side chain and expels it upwards (from the view of Figure 18) and hence the substrate positions its *meta*-hydroxyl group to the methyl donor SAM. This makes the variants T133M/A134N/T135Q, T133M/A134N and A134N strictly *meta*-specific towards caffeic acid. Interestingly, while these mutants are *meta*-selective against 3,4-dihydroxybenzaldehyde (**4a**), they are *para*-specific to 3,4-dihydroxybenzoic acid (**5a**).

2.2. Expanding the substrate scope and enhancing the regioselectivity of IeOMT against flavonoids (Article II)

In order to discover the factors determining the substrate discrimination between the FOMTs, IOMTs and POMTs, as discussed in Chapter 1.3.2, we chose 21 plant OMTs owning different substrate preferences and regioselectivities isolated from various plant species for comparison (Figure 19a). Meanwhile, we looked into the substrate binding residues in the crystal structure of the *Medicago sativa* 7/4′-IOMT which adopts an active conformation with the ligand isoformononetin (7-methoxy-4'-hydroxyisoflavone) bound in the enzyme active site (PDB ID: 1FP2; Figure 19b).[17] Met183, Met323 and Phe179, which are conserved throughout the plant OMT superfamily, have constrained the

aromatic A-ring and help positioning the 7-hydroxyl group to the catalytic residue His272 and SAM.[17] Interestingly, the residue Asn322 stabilizing the C-ring ketone group is conserved only among the 7-IOMTs, while the other OMTs rather have middle size hydrophobic residues, namely Ile, Val or Met, at this position. The Leu326 in the 7-IOMTs interacts with the C-ring of isoformononetin, and the residues at this position in other OMTs are quite diverse, varying from the hydrophobic residues Leu, Val, Ile, or Met, to the basic Arg or His. In order to investigate the influence of these substrate binding residues, we constructed the mutants L322H/N/M and Y326H/R/L based on the IeOMT-T133M, a variant which has been proved to increase enzyme activities against phenolic substrates.[120] The commonly known (iso)flavonoids eriodictyol and naringenin (flavanones), luteolin (flavone), quercetin (flavonol) and genistein (isoflavone) were chosen as substrates and the enzyme activities were tested (Figure 20).
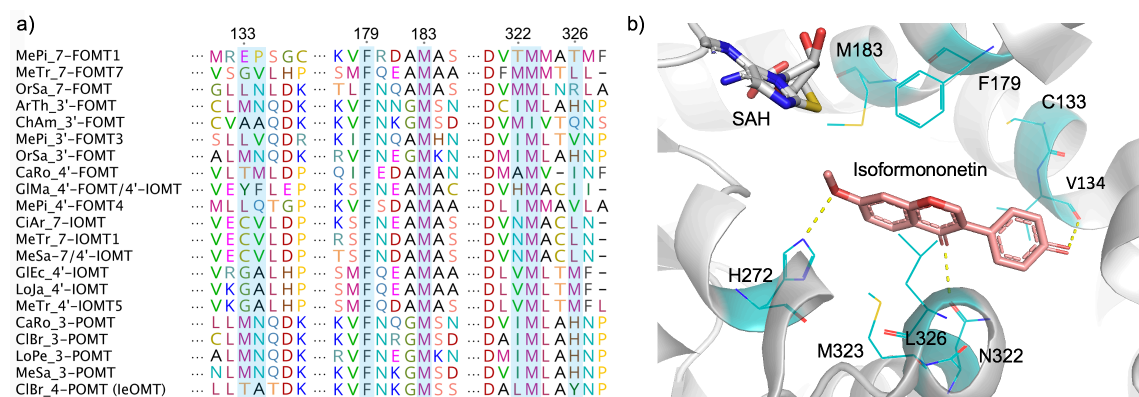
Figure 19. a) Sequence alignment of plant FOMTs, IOMTs and POMTs, performed with Geneious 10.0.2. Residues involved in substrate binding are highlighted. Numbering of residues are based on the sequence of the *Clarkia breweri* IeOMT. OMTs are named by their original organisms and regioselectivity. MePi, *Mentha piperita;* MeTr, *Medicago truncatula*; OrSa, *Oryza sativa*; ArTh, *Arabidopsis thaliana*; ChAm, *Chrysosplenium americanum*; CaRo, *Catharanthus roseus*; GlMa, *Glycine max*; CiAr, *Cicer arietinum*; MeSa, *Medicago sativa*; GlEc, *Glycyrrhiza echinate*; LoJa, *Lotus japonicus*; ClBr, *Clarkia breweri*; LoPe, *Lolium perenne*. MeTr_7-FOMT7 is a putative IOMT but it has higher preference to naringenin (flavanone) than isoflavonoids. b) Enzyme active site of MeSa-7/4′-IOMT with the ligand isoformononetin (7-methoxy-4'-hydroxyisoflavone). The substrate binding residues are shown in lines with elemental colouring. Figures from Tang *et al*., 2020.[123]
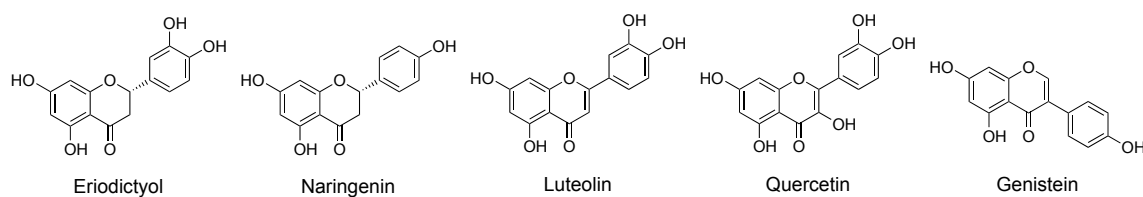
Figure 20. Structures of the flavanones eriodictyol and naringenin, the flavone luteolin, the flavonol quercetin and the isoflavone genistein.

Although IeOMT is a phenylpropanoid OMT, it showed high activities against the flavonoids eriodictyol, luteolin and quercetin, regiospecifically methylating the 3′-hydroxyl groups (Figure 21). In the absence of the 3′-hydroxyl group, the wild-type IeOMT displayed very low activity, as seen against naringenin. The variant T133M has increased the 3'-methylation activity against eriodictyol, luteolin and quercetin and the mutants T133M/Y326H and T133M/Y326L further methylates the 4′-hydroxyl groups, producing the 3′,4′-dimethylated products. Interestingly, these mutants turned out to methylate the 7-OH of naringenin instead of the 4′-OH when the 3′-OH is absent. On the other hand, the wild-type IeOMT displayed very minor activity towards the isoflavone genistein and the mutants showed different regioselectivity. While most of the mutants exhibited higher selectivity against the 4′-hydroxyl group, T133M/Y326H and T133M/Y326L also produced the 7-methylated genistein.

In silico analysis also proved regioselectivity of the tested mutants. As seen in Figure 22A, the wild-type IeOMT accommodated eriodictyol, luteolin and quercetin in an orientation that the 3′-OH is 3.2 to 3.5 Å to the methyl group of the SAM and was favoured for methylation. Introduction of mutations T133M and Y326L enabled the inverse binding of these three flavonoids in the binding pocket, bringing the 4′-hydroxyl group in the proximity of the methyl group of the SAM (Figure 22b). In the case of genistein, the variant T133M/Y326H has a shift of regioselectivity to position 7 compared to other mutants. As seen in Figure 22C, the 7-OH showed a distance of 3.7 Å to the methyl group of SAM and the His326 stabilized the 5-OH of genistein. As for naringenin, the ring B of the substrate was pushed away because of steric clashes with Met322 in the mutant T133M/L322M and thus enabled a binding pattern where the 7-OH of the ring A was accessible to the SAM at a catalytic distance of 3.4 Å (Figure 22D).
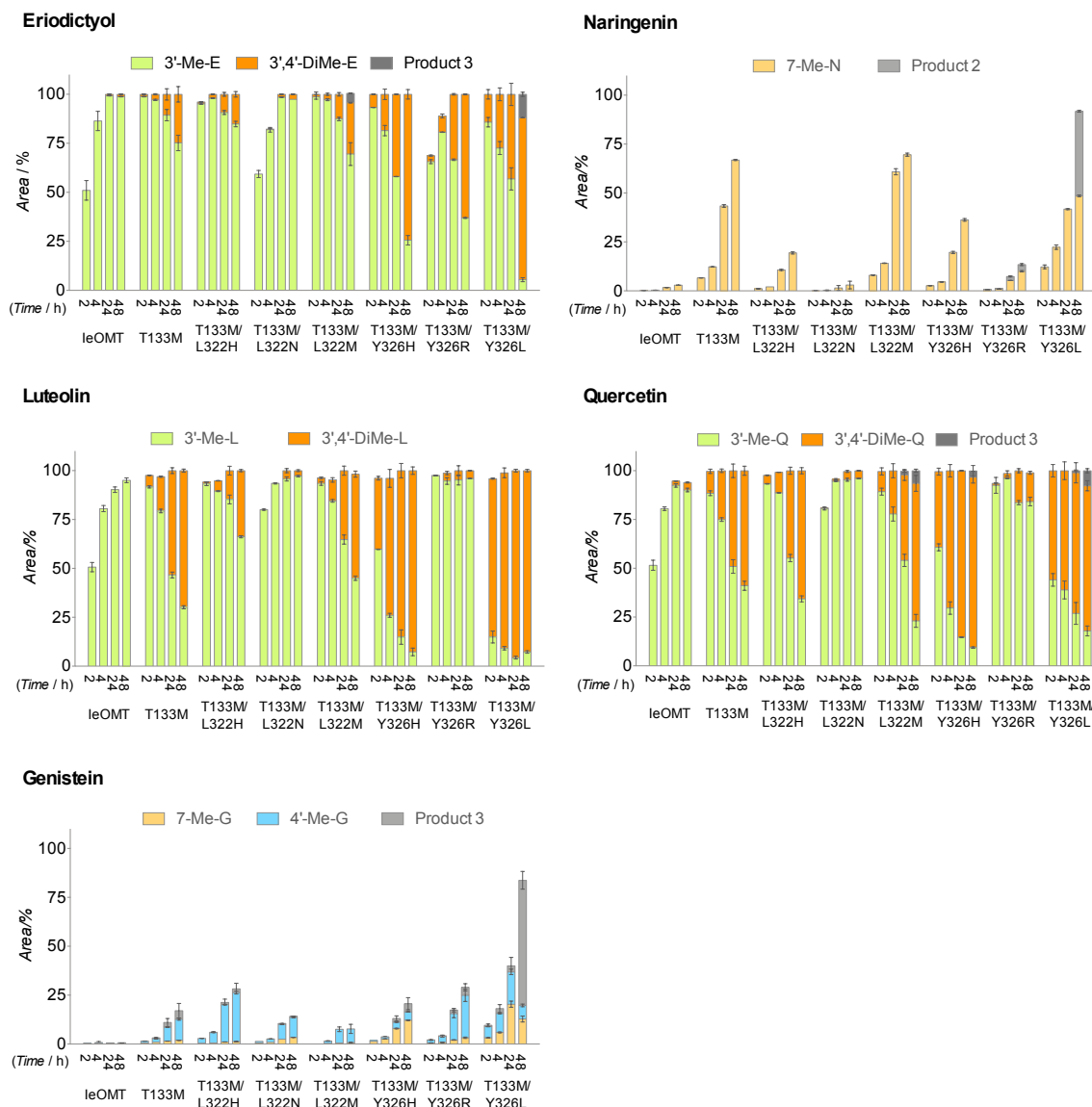
Figure 21. Product composition of each reaction catalysed by the wild-type IeOMT and its mutants. Area percentages were calculated by peak area of both substrates and products as determined by HPLC. 3′-Me-E, 3′-methyleriodictyol; 3′,4′-DiMe-E, 3′,4′-dimethyleriodictyol; 3′-Me-L, 3′-methylluteolin; 3′,4′-DiMe-L, 3′,4′-dimethylluteolin; 3′-Me-Q, 3′-methylquercetin; 3′,4′-DiMe-Q, 3′,4′-dimethylquercetin. Structures of these products are confirmed by NMR and LC-MS. 7 Me-N, 7-methylnaringenin; 7-Me-G, 7-methylgenistein; 4′-Me-G, 4′-methylgenistein. These products were confirmed by comparing their retention times on HPLC to commercial standards as well as by LC-MS. Racemic eriodictyol and naringenin were used as substrates and racemic products were obtained. Data reported are the means and standard deviations calculated from three independent measurements. Figures from Tang *et al.,* 2020.[123]
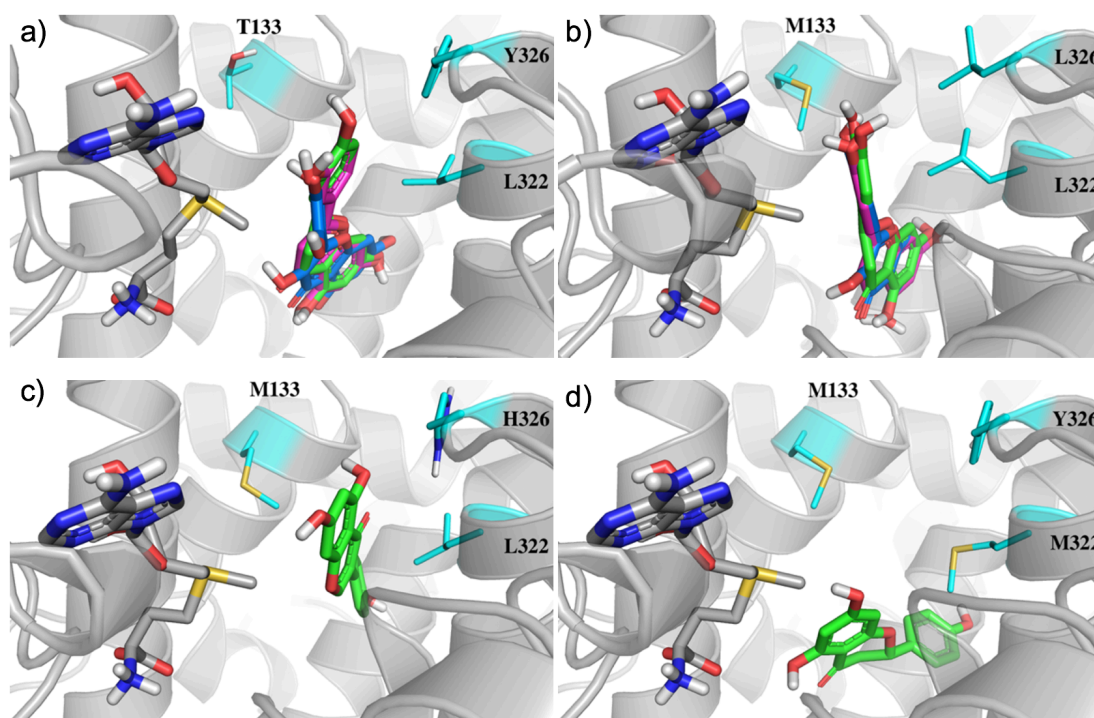
Figure 22. Models of the *Clarkia breweri* IeOMT variants with SAM (grey sticks with elemental colouring at the left side) and the substrates in the binding pocket. The three targeted residues are shown as lines. For clarity, non-polar hydrogens are removed. a) (*S*)-eriodictyol (green), luteolin (magenta) and quercetin (blue) docked in the binding pocket of wild-type IeOMT, exposing the 3′-hydroxyl group for methylation. b) the same substrates in the binding pocket of IeOMT_T133M/Y326L, exposing the 4′-hydroxyl group for methylation. c) Genistein binding on IeOMT_T133M/Y326H provides the 7-position for methylation, d) (*S*)-Naringenin binding pattern in IeOMT_T133M/L322M also exposes the 7-position, however, with a totally different binding pattern. Figures from Tang *et al*., 2020.[123]

2.3. Directed evolution of HMT for the biocatalytic synthesis of diverse SAM analogues (Article III)

As discussed in Chapter 1.6.3, it is of great potential to discover promiscuous HMTs to synthesize SAM analogues from SAH and alkyl iodides. We first assessed the SAH alkylation activities of the *Arabidopsis thaliana* HMT (*At*HMT), *Chloracidobacterium thermophilum* HMT (*Ct*HMT) and *Raphanus sativus* HMT (*Rs*HMT), using methyl iodide (MeI), ethyl iodide (EtI), propyl iodide (PrI), and butyl iodide (BuI) as alkyl donors.[91-92, 124] These enzymes were expressed in an SAH nucleosidase-deficient strain of *E. coli* BL21(DE3) which avoids contamination of recombinant enzymes with this SAH-degrading enzyme.[91] All three HMTs were highly active towards MeI and displayed minor activities against EtI (Table 1). However, none of the enzymes had significant activity towards PrI

or BuI (less than 0.01 µmol/min/mg). *At*HMT had the highest ethyltransferase activity (2.69 ± 0.15 µmol/min/mg) and a solved crystal structure and was therefore chosen as the starting point for semi-rational protein engineering.

Table 1. Specific activities (µmol/min/mg) of HMTs towards different alkyl iodides for synthesizing the corresponding SAM analogues.[a] Table from Tang *et al.*, 2021.[125]

Enzyme	MeI	EtI	PrI	BuI
*At*HMT	5942 ± 204	2.69 ± 0.15	<0.01	<0.01
*Ct*HMT	132 ± 9	0.11 ± 0.07	<0.01	<0.01
*Rs*HMT	8457 ± 361	1.18 ± 0.11	<0.01	<0.01

[a] Reaction times were 10 min for MeI and EtI, and 4h for PrI and BuI. Negative controls performed without enzyme were subtracted. Activities were determined using the iodide assay described in Figure 23. Data reported are the means and standard deviations calculated from three independent measurements.

In order to perform directed evolution and investigate all randomized residues of HMT libraries, we established a high-throughput assay to determine the amount of iodide released for each molecule of SAM analogue formed. We used a recombinant *Curvularia inaequalis* vanadium-dependent chloroperoxidase (*Ci*VCPO) to oxidize iodide to hypoiodous acid (HOI), which subsequently oxidizes the chromogen 3,3′,5,5′-tetramethylbenzidine (TMB), resulting in the formation of blue colour (Figure 23).[126] This assay has been developed from the ultrasensitive halide assay previously established in our group but is highly sensitive to iodide and much less sensitive to chloride (Figure 23b).[127] This enabled detection of iodide in a wide concentration range regardless of the presence of chloride in crude cell lysates. Most importantly, our assay is general and suitable for monitoring the release of iodide from various alkyl iodides. Activities of HMTs can be determined by measuring the change in absorbance at 570 nm which is proportional to iodide concentration in the range from 5 µM to 400 µM and were demonstrated by HPLC.

a)

b)

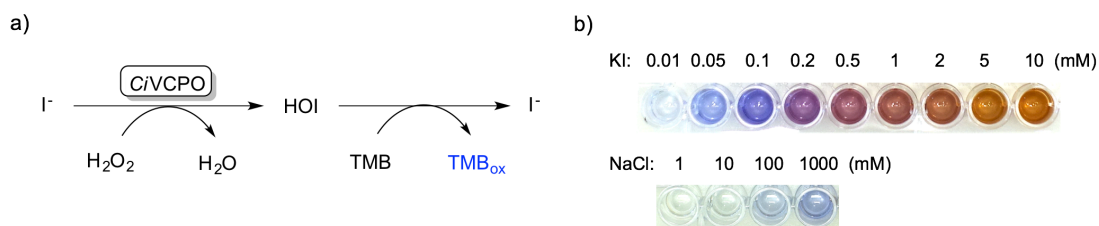KI: 0.01 0.05 0.1 0.2 0.5 1 2 5 10 (mM)

NaCl: 1 10 100 1000 (mM)

Figure 23. a) The iodide produced after the alkyl group transfer from alkyl iodide to SAH can be detected using hydrogen peroxide, a haloperoxidase *Ci*VCPO, and TMB. b) Colour of reaction mixtures of of KI and NaCl (1 μL; various concentrations) with the iodide assay solution (50 μL) after 30 min of reaction. Figures from Tang *et al.*, 2021.[125]

We docked EtI, PrI, and BuI into the active site of the *At*HMT crystal structure (PDB ID: 3LCC) with bound SAH (Figure 24a).[128] The alkyl-binding pocket is quite spacious and could accommodate even the butyl iodide. We selected ten residues forming the alkyl-binding site (P20, V23, L27, W36, W47, Y139, V140, C143, Y172 and R214), constructed NNK site-saturation libraries and screened them by incubating crude lysate with SAH and EtI.[129] Based on the four hits (V23T, W36F, V140C, and V140T) with improved activities relative to the wild-type *At*HMT, we constructed NNK libraries at each other hit position and found out more hits with double-mutant combinations of V23K/L/T, W36F, and V140C/T. We also constructed triple mutants based on the single and double mutants. Finally, we expressed and purified the wild-type *At*HMT and all these variants and determined their specific activities towards EtI (Figure 24b). The specific activity of the most active V140T variant (14.02 μmol/min/mg) towards EtI was fivefold higher than that of the wild-type *At*HMT (2.69 μmol/min/mg) and its k_{cat} value was 6-fold higher than that for the wild type (data not shown).

In addition to increased ethyltransferase activity, the V140T variant was also the most active one towards PrI and allyl iodide (Allyl-I), compared to the wild type and other mutants (data not shown). Therefore, the V140T variant was used to convert SAH (15 mg) to *S*-adenosyl-L-ethionine (SAE, **8**), *S*-adenosyl-L-propionine (SAP, **9**), and *S*-adenosyl-L-allylionine (SAA, **10**) using EtI, PrI, and Allyl-I, respectively. Conversions were 90%, 50%, and 70% after 14, 24, and 14 h, respectively (Figure 25a). To achieve biocatalytic alkylation via the HMT-MT cascade reaction, we applied the V140T *At*HMT for *in vitro* regeneration of SAE, SAP, and SAA from catalytic amounts of SAH, and MTs transferring the alkyl groups to target substrates (Figure 25b). The V140T-*At*HMT and the T133M-Y326L variant of IeOMT (identified in Chapter 2.2) catalysed regioselective mono-ethylation of luteolin to 3′-*O*-ethylluteolin (**11**) with 41% conversion.[123] Similarly, human

COMT (the soluble form) in cascade with V140T-*At*HMT catalysed regioselective mono-allylation of 3,4-dihydroxybenzaldehyde to produce 4-allyloxy-3-hydroxybenz-aldehyde (**12**) with 48 % conversion.[33] The conversions achieved correspond to 41 and 48 SAH regeneration cycles. The identities of the products were confirmed by ^{1}H and ^{13}C NMR spectroscopy. While a cascade of V140T-*At*HMT and COMT catalysed the propylation of 3,4-dihydroxybenzaldehyde, the conversion was low (~5%) and the product could not be purified for NMR spectroscopy. Therefore, we isolated SAP synthesized using the V140T-*At*HMT and confirmed its structure by ^{1}H and ^{13}C NMR.
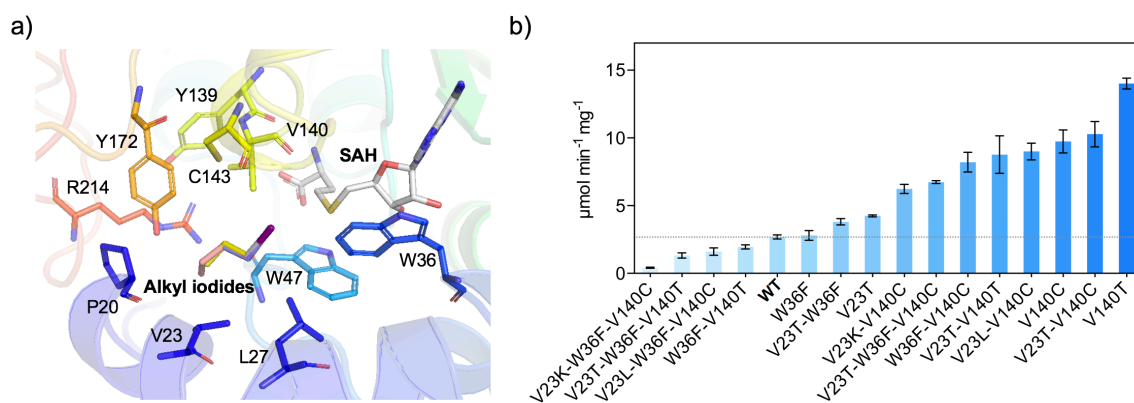
Figure 24. a) The crystal structure of *At*HMT and the modelled alkyl iodides in the active site. The selected mutated sites, SAH and the alkyl iodides which included MeI, EtI, PrI, and BuI are shown as sticks with elemental colouring. The carbons of SAH, MeI, EtI, PrI and BuI are shown in white, grey, purple, yellow and pink, respectively. b) Specific activities of wild-type *At*HMT and several single, double, and triple mutants, determined using EtI as substrate. Data plotted are the means, with standard deviation, of three independent measurements. Figures from Tang *et al.*, 2021.[125]

Thus, we have demonstrated that the scope of biocatalytic alkylations can be rapidly expanded by harnessing the promiscuity of an engineered halide methyltransferase. The HMT from *Arabidopsis thaliana* enabled us to produce three SAM analogues from cheap and readily available alkyl iodides. Importantly, we also demonstrated application of the engineered HMT in bioalkylation cascades. It could catalyse over 40 cycles of alkylation and SAH regeneration, allowing SAH to be used in catalytic rather than stoichiometric amounts. As noted by Liao *et al.*, the use of toxic alkyl iodides as reagents might be considered a drawback of this approach.[130] However, the same applies to chemical alkylation and the biocatalytic alternative functions under milder conditions and offers chemo-, regio-, and stereoselectivity.

Figure 25. a) Preparative-scale (15 mg) synthesis of SAE (**8**), SAP (**9**), and SAA (**10**) catalysed by the V140T-*At*HMT. The "**A**" represents the adenosyl moiety. Conversions shown are calculated from the peak areas of substrates and products from the HPLC chromatograms. b) Preparative-scale (10 mg) synthesis of alkylated products using cyclic MT-HMT cascades, employing 100 µM SAH and 80 mM alkyl iodide. The IeOMT variant T133M-Y326L and the V140T-*At*HMT catalysed the ethylation of luteolin and produced 4′-*O*-ethylluteolin (**11**) with 41% conversion. Human COMT and V140T-*At*HMT catalysed the allylation of 3,4-dihydroxybenzaldehyde and produced 4-allyloxy-3-hydroxybenzaldehyde (**12**) with 48% conversion. Insignificant conversion took place if the HMT and MT were not added. Conversions are calculated from the peak areas of substrates and products from the HPLC chromatograms and numbers of SAH regeneration cycles were calculated as $[product]/[SAH]_{t=0}$. Figures from Tang *et al.*, 2021.[125]

The best variant V140T had the highest activity against all alkyl iodides tested, demonstrating that screening using EtI can increase activity against other alkyl iodides of interest. However, the ethyltransferase activity of V140T is significantly higher than its propyltransferase activity, as 'you get what you screen for'.[131] We screened using EtI to maximize chances of finding improved variants, but the libraries and screening methodology reported here could be used to identify variants with improved activity towards other alkyl iodides. Exploring the intrinsic alkyl-substituent promiscuity of naturally occurring HMTs presents a very promising avenue for future research. Not only might a highly promiscuous variant already exist in nature, but also chances of finding an optimal starting point for directed evolution would be improved by characterizing a larger number of extant proteins. After our work was published, Bengel *et al.* reported the discovery of promiscuous natural HMTs that can transfer up to the butyl group, which has confirmed our prediction.[114]

3. Summary

S-adenosyl-L-methionine (SAM)-dependent methyltransferases (MTs) are important biocatalysts to achieve chemo-, regio- and stereoselective methylation. They display diverse substrate preferences and selectivity according to their functions. By studying their structure-function relationship, we were able to explore their catalytic promiscuity and to engineer them for desired functions. Apart from methylation, biocatalytic alkylations which expand the structural and functional diversity of chemicals are highly desirable.

Article I investigated effects of three specific residues which are involved in substrate discrimination in the *Clarkia breweri* isoeugenol 4-OMT (IeOMT) in the methylation of phenolic compounds. Specifically, the variant T133M improved the enzymatic activities against all tested substrates while the enhancement of regioselectivity was driven by the mutant A134N. The mutants T133M/A134N and T133M/A134N/T135Q benefited from these two mutations, which not only expand the substrate scope but also enhance the regioselectivity of IeOMT. **Article II** describes how the substrate binding residues influence the substrate discrimination between flavonoid OMTs (FOMTs), isoflavonoid OMTs (IOMTs) and phenylpropanoid OMTs (POMTs). The results showed that a mutation T133M of the POMT IeOMT, increased enzyme activity against the flavonoids eriodictyol, naringenin, luteolin, quercetin and even the isoflavonoid genistein, while mutations of the non-conserved residues at positions 322 and 326 affected both the activity and the regioselectivity of the enzyme. On the basis of this work, methylated flavonoids that are rare in nature were produced in high purity.

The work in **Article III** has explored the promiscuity of halide methyltransferases (HMTs) in the synthesis of SAM analogues from *S*-adenosyl-L-homocysteine (SAH) and diverse alkyl iodides. Based on a sensitive and specific colorimetric assay for iodide detection, directed evolution of the *Arabidopsis thaliana* HMT was performed to target the desired mutants. The V140T variant, apart from a five-fold improvement of ethyl transfer activity compared to the wild type, also accepts propyl- and allyl iodide as substrates to produce the corresponding SAM analogues. The regioselective ethylation of luteolin and allylation of 3,4-dihydroxybenzaldehyde, and simultaneous SAM analogue regeneration, were achieved through the HMT-MT one-pot cascade reaction. The production of SAM analogues provides the key cofactors to MTs for carrying out diverse alkylation reactions and enables discovery of promiscuous MTs.

4. References

[1]   K. Bauer, D. Garbe, H. Surburg, *Common fragrance and flavor materials: preparation, properties and uses, 3rd Edition*, John Wiley & Sons, **2008**.

[2]   S. R. Rao, G. A. Ravishankar, *J. Sci. Food Agric.* **2000**, *80*, 289.

[3]   H. Schönherr, T. Cernak, *Angew. Chem., Int. Ed.* **2013**, *52*, 12256.

[4]   E. J. Barreiro, A. E. Kümmerle, C. A. M. Fraga, *Chem. Rev.* **2011**, *111*, 5215.

[5]   J. Quancard, B. Bollbuck, P. Janser, D. Angst, F. Berst, P. Buehlmayer, M. Streiff, C. Beerli, V. Brinkmann, D. Guerini, P. A. Smith, T. J. Seabrook, M. Traebert, K. Seuwen, R. Hersperger, C. Bruns, F. Bassilana, M. Bigaud, *Chem. Biol.* **2012**, *19*, 1142.

[6]   Z. D. Smith, A. Meissner, *Nat. Rev. Genet.* **2013**, *14*, 204.

[7]   M. V. C. Greenberg, D. Bourc'his, *Nat. Rev. Mol. Cell Biol.* **2019**, *20*, 590.

[8]   E. L. Greer, Y. Shi, *Nat. Rev. Genet.* **2012**, *13*, 343.

[9]   C. D. Davis, E. O. Uthus, *Exp. Biol. Med.* **2004**, *229*, 988.

[10]   D. H. K. Lim, E. R. Maher, *Obstet. Gynaecol.* **2010**, *12*, 37.

[11]   Y. Fu, D. Dominissini, G. Rechavi, C. He, *Nat. Rev. Genet.* **2014**, *15*, 293.

[12]   D. Y. Lee, C. Teyssier, B. D. Strahl, M. R. Stallcup, *Endocr. Rev.* **2005**, *26*, 147.

[13]   D. S. Goldstein, *Clin. Auton. Res.* **2010**, *20*, 331.

[14]   I. J. Kopin, *Pharmacol. Rev.* **1985**, *37*, 333.

[15]   C. R. Creveling, N. Dalgard, H. Shimizu, J. W. Daly, *Mol. Pharmacol.* **1970**, *6*, 691.

[16]   S. G. Lee, Y. Kim, T. D. Alpert, A. Nagata, J. M. Jez, *J. Biol. Chem.* **2012**, *287*, 1426.

[17]   C. Zubieta, X. Z. He, R. A. Dixon, J. P. Noel, *Nat. Struct. Mol. Biol.* **2001**, *8*, 271.

[18]   J. L. Martin, J. Begun, M. J. McLeish, J. M. Caine, G. L. Grunewald, *Structure* **2001**, *9*, 977.

[19]   C. J. Liu, B. E. Deavours, S. B. Richard, J. L. Ferrer, J. W. Blount, D. Huhman, R. A. Dixon, J. P. Noel, *Plant Cell* **2006**, *18*, 3656.

[20]   A. W. Struck, M. L. Thompson, L. S. Wong, J. Micklefield, *ChemBioChem* **2012**, *13*, 2642.

[21]   R. W. Woodard, M. D. Tsai, H. G. Floss, P. A. Crooks, J. K. Coward, *J. Biol. Chem.* **1980**, *255*, 9124.

[22]   D. O'Hagan, J. W. Schmidberger, *Nat. Prod. Rep.* **2010**, *27*, 900.

[23]   T. Pavkov-Keller, K. Steiner, M. Faber, M. Tengg, H. Schwab, M. Gruber-Khadjawi, K. Gruber, *Plos One* **2017**, *12*, 1.

[24]   P. M. Dewick, *Medicinal natural products: a biosynthetic approach, 3rd Edition*, John Wiley & Sons, **2002**.

[25]   D. K. Liscombe, G. V. Louie, J. P. Noel, *Nat. Prod. Rep.* **2012**, *29*, 1238.

[26]   L. Wessjohann, A.-K. Bauer, M. Dippe, J. Ley, T. Geißler, in *Applied Biocatalysis: From*

Fundamental Science to Industrial Applications (Eds.: L. Hilterhaus, A. Liese, U. Kettling and G. Antranikian), John Wiley & Sons, **2016**, pp. 121–146.

[27]   B. G. Kim, S. H. Sung, Y. Chong, Y. Lim, J. H. Ahn, *J. Plant Biol.* **2010**, *53*, 321.

[28]   L. C. Ward, H. V. McCue, A. J. Carnell, *ChemCatChem* **2020**, DOI: 10.1002/cctc.202001316.

[29]   R. K. Ibrahim, A. Bruneau, B. Bantignies, *Plant Mol. Biol.* **1998**, *36*, 1.

[30]   K. C. Lam, R. K. Ibrahim, B. Behdad, S. Dayanandan, *Genome* **2007**, *50*, 1001.

[31]   J. P. Noel, R. A. Dixon, E. Pichersky, C. Zubieta, J. L. Ferrer, in *Recent advances in phytochemistry* (Ed.: J. T. Romeo), Elsevier, **2003**, pp. 37–58.

[32]   C. P. Joshi, V. L. Chiang, *Plant Mol. Biol.* **1998**, *37*, 663.

[33]   P. T. Männistö, S. Kaakkola, *Pharmacol. Rev.* **1999**, *51*, 593.

[34]   P. T. Männistö, in *Advances in pharmacology* (Eds.: D. S. Goldstein, G. Eisenhofer, and R. McCarty), Elsevier, **1997**, pp. 324–328.

[35]   T. Lotta, J. Vidgren, C. Tilgmann, I. Ulmanen, K. Melen, I. Julkunen, J. Taskinen, *Biochemistry* **1995**, *34*, 4202.

[36]   D. Guo, F. Chen, K. Inoue, J. W. Blount, R. A. Dixon, *Plant Cell* **2001**, *13*, 73.

[37]   J. Ralph, K. Lundquist, G. Brunow, F. Lu, H. Kim, P. F. Schatz, J. M. Marita, R. D. Hatfield, S. A. Ralph, J. H. Christensen, *Phytochem. Rev.* **2004**, *3*, 29.

[38]   L. B. Davin, N. G. Lewis, *Curr. Opin. Biotechnol.* **2005**, *16*, 398.

[39]   E. Pichersky, J. P. Noel, N. Dudareva, *Science* **2006**, *311*, 808.

[40]   E. Pichersky, J. Gershenzon, *Curr. Opin. Plant Biol.* **2002**, *5*, 237.

[41]   T. S. Kaufman, *J. Braz. Chem. Soc.* **2015**, *26*, 1055.

[42]   J. Wang, E. Pichersky, *Arch. Biochem. Biophys.* **1998**, *349*, 153.

[43]   J. Wang, E. Pichersky, *Arch. Biochem. Biophys.* **1999**, *368*, 172.

[44]   C. Kandaswami, E. Middleton, in *Free radicals in diagnostic medicine* (Ed.: D. Armstrong), Springer, **1994**, pp. 351–376.

[45]   D. Procházková, I. Boušová, N. Wilhelmová, *Fitoterapia* **2011**, *82*, 513.

[46]   S. Kaur, P. Mondal, *J. Microbiol. Exp.* **2014**, *1*, 23.

[47]   A. García-Lafuente, E. Guillamón, A. Villares, M. A. Rostagno, J. A. Martínez, *Inflammation Res.* **2009**, *58*, 537.

[48]   M. H. Pan, C. S. Lai, C. T. Ho, *Food Funct.* **2010**, *1*, 15.

[49]   M. K. Chahar, N. Sharma, M. P. Dobhal, Y. C. Joshi, *Pharmacogn. Rev.* **2011**, *5*, 1.

[50]   D. Raffa, B. Maggio, M. V. Raimondi, F. Plescia, G. Daidone, *Eur. J. Med. Chem.* **2017**, *142*, 213.

[51]   I. V. Pavlidis, M. Gall, T. Geißler, E. Gross, U. T. Bornscheuer, in *Applied Biocatalysis: From Fundamental Science to Industrial Applications* (Eds.: L. Hilterhaus, A. Liese, U. Kettling and G. Antranikian), John Wiley & Sons, **2016**, pp. 179–198.

[52] A. N. Panche, A. D. Diwan, S. R. Chandra, *J. Nutr. Sci.* **2016**, *5*, 1.

[53] M. L. Falcone Ferreyra, S. Rius, P. Casati, *Front. Plant Sci.* **2012**, *3*, 222.

[54] T. Walle, *Mol. Pharm.* **2007**, *4*, 826.

[55] S. H. Thilakarathna, H. P. Rupasinghe, *Nutrients* **2013**, *5*, 3367.

[56] N. Koirala, N. H. Thuan, G. P. Ghimire, D. V. Thang, J. K. Sohng, *Enzyme Microb. Technol.* **2016**, *86*, 103.

[57] H. Takemura, T. Itoh, K. Yamamoto, H. Sakakibara, K. Shimoi, *Bioorg. Med. Chem.* **2010**, *18*, 6310.

[58] H. Takemura, H. Uchiyama, T. Ohura, H. Sakakibara, R. Kuruto, T. Amagai, K. Shimoi, *J. Steroid Biochem. Mol. Biol.* **2010**, *118*, 70.

[59] J. P. Ley, U.S. Patent 8,685,436, **2014**.

[60] J. P. Ley, G. Krammer, G. Reinders, I. L. Gatfield, H. J. Bertram, *J. Agric. Food Chem.* **2005**, *53*, 6061.

[61] C. Zubieta, P. Kota, J. L. Ferrer, R. A. Dixon, J. P. Noel, *Plant Cell* **2002**, *14*, 1265.

[62] D. H. Kim, B. G. Kim, Y. Lee, J. Y. Ryu, Y. Lim, H. G. Hur, J. H. Ahn, *J. Biotechnol.* **2005**, *119*, 155.

[63] B. G. Kim, B. R. Jung, Y. Lee, H. G. Hur, Y. Lim, J. H. Ahn, *J. Agric. Food Chem.* **2006**, *54*, 823.

[64] N. J. Walton, A. Narbad, C. Faulds, G. Williamson, *Curr. Opin. Biotechnol.* **2000**, *11*, 490.

[65] J. T. Nelson, J. Lee, J. W. Sims, E. W. Schmidt, *Appl. Environ. Microbiol.* **2007**, *73*, 3575.

[66] J. Siegrist, J. Netzer, S. Mordhorst, L. Karst, S. Gerhardt, O. Einsle, M. Richter, J. N. Andexer, *FEBS Lett.* **2017**, *591*, 312.

[67] B. J. C. Law, M. R. Bennett, M. L. Thompson, C. Levi, S. A. Shepherd, D. Leys, J. Micklefield, *Angew. Chem., Int. Ed.* **2016**, *55*, 2683.

[68] K. Goedecke, M. Pignot, R. S. Goody, A. J. Scheidig, E. Weinhold, *Nat. Struct. Biol.* **2001**, *8*, 121.

[69] A. Jeltsch, *Biol. Chem.* **2001**, *382*, 707.

[70] P. A. Boriack-Sjodin, K. K. Swinger, *Biochemistry* **2016**, *55*, 1557.

[71] A. A. McCarthy, J. G. McCarthy, *Plant Physiol.* **2007**, *144*, 879.

[72] H. Uefuji, Y. Tatsumi, M. Morimoto, P. Kaothien-Nakayama, S. Ogita, H. Sano, *Plant Mol. Biol.* **2005**, *59*, 221.

[73] H. Minami, J. S. Kim, N. Ikezawa, T. Takemura, T. Katayama, H. Kumagai, F. Sato, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 7393.

[74] J. M. Hagel, P. J. Facchini, *Plant Cell Physiol.* **2013**, *54*, 647.

[75] M. R. Bennett, M. L. Thompson, S. A. Shepherd, M. S. Dunstan, A. J. Herbert, D. R.

References

M. Smith, V. A. Cronin, B. R. K. Menon, C. Levy, J. Micklefield, *Angew. Chem., Int. Ed.* **2018**, *57*, 10600.

[76] B. Giese, *Angew. Chem., Int. Ed.* **1983**, *22*, 753.

[77] A. V. Subrahmanyam, S. Thayumanavan, G. W. Huber, *ChemSusChem* **2010**, *3*, 1158.

[78] C. Mahlert, F. Kopp, J. Thirlway, J. Micklefield, M. A. Marahiel, *J. Am. Chem. Soc.* **2007**, *129*, 12011.

[79] Y. T. Huang, S. Y. Lyu, P. H. Chuang, N. S. Hsu, Y. S. Li, H. C. Chan, C. J. Huang, Y. C. Liu, C. J. Wu, W. B. Yang, T. L. Li, *ChemBioChem* **2009**, *10*, 2480.

[80] M. A. Skiba, A. P. Sikkema, W. D. Fiers, W. H. Gerwick, D. H. Sherman, C. C. Aldrich, J. L. Smith, *ACS Chem. Biol.* **2016**, *11*, 3319.

[81] S. Kishimoto, Y. Tsunematsu, T. Matsushita, K. Hara, H. Hashimoto, Y. Tang, K. Watanabe, *Biochemistry* **2019**, *58*, 3933.

[82] I. Crnovcic, R. Sussmuth, U. Keller, *Biochemistry* **2010**, *49*, 9698.

[83] S. C. Chen, C. H. Huang, S. J. Lai, J. S. Liu, P. K. Fu, S. T. Tseng, C. S. Yang, M. C. Lai, T. P. Ko, Y. Chen, *Sci. Rep.* **2015**, *5*, 1.

[84] M. Tengg, H. Stecher, L. Offner, K. Plasch, F. Anderl, H. Weber, H. Schwab, M. Gruber-Khadjawi, *ChemCatChem* **2016**, *8*, 1354.

[85] M. Pacholec, J. H. Tao, T. W. Walsh, *Biochemistry* **2005**, *44*, 14969.

[86] H. Stecher, M. Tengg, B. J. Ueberbacher, P. Remler, H. Schwab, H. Griengl, M. Gruber-Khadjawi, *Angew. Chem., Int. Ed.* **2009**, *48*, 9546.

[87] J. K. Coward, M. D'Urso-Scott, W. D. Sweet, *Biochem. Pharmacol.* **1972**, *21*, 1200.

[88] M. Baudry, F. Chast, J. C. Schwartz, *J. Neurochem.* **1973**, *20*, 13.

[89] S. C. Lu, *Int. J. Biochem. Cell Biol.* **2000**, *32*, 391.

[90] S. Mordhorst, J. Siegrist, M. Müller, M. Richter, J. N. Andexer, *Angew. Chem., Int. Ed.* **2017**, *56*, 4037.

[91] C. Liao, F. P. Seebeck, *Nat. Catal.* **2019**, *2*, 696.

[92] Y. Nagatoshi, T. Nakamura, *Plant Biotechnol.* **2007**, *24*, 503.

[93] I. R. McDonald, K. L. Warner, C. McAnulla, C. A. Woodall, R. S. Oremland, J. C. Murrell, *Environ. Microbiol.* **2002**, *4*, 193.

[94] J. M. Attieh, A. D. Hanson, H. S. Saini, *J. Biol. Chem.* **1995**, *270*, 9250.

[95] I. J. McKean, P. A. Hoskisson, G. A. Burley, *ChemBioChem* **2020**, *21*, 2890.

[96] T. D. Huber, B. R. Johnson, J. Zhang, J. S. Thorson, *Curr. Opin. Biotechnol.* **2016**, *42*, 189.

[97] K. Hartstock, B. S. Nilges, A. Ovcharenko, N. V. Cornelissen, N. Pullen, A. M. Lawrence-Dorner, S. A. Leidel, A. Rentmeister, *Angew. Chem., Int. Ed.* **2018**, *57*, 6342.

[98] W. Peters, S. Willnow, M. Duisken, H. Kleine, T. Macherey, K. E. Duncan, D. W. Litchfield, B. Lüscher, E. Weinhold, *Angew. Chem., Int. Ed.* **2010**, *49*, 5170.

[99]   S. Wu, R. Snajdrova, J. C. Moore, K. Baldenius, U. T. Bornscheuer, *Angew. Chem., Int. Ed.*, **2020**, DOI: 10.1002/anie.202006648.

[100]   B. Hauer, *ACS Catal.* **2020**, *10*, 8418.

[101]   C. Dalhoff, G. Lukinavičius, S. Klimašauskas, E. Weinhold, *Nat. Chem. Biol.* **2006**, *2*, 31.

[102]   B. J. Law, A.-W. Struck, M. R. Bennett, B. Wilkinson, J. Micklefield, *Chem. Sci.* **2015**, *6*, 2885.

[103]   J. N. Andexer, A. Rentmeister, *Nat. Chem.* **2020**, *12*, 791.

[104]   A. J. Herbert, S. A. Shepherd, V. A. Cronin, M. R. Bennett, R. Sung, J. Micklefield, *Angew. Chem., Int. Ed.* **2020**, *59*, 14950.

[105]   J. Kim, H. Xiao, J. B. Bonanno, C. Kalyanaraman, S. Brown, X. Tang, N. F. Al-Obaidi, Y. Patskovsky, P. C. Babbitt, M. P. Jacobson, Y.-S. Lee, S. C. Almo, *Nature* **2013**, *498*, 123.

[106]   I. J. W. McKean, J. C. Sadler, A. Cuetos, A. Frese, L. D. Humphreys, G. Grogan, P. A. Hoskisson, G. A. Burley, *Angew. Chem., Int. Ed.* **2019**, *58*, 17583.

[107]   R. Borchardt, Y. S. Wu, *J. Med. Chem.* **1976**, *19*, 1099.

[108]   S. Khani-Oskouee, J. P. Jones, R. W. Woodard, *Biochem. Biophys. Res. Commun.* **1984**, *121*, 181.

[109]   F. Wang, S. Singh, J. Zhang, T. D. Huber, K. E. Helmich, M. Sunkara, K. A. Hurley, R. D. Goff, C. A. Bingman, A. J. Morris, *FEBS J.* **2014**, *281*, 4224.

[110]   M. Thomsen, S. B. Vogensen, J. Buchardt, M. D. Burkart, R. P. Clausen, *Org. Biomol. Chem.* **2013**, *11*, 7606.

[111]   M. Dippe, W. Brandt, H. Rost, A. Porzel, J. Schmidt, L. A. Wessjohann, *Chem. Commun.* **2015**, *51*, 3637.

[112]   F. Michailidou, N. Klocker, N. V. Cornelissen, R. K. Singh, A. Peters, A. Ovcharenko, D. Kummel, A. Rentmeister, *Angew. Chem., Int. Ed.* **2020**, DOI: 10.1002/anie.202012623.

[113]   S. Singh, J. Zhang, T. D. Huber, M. Sunkara, K. Hurley, R. D. Goff, G. Wang, W. Zhang, C. Liu, J. Rohr, *Angew. Chem., Int. Ed.* **2014**, *53*, 3965.

[114]   L. L. Bengel, B. Aberle, A.-N. Egler-Kemmerer, S. Kienzle, B. Hauer, S. C. Hammer, **2020**, DOI: 10.1002/anie.202014239.

[115]   J. Kim, H. Xiao, J. B. Bonanno, C. Kalyanaraman, S. Brown, X. Tang, N. F. Al-Obaidi, Y. Patskovsky, P. C. Babbitt, M. P. Jacobson, *Nature* **2013**, *498*, 123.

[116]   R. T. Byrne, F. Whelan, P. Aller, L. E. Bird, A. Dowle, C. M. C. Lobley, Y. Reddivari, J. E. Nettleship, R. J. Owens, A. A. Antson, D. G. Waterman, *Acta Crystallogr., Sect. D: Struct. Biol.* **2013**, *69*, 1090.

[117]   J. Deen, C. Vranken, V. Leen, R. K. Neely, K. P. F. Janssen, J. Hofkens, *Angew. Chem.,*

Int. Ed. **2017**, *56*, 5182.

[118] R. Wang, M. K. Luo, *Curr. Opin. Chem. Biol.* **2013**, *17*, 729.

[119] K. Islam, *Cell Chem. Biol.* **2018**, *25*, 1171.

[120] Q. Tang, U. T. Bornscheuer, I. V. Pavlidis, *ChemCatChem* **2019**, *11*, 3227.

[121] G. V. Louie, M. E. Bowman, Y. Tu, A. Mouradov, G. Spangenberg, J. P. Noel, *Plant Cell* **2010**, *22*, 4114.

[122] M. W. Bhuiya, C. J. Liu, *J. Biol. Chem.* **2010**, *285*, 277.

[123] Q. Tang, Y. M. Vianney, K. Weisz, C. W. Grathwol, A. Link, U. T. Bornscheuer, I. V. Pavlidis, *ChemCatChem* **2020**, *12*, 3721.

[124] N. Itoh, H. Toda, M. Matsuda, T. Negishi, T. Taniguchi, N. Ohsawa, *BMC Plant Biol.* **2009**, *9*, 116.

[125] Q. Tang, C. Grathwol, A. S. Aslan-Üzel, S. Wu, A. Link, I. V. Pavlidis, C. P. S. Badenhorst, U. T. Bornscheuer, *Angew. Chem., Int. Ed.* **2021**, *60*, 1524.

[126] P. M. Bozeman, D. B. Learn, E. L. Thomas, *J. Immunol. Methods* **1990**, *126*, 125.

[127] A. S. Aslan-Üzel, A. Beier, D. Kovář, C. Cziegler, S. K. Padhi, E. D. Schuiten, M. Dörr, D. Böttcher, F. Hollmann, F. Rudroff, M. D. Mihovilovic, T. Buryška, J. Damborský, Z. Prokop, C. P. S. Badenhorst, U. T. Bornscheuer, *ChemCatChem* **2020**, *12*, 2032.

[128] J. W. Schmidberger, A. B. James, R. Edwards, J. H. Naismith, D. O'Hagan, *Angew. Chem., Int. Ed.* **2010**, *49*, 3646.

[129] M. T. Reetz, J. D. Carballeira, *Nat. Protoc.* **2007**, *2*, 891.

[130] C. Liao, F. P. Seebeck, *Angew. Chem., Int. Ed.* **2020**, *59*, 7184.

[131] L. You, F. Arnold, *Protein Eng., Des. Sel.* **1994**, *9*, 77.

Author contributions

Article I   **Specific residues expand the substrate scope and enhance the regioselectivity of a plant *O*-methyltransferase**

Q. Tang, U. T. Bornscheuer, I. V. Pavlidis, *ChemCatChem* **2019**, *11*, 3227.

I. V. P. and U. T. B. initiated the study and supervised the project. Q. T. planned and performed the experiments. I. V. P. performed the computational analysis. Q. T. and I. V. P. wrote the manuscript with input from U. T. B. The manuscript was revised and approved by all authors.

Article II   **Influence of substrate binding residues on the substrate scope and regioselectivity of a plant *O*-methyltransferase against flavonoids**

Q. Tang, Y. M. Vianney, K. Weisz, C. W. Grathwol, A. Link, U. T. Bornscheuer, I. V. Pavlidis, *ChemCatChem* **2020**, *12*, 3721.

Q. T., I. V. P. and U. T. B. initiated the study, I. V. P. and U. T. B. supervised the project. Q. T. planned and performed the biochemical experiments. Y. M. V. and K. W. conducted the NMR experiments and analysed the data. C. W. G. and A. L. performed the preparative HPLC. I. V. P. performed the computational analysis. Q. T. and I. V. P. wrote the manuscript with input from U. T. B. The manuscript was revised and approved by all authors.

Article III   **Directed evolution of a halide methyltransferase enables biocatalytic synthesis of diverse SAM analogues**

Q. Tang, C. W. Grathwol, A. S. Aslan-Üzel, S. Wu, A. Link, I. V. Pavlidis, C. P. S. Badenhorst, U. T. Bornscheuer, *Angew. Chem., Int. Ed.* **2021**, *60*, 1524; *Angew. Chem.* **2021**, *133*, 1547.

C. P. S. B. and Q. T. developed the concept for this work and C. P. S. B., U. T. B. and I. V. P. supervised the project. Q. T. and C. P. S. B. planned, and Q. T. performed the biochemical experiments. C. W. G. and A. L. performed the preparative HPLC, conducted the NMR experiments and analysed the data. A. S. A.-Ü. and C. P. S. B. established the

iodide assay. S. W. performed the SAH nucleosidase gene knockout in *E. coli*. I. V. P. performed the computational analysis. Q. T. and C. P. S. B. wrote the manuscript with input from U. T. B. and I. V. P., The manuscript was revised and approved by all authors.

Qingyun Tang   Prof. Dr. Uwe T. Bornscheuer

Articles

Article I

Specific Residues Expand the Substrate Scope and Enhance the Regioselectivity of a Plant O-Methyltransferase

Qingyun Tang,[a] Uwe T. Bornscheuer,*[a] and Ioannis V. Pavlidis*[b]

An isoeugenol 4-O-methyltransferase (IeOMT), isolated from the plant *Clarkia breweri*, can be engineered to a caffeic acid 3-O-methyltransferase (CaOMT) by replacing three consecutive residues. Here we further investigated functions of these residues by constructing the triple mutant T133M/A134N/T135Q as well as single mutants of each residue. Phenolics with different chain lengths and different functional groups were investigated. The variant T133M improves the enzymatic activities against all tested substrates by providing beneficial interactions to residues which directly interact with the substrate. Mutant A134N significantly enhanced the regioselectivity. It is *meta*-selective or even specific against most of the tested substrates but *para*-specific towards 3,4-dihydroxybenzoic acid. The triple mutant T133M/A134N/T135Q benefits from these two mutations, which not only expand the substrate scope but also enhance the regioselectivity of IeOMT. On the basis of our work, regiospecific methylated phenolics can be produced in high purity by different IeOMT variants.

Introduction

Methylation is an important and ubiquitous process both in biological activities and chemical synthesis. Through methylation, the physicochemical properties (hydrophobicity, solubility, pK values), odor and taste, or the bioactivity of a compound can be altered.[1] Various biological activities and diseases are regulated through the methylation of nucleotides, proteins and hormones.[2] Methylation can be achieved by a specific enzyme class, methyltransferases (MTs, E.C. 2.1.1.x), using the biomolecule *S*-adenosyl-L-methionine (SAM) as methyl donor. Depending on the atoms to be methylated, MTs can be classified into *C*-, *O*-, *N*-, *S*-, *Se*-, *As*- and also halide ion MTs, among which *O*-MTs are the most abundant class.[3] Compared to chemical methylation methods, enzymatic methylation processes are environmentally friendly alternatives and can be superior due to their chemo-, regio- and stereoselectivity. Thus, methyltransferases are an attractive class of biocatalysts for industrial applications. So far, their potential is not utilized in industry, mainly due to two reasons. The first is relevant to the methyl donor: SAM is expensive to be provided in equimolar amounts to the substrate, while it is unstable under mild conditions. Studies on efficient *in vitro* regeneration of SAM, or the usage of other methyl donors were published recently.[4] The second reason is relevant to the reactions catalyzed: MTs were considered able to only perform methylations. However, it was shown recently that either the wild-type enzymes can have catalytic promiscuity or protein engineering could help to expand the range of substrates to be accepted.[5]

Mammalian catechol *O*-MTs (COMTs) are widely studied due to their importance in the regulation of neurotransmitters and hormones (L-DOPA, dopamine, noradrenaline, adrenaline etc.), as well as their roles as drug targets.[2c,6] COMT accepts a wide range of catechol derivatives as substrates and the majority of COMTs preferably methylate the *meta*-hydroxyl groups (*meta*-OH or 3-OH; Scheme 1).[2c,7] They have been utilized as catalysts in the production of the flavor compound vanillin (**4b**).[8] Only a few COMTs were characterized to have *para*-preference (4-OH), which serves as a regio-complementary catalyst to the common COMTs and enables the synthesis of uncommon methylated products.[9] The regioselectivity of COMT can also be altered through protein engineering. For instance, Law and coworkers engineered several regiocomplementary mutants of the rat COMT to provide access to alkylated catechols (i.e. using 3,4-dihydroxybenzoic acid, 3,4-dihydroxyphenylacetic acid and 4-nitrocatechol as substrates) while they could explain the reason for the altered regioselectivity by binding models.[5d]

Caffeic acid/5-hydroxyferulic acid 3/5—O—MTs (CaOMTs) are key enzymes in the biosynthesis of lignin precursors found in all lignin-producing plants.[10] They catalyze methylation of the lignin precursors caffeyl alcohol, caffeic aldehyde, caffeic acid (**3a**) and 5-hydroxyconiferyl alcohol/aldehyde/acid exclusively at the *meta*-OH, leaving the *para*-OH free to form lignin polymers and to avoid oxidative radical coupling.[11] These compounds are also the building blocks of chalcones and flavonoids.[12]

Wang et al. have isolated a CaOMT and an isoeugenol *O*-methyltransferase (IeOMT) from *Clarkia breweri*.[13] Though the two OMTs share 83% amino acid sequence identity, they display distinct substrate preference as well as regioselectivity. While CaOMT serves to produce *meta*-methylated lignin precursors, as the other CaOMTs do, IeOMT only accepts isoeugenol (**1b**) or eugenol (**2b**) as substrates and methylates

[a] Q. Tang, Prof. U. T. Bornscheuer
     Dept. of Biotechnology and Enzyme Catalysis
     Institute of Biochemistry, University of Greifswald
     Felix-Hausdorff-Straße 4, 17487 Greifswald (Germany)
     E-mail: uwe.bornscheuer@uni-greifswald.de
[b] Prof. I. V. Pavlidis
     Dept. of Chemistry, University of Crete
     Voutes University Campus, 70013 Heraklion (Greece)
     E-mail: ipavlidis@uoc.gr
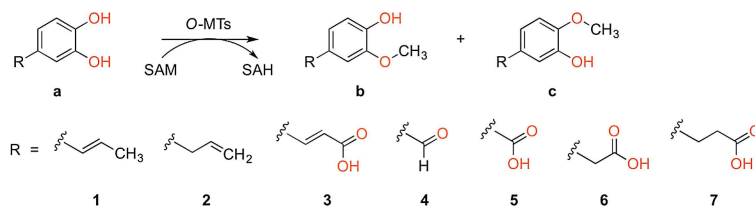
Scheme 1. *O*-MTs catalyze methylation of phenolic substrates (**a**) to give *meta*- (**b**) or *para*- (**c**) methylated products. *S*-Adenosyl-L-methionine (SAM) is provided as methyl donor and SAH (*S*-adenosyl-L-homocysteine) is generated after the methyl group donation. Structures of different substrates are shown, based on the R groups and the phenolic moieties **a**, **b** or **c**.

the sole free *para*-OH to produce the floral scent compounds methyl isoeugenol and methyl eugenol. Substitution of the amino acid cluster T133-A134-T135 in IeOMT with the corresponding residues in CaOMT (M135-N136-Q137) enabled IeOMT to acquire CaOMT character, regarding the substrate specificity as well as the methylation regiospecificity.[14] However, the mechanism of the substrate discrimination has not been elucidated yet.

Herein, we investigated in detail the effects of these residues on the enzymatic function of IeOMT. Our results show that the substrate scope of IeOMT can be expanded and the regioselectivity can be enhanced by specific mutations.

Results and Discussion

Investigation of the Critical Residues in IeOMT

As pointed out above, we aimed to study the effects of residues 133–135 in IeOMT and thus the triple mutant T133M/A134N/T135Q and the respective single mutants at each position were constructed and expressed in *E. coli.* Activities of purified enzymes against isoeugenol (**1b**), eugenol (**2b**) and caffeic acid (**3a**) were studied. As shown in Figure 1, the triple mutant T133M/A134N/T135Q of IeOMT has increased methylation activities against all substrates. Specifically, the single mutant T133M has a significant effect on the enhancement of the activities, while the T135Q largely impairs enzymatic activity. The double mutant T133M/A134N lacking T135Q was constructed to avoid its adverse effects on the enzyme activity. Nevertheless, the double mutant did not exhibit improved activity compared to the triple mutant.

IeOMT is considered a 4-*O*-MT as it displays significant activity towards the solely available *para*-OH of isoeugenol and eugenol.[14] However, it is not regioselective against caffeic acid (**3a**) which contains both free *meta*- and *para*-hydroxyl groups (Figure 1). The triple mutant T133M/A134N/T135Q has conferred *meta*-specificity to IeOMT against caffeic acid, as reported by Wang et al.[14] Interestingly, the mutants T133M and A134N at adjacent positions gave different regioselectivities against caffeic acid. T133 M increases the preference for the *para*-OH, leading to the major production of isoferulic acid (**3c**), while A134N is strictly *meta*-specific. This selectivity is not substrate concentration-dependent, as the $K_{M,obs}$ of IeOMT_T133M is lower

than the working concentration (Table 1). Especially, for the reaction with caffeic acid (**3a**), the ratio of products **3c** to **3b** was always between 3.1 to 3.4, in substrate concentrations ranging from 0.05 mM to 4 mM (data not shown). The effects of increasing enzyme activity observed for T133M and the regiospecificity of A134N are complementary in mutants T133M/A134N/T135Q and T133M/A134N, which produce pure ferulic acid (**3b**) with considerable conversions. A CaOMT from *Catharanthus roseus* (CaRo_MT) containing the same M-N-Q amino acid cluster in the corresponding positions was also investigated. It is highly active and *meta*-specific against its intrinsic substrate caffeic acid while showing minor activity against isoeugenol.

Table 1. Observed kinetic constants of IeOMT_T133M against isoeugenol (**1b**), caffeic acid (**3a**) and 3,4-dihydroxybenzaldehyde (**4a**).

Substrate	$K_{M,obs}$ [mM]	$k_{cat,obs}$ [s^{-1}]	$k_{cat,obs}/K_{M,obs}$ [s$^{-1} \cdot$ mM^{-1}]
1b	0.0058 ± 0.0003	479 ± 7	82586
3a	0.719 ± 0.035	64 ± 1	89
4a	0.119 ± 0.014	226 ± 12	1899

Investigation of the Substrate Scope

Since some of the IeOMT mutations led to increased enzyme activity and regioselectivity, we further investigated their functions against other phenolics. 3,4-Dihydroxyphenolic compounds having different chain lengths and different functional groups were chosen (**4a**–**7a**, Scheme 1, Figure 1). Similar to the reactions with isoeugenol and caffeic acid, T133M/A134N/T135Q, T133M/A134N and the T133M variant show improved activities against all substrates compared to the wild-type IeOMT. This is more obvious for 3,4-dihydroxybenzaldehyde (**4a**) and 3,4-dihydroxybenzoic acid (**5a**) which have a one-carbon aldehyde and a carboxylic group attached to the phenyl ring, respectively. Activities against 3,4-dihydroxyphenylacetic acid (**6a**) and 3,4-dihydroxyphenylpropanoic acid (**7a**), having longer carboxylic side chains, were not much improved. While CaRo_MT is highly active to its intrinsic substrate caffeic acid (**3a**), it also shows high activity against 3,4-dihydroxyphenylpropanoic acid (**7a**) with the three-carbon saturated carboxylic side chain.
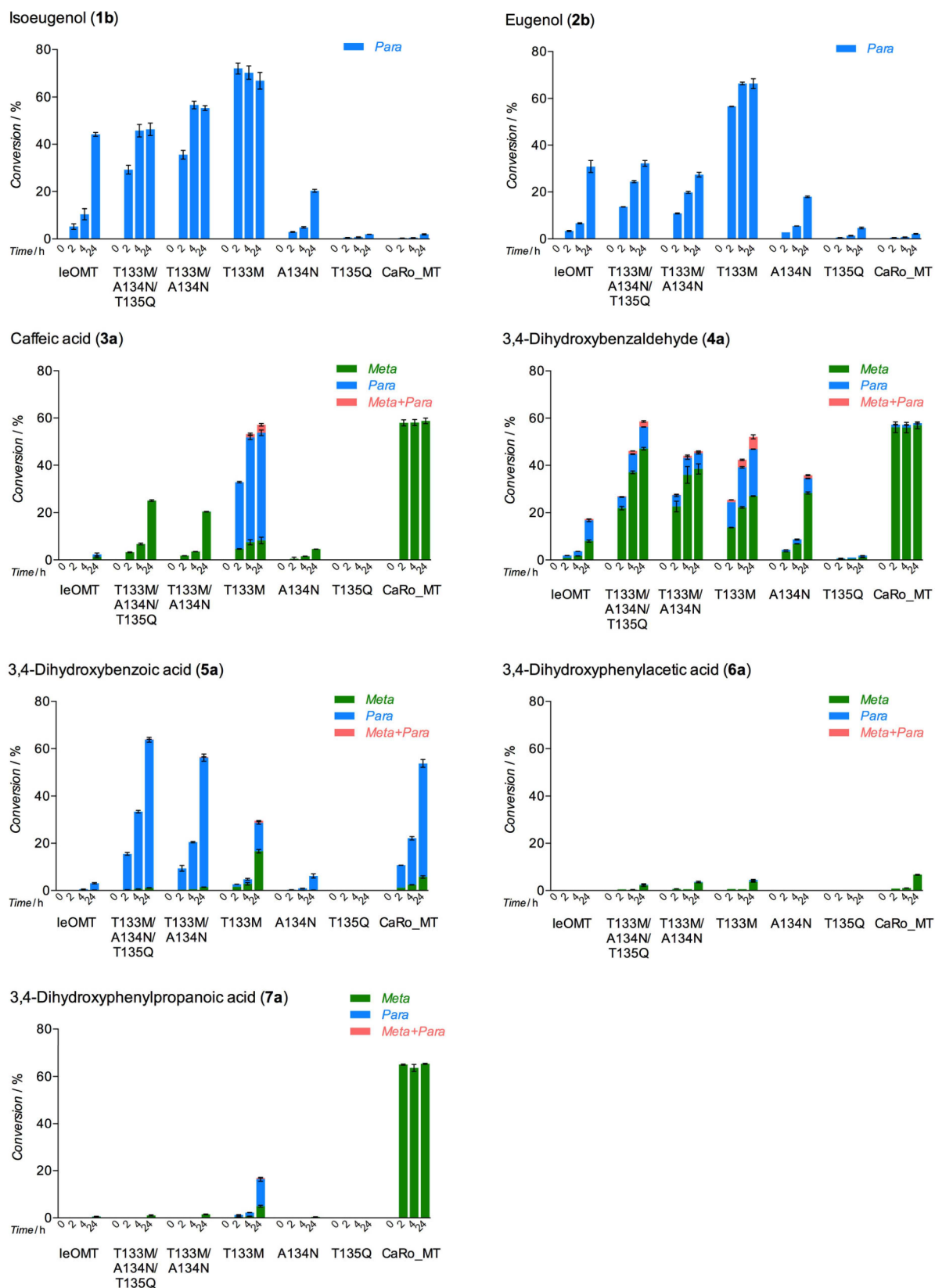
Figure 1. Methylation efficiency and regioselectivity of IeOMT and its mutants against different substrates. Values obtained for negative controls (without enzyme) were subtracted.

Interestingly, while T133M/A134N/T135Q, T133M/A134N and A134N variants are *meta*-selective against 3,4-dihydroxybenzaldehyde (**4a**) and produce mainly vanillin (**4b**), they show *para*-specificity towards 3,4-dihydroxybenzoic acid (**5a**). CaRo_MT also shows a totally different regiospecificity against these two substrates.

Structural Analysis

The T133M Variant Exhibits Increased Methylation Activity by Providing Beneficial Interactions to the First Sphere of Interaction Residues

The crystal structure of a *Clarkia breweri* IeOMT variant (T133L, E165I and F175I) with isoeugenol and SAH bound in the active site has been determined by Bhuiya and coworkers.[15] The structure is in the closed conformation, referring to the active catalytic states.[16] To perform bioinformatic analysis on the wild-type IeOMT, as well as our mutants, we performed a structure refinement (25 °C, pH 7.5) and back mutations to the wild-type sequence *in silico* and a subsequent energy minimization. As shown in Figure 2a, isoeugenol is surrounded and
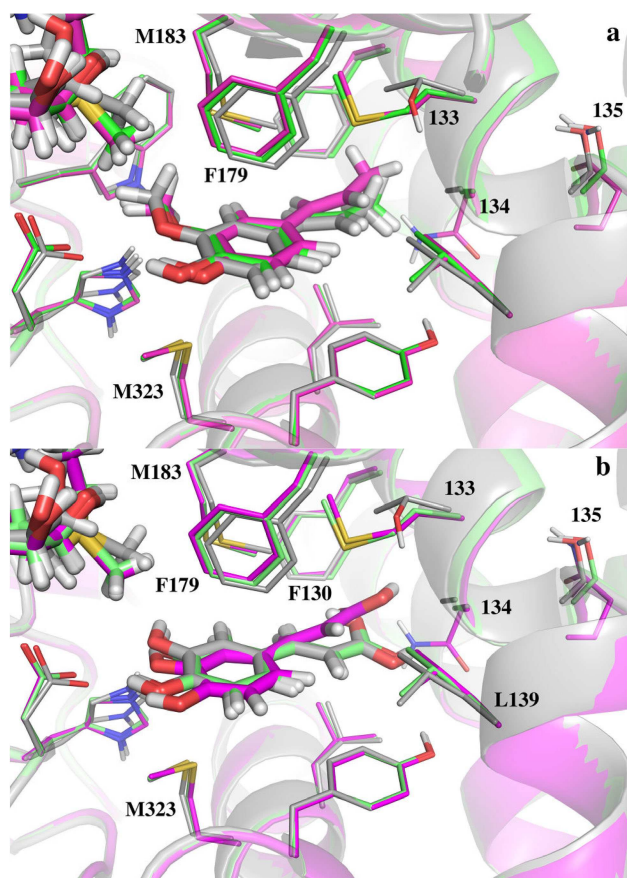
Figure 2. Models of *Clarkia breweri* IeOMT wild-type (grey), T133M (green) and T133M/A134N/T135Q (magenta) with SAM and isoeugenol (a) and caffeic acid (b) in the active site. The residues interacting with the substrate are shown as lines, the most important ones are labeled.

stabilized by a series of hydrophobic residues. Among them, two Met residues (M183 and M323) are highly conserved in plant OMTs and play a major role in encapsulating and orienting the hydrophobic and aromatic substrates.[17] It is reported that the -SH/π interaction between these methionines and the aromatic ring plays an important role in stabilizing the structure,[19] among with aromatic residues such as F179.

However, the methionine at position 133 (T133M) located above the propenyl chain of isoeugenol is far away from this system. It seems that its hydrophobicity among with the higher flexibility of the side chain of methionine compared to threonine, provides a beneficial environment for this hydrophobic chain of isoeugenol (Figure 2a). Moreover, through a residue interaction network (RIN) analysis we observed the interactions between the three residues of interest (133–135) and the substrate. In this analysis we observed that the residue at position 133 (T or M) does not directly interact with the substrate, but it acts supportively to the residues that directly interact with the substrate. The most notable interactions established are the ones with the F179 residue, which as we mentioned stabilizes the aromatic substrates at the active site. Thus, it seems that the T133M mutation provides the proper orientation of the phenylalanine to bring the substrate to the proper position. The distance of the methyl group of SAM to the oxygen of the hydroxyl group of isoeugenol decreased from 6.1 Å in the wild-type to 5.0 Å in the T133M mutant. This is also supported by literature, as it has been reported that T133M of IeOMT increases catalytic efficiency against coniferyl alcohol by improving substrate affinity.[19]

In our case we observed the increased enzymatic activities of the T133M mutant also against caffeic acid (**3a**), 3,4-dihydroxybenzaldehyde (**4a**) and 3,4-dihydroxybenzoic acid (**5a**). It is interesting to note that CaRo_MT, which also has the "MNQ" motif, is highly active against caffeic acid (**3a**), 3,4-dihydroxybenzaldehyde (**4a**), 3,4-dihydroxybenzoic acid (**5a**) and 3,4-dihydroxyphenylpropanoic acid (**7a**). This supportive role of the position 133 can also explain the reason why the beneficial effect was also observed with substrates with more polar R-groups, such as caffeic acid (**3a**) that has a carboxylic function. As seen in Figure 2b, M133 interacts better with F130, L139, F179, and even establishes interactions with M183, while in the wild-type T133 interacts weakly only with the first three residues.

Interesting is the effect of the position 135, which is further away from the substrate and its side chain does not even point towards the binding site. It seems that the negative effect the T135Q mutation has on the activity is primarily related to secondary interactions, which potentially break the interactions established with the substrate. We hypothesize that the position 134, due to backbone interactions with the position 135, can negate the negative effect and that this is the reason why we did not observe this inactivation in the triple mutant.

Mechanism of Regioselectivity

Both IeOMT and CaOMT belong to the type I plant *O*-MT family and catalyze methylation via a general acid/base-mediated S_N2-type nucleophilic substitution.[15–16,20] The substrate hydroxyl groups will be deprotonated by the catalytic residues and carry out a nucleophilic attack on the reactive methyl group of SAM afterwards. The mechanism of regioselectivity can be better explained by the superimposed active sites of the *Clarkia breweri* IeOMT (4-*O*-MT) and *Perennial Ryegrass* CaOMT (3-*O*-MT; Figure 3). Since the substrate is stabilized by a series of
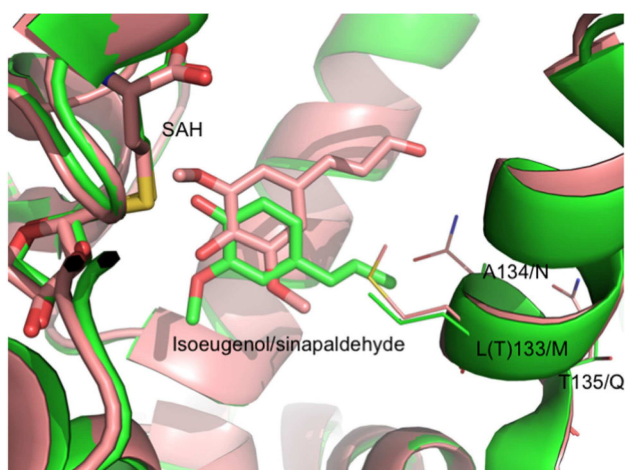
Figure 3. Superimposed active sites of the *Clarkia breweri* IeOMT (green) and *Perennial Ryegrass* CaOMT (PDB ID: 3p9i; pink; with SAH and sinapaldehyde bound). The mutated sites in IeOMT and the corresponding residues in *Perennial Ryegrass* CaOMT are shown. Interactions between the substrate side chain and residues 133 and 134 affect the orientation of the substrate. The hydroxyl group which is closer to the sulfur atom of SAH (elemental coloring) will be preferably methylated.

hydrophobic residues in the enzyme active site, the interaction between the substrate side chain and the adjacent residues affects the orientation of the substrate.

Depending on whether the *meta*- or the *para*-hydroxyl group is closer to the methyl group of SAM they are preferably methylated and thus determine the regioselectivity.[20c] Since the added Met residue in mutant T133M is located above the isoeugenol side chain and stabilizes it through hydrophobic interaction, it attracts the substrate side chain to move downwards (from the view of Figure 3) and makes the *para*-hydroxyl group well-exposed to SAM. Similarly, T133M is also *para*-selective to caffeic acid. However, T133M does not show obvious regioselectivity against 3,4-dihydroxy-benzaldehyde (**4a**) or 3,4-dihydroxybenzoic acid (**5a**), probably because these substrates have shorter side chains and therefore are more flexible in the enzyme active site.

In mutant A134N, the mutated residue has steric hindrance with the substrate side chain and expels it upwards (from the view of Figure 3) and hence the substrate positions its *meta*-hydroxyl group to the methyl donor SAM. This makes the variants T133M/A134N/T135Q, T133M/A134N and A134N strictly

meta-specific towards caffeic acid. Interestingly, while these mutants are *meta*-selective against 3,4-dihydroxybenzaldehyde (**4a**), they are *para*-specific to 3,4-dihydroxybenzoic acid (**5a**). In our model (Figure 2b) the residue N134 of the triple mutant is the only one that directly interacts with the caffeic acid, forming a hydrogen bond and thus probably orienting it in the proper position for the observed *para*-selectivity.

The regioselectivity of the mammalian COMT is regulated through a different mechanism. Unlike IeOMT, in which the substrate can rotate in the active site pocket, the two hydroxyl groups of the catechol derivative substrate are coordinated and stabilized by the COMT-bound Mg^{2+} ion and only one hydroxyl group is exposed to SAM and being methylated.[5d,21] The substituent on the catechol moiety can be located at either the *meta*- or *para*-position of the methylated hydroxyl group. Law et al. have reported that polar and charged substrate side chains tend to orientate towards the solvent and it results in *meta*-methylation, while neutral or hydrophobic substituents are more likely to approach the hydrophobic part of the active site and promote *para*-methylation.[5d] Therefore, the regioselectivity of COMT depends on the property of the substrate side chain and its relative position to the methylated hydroxyl group.

Conclusions

In our study, we discovered that a triple mutation T133M/A134N/T135Q in an isoeugenol *O*-MT, which is involved in substrate discrimination, exhibits expanded substrate scope and also enhanced regioselectivity. Further investigations showed that variant T133M improved the enzymatic activities against all tested compounds by providing beneficial interactions to residues which directly interact with the substrate. The enhancement of regioselectivity derives from mutant A134N, which drives the enzyme *meta*-selective to most of the tested phenolics and *para*-specific towards 3,4-dihydroxybenzoic acid. As a result, ferulic acid (**3b**), isoferulic acid (**3c**), vanillin (**4b**) and isovanillic acid (**5c**) can be produced in high purity by different IeOMT variants.

Experimental Section

Materials

*Dpn*I and *S*-adenosyl-L-methionine (32 mM solution) were purchased from New England Biolabs, Inc. All oligonucleotides were purchased from ThermoFisher Scientific. Competent cells were self prepared. All other chemicals were purchased from Sigma or TCI and were of analytical grade or higher purity.

Gene Construction, Mutagenesis, Expression and Purification

The codon optimized genes (for expression in *E. coli*) for IeOMT from *Clarkia breweri* (accession number of protein: O04385) and for CaOMT from *Catharanthus roseus* (accession number of protein: Q8W013) with C-terminal His6-tag were synthesized and subcloned

into the *E. coli* expression vector pET-21a(+) by GenScript USA Inc.. Mutagenesis to generate the triple mutant T133M/A134N/T135Q was carried out with a modified version of the FastCloning method, by amplifying the whole pET-21a(+)-IeOMT vector using primers with 18 bp overlapping at their 5′ ends.[22] Other mutants were constructed following the QuikChange (Stratagene) protocol. Primers are as below:

T133M/A134N/T135Q
fw: 5′-ATGAATCAAGACAAGGTTCTGCTGGAGCCGTGGTTC-3′;
rv: 5′-AACCTTGTCTTGATTCATCAGCAGAAACGGCGCCAG-3′;
T133M/A134N
fw: 5′-GCCGTTTCTGCTGATGAACACCGACAAGGTTCTG-3′;
rv: 5′-CAGAACCTTGTCGGTGTTCATCAGCAGAAACGGC-3′;
T133M
fw: 5′-CGTTTCTGCTGATGGCGACCGACAAGG-3′;
rv: 5′-CCTTGTCGGTCGCCATCAGCAGAAACG-3′;
A134N
fw: 5′-CTGCTGACCAACACCGACAAGGTTC-3′;
rv: 5′-GAACCTTGTCGGTGTTGGTCAGCAG-3′;
T135Q
fw: 5′-GCTGACCGCGCAAGACAAGGTTCTGC-3′;
rv: 5′-GCAGAACCTTGTCTTGCGCGGTCAGC-3′.

The wild-type IeOMT templates in PCR products were digested by *Dpn*I before transforming into *E. coli* Top 10 chemically competent cells for plasmid amplification. After confirming the correct constructs by DNA sequencing (Eurofins, Germany), the newly constructed mutant plasmids were transformed into *E. coli* Bl21 (DE3) chemically competent cells for protein expression. The *E. coli* containing IeOMT or mutant plasmids were grown in terrific broth (TB) containing 100 µg/mL ampicillin at 37 °C until A_{600nm} reached 0.8–1.0. After induction with 0.2 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), the cultures were grown overnight (18–22 h) at 20 °C. Then the cells were harvested by centrifugation, suspended in lysis buffer (20 mM NaH_2PO_4, 500 mM NaCl, 40 mM imidazole, pH 7.5) and lysed by French press. After centrifugation, the supernatant was loaded to a Ni-NTA gravity flow column (Carl Roth, Germany), washed with 10 bed volumes of lysis buffer and the targeted proteins were eluted with elution buffer (20 mM NaH_2PO_4, 500 mM NaCl, 300 mM imidazole, pH 7.5). Then the purified proteins were desalted by Sephadex G-25 in PD-10 desalting columns (GE healthcare, USA) and stored in 50 mM sodium phosphate buffer, pH 7.5. The protein purities were confirmed by SDS-PAGE.

Enzyme Assays and HPLC Analysis

A typical reaction mixture was composed of 0.1 mg/mL purified enzyme, 5 mM DTT, 1 mM substrate, 1 mM SAM and 25 % *E. coli* cell lysate, in 50 mM phosphate buffer, pH 7.5. Negative controls were also performed without adding MTs. Assays were performed in triplicate. Reactions were incubated at 28 °C with 1000 rpm agitation in an Eppendorf Thermomixer. Samples were taken after 2, 4 and 24 h reaction time and an equal volume of acetonitrile was added to quench the reactions. The samples were centrifuged at full speed for 30 min to remove protein precipitate and then 200 µL supernatant were transferred to an HPLC sample vial inserts prior to HPLC analysis. Analyses were performed on a VWR Hitachi Elite LaChrom system equipped with a Kinetex EVO C18 (4.6×250 mm column, 5 µm particle size, Phenomenex) reversed-phase column. The mobile phase for the separation of isoeugenol (**1 b**) and the methylated product was 10 mM sodium 1-heptanesulfonate, 20 mM phosphate, pH 4.4 and 50 % (v/v) acetonitrile; 50 % water and methanol were used for detection of eugenol (**2 b**) and methyl eugenol; 0.1 % acetic acid and acetonitrile were used for the separation of caffeic acid (**3 a**) (18 % acetonitrile), 3,4-dihydroxyben-

zoic acid (**5 a**) (10 % acetonitrile), 3,4-dihydroxyphenylacetic acid (**6 a**) (20 % acetonitrile) and 3,4-dihydroxyphenylpropanoic acid (**7 a**) (20 % acetonitrile) and their methylated products. A modified linear gradient with 0–2 min 5 %–20 % methanol, 2–11 min 23 % methanol, 11–13 min 95 % methanol, 13–16 min 5 % methanol and lasted until 20 min, mixed with 0.1 % acetic acid was applied for the separations of 3,4-dihydroxybenzaldehyde (**4 a**) and products.[5d] Wavelengths for the detections of isoeugenol, eugenol, caffeic acid, 3,4-dihydroxybenzaldehyde and the phenolic acids were 260, 280, 320, 320 and 280 nm, respectively. All analyses were performed at a flow rate 1 mL/min and the column temperature was 35 °C. The identities of *meta*-, *para*- and double-methylated products were confirmed by comparison with chromatographic elution times of commercial standards. Substrate and product concentrations were determined by comparing the peak areas to the calibration curves of each compound and the conversions were calculated. The kinetic constants ($K_{M,obs}$) of IeOMT_T133 M were determined at various substrate concentrations and 1 mM SAM under the same reaction condition as in the enzyme assays using appropriate amounts of purified enzyme. Initial reaction rates were measured and fitted to the Michaelis-Menten model using GraphPad Prism 6.0 (GraphPad Software Inc.) to determine the $V_{max, obs}$ and $K_{M,obs}$ values for each substrate. Then $k_{cat,obs}$ was defined by dividing $V_{max,obs}$ by the enzyme concentration and $k_{cat,obs}/K_{M,obs}$ was calculated.

Bioinformatic Analysis

The bioinformatic analysis was performed with YASARA 18.11.21. First the structure of 3REO was back-mutated to its wild-type sequence and the SAH was transformed to SAM by the addition of the methyl group and this structure was refined at pH 7.5, 25 °C for 500 ps, taking a snapshot every 25 ps. The structure with the lowest energy was selected for the further experiments. For the mutants, the respective amino acids were swapped with subsequent energy minimization. The same was performed for the exchange of substrate. As we knew the binding pattern of isoeugenol, we swapped the atoms of the R-group to form the caffeic acid and subsequently minimized the energy. The figures were prepared with Pymol. The RIN analysis was performed with RINalyzer according to literature, with the use of Cytoscape 3.7 and Chimera 1.13.[23]

Conflict of Interest

The authors declare no conflict of interest.

[1] a) N. Koirala, N. H. Thuan, G. P. Ghimire, D. V. Thang, J. K. Sohng, *Enzyme Microb. Technol.* **2016**, *86*, 103–116; b) S. A. Heleno, A. Martins, M. J. R. Queiroz, I. C. Ferreira, *Food Chem.* **2015**, *173*, 501–513.

[2] a) C. D. Davis, E. O. Uthus, *Exp. Biol. Med.* **2004**, *229*, 988–995; b) A. E. McBride, P. A. Silver, *Cell* **2001**, *106*, 5–8; c) P. T. Männistö, S. Kaakkola, *Pharmacol. Rev.* **1999**, *51*, 593–628.

[3] A. W. Struck, M. L. Thompson, L. S. Wong, J. Micklefield, *ChemBioChem* **2012**, *13*, 2642–2655.

[4] a) S. Mordhorst, J. Siegrist, M. Müller, M. Richter, J. N. Andexer, *Angew. Chem. Int. Ed.* **2017**, *56*, 4037–4041; *Angew. Chem.* **2017**, 129, 4095–4099; b) J. C. Sadler, L. D. Humphreys, R. Snajdrova, G. A. Burley, *ChemBioChem* **2017**, *18*, 992–995; c) J. E. Farnberger, N. Richter, K. Hiebler, S. Bierbaumer, M. Pickl, W. Skibar, F. Zepeck, W. Kroutil, *Comm. Chem.*, **2018**, 1, Art. No. 82; d) N. Richter, J. Farnberger, S. Pompei, C. Grimm, W. Skibar, F. Zepeck, W. Kroutil, *Adv. Synth. Catal.*, **2019**, doi: 10.1002/adsc.201801590.

[5] a) S. Singh, J. Zhang, T. D. Huber, M. Sunkara, K. Hurley, R. D. Goff, G. Wang, W. Zhang, C. Liu, J. Rohr, *Angew. Chem. Int. Ed.* **2014**, *53*, 3965–3969; *Angew. Chem.*, **2014**, 126, 4046–4050; b) B. W. K. Lee, H. G. Sun, T. Zang, B. J. Kim, J. F. Alfaro, Z. S. Zhou, *J. Am. Chem. Soc.* **2010**, *132*, 3642–3643; c) M. Tengg, H. Stecher, L. Offner, K. Plasch, F. Anderl, H. Weber, H. Schwab, M. Gruber-Khadjawi, *ChemCatChem* **2016**, *8*, 1354–1360; d) B. J. C. Law, M. R. Bennett, M. L. Thompson, C. Levy, S. A. Shepherd, J. Leys, J. Micklefield, *Angew. Chem. Int. Ed.* **2016**, *55*, 2683–2687; *Angew. Chem.* **2016**, *128*, 2733–2737.

[6] I. J. Kopin, *Pharmacol. Rev.* **1985**, *37*, 333–364.

[7] C. R. Creveling, N. Dalgard, H. Shimizu, J. W. Daly, *Mol. Pharmacol.* **1970**, *6*, 691–696.

[8] N. J. Walton, A. Narbad, C. Faulds, G. Williamson, *Curr. Opin. Biotechnol.* **2000**, *11*, 490–496.

[9] a) J. T. Nelson, J. Lee, J. W. Sims, E. W. Schmidt, *Appl. Environ. Microbiol.* **2007**, *73*, 3575–3580; b) J. Siegrist, J. Netzer, S. Mordhorst, L. Karst, S. Gerhardt, O. Einsle, M. Richter, J. N. Andexer, *FEBS Lett.* **2017**, *591*, 312–321.

[10] D. Guo, F. Chen, K. Inoue, J. W. Blount, R. A. Dixon, *Plant Cell* **2001**, *13*, 73–88.

[11] a) J. Ralph, K. Lundquist, G. Brunow, F. Lu, H. Kim, P. F. Schatz, J. M. Marita, R. D. Hatfield, S. A. Ralph, J. H. Christensen, *Phytochem. Rev.*

2004, *3*, 29–60; b) L. B. Davin, N. G. Lewis, *Curr. Opin. Biotechnol.* **2005**, *16*, 398–406.

[12] P. M. Dewick, *Medicinal natural products: a biosynthetic approach*, John Wiley & Sons, **2002**.

[13] J. Wang, E. Pichersky, *Arch. Biochem. Biophys.* **1998**, *349*, 153–160.

[14] J. Wang, E. Pichersky, *Arch. Biochem. Biophys.* **1999**, *368*, 172–180.

[15] M. W. Bhuiya, C. J. Liu, .*http://www.rcsb.org/pdb, pdb-code: 3REO*

[16] G. V. Louie, M. E. Bowman, Y. Tu, A. Mouradov, G. Spangenberg, J. P. Noel, *Plant Cell* **2010**, *22*, 4114–4127.

[17] C. Zubieta, X. Z. He, R. A. Dixon, J. P. Noel, *Nat. Struct. Mol. Biol.* **2001**, *8*, 271.

[18] a) C. C. Valley, A. Cembran, J. D. Perlmutter, A. K. Lewis, N. P. Labello, J. Gao, J. N. Sachs, *J. Biol. Chem.* **2012**, *287*, 34979–34991; b) C. R. Forbes, S. K. Sinha, H. K. Ganguly, S. Bai, G. P. Yap, S. Patel, N. J. Zondlo, *J. Am. Chem. Soc.* **2017**, *139*, 1842–1855.

[19] M. W. Bhuiya, C. Liu, *J. Biol. Chem.* **2010**, *285*, 277–285.

[20] a) J. P. Noel, R. A. Dixon, E. Pichersky, C. Zubieta, J. L. Ferrer, in *Recent Advances in Phytochemistry, Vol. 37*, Elsevier, **2003**, pp. 37–58; b) D. O'Hagan, J. W. Schmidberger, *Nat. Prod. Rep.* **2010**, *27*, 900–918; c) D. K. Liscombe, G. V. Louie, J. P. Noel, *Nat. Prod. Rep.* **2012**, *29*, 1238–1250.

[21] J. Vidgren, L. A. Svensson, A. Liljas, *Nature* **1994**, *368*, 354.

[22] C. Li, A. Wen, B. Shen, J. Lu, Y. Huang, Y. Chang, *BMC Biotechnol.* **2011**, *11*, 92.

[23] N. T. Doncheva, Y. Assenov, F. S. Domingues, M. Albrecht, *Nat. Protoc.* **2012**, *7*, 670–685.

Article II

ChemCatChem

Full Papers
doi.org/10.1002/cctc.202000471

Chem
Europ
European (
Societies P

Check for
updates

Influence of Substrate Binding Residues on the Substrate Scope and Regioselectivity of a Plant *O*-Methyltransferase against Flavonoids

Qingyun Tang,[a] Yoanes M. Vianney,[a] Klaus Weisz,[a] Christoph W. Grathwol,[b] Andreas Link,[b] Uwe T. Bornscheuer,*[a] and Ioannis V. Pavlidis*[c]

Dedication to Prof. Marko Mihovilovic on the occasion of his 50th birthday.

Methylation of free hydroxyl groups is an important modification for flavonoids. It not only greatly increases absorption and oral bioavailability of flavonoids, but also brings new biological activities. Flavonoid methylation is usually achieved by a specific group of plant *O*-methyltransferases (OMTs) which typically exhibit high substrate specificity. Here we investigated the effect of several residues in the binding pocket of the *Clarkia breweri* isoeugenol OMT on the substrate scope and regioselectivity against flavonoids. The mutation T133M, identi-fied as reported in our previous publication, increased the activity of the enzyme against several flavonoids, namely eriodictyol, naringenin, luteolin, quercetin and even the iso-flavonoid genistein, while a reduced set of amino acids at positions 322 and 326 affected both, the activity and the regioselectivity of the methyltranferase. On the basis of this work, methylated flavonoids that are rare in nature were produced in high purity.

Introduction

Flavonoids are a large group of natural polyphenols and secondary metabolites from plants. They attract a lot of attention due to their nutritional, health-beneficial and pharma-cological properties including free radical-scavenging antioxida-tive activities, anti-inflammatory, antimicrobial and anticancer activities.[1] Flavonoids are classified into flavonoids (2-phenyl-benzopyrans), isoflavonoids (3-phenylbenzopyrans) and neo-flavonoids (4-phenylbenzopyrans) based on the position of the phenyl ring. Flavonoids are further divided into flavanes, flavanones, flavonols, flavones and anthocyanins (Figure 1).[2] Various modification reactions such as oxidation, hydroxylation, glycosylation and methylation lead to a huge variety of flavonoids.[3] In particular, *O*-methylation of free hydroxyl

groups on dietary flavonoids greatly increases their absorption and oral bioavailability through the improvement of their metabolic stability and better membrane transportation.[4] Methylation of flavonoids can also bring new biological activities. For example, chrysoeriol (4′,5,7-trihydroxy-3′-meth-oxyflavone, 3′-methylluteolin) and isohamnetin (3,4′,5,7-tetrahy-droxy-3′-methoxyflavone, 3′-methylquercetin) show strong and selective inhibition on the formation of a carcinogenic estrogen metabolite related to breast cancer.[5] Besides, eriodictyol (3′,4′,5,7-tetrahydroxyflavanone) and homoeriodictyol (4′,5,7-trihydroxy-3′-methoxyflavanone, 3′-methyleriodictyol) are known by their remarkable bitter masking effect.[6]

Methylation of hydroxyl moieties can be achieved by *S*-adenosyl-ʟ-methionine (SAM)-dependent *O*-methyltransferases (OMTs). The methyl group provided by SAM is transferred to the free hydroxyl group and the methyl ether derivative is produced. According to Noel *et al.*, OMTs isolated from plants are categorized into three classes based on sequence align-ments and structural studies.[7] Type I OMTs accept phenyl-propanoids, flavonoids, isoflavonoids or chalcones as substrates with molecular weights ranging from 40 to 43 kDa and are not bivalent cation-dependent. Type II and type III OMTs methylate the phenylpropanoid-CoA derivatives and the carboxylic acid, respectively. Although phenylpropanoid OMTs (POMTs), flavo-noid OMTs (FOMTs) and isoflavonoid OMTs (IOMTs) which all belong to type I OMTs are highly similar regarding their sequences and protein structures, they are divided into differ-ent OMT groups based on their substrate specificity. POMTs catalyze methylation of caffeic alcohol/aldehyde/acid which products then serve for lignin and flavonoid biosynthesis.[8] FOMTs accept a wide range of flavonoid substrates and show strict regioselectivity. Isoflavonoids are exclusively produced in leguminous plants and function as phytoalexins. Only SOMT2

[a] Q. Tang, Y. M. Vianney, Prof. K. Weisz, Prof. U. T. Bornscheuer
     Institute of Biochemistry, University of Greifswald
     Felix-Hausdorff-Str. 4
     17489 Greifswald (Germany)
     E-mail: uwe.bornscheuer@uni-greifswald.de
[b] C. W. Grathwol, Prof. Dr. A. Link
     Institute of Pharmacy, University of Greifswald
     Friedrich-Ludwig-Jahn-Str. 17
     17489 Greifswald (Germany)
[c] Prof. I. V. Pavlidis
     Dept. of Chemistry
     University of Crete
     Voutes University Campus
     70013 Heraklion (Greece)
     E-mail: ipavlidis@uoc.gr

ChemCatChem

Full Papers
doi.org/10.1002/cctc.202000471

Chemistry Europe
European Chemical
Societies Publishing

Figure 1. Structures of the flavanones eriodictyol and naringenin, the flavone luteolin, the flavonol quercetin and the isoflavone genistein. Ring assignment and backbone atoms numbering are shown in the structure of eriodictyol.

from soy bean (*Glycine max*) and SaOMT2 from *Streptomyces avermitilis* were discovered to show both flavonoids and isoflavonoids methylation activities.[9] The crystal structure of an IOMT with isoformononetin (4′-hydroxy-7-methoxyisoflavone) in the active site shows that some residues are critical for the stabilization of isoflavonoids.[10] However, no crystal structure of FOMT is yet available.

In order to study the substrate discrimination between the plant OMTs, we compared the substrate binding residues of POMT, FOMT and IOMT with different substrate preferences and regioselectivities. Based on the differences of these residues, we have designed and constructed mutants based on the isoeugenol POMT from *Clarkia breweri* including the mutation T133M (IeOMT_T133M). In a previous work of ours,[11] we identified this variant to have expanded substrate scope and altered regioselectivity against phenolic compounds. Thus, it was selected as the starting point for this study. We discovered that some mutants specifically methylated the 3′-OH of the tested flavonoids and others further methylated the 4′-OH, producing dimethylated flavonoids. The tested mutants also brought activities and different regioselectivities to the isoflavonoid genistein.

Figure 2. Part of the sequence alignment of plant flavonoid OMTs (FOMTs), isoflavonoid OMTs (IOMTs) and phenylpropanoid OMTs (POMTs) performed with Geneious 10.0.2 (the whole sequence alignment is shown in Figure S1). Residues involved in substrate binding are highlighted in blue. Numbering of residues are based on the sequence of the *Clarkia breweri* isoeugenol OMT (IeOMT). OMTs are named by their original organisms and regioselectivity. MePi, *Mentha piperita;* MeTr, *Medicago truncatula*; OrSa, *Oryza sativa*; ArTh, *Arabidopsis thaliana*; ChAm, *Chrysosplenium americanum*; CaRo, *Catharanthus roseus*; GlMa, *Glycine max*; CiAr, *Cicer arietinum*; MeSa, *Medicago sativa*; GlEc, *Glycyrrhiza echinate*; LoJa, *Lotus japonicus*; ClBr, *Clarkia breweri*; LoPe, *Lolium perenne*. MeTr_7-FOMT7 is a putative IOMT but it has higher preference against naringenin (flavanone) than isoflavonoids.

Results and Discussion

Substrate binding residues of different plant OMTs

Methylation reactions mostly take place at the 7-, 3′- and 4′-hydroxyl groups of flavonoids, the 7- and 4′-hydroxyl groups of isoflavonoids and the 3- and 4-hydroxyl groups of phenylpropanoids. In order to discover the factors determining the substrate discrimination of plant OMTs, we chose 21 plant OMTs from different plant species with different substrate preferences and regioselectivities for comparison. Sequence alignment shows that the sequences are extremely diverse between different plant OMTs, with only 4.6% identity (Figure 2, Figure S1). Since several IOMT crystal structures have been resolved but no FOMT structure is available, we chose MeSa-7/4′-IOMT which has been obtained in a catalytic conformation (PDB code: 1FP2) and investigated based on it the substrate-enzyme interactions.[10] The ligand isoformononetin, the 7-methylated product of the isoflavonone daidzein, is situated in the enzyme active site and well-stabilized by multiple interactions. Residues critical for substrate binding are

highlighted in Figure 1. Met183 and Met323 constrain the aromatic A-ring and help positioning the 7-hydroxyl group to the catalytic residue His272 and SAM. These two residues are conserved throughout the plant OMT superfamily. Zubieta *et al.* suggested that the interaction of the ketone group in the C-ring is stabilized by the amide side chain of Asn322, a residue that is only conserved among the 7-IOMTs. Other OMTs rather have middle size hydrophobic residues, namely Ile, Val or Met, at this position. Leu326 also interacts with the C-ring of isoformononetin, but it locates closer to the ether oxygen. Residues at this position are quite different between the selective OMTs. They are either the hydrophobic Leu, Val, Ile or Met, or the basic residues Arg or His. In the absence of Asn322, the basic residues might play an important role in stabilizing the C-ring ketone group. The accommodation of the isoflavone B-ring is achieved by Cys133 and Val134. Although these two residues in other plant OMTs are quite diverse, they mostly have Gly/Leu/Met and Ala/Val/Asn at these two positions, respectively. These substrate binding residues bring proper binding patterns to their preferred substrates and thus determine substrate specificity and regioselectivity of different

ChemCatChem

Full Papers
doi.org/10.1002/cctc.202000471

**Chemistry
Europe**
European Chemical
Societies Publishing

plant OMTs. In order to investigate the influence of these residues, we constructed mutants L322H/N/M and Y326H/R/L using the IeOMT-T133M variant as template. IeOMT is a phenyl-propanoid 4-OMT isolated from *Clarkia breweri* and the variant T133M has been proved to expand the substrate scope and to enhance the regioselectivity.[11]

Substrate scope and enzyme activities

Flavonoids are classified into flavane, flavanone, flavone and flavonol, while isoflavonoids are divided into isoflavone and isoflavanone, depending on their structures. We have chosen several commonly known compounds eriodictyol and naringenin (flavanones), luteolin (flavone), quercetin (flavonol) and genistein (isoflavone) as substrates (Figure 1). Since optically pure flavanones will racemize in aqueous solution, we used racemic eriodictyol and naringenin as substrates and obtained racemic products.[12] Activities of the wild type IeOMT and designed mutants were tested against these substrates. In each reaction, a molar excess of SAM and 25% (v/v) *Escherichia coli* (*E. coli*) cell lysate, which contains *S*-adenosyl-L-homocysteine (SAH) nucleosidase, were provided in order to increase the yield.[13] Products were confirmed by either comparing their retention times on HPLC to commercial standards or structurally characterized via NMR and MS. Catalytic performance of the mutants are presented in area percentages calculated by peak area of both substrates and products measured by HPLC (Figure 3, Figure S2).

It needs to be stated that in an initial screening performed, the IeOMT_T133M exhibited an expanded substrate scope compared to the wild type (converting naringenin and genistein), while it also exhibited altered regioselectivity. For this reason, the variant IeOMT_T133M was selected as the template for the designed mutations.

Although IeOMT is a phenylpropanoid OMT, it shows high activities against the flavonoids eriodictyol, luteolin and quercetin and is highly regiospecific to the 3′-hydroxyl group. In the absence of the 3′-hydroxyl group, the wild-type IeOMT displays very low activity, as seen against naringenin. The variant T133M has a universal effect of increasing the enzymatic activity against these substrates. Moreover, this variant further methylates the 4′-hydroxyl groups, producing 3′,4′-dimethylated products. Mutants T133M/Y326H and T133M/Y326L further enabled increased production of 3′,4′-dimethylated eriodictyol, luteolin and quercetin. On the other hand, the variant T133M/L322N retains the regiospecificity against the 3′-hydroxyl group of all tested flavonoids and so does the mutant T133M/Y326R towards luteolin and quercetin. It is interesting to note that the T133M, T133M/L322M as well as other variants of IeOMT, instead of methylating the 4′-OH of naringenin, they methylate the 7-OH. We speculate that the product 2 from naringenin, further produced by variant T133M/Y326L after 48 h, could be another methylated product or further methylation on the 7-methylnaringenin has occurred.

The wild-type IeOMT displayed very minor activity towards the isoflavone genistein and the mutants showed different regioselectivity. While T133M, T133M/L322H (or N or M-substitutions) and T133M/Y326R exhibited higher selectivity against the 4′-hydroxyl group, T133M/Y326H and T133M/Y326L also produced the 7-methylated genistein. Product 3 was produced by mutants which displayed both 7- and 4′-methylation activities especially by variant T133M/Y326L after 48 h. According to the RP-HPLC chromatograms, it was eluted later than the single methylated genistein (Figure S2). Therefore, we assume that this is the 7,4′-dimethylated genistein.

Interestingly, both naringenin and genistein missing the 3′-OH group lead to 7-methylation, but when the B-ring shifts from position 2 to position 3 of the C-ring, the regioselectivity of some mutants shifted from the 7-hydroxyl group to the 4′-hydroxyl group.

Structural analysis

To gain an insight into the structural differences that lead to the different regioselectivity of the variants, we performed *in silico* analysis. As seen in Figure 4A, the wild-type IeOMT can accommodate eriodictyol, luteolin and quercetin in an orientation that the methylation of the 3′-position is favored. In all three cases the distance of the oxygen of 3′-hydroxyl group to the methyl group to be transferred is between 3.2 and 3.5 Å. It is interesting to note that eriodictyol and luteolin bind in a similar orientation, however, quercetin seems to be a little tilted in comparison to SAM, which brings also the 4′-position in closer proximity to the transfer group (3.3 Å to the 3′-OH and 3.4 Å to the 4′-OH group) and thus the wild-type IeOMT can also produce some dimethylated quercetin. In the case of the double mutant T133M/Y326L, the double methylation is increased for all three substrates. As seen in Figure 4B, the double mutation enabled the inverse binding of these three flavonoids in the binding pocket, bringing the 4′-hydroxyl group in the proximity of the methyl group of the SAM and thus the double methylation is favored. The reason for this inversed binding seems to be the more hydrophobic character of the introduced Y326L, which cannot accommodate the carbonyl of 5-position.

In the case of genistein, the T133M mutation increased the activity to a detectable level, but the mutations at positions 322 and 326 do not further increase the activity of the enzyme. However, the double mutant T133M/Y326H has a shift of its regioselectivity to position 7. As seen in Figure 4C, the substrate is bound with ring A facing SAM, and the position 7 is closer to the methyl group for the transfer (3.7 Å). It seems that the histidine at position 326 can interact with the hydroxyl group at position 5 of genistein (3.0 Å distance) and thus stabilizes the substrate in this orientation to complete the catalysis.

Naringenin differs from eriodictyol only by the lack of the 3′-hydroxyl group. Thus, although it can bind the same way in the active site of the wild type, the B-ring cannot be methylated to a 3′-methoxy derivative and thus the wild type is almost inactive. However, the mutation T133M (and the T133M/L322M mutations) enabled a different binding pattern, where the position 7 of the ring A is accessible to the SAM and the 7-
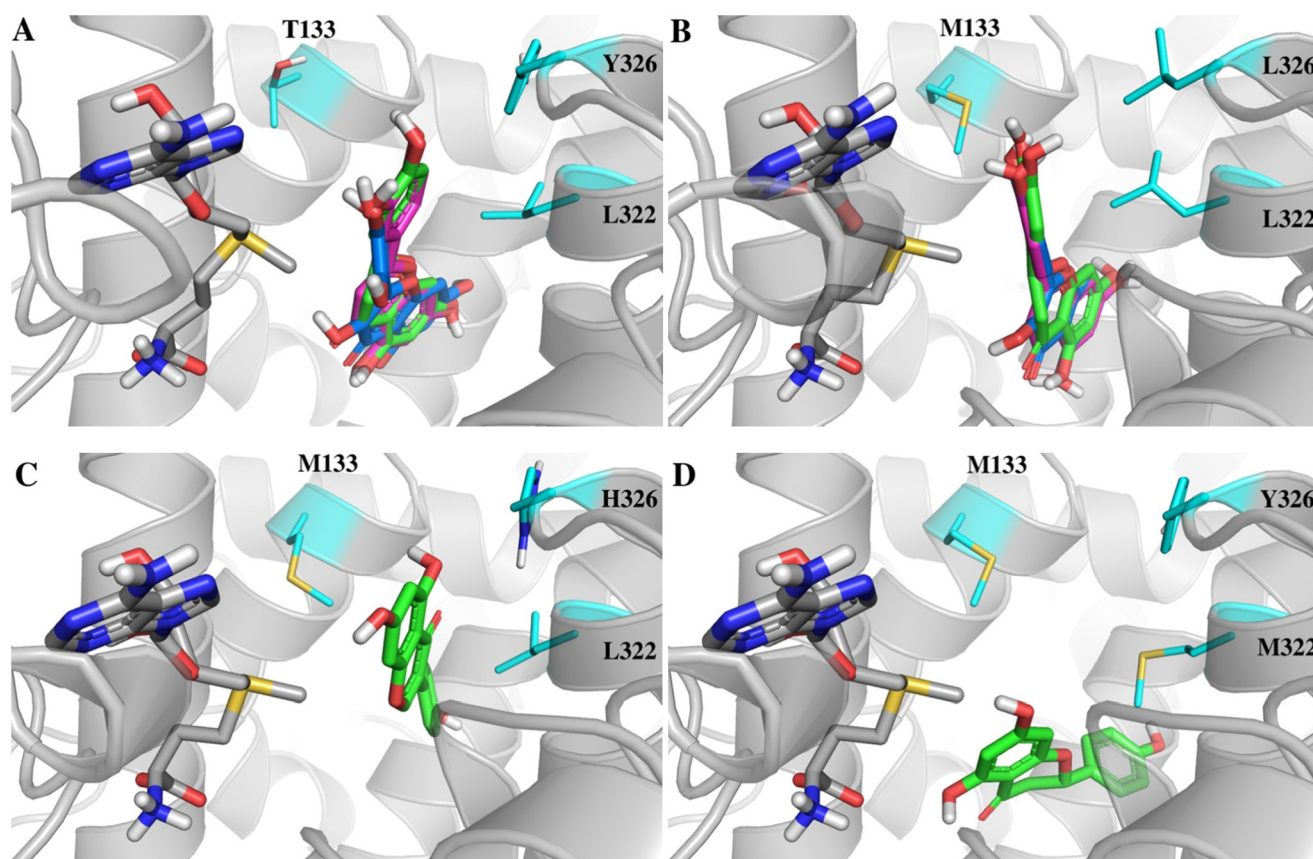
ChemCatChem

Full Papers
doi.org/10.1002/cctc.202000471

Chemistry Europe
European Chemical
Societies Publishing

Figure 3. Product composition of each reaction catalyzed by the wild-type IeOMT and its mutants. Area percentages were calculated by the peak areas of both substrates and products as determined by HPLC (Figure S2). 3′-Me-E, 3′-methyleriodictyol; 3′,4′-DiMe-E, 3′,4′-dimethyleriodictyol; 3′-Me-L, 3′-methylluteolin; 3′,4′-DiMe-L, 3′,4′-dimethylluteolin; 3′-Me-Q, 3′-methylquercetin; 3′,4′-DiMe-Q, 3′,4′-dimethylquercetin. Structures of these products are confirmed by NMR and LC-MS (see supporting information). 7-Me-N, 7-methylnaringenin; 7-Me-G, 7-methylgenistein; 4′-Me-G, 4′-methylgenistein. These products were confirmed by comparing their retention times on HPLC to commercial standards as well as LC-MS. Racemic eriodictyol and naringenin were used as substrates and racemic products were obtained. All experiments were performed in triplicates and standard deviations were provided.

hydroxyl group at catalytic distance (3.4 Å). The binding pattern differs for genistein, for which the enzyme exhibited the same regioselectivity. In the case of variant T133M/L322M (Figure 4D), the reason for this can be the lack of the histidine at position 326 that could interact with the 5-OH, in combination with the methionine in position 322 that pushes away the substrate. As naringenin is not planar, the ring B may cause steric clashes with 322 and thus the substrate binds in a different orientation.

Conclusions

In this study we highlighted the potential of IeOMT, a phenyl-propanoid OMT to be engineered to catalyze the methylation of a range of flavonoids and isoflavonoids. Three positions, namely 133, 322 and 326 had a significant impact on the binding of the substrates and variants that were produced with semi-rational design in these positions had altered catalytic activity and/or regioselectivity compared to the wild-type enzyme. These variants have provided access to methylated flavonoids that are rare in nature and may have interesting biological activities due to their methylation patterns.

ChemCatChem

Full Papers
doi.org/10.1002/cctc.202000471

Chemistry Europe
European Chemical
Societies Publishing

Figure 4. Models of *Clarkia breweri* IeOMT variants with SAM (grey sticks with elemental coloring at the left side) and substrates in the binding pocket. The three targeted residues are given as lines. For clarity, non-polar hydrogens are removed. A) (*S*)-eriodictyol (green), luteolin (magenta) and quercetin (blue) docked in the binding pocket of the wild-type IeOMT, exposing the 3′-hydroxyl group for methylation. B) the same substrates in the binding pocket of IeOMT_T133M/Y326L, exposing the 4′-hydroxyl group for methylation. C) Genistein binding on IeOMT_T133M/Y326H provides the 7-position for methylation. D) (*S*)-Naringenin binding pattern in IeOMT_T133M/L322M also exposes the 7-position, however, with a totally different binding pattern.

Experimental Section

Materials

*Dpn*I was purchased from New England Biolabs, Inc.. All primers were purchased from Thermo Fisher Scientific, Inc.. Competent cells were self-prepared. *S*-Adenosyl-ʟ-methionine was purchased from J&K Chemical Ltd.. Acetic acid, trifluoroacetic acid (TFA) and acetonitrile were purchased from Carl Roth GmbH and VWR chemicals and were of HPLC grade. (±)-Eriodictyol was purchased from Carl Roth GmbH, (±)-naringenin and quercetin were purchased from Sigma, luteolin, genistein and biochanin A (4′-methoxygenistein) were purchased from TCI GmbH, prunetin (7-methoxygenistein) and sakuranetin (7-methoxynaringenin) were purchased from Extrasynthese. All these chemicals have 95% or higher purity.

Mutagenesis, expression and purification

IeOMT from *Clarkia breweri* (accession number of protein: O04385.2) and the mutant T133M in pET-21a(+) with His₆-tag at the C-terminal were constructed as described in our previous work.[11] Further mutagenesis were performed following the QuikChange (Stratagene) protocol using pET-21a(+)−T133M as template. Primers are as below:

L322H fw: 5′-CACACCGATGCCCATATGCTGG-3′;

L322H rv: 5′-CCAGCATATGGGCATCGGTGTG-3′;
L322M fw: 5′-CACACCGATGCGATGATGCTGG-3′;
L322M rv: 5′-CCAGCATCATCGCATCGGTGTG-3′;
L322N fw: 5′-CACCGATGCGAACATGCTGGCGTATAAC-3′;
L322N rv: 5′-GTTATACGCCAGCATGTTCGCATCGGTG-3′;
Y326H fw: 5′-TGCTGGCGCATAACCCGGGCGGTAAAG-3′;
Y326H rv: 5′-CTTTACCGCCCGGGTTATGCGCCAGCA-3′;
Y326L fw: 5′-TGCTGGCGCTGAACCCGGGCGGTAAAG-3′;
Y326L rv: 5′-CTTTACCGCCCGGGTTCAGCGCCAGCA-3′;
Y326R fw: 5′-TGCTGGCGCGTAACCCGGGCGGTAAAG-3′;
Y326R rv: 5′-CTTTACCGCCCGGGTTACGCGCCAGCA-3′;

After PCR, the pET-21a(+)−T133M template was digested by *Dpn*I before transforming into *Escherichia coli* (*E. coli*) Top 10 chemically competent cells for plasmid amplification. After sequence confirmation (Eurofins, Germany), the newly constructed mutant plasmids were transformed into *E. coli* Bl21 (DE3) chemically competent cells for protein expression. Expression and purification of IeOMT and all mutants were carried out following the protocols given in our previous work.[11]

Enzyme assays and HPLC analysis

The reaction mixture consisted of 150 μM flavonoid substrate, 1 mM SAM, 5 mM DTT and 25% (v/v) *E. coli* Bl21 (DE3) cell lysate (not including the MTs of interest), in 50 mM sodium phosphate

ChemCatChem

Full Papers
doi.org/10.1002/cctc.202000471

Chemistry Europe
European Chemical
Societies Publishing

buffer, pH 7.5. Because the *S*-adenosyl- L -homocysteine (SAH) derived from SAM is a potent inhibitor for MTs, *E. coli* cell lysate which contains SAH nucleosidase was provided in each reaction in order to reduce the inhibition effect brought by the SAH.[13] 1.0 mg/mL (24.4 µM) purified enzymes were added into the reaction mixtures to initiate the enzyme reaction. Due to the instability of SAM, the same amount of SAM was supplied again after 24 h. The mixtures were incubated at 28 °C with 700 rpm agitation in an Eppendorf Thermomixer. Assays were performed in triplicate and negative controls were performed by replacing purified enzymes with buffer. Samples were taken at 2, 4, 24 and 48 h and an equal volume of acetonitrile was added to quench the reaction. Samples were vortexed vigorously and then centrifuged at full speed for 30 min to remove protein precipitate. 200 µL supernatant were transferred to HPLC sample vial inserts for analysis. HPLC analysis were performed on VWR Hitachi Elite LaChrom system equipped with the Kinetex EVO C18 (4.6×250 mm, 5 µm particle size, Phenomenex) reversed-phase column. 0.1% acetic acid and acetonitrile were used for the separation of flavonoid substrates and the corresponding methylated products, with the ratio 68:32 (v/v) for eriodictyol, luteolin and quercetin, and 60:40 (v/v) for naringenin and genistein. Wavelengths for the detection of flavanones (eriodictyol and naringenin), flavone and flavonol (luteolin and quercetin) and isoflavone (genistein) were 280, 260 and 260 nm, respectively. All analyses were performed at a flow rate of 1 mL/min and the column temperature was 35 °C. Identification of methylated products were confirmed by comparing their retention times on HPLC to commercial standards. For unknown products, preparative-scale reactions were performed with specific substrates and mutants and the products were isolated and identified by NMR and MS. Area percentages of substrate and each product were calculated to show the estimated yield for each product.

Biosynthesis, isolation and purification of unknown methylated products

Each preparative-scale reaction started with 30 mg flavonoid substrate (~100 µmol), 80 mg SAM (200 µmol), 20 mg purified enzyme, 0.1% NaN₃ and crude *E. coli* cell lysate in 100 mL sodium phosphate buffer (50 mM, pH 7.5). Purified IeOMT_T133M/Y326R was used for the biosynthesis of the 3′-methylated products of eriodictyol, luteolin and quercetin. IeOMT_T133M/Y326L was used to synthesize the 3′,4′-dimethylated products of eriodictyol, luteolin and quercetin. Reactions were performed in 500 mL round bottom flasks at 30 °C at 180 rpm for a total duration of 10 days, and the reaction progress was followed every 24 h by HPLC. The same amount of SAM and purified enzyme were resupplied at specific time intervals, when reaction seemed to slow down. When reactions were finished, substrates and products were extracted with 100 mL ethyl acetate (EtOAc) for three times, water was removed with dried MgSO₄ and then filtered. EtOAc were then removed using a rotary evaporator and the remaining dry substances were dissolved in 1 mL DMSO for subsequent purification. Separation of flavonoid substrates and products were performed firstly with the analytical column LiChrospher® 100 RP-18 (5 µm) LiChroCART® (250×4 mm, Merck) and then with the preparative column LiChrospher® 100 RP-18 (5 µm) Hibar® RT (250×25 mm, Merck), both are equipped with Shimadzu devices CBM-20A, LC-20A P, SIL-20A, FRC-10A and a SPD 20A UV/Vis detector. 0.1% TFA and acetonitrile (60/40, v/v) were used as mobile phase. Product fractions were collected, dried with rotary evaporator and lyophilizer before submission to NMR. All products had a purity of >99%, according to HPLC. The isolated yields (not optimized procedures) were the following: 3′-methyleriodictyol: 18.8% (5.9 mg), 3′,4′-dimethyleriodictyol: 35.5% (11.7 mg), 3′-meth-

ylluteolin 24.6% (15.5 mg), 3′,4′-dimethylluteolin 26.9% (17.7 mg), 3′-methylquercetin 6.8% (4.3 mg), 3′,4′-dimethylquercetin 16.8% (11.0 mg).

NMR and mass spectrometry analysis

NMR measurements were carried out on a Bruker Avance 600 MHz spectrometer equipped with an inverse ¹H/¹³C/¹⁵N/¹⁹F quadruple resonance cryoprobehead and z-field gradients. All compounds were dissolved in DMSO-*d*₆ and ¹H NMR, ¹³C NMR and DEPT-135 experiments were performed. Identification of single methylated products was done by comparing assignments against known standards. For the double-methylated products, additional 2D NMR experiments (NOESY, HSQC, HMBC) were performed. Data were processed and analyzed using TopSpin 4.0.7. All spectra were recorded at 25 °C. All peaks were referenced towards the DMSO-*d*₆ peak (¹H: 2.50 ppm; ¹³C: 39.51 ppm).

Mass spectrometry was performed using a Shimadzu LC-MS 8030 equipped with an ESI (electrospray ionization) source and a mass spectrometer, using the same column and mobile phase as in the HPLC analysis. Negative ionization mode was used and automatic MRM (multiple reaction monitoring) optimization was performed to acquire optimal fragmentation and maximal transmission of the desired product ions.

Bioinformatic analysis

The bioinformatic analysis was performed with YASARA 19.7.20. First, the structure of 3REO was back-mutated to its wild-type sequence and the SAH was transformed to SAM by the addition of the methyl group and this structure was refined at pH 7.5, 25 °C for 500 ps, taking a snapshot every 25 ps. The structure with the lowest energy was selected for further experiments. For the mutants, the respective amino acids were swapped with subsequent energy minimization. The same was performed for the preparation of the substrate molecules. The docking experiments were performed with VINA method, using the force field AMBER03 at 30 °C. Five receptor molecules were prepared for each experiment and in each receptor 25 dockings were performed, and the resulting structures were clustered when RMSD was <5 Å. The catalytic active conformation with the higher binding energy was selected. Figures were prepared with PyMol.

Accession numbers

The accession numbers of plant OMTs chosen for sequence alignment in Figure 1 are as follows: MePi_7-FOMT1 from *Mentha piperita* (AAR09598.1), MeTr_7-FOMT7 from *Medicago truncatula* (ABD83946.1), OrSa_7-FOMT from *Oryza sativa* (BAM13734.1), ArTh_3′-FOMT from *Arabidopsis thaliana* (AAB96879.1), ChAm_3′-FOMT from *Chrysosplenium americanum* (AAA80579.1), MePi_3′-FOMT3 from *Mentha piperita* (AAR09601.1), OrSa_3′-FOMT from *Oryza sativa* (XP_015650053.1), CaRo_4′-FOMT from *Catharanthus roseus* (AAR02420.1), GlMa_4′-FOMT/IOMT from *Glycine max* (C6TAY1.1), MePi_4′-FOMT4 from *Mentha piperita* (AAR09602.1), CiAr_7-IOMT from *Cicer arietinum* (XP_004489528.1), MeTr_7-IOMT1 from *Medicago truncatula* (AAY18582.1), MeSa-7/4′-IOMT from *Medicago sativa* (AAC49928.1), GlEc_4′-IOMT from *Glycyrrhiza echinate* (BAC58011.1), LoJa_4′-IOMT from *Lotus japonicus* (BAC58013.1), MeTr_4′-IOMT5 from *Medicago truncatula* (AAY18581.1), CaRo_3-POMT from *Catharanthus roseus* (AAK20170.1), ClBr_3-POMT from *Clarkia breweri* (O23760.1), LoPe_3-POMT from *Lolium perenne* (AAD10253.1), MeSa_3-POMT from *Medicago sativa* (AAB46623.1), ClBr_4-POMT from *Clarkia breweri* (O04385.2).

Acknowledgements

Conflict of Interest

The authors declare no conflict of interest.

[1] a) C. Kandaswami, E. Middleton, in *Free radicals in diagnostic medicine* (Ed.: D. Armstrong), Springer, **1994**, pp. 351–376; b) D. Procházková, I. Boušová, N. Wilhelmová, *Fitoterapia* **2011**, *82*, 513; c) S. Kaur, P. Mondal, *J. Microbiol. Exp.* **2014**, *1*, 23; d) A. García-Lafuente, E. Guillamón, A. Villares, M. A. Rostagno, J. A. Martínez, *Inflammation Res.* **2009**, *58*, 537; e) M. H. Pan, C. S. Lai, C. T. Ho, *Food Funct.* **2010**, *1*, 15; f) M. K. Chahar, N. Sharma, M. P. Dobhal, Y. C. Joshi, *Pharmacogn. Rev.* **2011**, *5*, 1; g) D. Raffa, B. Maggio, M. V. Raimondi, F. Plescia, G. Daidone, *Eur. J. Med. Chem.* **2017**, *142*, 213.
[2] A. N. Panche, A. D. Diwan, S. R. Chandra, *J. Nutr. Sci.* **2016**, *5*.
[3] a) K. Saito, K. Yonekura-Sakakibara, R. Nakabayashi, Y. Higashi, M. Yamazaki, T. Tohge, A. R. Fernie, *Plant Physiol. Biochem.* **2013**, *72*, 21; b) M. L. Falcone Ferreyra, S. Rius, P. Casati, *Front. Plant Sci.* **2012**, *3*, 222.
[4] a) T. Walle, *Mol. Pharm.* **2007**, *4*, 826; b) S. H. Thilakarathna, H. P. Rupasinghe, *Nutrients* **2013**, *5*, 3367.
[5] a) H. Takemura, T. Itoh, K. Yamamoto, H. Sakakibara, K. Shimoi, *Bioorg. Med. Chem.* **2010**, *18*, 6310; b) H. Takemura, H. Uchiyama, T. Ohura, H. Sakakibara, R. Kuruto, T. Amagai, K. Shimoi, *J. Steroid Biochem. Mol. Biol.* **2010**, *118*, 70.
[6] a) J. P. Ley, U.S. Patent 8,685,436, **2014**; b) J. P. Ley, G. Krammer, G. Reinders, I. L. Gatfield, H. J. Bertram, *J. Agric. Food Chem.* **2005**, *53*, 6061.
[7] J. P. Noel, R. A. Dixon, E. Pichersky, C. Zubieta, J. L. Ferrer, in *Recent advances in phytochemistry, Vol. 37*, Elsevier, **2003**, pp. 37–58.
[8] C. Zubieta, P. Kota, J. L. Ferrer, R. A. Dixon, J. P. Noel, *Plant Cell* **2002**, *14*, 1265.
[9] a) D. H. Kim, B. G. Kim, Y. Lee, J. Y. Ryu, Y. Lim, H. G. Hur, J. H. Ahn, *J. Biotechnol.* **2005**, *119*, 155; b) B. G. Kim, B. R. Jung, Y. Lee, H. G. Hur, Y. Lim, J. H. Ahn, *J. Agric. Food Chem.* **2006**, *54*, 823.
[10] C. Zubieta, X. Z. He, R. A. Dixon, J. P. Noel, *Nat. Struct. Mol. Biol.* **2001**, *8*, 271.
[11] Q. Tang, U. T. Bornscheuer, I. V. Pavlidis, *ChemCatChem* **2019**, *11*, 3227.
[12] J. A. Yáñez, N. D. Miranda, C. M. Remsberg, Y. Ohgami, N. M. Davies, *J. Pharm. Biomed. Anal.* **2007**, *43*, 255.
[13] N. Parveen, K. A. Cornell, *Mol. Microbiol.* **2011**, *79*, 7.

ChemCatChem

Supporting Information

Influence of Substrate Binding Residues on the Substrate Scope and Regioselectivity of a Plant *O*-Methyltransferase against Flavonoids

Qingyun Tang, Yoanes M. Vianney, Klaus Weisz, Christoph W. Grathwol, Andreas Link, Uwe T. Bornscheuer,* and Ioannis V. Pavlidis*

Supporting information

Influence of substrate binding residues on the substrate scope and regioselectivity of a plant *O*-methyltransferase against flavonoids

Qingyun Tang, Yoanes Maria Vianney, Klaus Weisz, Christoph W. Grathwol, Andreas Link, Uwe T. Bornscheuer* and Ioannis V. Pavlidis*

Table of contents

1. Sequence alignment of OMTs

```
                          1    5      15       25       35       45

MePi_7-FOMT1                    MAPEEDSLALAEAWNHGFGFIKTSIVKTAVELEIPDILESR
MeTr_7-FOMT7                 MDSSSNGSEESQLYHAQIHLYKHIYSFINSMALKSAVELGIADAIHNH
OrSa_7-FOMT     MGDMVSPVVHRHAAGGGSGGDDDDQACMYALELLGGSVVSMTLKAAIELGLVDELLA-
ArTh_3'-FOMT           MGSTAETQLTPVQVTDDEAALFAMQLASASVLPMALKSALELDLLEIMAKN
ChAm_3'-FOMT                      MLFAMQLASASVLPMVLKSAIELDLLEIIAS-
MePi_3'-FOMT3       MEASFENGRKRSSSSSSEEESAFSFAMELAAGSVLPMVIKSAIDLNLLELIKR-
OrSa_3'-FOMT           MGSTAADMAAAADEEACMYALQLASSSILPMTLKNAIELGLLETLQSA
CaRo_4'-FOMT            MDLQTAEFREAQAKIWSQAFSFANCAALKCAVKLGIADAIDNH
GlMa_4'-FOMT/4'-IOMT  MASPLNNGRKASEIFQGQALLYKHLLGFIDSKCLKWMVELDIPDIIHSH
MePi_4'-FOMT4          MVADEEVRVRAEAWNNAFGYIKPTAVATAVELGLPDILENH
CiAr_7-IOMT          MSNSLINGHNPREIFQAQALLYKHIFAFIDSMCLKWTVEMNIPNIIHNH
MeTr_7-IOMT1         MASSLNGRKPSEIFKAQALLYKHIFAFIDSMSLKWAVEMNIPNIIHNH
MeSa_7/4'-IOMT      MASSINGRKPSEIFKAQALLYKHIYAFIDSMSLKWAVEMNIPNIIQNH
GlEc_4'-IOMT          MAFSTNGSEEIELYHAQIHLYKHVYNFVSSMALKSAMELGIADVIHNH
LoJa_4'-IOMT        MDFSSSNGSEDTELSQAQIHLYKHVYNFVSSMALKSAMELGIADVIHSH
MeTr_4'-IOMT5         MAFSTNGSEESELYHAQIHLYKHVYNFVSSMALKSAMELGIADAIHNH
CaRo_3-POMT       MGSANPDNKNSMTKEEEEACLSAMRLASASVLPMVLKSAIELDLLELIKKS
ClBr_3-POMT      MGSTGNAETQLTPTHVSDEEANLFAMQLASASVLPMVLKAAIELDVLEIMAKS
LoPe_3-POMT            MGSTAADMAASADEDACMFALQLASSSVLPMTLKNAIELGLLEILVAA
MeSa_3-POMT          MGSTGETQITPTHISDEEANLFAMQLASASVLPMILKSALELDLLEIIAKA
ClBr_4-POMT (IeOMT) MGSTGNAEIQIIPTHSSDEEANLFAMQLASAAVLPMALKAAIELDVLEIMAKS


                          54       60       70       79       89       98

MePi_7-FOMT1        ------GAP--VSIPELATAVDCS-----ADRIYRVMRFLAYHGIFKRTKPPPESTEG
MeTr_7-FOMT7        ------GKP--MTLTELASSLKLH--PSKVSVLYRLLRLLTHNGFFAKTTL--MNGKE
OrSa_7-FOMT        ------AAGAAVTAEELAARLRLPAAVAAAAAVDRMLRLLASYGVVRCAT---EAGPD
ArTh_3'-FOMT       ------GSP--MSPTEIASKLPTK-NPEAPVMLDRILRLLTSYSVLTCSN---RKLSG
ChAm_3'-FOMT       -------QDTCMSPTEIASKLPTT-NPHAPTMIDRILRLLSSYSIVTCSV--RSVDD
MePi_3'-FOMT3      ------GGEEGASAYELAAQINAENPKAAAEMIDRILQLLAAHSVLTCRV----ETPP
OrSa_3'-FOMT       AVAGGGGKAALLTPAEVADKLPSKANPAAADMVDRMLRLLASYNVVRCEM---EEGAD
CaRo_4'-FOMT       --D---KKA--LTLSELTEELSIK--PSKSPFLQRLMRQLVNAGFFTEAKQLRDDNKD
GlMa_4'-FOMT/4'-IOMT --SH--GQP---ITFSELVSILQVP--PTKTRQVQSLMRYLAHNGFFEIVR------IH
MePi_4'-FOMT4      -----DGP---MSLLELSAATDCP-----AEPLHRLMRFLVFHGIFKKTA------KP
CiAr_7-IOMT        ------NKP---ITLSNLVSILQVP--SSKIGNVQRLMRYLAHNGFFEIIT----NQEE
MeTr_7-IOMT1       ------GKP---ISLSNLVSILQVP--SSKIGNVRRLMRYLAHNGFFEIIT------KE
MeSa_7/4'-IOMT     ------GKP---ISLSNLVSILQVP--SSKIGNVRRLMRYLAHNGFFEIIT------KE
GlEc_4'-IOMT       ------GKP---ITLPELASALKLH--PSKVGILYRFLRLLTHNGFFAKTTVPSQNGKD
LoJa_4'-IOMT       ------GKP---ITLPELATALNLR--PSKIGVLHRFLRLLTHNGFFAKTTVSRGEGAE
MeTr_4'-IOMT5      ------GKP--MTLSELASSLKLH--PSKVNILHRFLRLLTHNGFFAKTI---VKGKE
CaRo_3-POMT        ------GPGAYVSPSELAAQLPTQ-NPDAPVMLDRILRLLASYSVLNCTL----KDLPD
ClBr_3-POMT        --IPH-GSGAYISPAEIAAQLPTT-NPDAPVMLDRVLRLLASYSVVTCSL---RELPD
LoPe_3-POMT        --G---GKS--LTPTEVAAKLPSAANPEAPDMVDRILRLLASYNVVTCLV---EEGKD
MeSa_3-POMT        ------GPGAQISPIEIASQLPTT-NPDAPVMLDRMLRLLACYIILTCSV---RTQQD
ClBr_4-POMT (IeOMT) --V---PPSGYISPAEIAAQLPTT-NPEAPVMLDRVLRLLASYSVVTYTL---RELPS


                          106      114      124      132      142      152

MePi_7-FOMT1        GSV----YYAQTPVSRRLTRENLGPFVLLQGTMREPSGCVTAETLRTSKR--------
MeTr_7-FOMT7        GEEET--TYSLTPPSMLLISGKS---TCLSPFVSGVLHPCRLNVWHSSKKWLTEDKE-
OrSa_7-FOMT        GKARRS---EGAPVCKWLAAGSSSGEGSMAPLGLLNLDKVFMENWYYLKEAVS---EG
ArTh_3'-FOMT       DGVER--IYGLGPVCKYLTKNEDG--VSIAALCLMNQDKVLMESWYHLKDAIL---DG
ChAm_3'-FOMT       QRV-----YSPAPVCKYLTKNQDG--VSIAALCVAAQDKVLMECWYHMKDAVL---DG
MePi_3'-FOMT3      SRRRR---YSLAAVCKFLTRDEDG--ASLAPLSLVLQDRVFMEPWYHLKDVIV---EG
OrSa_3'-FOMT       GKLSR--RYAAAPVCKWLTPNEDG--VSMAALALMNQDKVLMESWYYLKDAVLD---G
CaRo_4'-FOMT       GRTTT--AYALTPVSRLLLKNEQ---WNLRGIVLTMLDPAELKAWSVLNDWFKN--DD
GlMa_4'-FOMT/4'-IOMT DNIE---AYALTAASELLVKSSE---LSLAPMVEYFLEPNCQGAWNQLKRWVHE--ED
MePi_4'-FOMT4      PLSNEAVYYARTALSRLFTRDE-------LGDFMLLQTGPLSQHPAGLTASSLRT--GK
CiAr_7-IOMT        ENEE---SYALTVASKLLVKGSD---LCLAPMVECVLDPTLSGSYHQLKKWIYE--ED
MeTr_7-IOMT1       EES-----YALTVASELLVRGSD---LCLAPMVECVLDPTLSGSYHELKKWIYE--ED
MeSa_7/4'-IOMT     EES-----YALTVASELLVRGSD---LCLAPMVECVLDPTLSGSYHELKKWIYE--ED
GlEc_4'-IOMT       GEEEEETAYALTPPSKLLVKGKP---TCLASIVRGALHPSSLDMWRSSEKWFKEDKE-
LoJa_4'-IOMT       EET-----AYGLTPPSKLLVKSNS---TCLAPIVKGALHPSSLDMWRSSKKWFLEDNEE
MeTr_4'-IOMT5      GDEEEEIAYSLTPPSKLLISGKP---TCLSSIVKGALHPSSLDMWRSSKKWFNEDKEQ
CaRo_3-POMT        GGIER--LYSLAPVCKFLTKNEDG--VSMAALLLMNQDKVLMESWYHLKDAVL---EG
ClBr_3-POMT        GKVER--LYGLAPVCKFLTKNEDG--VSLAPLCLMNQDKVLMESWYYLKDAIL---DG
LoPe_3-POMT        GRLSR--SYGAAPVCKFLTPNEDG--VSMAALALMNQDKVLMESWYYLKDAVL---DG
MeSa_3-POMT        GKVQR--LYGLATVAKYLVKNEDG--VSISALNLMNQDKVLMESWYHLKDAVL---DG
ClBr_4-POMT (IeOMT) GKVER--LYGLAPVCKFLTKNEDG--VSLAPFLLTATDKVLLEPWFYLKDAIL---EG


                          159      169      177      187      197      203

MePi_7-FOMT1        ----PGVVNENESDHLYEDPVF-SMKVFRDAMASHARMTTAAVIENYGEGFQ---GVG
MeTr_7-FOMT7        LSLFESARGETFWDYLNKDTESDELSMFQEAMAADSQIF-NLALKECNHVFE---GLE
OrSa_7-FOMT        GTAFDKAYGTSLFQYLGQDGNEPSNTLFNQAMASHSVVITNKLLQFFR-GFDAGAGVD
ArTh_3'-FOMT       GIPFNKAYGMSAFEYHGKDPA--FNKVFNNGMSNHSTITMKKILETYK-GFE---GLT
ChAm_3'-FOMT       GIPFNKAYGMPIFDYFAKDLG--SNKVFNKGMSDFSSMIIKKILETYK-GFQ---GLT
MePi_3'-FOMT3      GVAFERAYGVHAFEYHAKDPK--FNKIFNQAMHNQSIIFMKRILEIYK-GFE---GVK
OrSa_3'-FOMT       GIPFNKAYGMTAFEYHGKDTAR-FNRVFNGMKNHSVIITKKLLDLYT-GFDA---AS
CaRo_4'-FOMT       PTAFQTAHEKNYWDYTAENTQ--HCQIFEDAMANDSVLVSKLLVTEYKFLFE---GLT
GlMa_4'-FOMT/4'-IOMT LTVFGVSLGTPFWDFINKDPA--YNKSFNEAMACDSQML-NLAFRDCNWVFE---GLE
MePi_4'-FOMT4      PQFIRSVNGEDSWTDPVNGYH---MKVFSDAMAAHARETTAAIVRYCPAAFE---GIG
CiAr_7-IOMT        LTLFGVTLGSGFWDFLNKPE--YNKSFNEAMASDSKMI-NLALRDCNLVFE---GLE
MeTr_7-IOMT1       LTLFGVTLGSGFWDFLDKNPE--YNRSFNDAMASDSKLI-NLALRDCDFVFD---GLE
MeSa_7/4'-IOMT     LTLFGVTLGSGFWDFLDKNPE--YNTSFNDAMASDSKLI-NLALRDCDFVFD---GLE
GlEc_4'-IOMT       LTLFESATGESFWDFLNKDSESGTLSMFQEAMAADSQMF-KLALKECRHVFE---GLE
LoJa_4'-IOMT       LTLFESATGESFWEFLNKETESDTLSMFQEAMAADSHMF-KLALKECKHVFE---GLG
MeTr_4'-IOMT5      -TLFECATGESFWDFLNKDSESSTLSMFQDAMASDRMF-KLVLQENKRVFE---GLE
CaRo_3-POMT        GIPFNKAYGMTAFEYHGKDPR--FNKVFNQGMSNHSTIIMKKILEIYQ-GFQ---GLK
ClBr_3-POMT        GIPFNKAYGMTAFEYHGTDPR--FNKVFNRGMSDHSTITMKKIFEMYT-GFEA---LN
LoPe_3-POMT        GIPFNKAYGMTAFEYHGTDPR--FNRVFNGMKNHSIIITKKLLELYH-GFE---GLE
MeSa_3-POMT        GIPFNKAYGMTAFEYHGTDPR--FNKVFNKGMSDHSTITMKKILETYT-GFE---GLK
ClBr_4-POMT (IeOMT) GIPFNKAYGMNEFDYHGTDHR--FNKVFNKGMSSNSTITMKKILEMYN-GFE---GLT
```
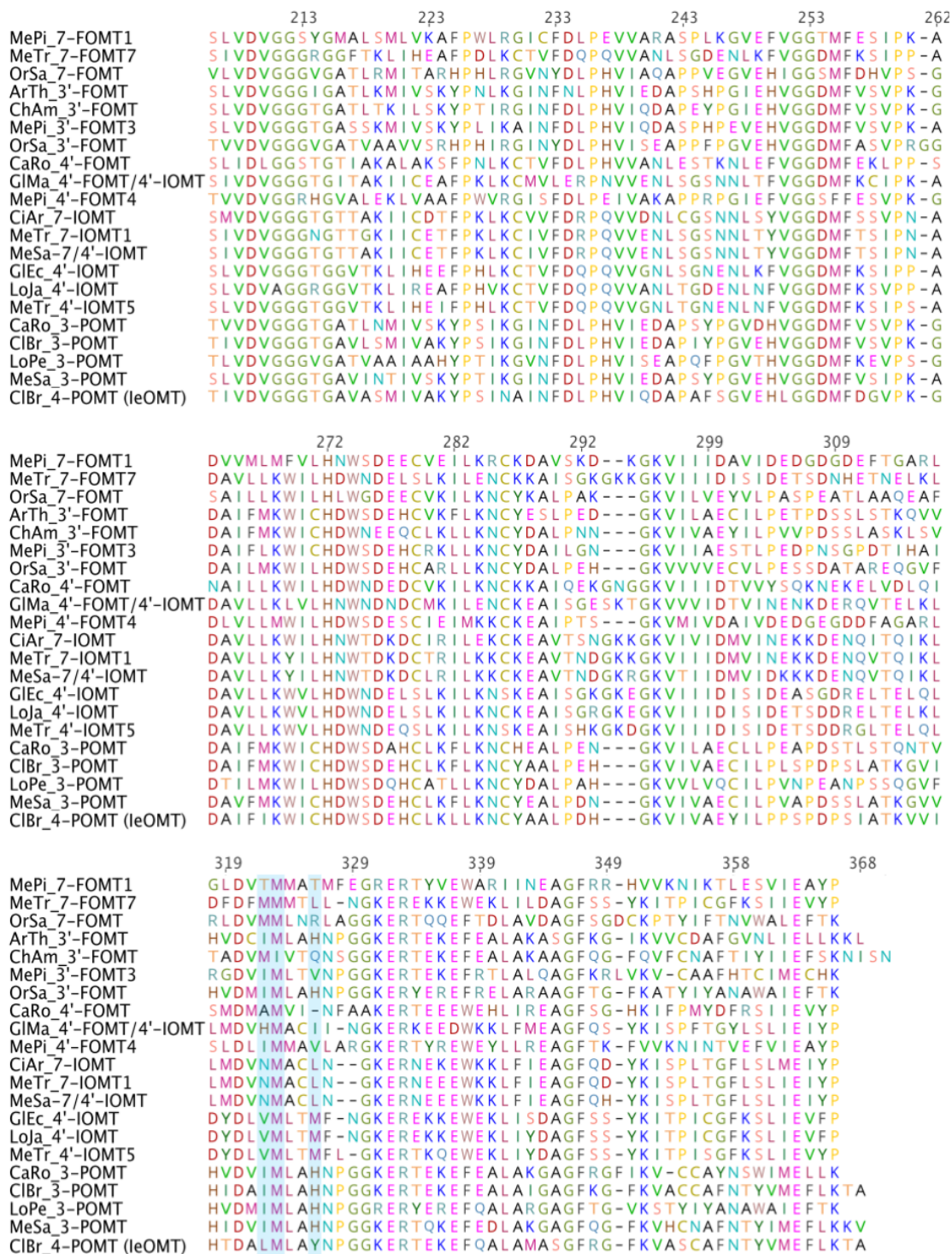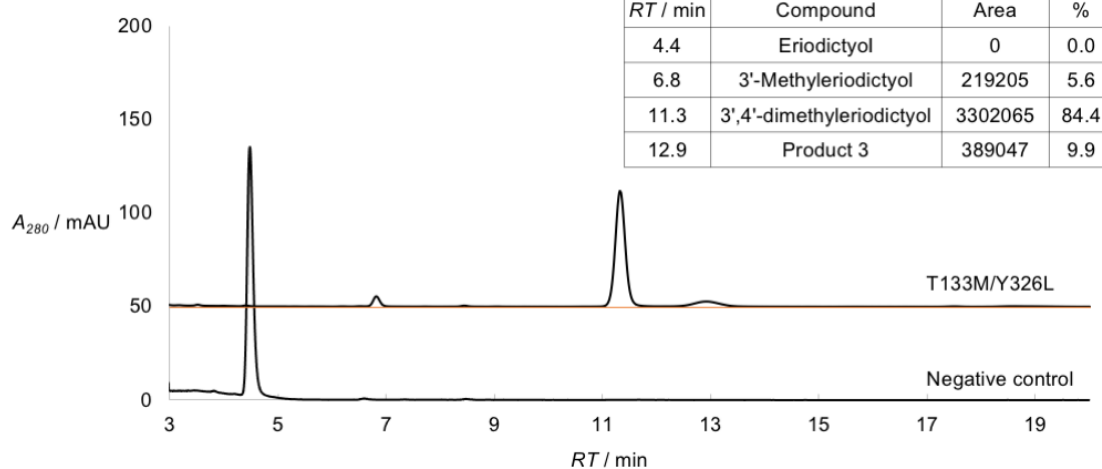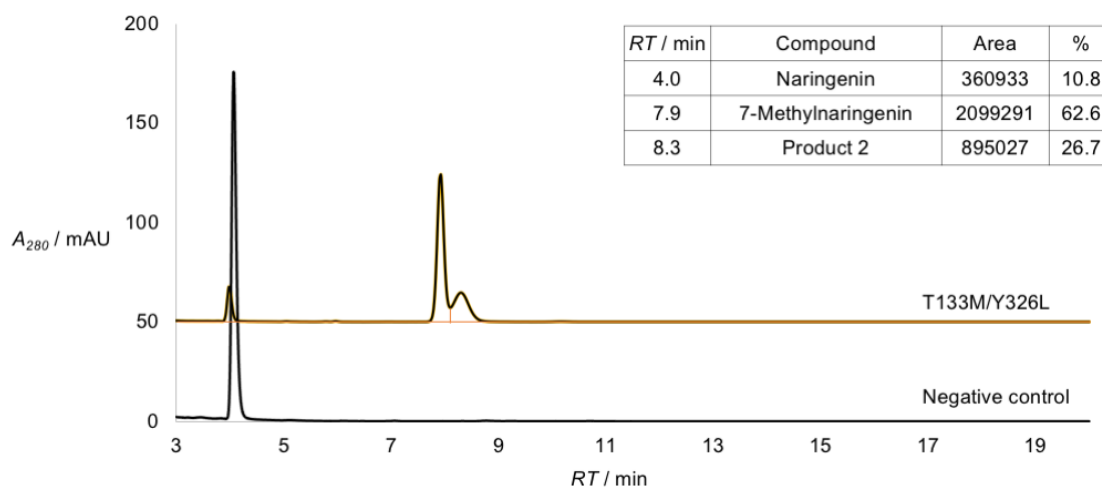
(to be continued)

```
                        213       223       233       243       253      262
MePi_7-FOMT1            SLVDVGGSYGMALSMLVKAFPWLRGICFDLPEVVARASPLKGVEFVGGTMFESIPK-A
MeTr_7-FOMT7            SIVDVGGGRGGFTKLIHEAFPDLKCTVFDQPQVVANLSGDENLKFVGGDMFKSIPP-A
OrSa_7-FOMT            VLVDVGGGVGATLRMITARHPHLRGVNYDLPHVIAQAPPVEGVEHIGGSMFDHVPS-G
ArTh_3'-FOMT           SLVDVGGGIGATLKMIVSKYPNLKGINFNLPHVIEDAPSHPGIEHVGGDMFVSVPK-G
ChAm_3'-FOMT           SLVDVGGGTGATLTKILSKYPTIRGINFDLPHVIQDAPEYPGIEHVGGDMFVSVPK-G
MePi_3'-FOMT3          SLVDVGGGTGASSKMIVSKYPLIKAINFDLPHVIQDASPHPEVEHVGGDMFVSVPK-A
OrSa_3'-FOMT           TVVDVGGGVGATVAAVVSRHPHIRGINYDLPHVISEAPPFPGVEHVGGDMFASVPRGG
CaRo_4'-FOMT           SLIDLGGSTGTIAKALAKSFPNLKCTVFDLPHVVANLESTKNLEFVGGDMFEKLPP-S
GlMa_4'-FOMT/4'-IOMT   SIVDVGGGTGITAKIICEAFPKLKCMVLERPNVVENLSGSNNLTFVGGDMFKCIPK-A
MePi_4'-FOMT4          TVVDVGGRHGVALEKLVAAFPWVRGISFDLPEIVAKAPPRPGIEFVGGSFFESVPK-G
CiAr_7-IOMT            SMVDVGGGTGTTAKIICDTFPKLKCVVFDRPQVVDNLCGSNNLSYVGGDMFSSVPN-A
MeTr_7-IOMT1           SIVDVGGGNGTTGKIICETFPKLKCIVFDRPQVVENLSGSNNLTYVGGDMFTSIPN-A
MeSa-7/4'-IOMT         SIVDVGGGTGTTAKIICETFPKLKCIVFDRPQVVENLSGSNNLTYVGGDMFTSIPN-A
GlEc_4'-IOMT           SLVDVGGGTGGVTKLIHEEFPHLKCTVFDQPQVVGNLSGNENLKFVGGDMFKSIPP-A
LoJa_4'-IOMT           SLVDVAGGRGGVTKLIREAFPHVKCTVFDQPQVVANLTGDENLNFVGGDMFKSVPP-A
MeTr_4'-IOMT5          SLVDVGGGTGGVTKLIHEIFPHLKCTVFDQPQVVGNLTGNENLNFVGGDMFKSIPS-A
CaRo_3-POMT            TVVDVGGGTGATLNMIVSKYPSIKGINFDLPHVIEDAPSYPGVDHVGGDMFVSVPK-G
ClBr_3-POMT            TIVDVGGGGAVLSMIVAKYPSIKGINFDLPHVIEDAPIYPGVEHVGGDMFVSVPK-G
LoPe_3-POMT            TLVDVGGGVGATVAAIAAHYPTIKGVNFDLPHVISEAPQFPGVTHVGGDMFKEVPS-G
MeSa_3-POMT            SLVDVGGGTGAVINTIVSKYPTIKGINFDLPHVIEDAPSYPGVEHVGGDMFVSIPK-A
ClBr_4-POMT (IeOMT)   TIVDVGGGTGAVASMIVAKYPSINAINFDLPHVIQDAPAFSGVEHLGGDMFDGVPK-G


                           272       282       292       299       309
MePi_7-FOMT1            DVVMLMFVLHNWSDEECVEILKRCKDAVSKD--KGKVIIIDAVIDEDGDGDEFTGARL
MeTr_7-FOMT7            DAVLLKWILHDWNDELSLKILENCKKAISGKGKKGKVIIIDISIDETSDNHETNELKL
OrSa_7-FOMT            SAILLKWILHLWGDEECVKILKNCYKALPAK---GKVILVEYVLPASPEATLAAQEAF
ArTh_3'-FOMT           DAIFMKWICHDWSDEHCVKFLKNCYESLPED---GKVILAECILPETPDSSLSTKQVV
ChAm_3'-FOMT           DAIFMKWICHDWNEEQCLKLLKNCYDALPNN---GKVIVAEYILPVVPDSSLASKLSV
MePi_3'-FOMT3          DAIFLKWICHDWSDEHCRKLLKNCYDAILGN---GKVIIAESTLPEDPNSGPDTIHAI
OrSa_3'-FOMT           DAILMKWILHDWSDEHCARLLKNCYDALPEH---GKVVVVECVLPESSDATAREQGVF
CaRo_4'-FOMT           NAILLKWILHDWNDEDCVKILKNCKKAIQEKGNGGKVIIIDTVVYSQKNEKELVDLQI
GlMa_4'-FOMT/4'-IOMT   DAVLLKLVLHNWNDNDCMKILENCKEAISGESKTGKVVVIDTVINENKDERQVTELKL
MePi_4'-FOMT4          DLVLLMWILHDWSDESCIEIMKKCKEAIPTS---GKVMIVDAIVDEDGEGDDFAGARL
CiAr_7-IOMT            DAVLLKWILHNWTDKDCIRILEKCKEAVTSNGKKGKVIVIDMVINEKKDENQITQIKL
MeTr_7-IOMT1           DAVLLKYILHNWTDKDCTRILKKCKEAVTNDGKKGKVIIIDMVINEKKDENQVTQIKL
MeSa-7/4'-IOMT         DAVLLKYILHNWTDKDCIRILKKCKEAVTNDGKRGKVTIIDMVIDKKKDENQVTQIKL
GlEc_4'-IOMT           DAVLLKWVLHDWNDELSLKILKNSKEAISGKGKEGKVIIIDISIDEASGDRELTELQL
LoJa_4'-IOMT           DAVLLKWVLHDWNDELSLKILKNCKEAISGRGKEGKVIIIDISIDETSDDRELTELKL
MeTr_4'-IOMT5          DAVLLKWVLHDWNDEQSLKILKNSKEAISHKGKDGKVIIIDISIDETSDDRGLTELQL
CaRo_3-POMT            DAIFMKWICHDWSDAHCLKFLKNCHEALPEN---GKVILAECLLPEAPDSTLSTQNTV
ClBr_3-POMT            DAIFMKWICHDWSDEHCLKFLKNCYAALPEH---GKVIVAECILPLSPDPSLATKGVI
LoPe_3-POMT            DTILMKWILHDWSDQHCATLLKNCYDALPAH---GKVVLVQCILPVNPEANPSSQGVF
MeSa_3-POMT            DAVFMKWICHDWSDEHCLKFLKNCYEALPDN---GKVIVAECILPVAPDSSLATKGVV
ClBr_4-POMT (IeOMT)   DAIFIKWICHDWSDEHCLKLLKNCYAALPDH---GKVIVAEYILPPSPDPSIATKVVI


                        319       329       339       349       358       368
MePi_7-FOMT1            GLDVTMMATMFEGRERTYVEWARIINEAGFRR-HVVKNIKTLESVIEAYP
MeTr_7-FOMT7            DFDFMMMTLL-NGKEREKKEWEKLILDAGFSS-YKITPICGFKSIIEVYP
OrSa_7-FOMT            RLDVMMLNRLAGGKERTQQEFTDLAVDAGFSGDCKPTYIFTNVWALEFTK
ArTh_3'-FOMT           HVDCIMLAHNPGGKERTEKEFEALAKASGFKG-IKVVCDAFGVNLIELLKKL
ChAm_3'-FOMT           TADVMIVTQNSGGKERTEKEFEALAKAAGFQG-FQVFCNAFTIYIIEFSKNISN
MePi_3'-FOMT3          RGDVIMLTVNPGGKERTEKEFRTLALQAGFKRLVKV-CAAFHTCIMECHK
OrSa_3'-FOMT           HVDMIMLAHNPGGKERYEREFRELARAAGFTG-FKATYIYANAWAIEFTK
CaRo_4'-FOMT           SMDMAMVI-NFAAKERTEEEWEHLIREAGFSG-HKIFPMYDFRSIIEVYP
GlMa_4'-FOMT/4'-IOMT   LMDVHMACII-NGKERKEEDWKKLFMEAGFQS-YKISPFTGYLSLIEIYP
MePi_4'-FOMT4          SLDLIMMAVLARGKERTYREWEYLLREAGFTK-FVVKNINTVEFVIEAYP
CiAr_7-IOMT            LMDVNMACLN--GKERNEKEWKKLFIEAGFQD-YKISPLTGFLSLMEIYP
MeTr_7-IOMT1           LMDVNMACLN--GKERNEEEWKKLFIEAGFQD-YKISPLTGFLSLIEIYP
MeSa-7/4'-IOMT         LMDVNMACLN--GKERNEEEWKKLFIEAGFQH-YKISPLTGFLSLIEIYP
GlEc_4'-IOMT           DYDLVMLTMF-NGKEREKKEWEKLISDAGFSS-YKITPICGFKSLIEVFP
LoJa_4'-IOMT           DYDLVMLTMF-NGKEREKKEWEKLIYDAGFSS-YKITPICGFKSLIEVFP
MeTr_4'-IOMT5          DYDLVMLTMFL-GKERTKQEWEKLIYDAGFSS-YKITPICGFKSLIEVYP
CaRo_3-POMT            HVDVIMLAHNPGGKERTEKEFEALAKGAGFRGFIKV-CCAYNSWIMELLK
ClBr_3-POMT            HIDAIMLAHNPGGKERTEKEFEALAIGAGFKG-FKVACCAFNTYVMEFLKTA
LoPe_3-POMT            HVDMIMLAHNPGGRERYEREFQALARGAGFTG-VKSTYIYANAWAIEFTK
MeSa_3-POMT            HIDVIMLAHNPGGKERTQKEFEDLAKGAGFQG-FKVHCNAFNTYIMEFLKKV
ClBr_4-POMT (IeOMT)   HTDALMLAYNPGGKERTEKEFQALAMASGFRG-FKVASCAFNTYVMEFLKTA
```

Figure S1. Sequence alignment of plant flavonoid OMTs (FOMTs), isoflavonoid OMTs (IOMTs) and phenylpropanoid OMTs (POMTs). Residues involved in substrate binding are highlighted in blue. Numbering of residues are based on the sequence of the *Clarkia breweri* isoeugenol OMT (IeOMT). OMTs are named by their original organisms and regioselectivity.

2. HPLC chromatograms

Eriodictyol

RT / min	Compound	Area	%
4.4	Eriodictyol	0	0.0
6.8	3'-Methyleriodictyol	219205	5.6
11.3	3',4'-dimethyleriodictyol	3302065	84.4
12.9	Product 3	389047	9.9

Naringenin

RT / min	Compound	Area	%
4.0	Naringenin	360933	10.8
7.9	7-Methylnaringenin	2099291	62.6
8.3	Product 2	895027	26.7

Luteolin

RT / min	Compound	Area	%
4.6	Luteolin	0	0.0
7.3	3'-Methylluteolin	404952	13.4
12.3	3',4'-Dimethylluteolin	2620537	86.6

(to be continued)

4

Quercetin

RT / min	Compound	Area	%
4.7	Quercetin	0	0.0
7.7	3'-Methylquercetin	422811	16.8
13.6	3',4'-Dimethylquercetin	1938446	76.9
16.3	Product 4	160798	6.4

Genistein

RT / min	Compound	Area	%
4.1	Genistein	968853	14.6
8.9	7-Methylgenistein	779440	11.7
9.4	4'-Methylgenistein	423013	6.4
11.0	Product 3	4486737	67.4

Figure S2. HPLC chromatograms of negative control and T133M/Y326L at 48h against the substrates. Area of both substrates and products were shown and area percentages were calculated. For example, Area / % of 3′-methyleriodictyol = Peak area$_{3'\text{-methyleriodictyol}}$ / (Peak area$_{\text{eriodictyol}}$ + Peak area$_{3'\text{-methyleriodictyol}}$ + Peak area$_{3'\text{-methyleriodictyol}}$ + Peak area$_{\text{Product 3}}$) × 100%.

3. NMR analysis of 3′-methyleriodictyol

^{1}H, ^{13}C and DEPT-135 NMR spectra of the methylated eriodictyol corresponds to the 3′-methyleriodictyol standard.[1] Both eriodictyol and its methylated products exist in racemic form.
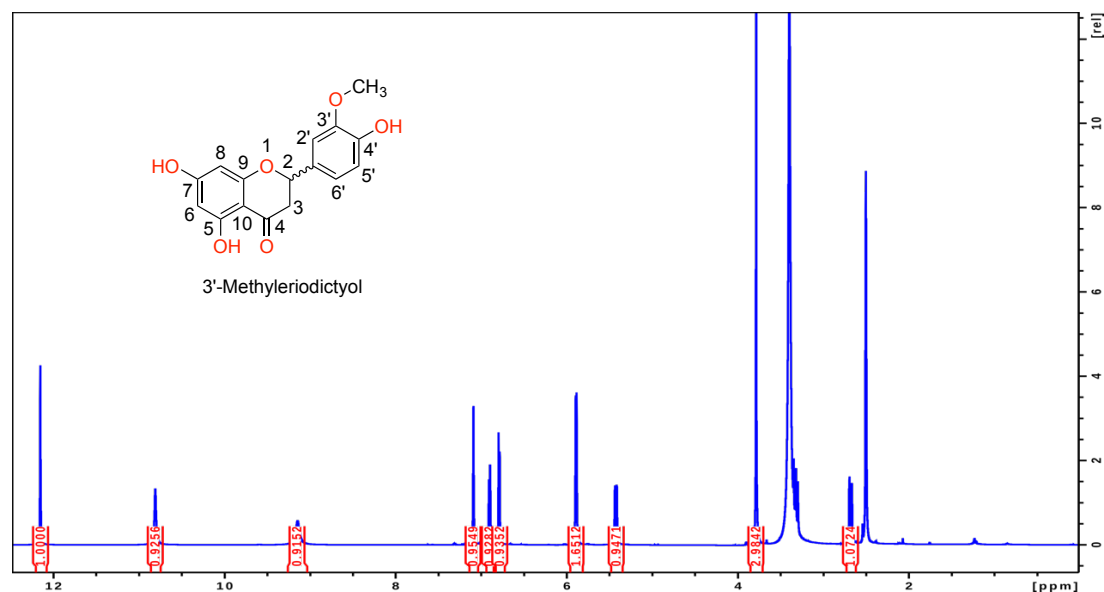
Figure S1. ^{1}H NMR spectrum of 3′-methyleriodictyol. ^{1}H NMR (600 MHz, DMSO-d_6): δ [ppm] = 12.15 (s, 1H, 5-OH), 10.81 (s, 1H, 7-OH), 9.14 (s, 1H, 3-OH), 7.09 (d, J = 2.0 Hz, 1H, 2′-H), 6.90 (dd, J = 8.1, 2.0 Hz, 1H, 6′-H), 6.79 (d, J = 8.0 Hz, 1H, 5′-H), 5.91–5.86 (m, 2H, 6-H, 8-H), 5.42 (dd, J = 12.9, 3.0 Hz, 1H, 2-H), 3.78 (s, 3H, 3′-OCH$_3$), 3.32 (dd, J = 17.1, 12.9 Hz, 1H, *trans* 3-H), 2.68 (dd, J = 17.1, 3.1 Hz, 1H, *cis* 3-H).
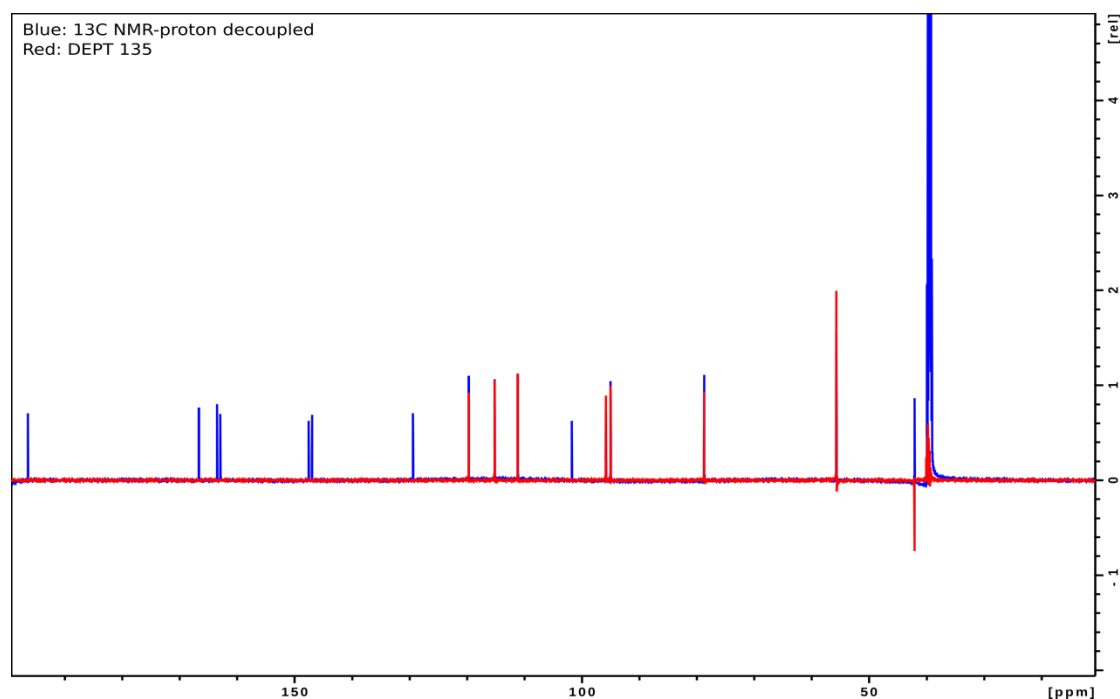
Figure S2. ^{13}C NMR and DEPT-135 spectra of 3′-methyleriodictyol. ^{13}C NMR (150 MHz, DMSO-d_6): δ [ppm] = 196.43 (4-C), 166.66 (7-C), 163.48 (5-C), 162.93 (9-C), 147.55 (4′-C), 146.97 (3′-C), 129.39 (1′-C), 119.69 (5′-C), 115.17 (2′-C), 111.17 (6′-C), 101.75 (10-C), 95.80 (6-C), 94.99 (8-C), 78.72 (2-C), 55.69 (3′-OCH$_3$), 42.10 (3-C).

6

4. NMR analysis of 3′,4′-dimethyleriodictyol

In addition to 1D ^{1}H NMR spectra, NOESY and ^{1}H–^{13}C HSQC experiments were performed to confirm the structure of 3′,4′-dimethyleriodictyol.
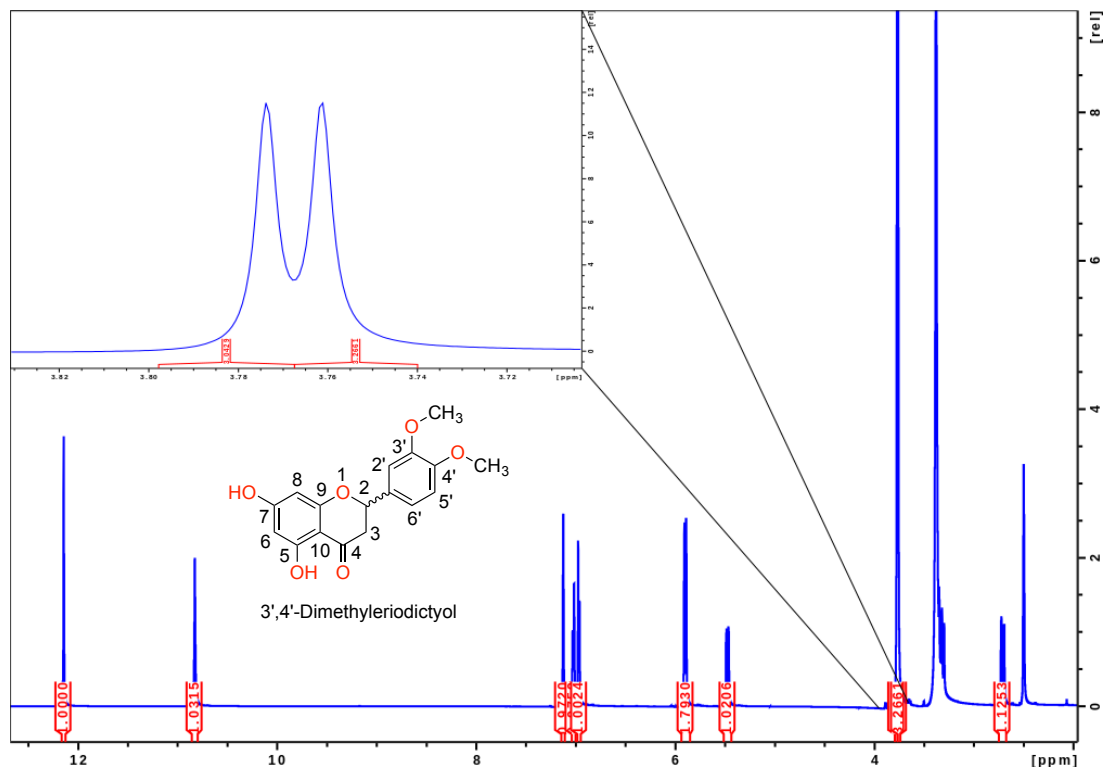
Figure S3. ^{1}H NMR spectrum of 3′,4′-dimethyleriodictyol. ^{1}H NMR (600 MHz, DMSO-d_6): δ [ppm] = 12.14 (s, 1H, 5-OH), 10.83 (s, 1H, 7-OH), 7.13 (d, J = 2.0 Hz, 1H, 2′-H), 7.02 (dd, J = 8.3, 2.0 Hz, 1H, 6′-H), 6.97 (d, J = 8.3 Hz, 1H, 5′-H), 5.91 (d, J= 2.2 Hz, 8-H), 5.89 (d, J= 1.9 Hz, 6-H), 5.48 (dd, J = 12.9, 3.1 Hz, 1H, 2-H), 3.77 (s, 3H, 3′-OCH$_3$) 3.76 (s, 3H, 4′-OCH$_3$), 3.32 (dd, J = 17.1, 12.8 Hz, 1H, *trans* 3-H), 2.71 (dd, J = 17.1, 3.1 Hz, 1H, *cis* 3-H).
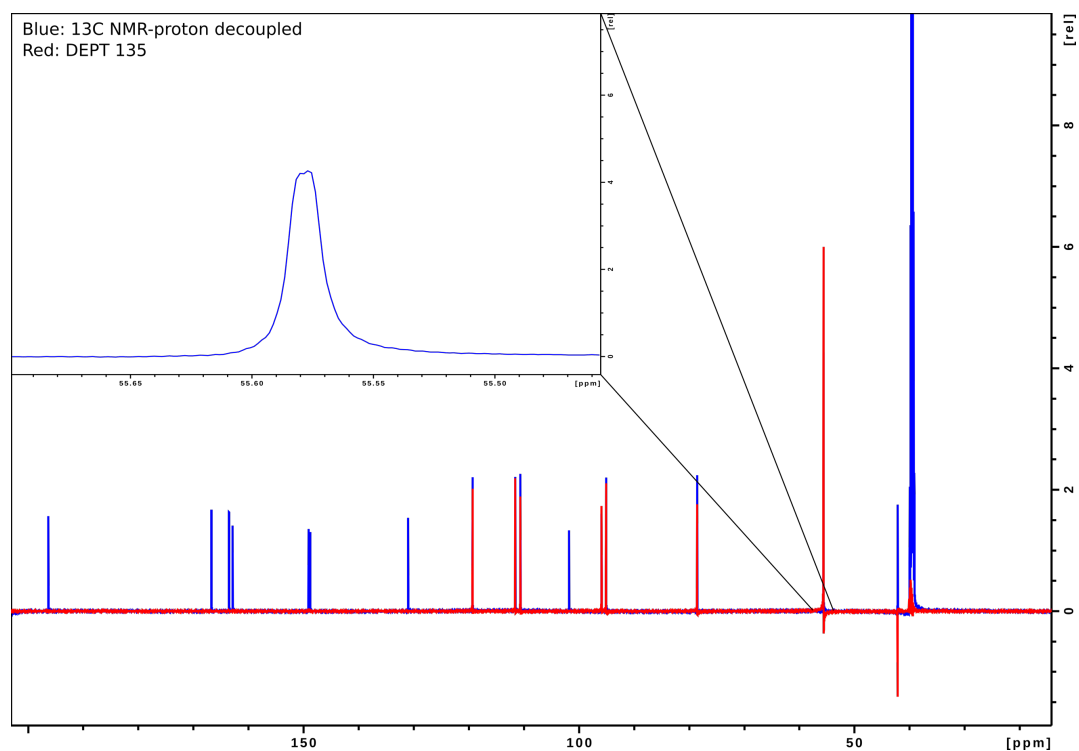
Figure S4. ^{13}C NMR and DEPT-135 spectra of 3′,4′-dimethyleriodictyol. ^{13}C NMR (150 MHz, DMSO-d_6): δ [ppm] = 196.28 (4-C), 166.69 (7-C), 163.50 (5-C), 162.85 (9-C), 149.06 (4′-C), 148.74 (3′-C), 130.98 (1′-C), 119.29 (6′-C), 111.53 (5′-C), 110.60 (2′-C), 101.77 (10-C), 95.87 (6-C), 95.04(8-C), 78.52 (3-C), 55.58 (OCH$_3$), 55.575 (OCH$_3$), 42.12 (2-C).
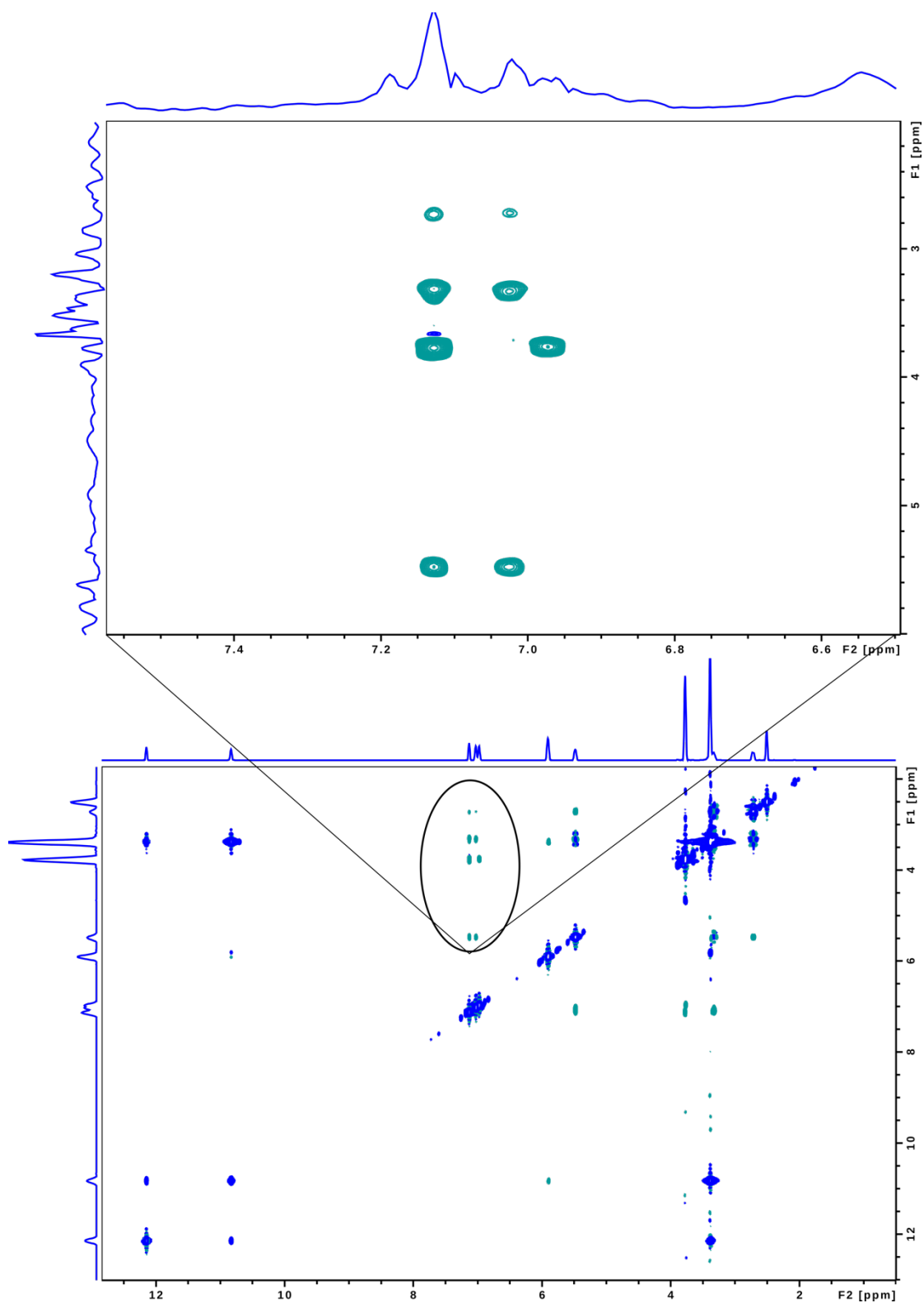
Figure S5. NOESY spectrum of 3′,4′-dimethyleriodictyol (300 ms mixing time). Enlarged are the NOESY cross-peaks: 7.13 (2′-H)–5.48 (2-H), 3.78 (3′-OCH₃), 3.31 (*trans* 3-H), & 2.73 (*cis* 3-H); 7.02 (6′-H)–5.48 (2-H), 3.33 (*trans* 3-H), & 2.73 (*cis* 3-H); 6.97 (5′-H)–3.76 (4′-OCH₃).
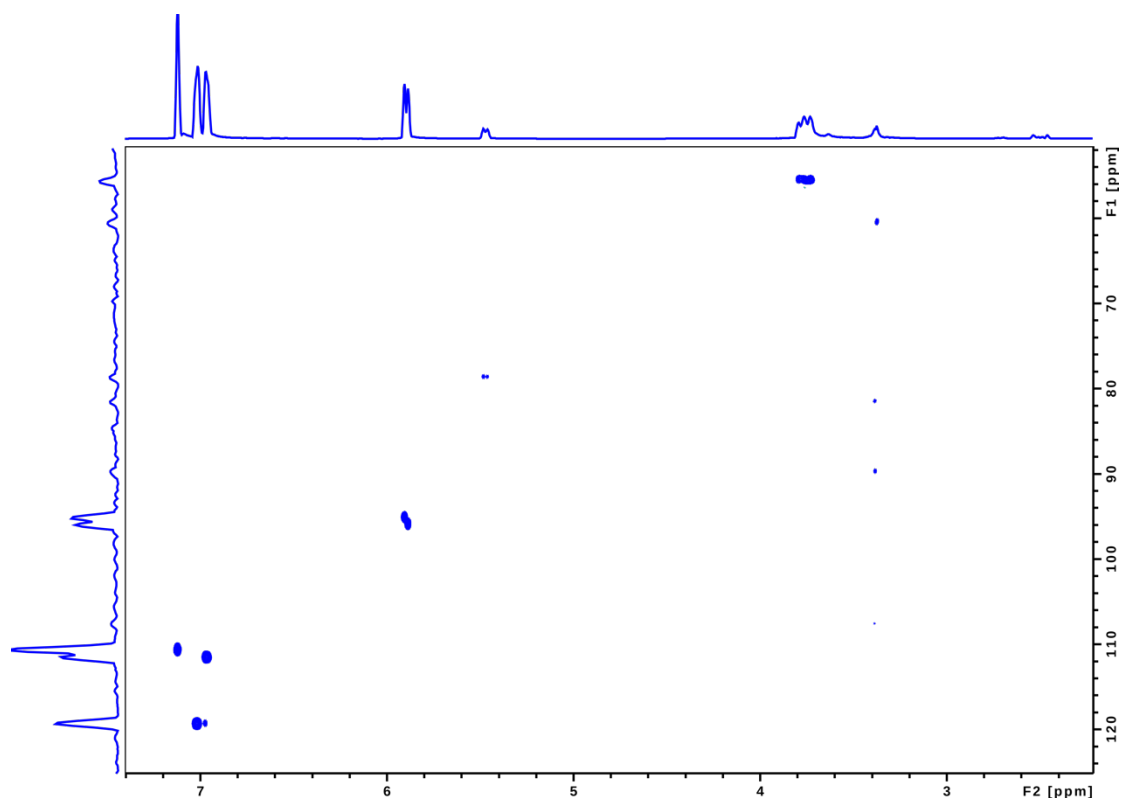
Figure S6. ^1H–^{13}C HSQC spectrum of 3′,4′-dimethyleriodictyol.

5. NMR analysis of 3′-methylluteolin

^1H, ^{13}C and DEPT-135 NMR spectra of the methylated luteolin corresponds to the 3′-methylluteolin standard.[2]
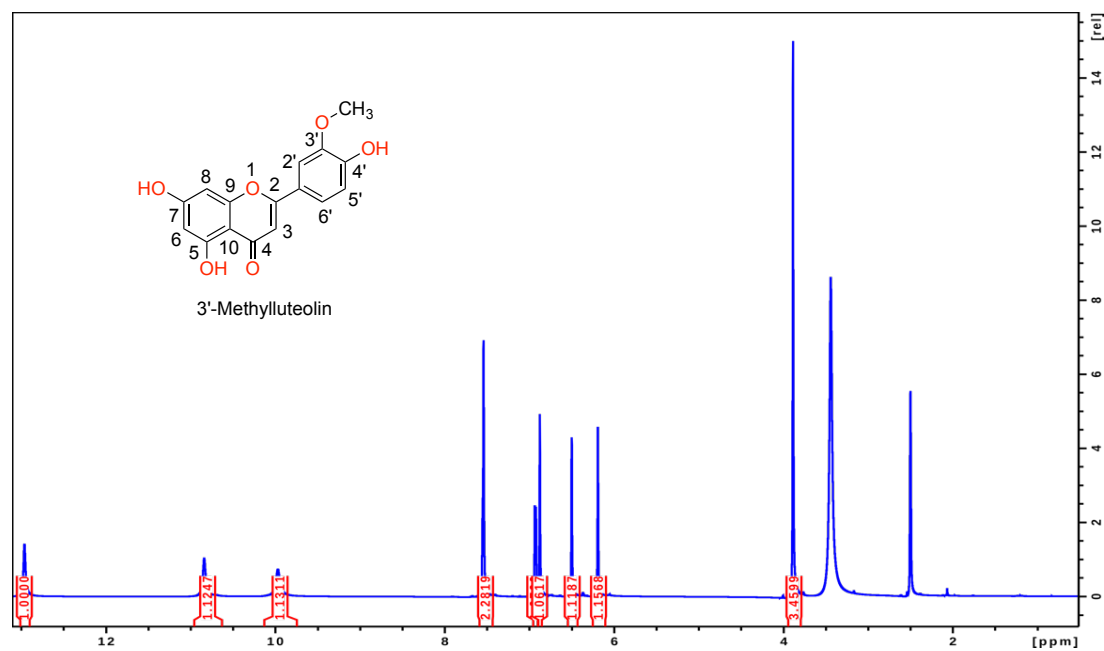
Figure S7. ^1H NMR spectrum of 3′-methylluteolin. ^1H NMR (600 MHz, DMSO-d_6): δ [ppm] = 12.96 (s, 1H, 5-OH), 10.84 (s, 1H, 7-OH), 9.97 (s, 1H, 4′-OH), 7.55 (d, 6′-H), 7.54 (s, 1H, 2′-H) 6.93 (d, J = 8.4 Hz, 1H, 5′-H), 6.88 (s, 1H, 3-H), 6.50 (d, J= 2.0 Hz, 1H, 8-H), 6.19 (m, 1H, 6-H), 3.89 (s, 3H, 3′-OCH₃).
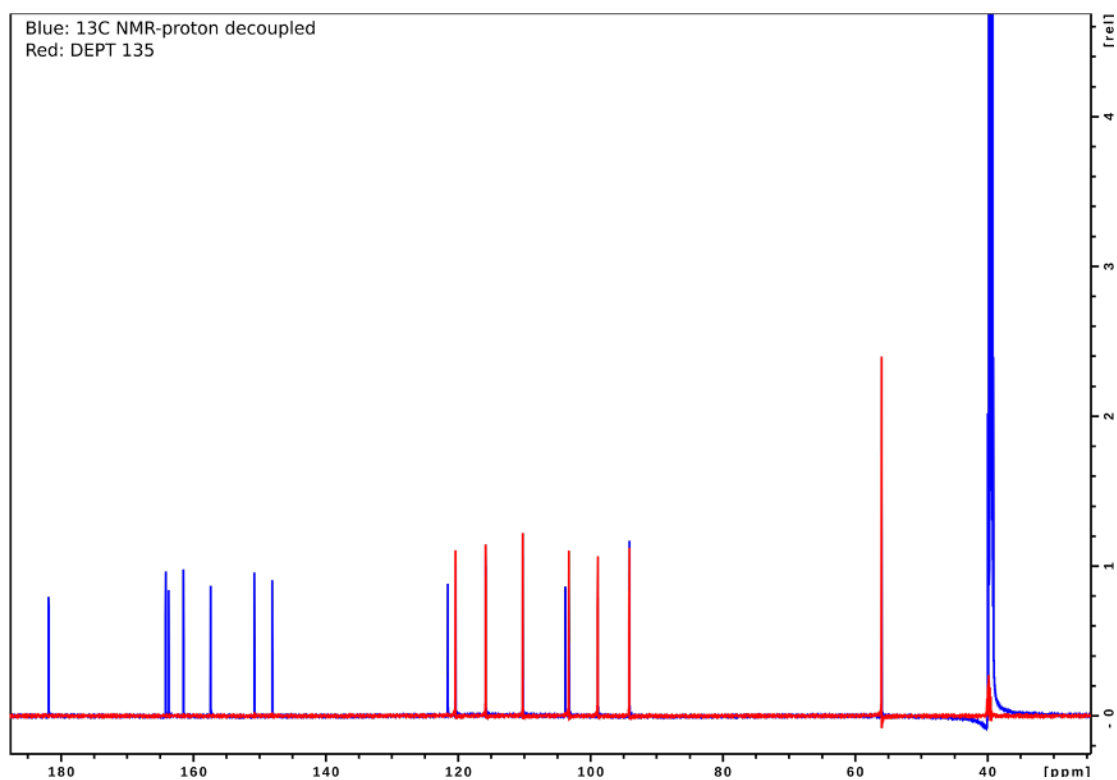
Figure S8. ^{13}C and DEPT-135 spectra of 3′-methylluteolin. ^{13}C NMR (150 MHz, DMSO-d_6): δ [ppm] = 181.82 (4-C), 164.15 (7-C), 163.68 (2-C), 161.45 (5-C), 157.35 (9-C), 150.74 (4′-C), 148.04 (3′-C), 121.54 (1′-C), 120.37 (6′-C), 115.78 (5′-C), 110.18 (2′-C), 103.74 (10-C), 103.22 (3-C), 98.84 (6-C), 94.07 (8-C), 55.96 (3′-OCH$_3$).

6. NMR analysis of 3′,4′-dimethylluteolin

In addition to 1D ^{1}H NMR spectra, NOESY, ^{1}H–^{13}C HSQC and ^{1}H–^{13}C HMBC experiments were also performed to confirm the structure of 3′,4′-dimethylluteolin. The 1D ^{1}H NMR spectrum corresponds to the 3′,4′-dimethylluteolin standard.[2]

Figure S9. ^{1}H NMR spectrum of 3′,4′-dimethylluteolin. ^{1}H NMR (600 MHz, DMSO-d_6): δ [ppm] = 12.91 (s, 1H, 5-OH), 10.85 (s, 1H, 7-OH), 7.64 (d, J = 8.1 Hz, 1H, 6′-H), 7.52 (s, 1H, 2′-H), 7.09 (d, J = 8.2 Hz, 1H, 5′-H), 6.92 (s, 1H, 3-H), 6.50 (s, 1H, 8-H), 6.19 (s, 1H, 6-H), 3.86 (s, 3H, 3′-OCH$_3$), 3.84 (s, 3H, 4′-OCH$_3$).

Figure S10. ^{13}C and DEPT-135 spectra of 3′,4′-dimethylluteolin. ^{13}C NMR (150 MHz, DMSO-d_6): δ [ppm] = 181.82 (4-C), 164.21 (7-C), 163.29 (2-C), 161.44 (5-C), 157.35 (9-C), 152.11 (4′-C), 149.0 (3′-C), 122.89 (1′-C), 119.99 (6′-C), 111.62 (5′-C), 109.36 (2′-C), 103.81 (3-C), 103.77 (10-C), 98.87 (6-C), 94.10 (8-C), 55.83 (3′-OCH$_3$), 55.71 (4′-OCH$_3$).

Figure S11. NOESY spectrum of 3′,4′-dimethylluteolin (300 ms mixing time). Circled NOESY cross-peak: 3.87 (3′-OCH$_3$)–7.52 (2′-H).
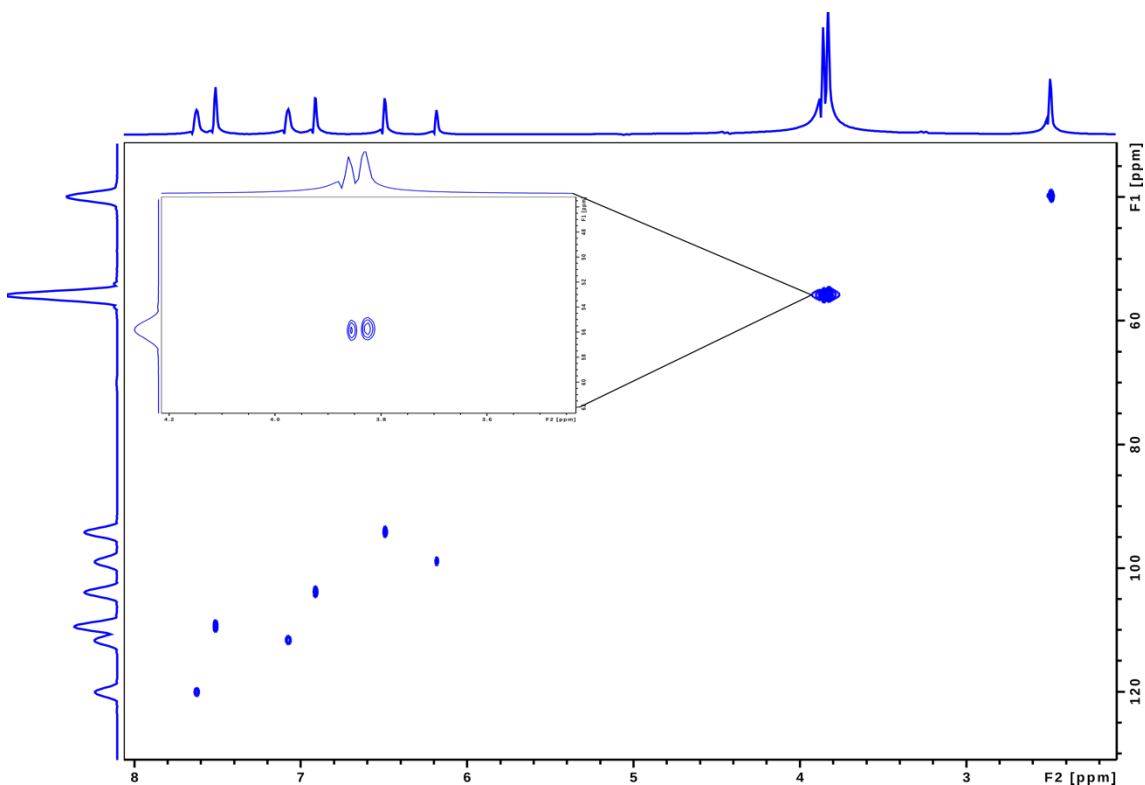
Figure S12. ^{1}H–^{13}C HSQC spectrum. HSQC cross-peaks: 3.86–55.83 (3′-OCH$_3$); 3.84–55.72 (3′-OCH$_3$).
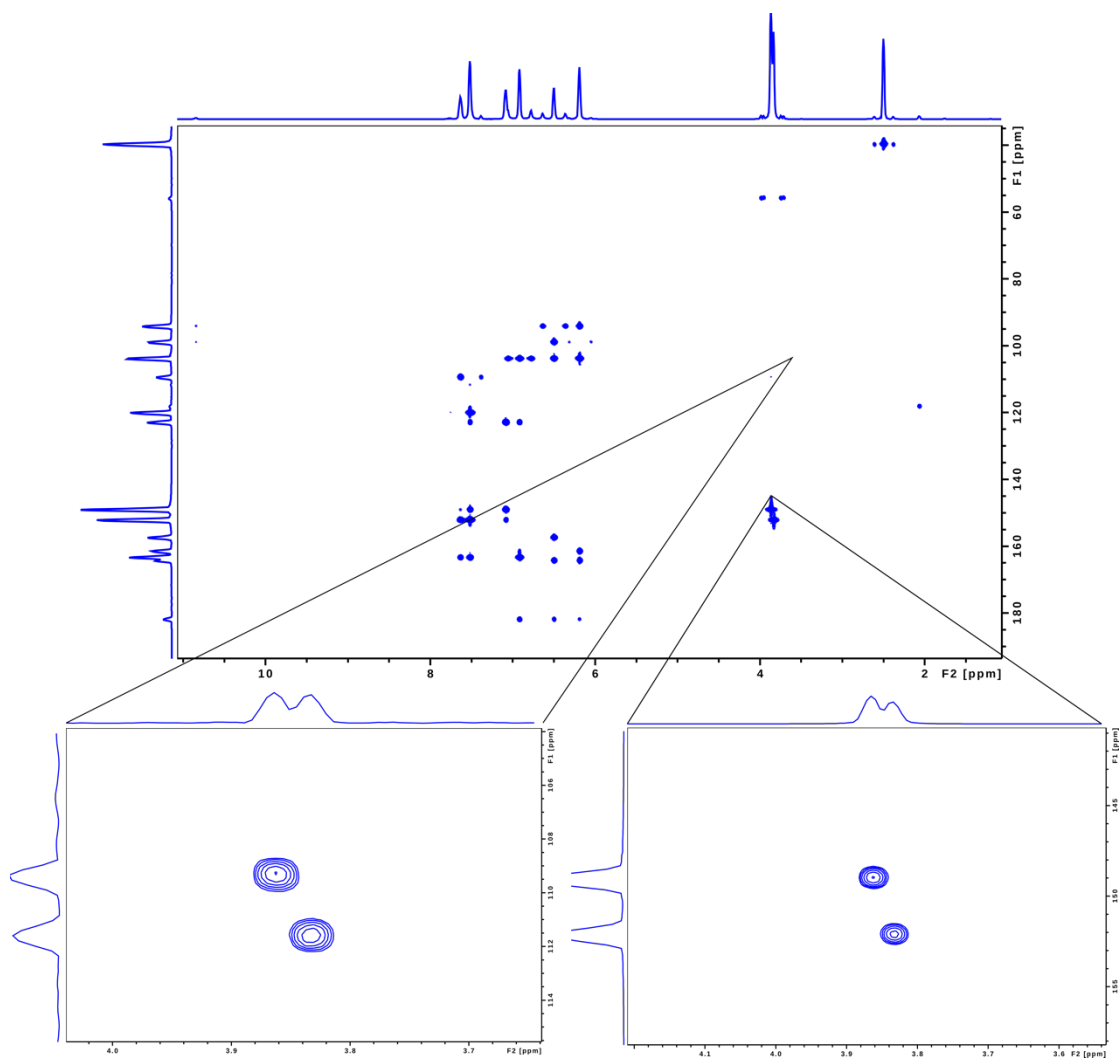
Figure S13. ^{1}H–^{13}C HMBC spectrum. HMBC cross-peaks: 3.86 (3′-OCH$_3$)–148.96 (3′-C) & 109.35 (2′-C); 3.83 (4′-OCH$_3$)–152.13 (4′-C) & 111.62 (5′-C).

7. NMR analysis of 3′-methylquercetin

^{1}H, ^{13}C and DEPT-135 NMR spectra of the methylated quercetin corresponds to the 3′-methylquercetin standard.[3]
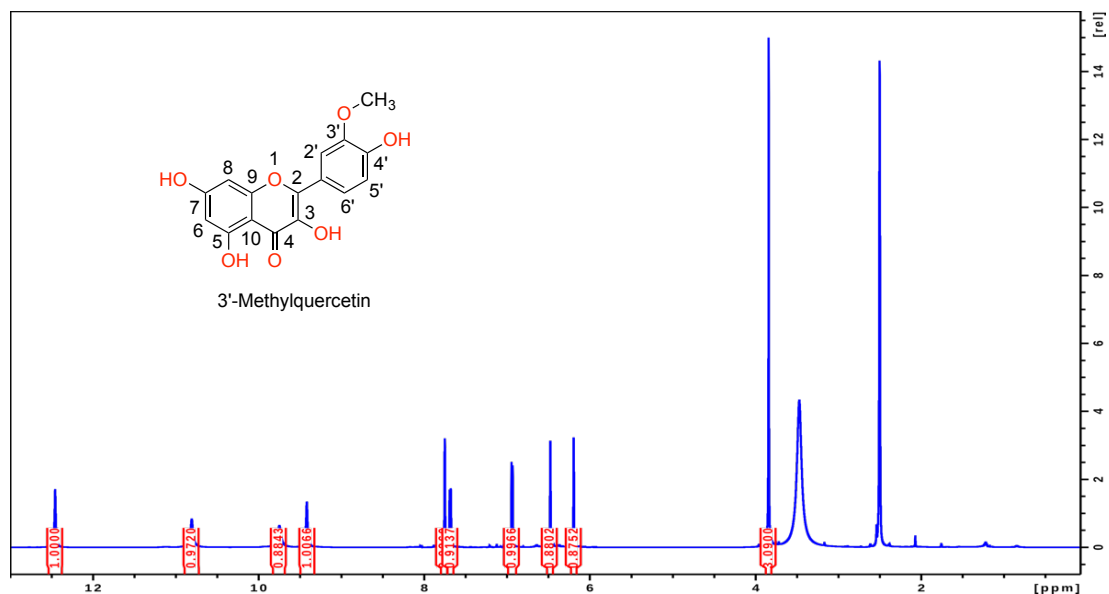
Figure S14. ^{1}H NMR spectrum of 3′-methylquercetin. ^{1}H NMR (600 MHz, DMSO-d_6): δ [ppm] = 12.46 (s, 1H, 5-OH), 10.81 (s, 1H, 7-OH), 9.75 (s, 1H, 4′-OH), 9.42 (s, 1H, 3-OH), 7.75 (d, J = 2.1 Hz, 1H, 2′-H), 7.69 (dd, J = 8.4, 2.1 Hz, 1H, 6′-H), 6.94 (d, J = 8.5 Hz, 1H, 5′-H), 6.48 (d, J = 2.0 Hz, 1H, 8-H), 6.20 (d, J = 2.1 Hz, 1H, 6-H), 3.84 (s, 3H, 3′-OCH$_3$).
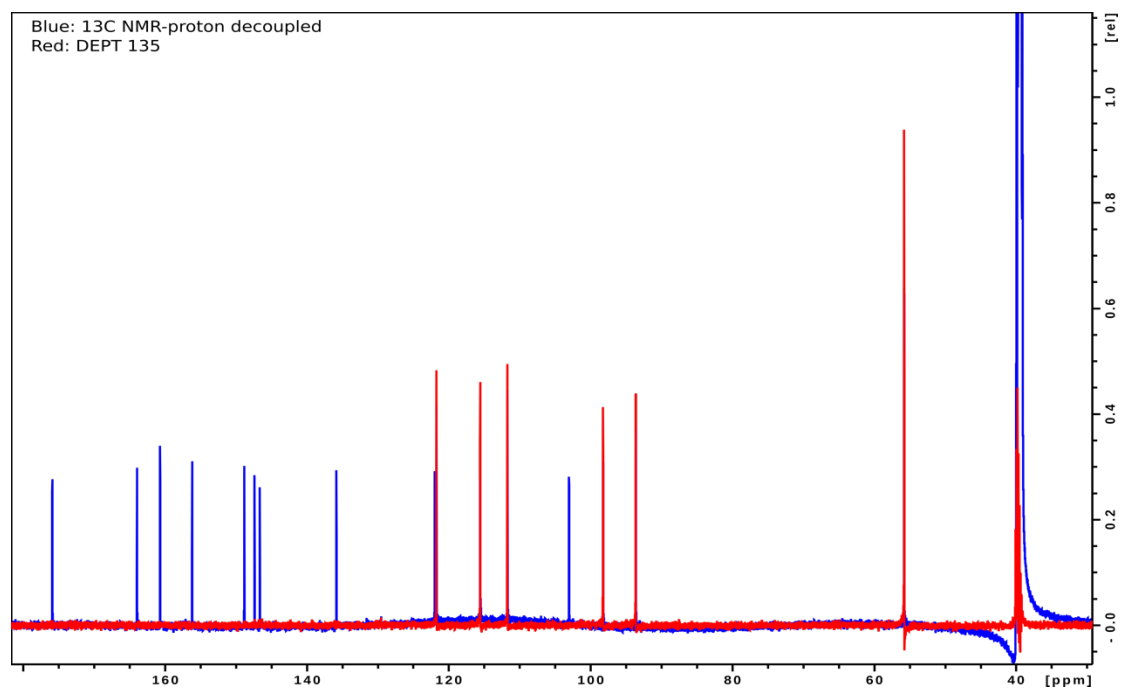
Figure S15. ^{13}C and DEPT-135 spectra of 3′-methylquercetin. ^{13}C NMR (150 MHz, DMSO-d_6): δ [ppm] = 175.87 (4-C), 163.93 (7-C), 160.68 (5-C), 156.15 (9-C), 148.81 (4′-C), 147.36 (3′-C), 146.62 (2-C), 135.81 (3-C), 121.96 (1′-C), 121.71 (6′-C), 115.54 (5′-C), 111.70 (2′-C), 103.02 (10-C), 98.21 (6-C), 93.60 (8-C), 55.76 (3′-OCH$_3$).

8. NMR analysis of 3′,4′-dimethylquercetin

In addition to 1D ^{1}H NMR spectra, NOESY, and ^{1}H–^{13}C HMBC experiments were performed to confirm the structure of 3′,4′-dimethylquercetin.
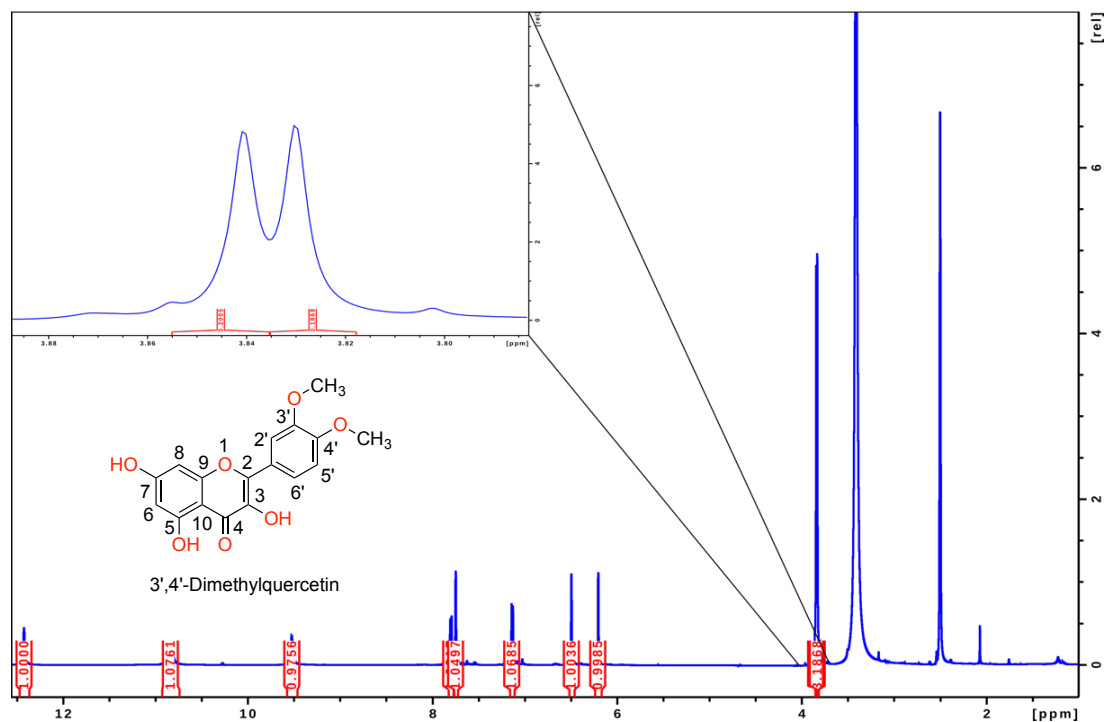
Figure S16. ^{1}H NMR spectrum of 3′,4′-dimethylquercetin. ^{1}H NMR (600 MHz, DMSO-d_6): δ [ppm] = 12.42 (s, 1H, 5-OH), 10.83 (s, 1H, 7-OH), 9.52 (3-OH s, 1H, 3-OH), 7.80 (d, J = 8.5 Hz, 1H, 5′-H), 7.75 (s, 1H, 2′-H), 7.13 (d, J = 8.6 Hz, 1H, 6′-H), 6.49 (d, J= 1.8 Hz, 1H, 8-H), 6.20 (d, J = 2.0 Hz, 1H, 6-H), 3.84 (s, 3H, 4′-OCH$_3$), 3.83 (s, 3H, 3′-OCH$_3$).
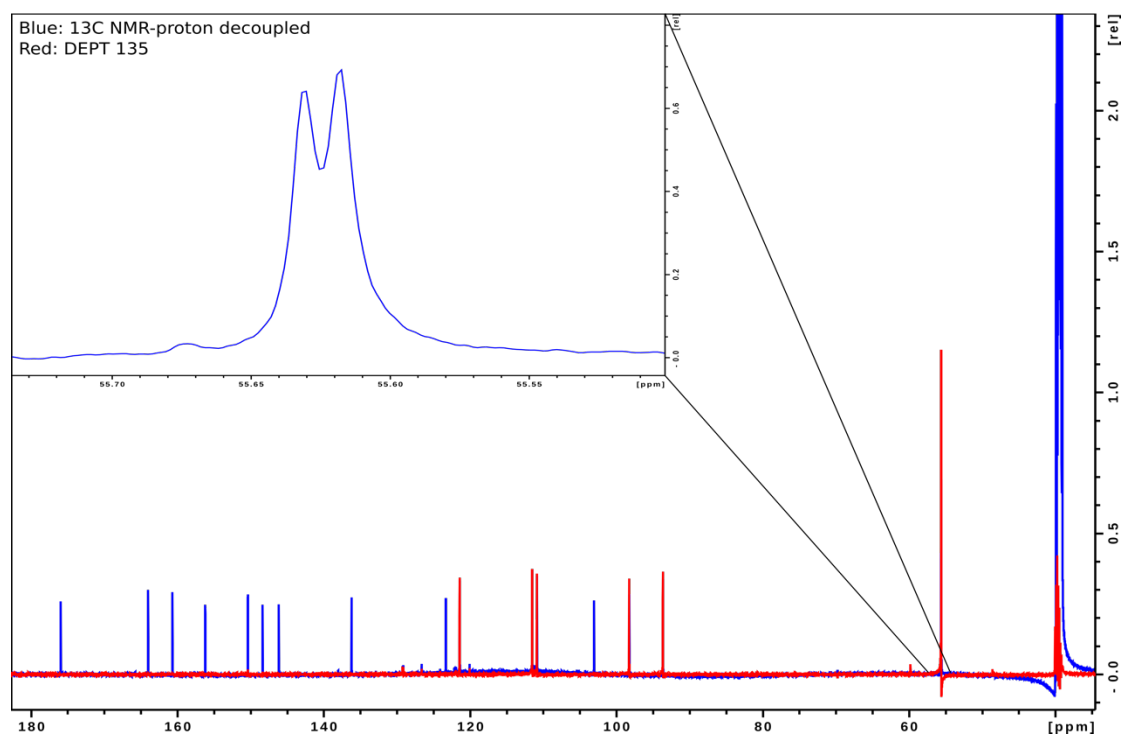
Figure S17. ^{13}C and DEPT-135 spectra of 3′,4′-dimethylquercetin. ^{13}C NMR (150 MHz, DMSO-

d₆): δ [ppm] = 175.98 (4-C), 164.03 (7-C), 160.70 (5-C), 156.22 (9-C), 150.39 (4′-C), 148.38 (3′-C), 146.17 (2-C), 136.22 (3-C), 123.33 (1′-C), 121.44 (6′-C), 111.53 (5′-C), 110.89 (2′-C), 103.08 (10-C), 98.26 (6-C), 93.66 (8-C), 55.63 (OCH₃), 55.62 (OCH₃).
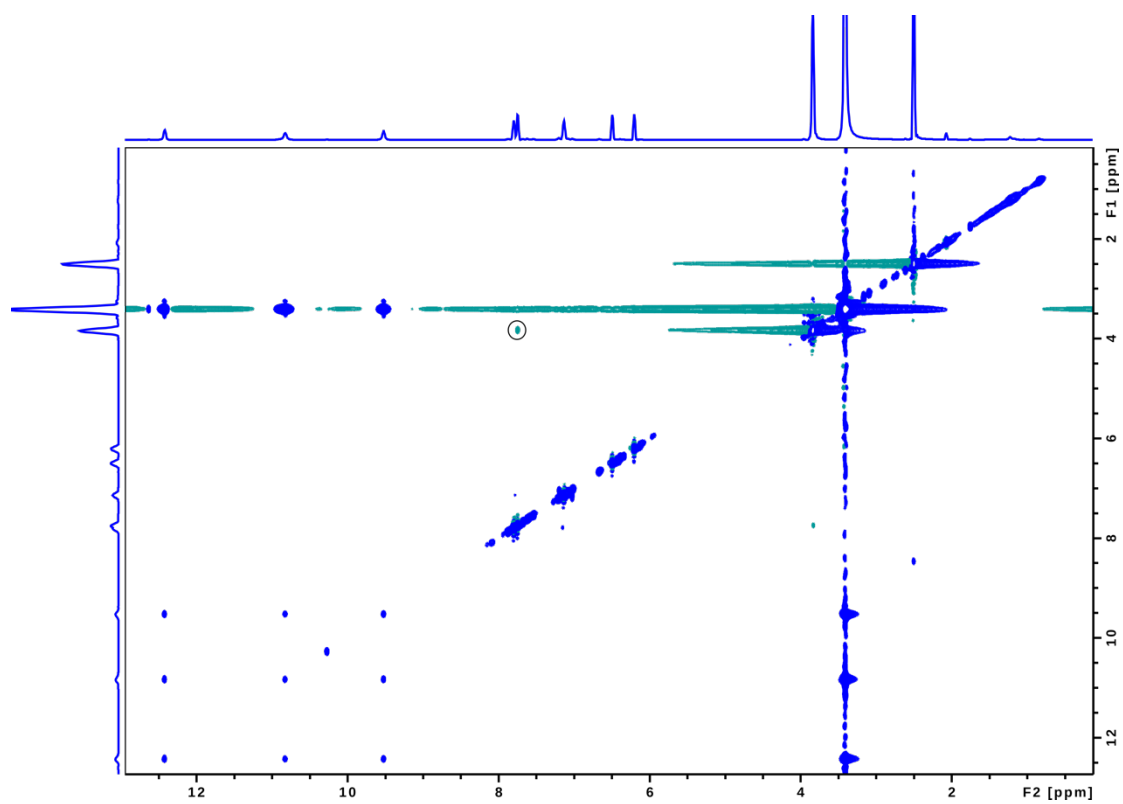
Figure S18. NOESY spectrum of 3′,4′-dimethylquercetin (300 ms mixing time). Circled NOESY cross-peak: 3.83 (3′-OCH₃)–7.75 (2′-H).
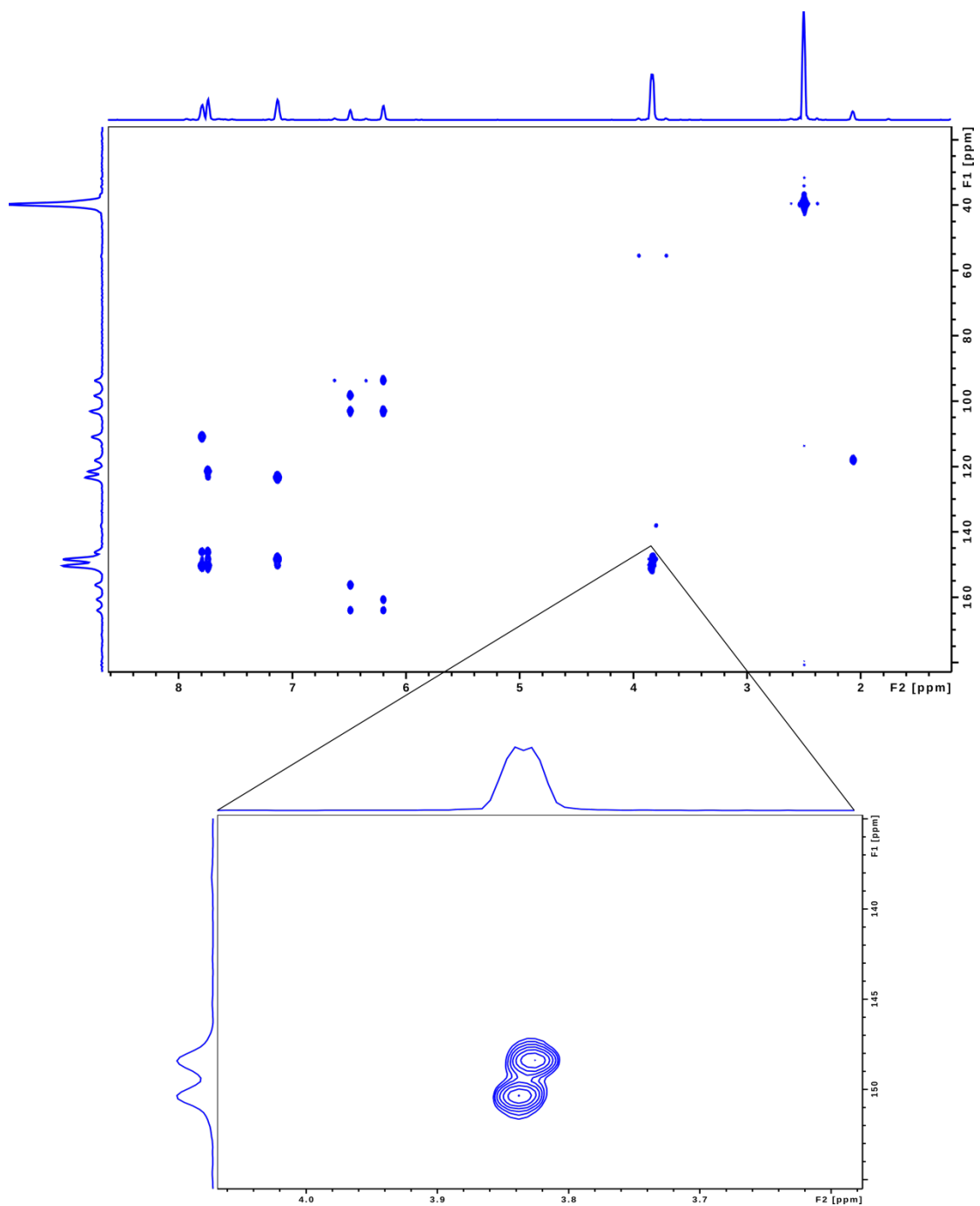
Figure S19. ^{1}H–^{13}C HMBC spectrum. HMBC cross-peak: 3.84 (4′-OCH$_3$)–150.38 (4′-C); 3.83 (3′-OCH$_3$)–148.37 (3′-C).

9. Mass spectrometry

[M-H]⁻ ions of precursors and products of each flavonoid substrate and product were obtained by LC-MS through MRM optimization in the negative ionization mode, following typical fragmentation modes for flavonoids and isoflavonoids (Table S1 and Figure S20).[4]

Table S1. Precursor and product ions of flavonoid substrates and products, obtained by the MRM optimization method in the negative ionization mode.

Substance	M. W. (g/mol)	Precursor ion* m/z (-)	Product ion m/z (-)
Eriodictyol	288.3	287	151, ^{1,3}A⁻; 135, ^{1,3}B⁻
3′-Methyleriodictyol	302.3	301	151, ^{1,3}A⁻
3′,4′-Dimethyleriodictyol	316.3	315	165, ^{0,2}B⁻
Luteolin	286.2	285	107, ^{0,4}A⁻; 133, ^{1,3}B⁻
3′-Methylluteolin	300.3	299	285, -CH_3; 257, -CH_3-CO
3′,4′-Dimethylluteolin	314.3	313	-
Quercetin	302.2	301	151, ^{1,3}A⁻
3′-Methylquercetin	316.3	315	151, ^{1,3}A⁻; 301, -CH_3
3′,4′-Dimethylquercetin	330.3	329	-
Naringenin	272.3	271	119, ^{1,3}B⁻
7-Methylnaringenin	286.3	285	119, ^{1,3}B⁻
Genistein	270.2	269	133, ^{0,3}B⁻
7-Methylgenistein	284.3	283	269, -CH_3
		283	133, ^{0,3}B⁻
4′-Methylgenistein	284.3	283	269, -CH_3
		283	240, -CH_3-CO

*Error for the mass detection is ± 1.
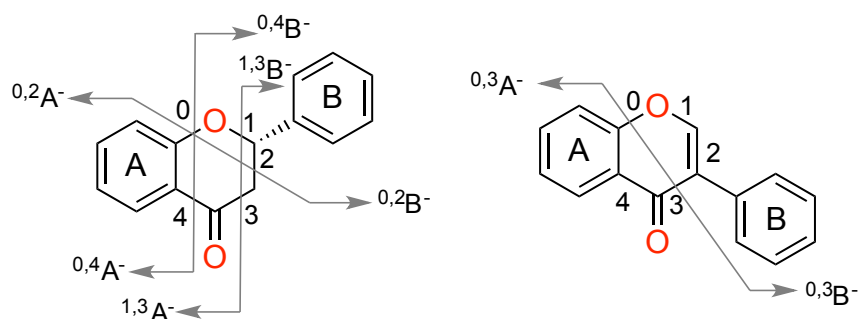
Figure S20. Fragmentation modes of flavonoids (left) and isoflavonoids (right).

References

[1]   A. S. Ibrahim, A. M. Galal, M. S. Ahmed, G. S. Mossa, *Chem. Pharm. Bull.* **2003**, *51*, 203.

[2]   E. Franco, M. Mélo, B. Jatobá, A. Santana, A. Silva, T. Silva, M. Nascimento, M. Maia, *Evid.-Based Complementary Altern. Med.* **2015**, *2015*.

[3]   E. J. Lee, B. H. Moon, Y. Park, S. W. Hong, S. H. Lee, Y. G. Lee, Y. H. Lim, *Bull. Korean Chem. Soc.* **2008**, *29*, 507.

[4]   a) N. Fabre, I. Rustan, E. de Hoffmann, J. Quetin Leclercq, *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 707; b) M. A. Farag, D. V. Huhman, R. A. Dixon, L. W. Sumner, *Plant Physiol.* **2008**, *146*, 387.

Article III

SAM Analogs | Hot Paper

Directed Evolution of a Halide Methyltransferase Enables Biocatalytic Synthesis of Diverse SAM Analogs

Qingyun Tang, Christoph W. Grathwol, Aşkın S. Aslan-Üzel, Shuke Wu, Andreas Link, Ioannis V. Pavlidis,* Christoffel P. S. Badenhorst,* and Uwe T. Bornscheuer*

This article is dedicated to Paul Kamer, an outstanding scientist

Abstract: *Biocatalytic alkylations are important reactions to obtain chemo-, regio- and stereoselectively alkylated compounds. This can be achieved using S-adenosyl-L-methionine (SAM)-dependent methyltransferases and SAM analogs. It was recently shown that a halide methyltransferase (HMT) from Chloracidobacterium thermophilum can synthesize SAM from SAH and methyl iodide. We developed an iodide-based assay for the directed evolution of an HMT from Arabidopsis thaliana and used it to identify a V140T variant that can also accept ethyl-, propyl-, and allyl iodide to produce the corresponding SAM analogs (90, 50, and 70% conversion of 15 mg SAH). The V140T AtHMT was used in one-pot cascades with O-methyltransferases (IeOMT or COMT) to achieve the regioselective ethylation of luteolin and allylation of 3,4-dihydroxybenzaldehyde. While a cascade for the propylation of 3,4-dihydroxybenzaldehyde gave low conversion, the propyl-SAH intermediate could be confirmed by NMR spectroscopy.*

The "magic methyl effect" refers to the ability of an appropriately placed methyl group to dramatically alter the biological properties of a compound.[1] This makes selective alkylation of molecules highly desirable.[2] Bioalkylations are important because of the exquisite chemo-, regio- and

stereospecificity achievable using enzymes.[2,3] In nature, selective methylation is catalyzed by methyltransferases (MT, E.C. 2.1.1.X) that use *S*-adenosyl-L-methionine (SAM, "nature's methyl iodide") as methyl donor.[4] Interestingly, many methyltransferases are insensitive to the size of the alkyl substituent of SAM and could catalyze other alkylation reactions if the necessary SAM analogs were available.[5] This promiscuity can be harnessed to expand the structural and functional diversity of chemicals and enable various applications, such as site-selective modification of molecules with fluorescent or "clickable" groups.[4a,5a]

SAM analogs enabling these diverse alkylation reactions are crucial not only for expanding the industrial relevance of biocatalytic alkylation but also for discovering promiscuous MTs.[1c,6] However, very few naturally occurring SAM analogs[5a,6,7] are known, making limited access to SAM analogs one of the most serious impediments to progress in the field. With the exception of *S*-adenosyl-L-ethionine (SAE), SAM analogs are not readily available from commercial suppliers and have to be prepared on demand. Alkylation of *S*-adenosyl-L-homocysteine (SAH) using alkyl halides results in low yields and contamination with the biologically inactive (*R,S*) diastereomers.[5b] These side products can be potent MT inhibitors and are not easily separated from the desired SAM analogs.[8] Enzymatic synthesis of SAM analogs is more specific and yields only the desired isomers. Methionine adenosyltransferases (MATs) or halogenases catalyze the production of SAM analogs by combining the adenosyl moiety of ATP or 5′-chloro-5′-deoxyadenosine, respectively, to the sulfur atom of methionine analogs.[9] The major drawback of these approaches is that they require methionine analogs, which are expensive, if at all commercially available (Figure S1).

Liao and Seebeck recently reported the enzymatic synthesis of SAM from SAH and methyl iodide, using a halide methyltransferase (HMT) from *Chloracidobacterium thermophilum* (*Ct*HMT).[10] We realized that, like many MTs, some HMTs would be promiscuous and that this would allow a diverse set of SAM analogs to be enzymatically synthesized using cheap and readily available alkyl iodides, without having to chemically synthesize methionine analogs.

We expressed and purified the HMT from *Chloracidobacterium thermophilum* (*Ct*HMT) originally described by Liao and Seebeck.[10] We constructed an SAH nucleosidase-deficient strain of *E. coli* BL21(DE3) to avoid contamination of recombinant enzymes with this SAH-degrading enzyme

[*] M. Eng. Q. Tang, M. Sc. A. S. Aslan-Üzel, Dr. S. Wu,
Dr. C. P. S. Badenhorst, Prof. Dr. U. T. Bornscheuer
Institute of Biochemistry, University of Greifswald
Felix-Hausdorff-Strasse 4, 17489 Greifswald (Germany)
E-mail: chris.badenhorst@uni-greifswald.de
uwe.bornscheuer@uni-greifswald.de

Dr. C. W. Grathwol, Prof. Dr. A. Link
Institute of Pharmacy, University of Greifswald
Friedrich-Ludwig-Jahn-Strasse 17, 17489 Greifswald (Germany)

Prof. Dr. I. V. Pavlidis
Department of Chemistry, University of Crete
Voutes University Campus
70013 Heraklion (Greece)
E-mail: ipavlidis@uoc.gr

(Method S1.3). We also prepared HMTs from *Arabidopsis thaliana* (*At*HMT) and *Raphanus sativus* (*Rs*HMT) as these have been well characterized and have high activities towards iodide.[11] We assessed the abilities of the three recombinant HMTs to alkylate SAH, using methyl iodide (MeI), ethyl iodide (EtI), propyl iodide (PrI), and butyl iodide (BuI) as alkyl donors. All three HMTs were active using methyl iodide, although *Ct*HMT was significantly less active than *At*HMT and *Rs*HMT (Table S1). As expected, all three enzymes had promiscuous activity against ethyl iodide, although this was three orders of magnitude lower than methyltransferase activity for each of the enzymes (Table S1). However, none of the enzymes had significant activity towards propyl iodide or butyl iodide (less than 0.01 µmol min^{-1}mg^{-1}). *At*HMT had the highest ethyltransferase activity (2.69 ± 0.15 µmol min^{-1}mg^{-1}) and a solved crystal structure and was therefore chosen as the starting point for semi-rational protein engineering.
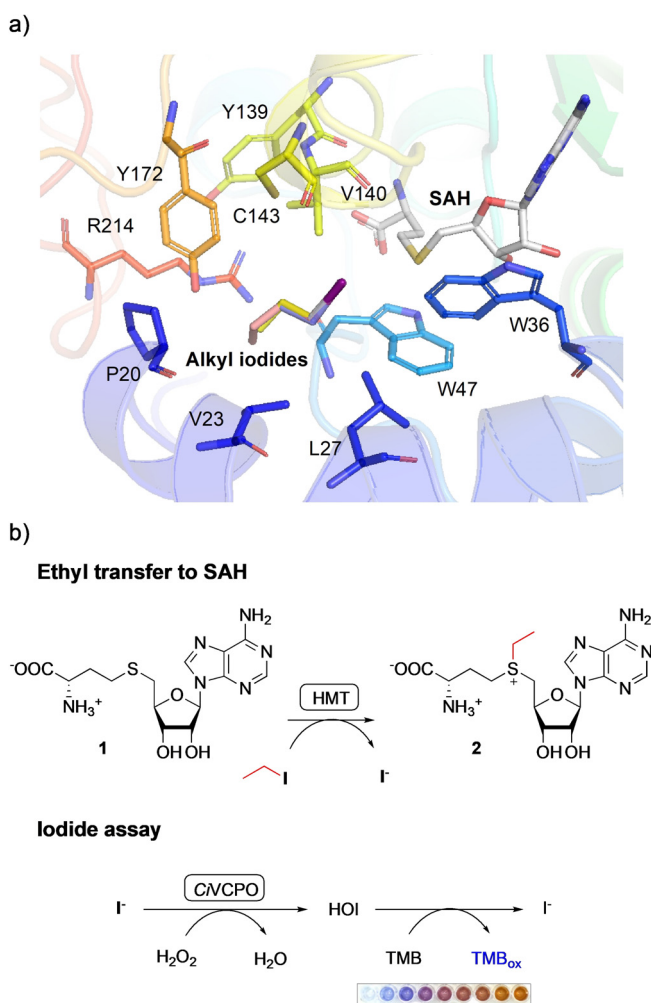
Figure 1. a) The crystal structure of *At*HMT and the modelled alkyl iodides in the active site. The selected mutated sites, SAH, and the alkyl iodides, which include MeI, EtI, PrI, and BuI, are shown as sticks with elemental coloring. The carbons of SAH, MeI, EtI, PrI, and BuI are shown in white, grey, purple, yellow, and pink, respectively. b) Transfer of the ethyl group from ethyl iodide to SAH (**1**) and the production of SAE (**2**) and iodide. The iodide produced can be detected using hydrogen peroxide, a haloperoxidase, and TMB.

We docked EtI, PrI, and BuI into the active site of the *At*HMT crystal structure (PDB ID: 3LCC) with bound SAH (Figure 1 a).[12] The alkyl-binding pocket is quite spacious and could accommodate even butyl iodide (in agreement with alkyl promiscuity of many methyltransferases). The iodine atoms of all docked substrates were superposed. We selected the residues forming the alkyl-binding site (P20, V23, L27, W36, W47, Y139, V140, C143, Y172, and R214) for further investigation. For each of these positions, an NNK site-saturation library was created using mutagenesis PCR. For each NNK library, 96 clones were screened by incubating crude lysate with 1 mM SAH and 5 mM ethyl iodide (Supporting Method S2). While the formation of SAE could be monitored by HPLC, this would require 20 hours per library, making it very tedious to investigate all randomized residues. Therefore, we needed a high-throughput method for determining the amount of product formed in each assay. Because an iodide ion is released for each molecule of SAE formed (Figure 1 b), we initially considered using our recently published ultrasensitive halide assay.[13] However, we were concerned that background signals originating from chloride in crude cell lysates would complicate the interpretation of results. Therefore, we developed a modified assay that is insensitive to chloride, but highly sensitive to iodide.

We used a recombinant *Curvularia inaequalis* vanadium-dependent chloroperoxidase (*Ci*VCPO) to oxidize iodide to hypoiodous acid (HOI). The HOI formed oxidizes the chromogen 3,3′,5,5′-tetramethylbenzidine (TMB), resulting in the formation of blue color (Figure S2a).[14] In the process, iodide is released again, so that one iodide anion can catalyze the oxidation of multiple TMB molecules (Figure 1 b). This results in very high sensitivity, and only 1 µL of sample is required for the assay. The change in absorbance at 570 nm is directly proportional to iodide concentration in the range from 5 to 400 µM (Figure S3a). Chloride concentrations up to 1 M result in only slight color changes (Figure S2b). Most importantly, our assay is general and suitable for monitoring the release of iodide from a range of alkyl iodides. This enables both high-throughput screening of mutant libraries and screening of HMT variants against various alkyl iodides. Activities of HMTs determined using the iodide assay were comparable to those determined using HPLC (Figure S3c), demonstrating the validity and reliability of the assay.

The 10 NNK libraries at the selected positions were screened and four hits (V23T, W36F, V140C, and V140T) with improved activities relative to the wild-type *At*HMT were identified (Figure S4). Based on these hits, we constructed NNK libraries at each other hit position, for a total of six additional libraries (Table S2). Hits from screening these libraries were double-mutant combinations of V23K/L/T, W36F, and V140C/T. Finally, we constructed triple mutants based on the single and double mutants. We expressed and purified the wild-type *At*HMT and all these variants and determined their specific activities towards ethyl iodide (Figure 2). The specific activity of the most active V140T variant (14.02 µmol min^{-1}mg^{-1}) towards ethyl iodide was fivefold higher than that of the wild-type *At*HMT (2.69 µmol min^{-1}mg^{-1}).
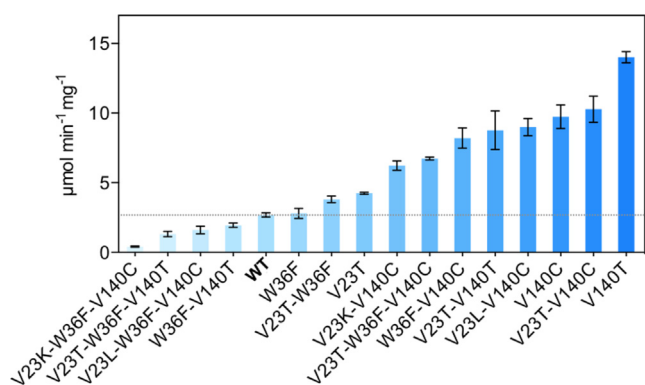
Figure 2. Specific activities of wild-type *At*HMT and several single, double, and triple mutants, determined using ethyl iodide as substrate. Data plotted are the means, with standard deviation, of three independent measurements. The data used for plotting this graph are shown in Table S3.

Kinetic analysis (Figures S5 and S6) showed that the activity improvement relied on an increased k_{cat} and not decreased K_M for ethyl iodide (Table 1). The k_{cat} value for the V140T was 6-fold higher than that for the wild type. The relative preference for ethyl transfer was 36-fold higher for the V140T variant than for the wild-type *At*HMT.

In addition to increased ethyltransferase activity, the V140T variant was also more active towards propyl iodide and allyl iodide compared to the wild type (Figure S7). The V140T variant was used to convert SAH (15 mg) to SAE (**2**), *S*-propyl-L-homocysteine (SAP (**3**)), and *S*-allyl-L-homocysteine (SAA (**4**)) using ethyl iodide, propyl iodide, and allyl iodide. Conversions were 90, 50, and 70% after 14, 24, and 14 h, respectively (Figure 3a and Table S4).

The V140T variant is therefore valuable for the synthesis of various SAM analogs. However, using stoichiometric amounts of these alkylating agents would be expensive due to poor atom economy and the high cost of SAH. Therefore, we used the V140T *At*HMT for in vitro regeneration of SAE, SAP, and SAA, from catalytic amounts of SAH, in one-pot biocatalytic alkylation cascades (Figure 3b). The T133M–Y326L variant of isoeugenol *O*-MT (IeOMT) and the V140T *At*HMT catalyzed regioselective mono-ethylation of luteolin to 3′-*O*-ethylluteolin (**6**) with 41% conversion.[15] Similarly, human catechol *O*-MT (COMT) in cascade with V140T *At*HMT catalyzed regioselective mono-allylation of 3,4-dihydroxybenzaldehyde to produce 4-allyloxy-3-hydroxybenzaldehyde (**8**) with 48% conversion (Table S5).[16] The identities of the products were confirmed
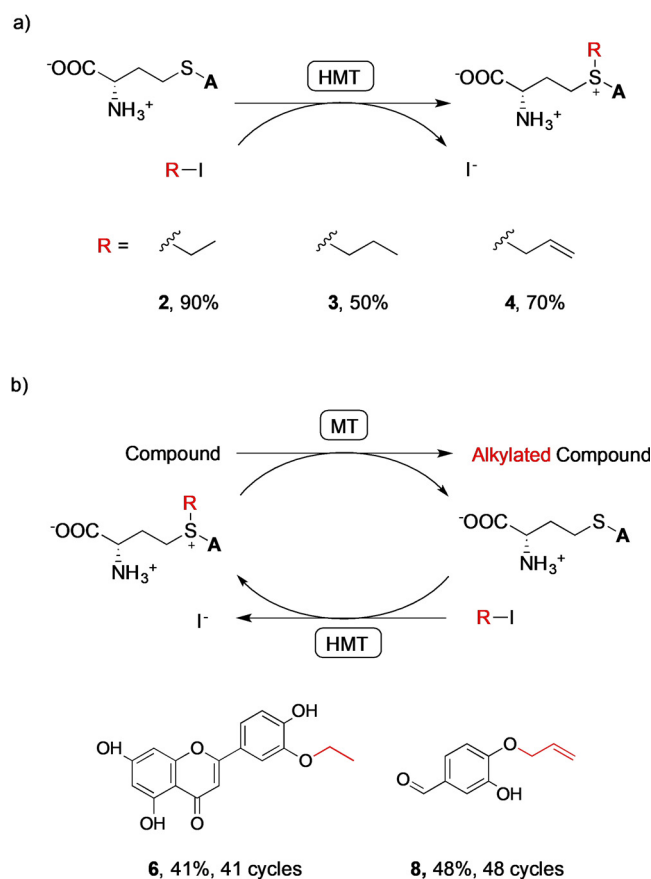
Figure 3. a) Preparative-scale (15 mg) synthesis of SAE (**2**), SAP (**3**), and SAA (**4**) catalyzed by the V140T *At*HMT. The "A" represents the adenosyl moiety. Conversions shown are from Table S4. b) Production (10–20 mg scale) of alkylated products using cyclic MT-HMT cascades, employing 100 µM SAH and 80 mM alkyl iodide. The IeOMT variant T133M-Y326L and the V140T *At*HMT catalyzed the ethylation of luteolin and produced 4′-*O*-ethylluteolin (**6**) with 41% conversion. Human COMT and V140T *At*HMT catalyzed the allylation of 3,4-dihydroxybenzaldehyde and produced 4-allyloxy-3-hydroxybenzaldehyde (**8**) with 48% conversion. Insignificant conversion took place if the HMT and MT were not added (Figures S9 and S11). Conversions are from Table S5 and numbers of SAH regeneration cycles were calculated as [product]/[SAH]$_{t=0}$.

by ^{1}H and ^{13}C NMR spectroscopy (Figures S12–S17). The conversions achieved correspond to 41 and 48 SAH regeneration cycles. While a cascade of human COMT and V140T *At*HMT catalyzed the propylation of 3,4-dihydroxybenzaldehyde, the conversion was low (ca. 5%) and the product could not be purified for NMR spectroscopy. Therefore, we isolated

Table 1: Kinetic parameters of wild-type *At*HMT and the V140T variant.

Variant	Methyltransferase activity			Ethyltransferase activity			Substrate preference [EtI]/[MeI][a]
	k_{cat} [min^{-1}]	$K_{M, MeI}$ [mM]	$k_{cat}/K_{M, MeI}$ [min^{-1}/mM]	k_{cat} [min^{-1}]	$K_{M, EtI}$ [mM]	$k_{cat}/K_{M, EtI}$ [min^{-1}/mM]	
WT	364.4±12.7	1.43±0.19	254.5	0.36±0.03	9.36±1.66	0.038	1.5×10^{-4}
V140T	182.4±11.7	4.70±0.76	38.7	2.38±0.16	11.14±1.52	0.214	5.5×10^{-3}

[a] Substrate preference for EtI relative to MeI is the ratio of k_{cat}/K_M values for the two substrates. This preference is ca. 36-fold higher for the V140T variant.

SAP synthesized using the V140T *At*HMT (Figure S9b) and confirmed its structure by ^{1}H and ^{13}C NMR (Figures S18 and S19).

We have demonstrated that the scope of biocatalytic alkylations can be rapidly expanded by harnessing the promiscuity of an engineered halide methyltransferase. The HMT from *Arabidopsis thaliana* enabled us to produce three SAM analogs from cheap and readily available alkyl iodides. Importantly, we also demonstrated application of the engineered HMT in bioalkylation cascades. It could catalyze over 40 cycles of alkylation and SAH regeneration, allowing SAH to be used in catalytic rather than stoichiometric amounts (Figure 3b). As noted by Liao and Seebeck, the use of toxic alkyl iodides as reagents might be considered a drawback of this approach. However, the same applies to chemical alkylation and the biocatalytic alternative functions under milder conditions and offers chemo-, regio-, and stereoselectivity.[17]

Our best variant (V140T) had the highest activity against all alkyl iodides tested (Figure S7), demonstrating that screening using ethyl iodide can increase activity against other alkyl iodides of interest. However, the ethyltransferase activity of V140T is significantly higher than its propyltransferase activity, as "you get what you screen for".[18] We screened using ethyl iodide to maximize chances of finding improved variants, but the libraries and screening methodology reported here could be used to identify variants with improved activity towards other alkyl iodides. Exploring the intrinsic alkyl-substituent promiscuity of naturally occurring HMTs presents a very promising avenue for future research. Not only might a highly promiscuous variant already exist in nature, but also chances of finding an optimal starting point for directed evolution would be improved by characterizing a larger number of extant proteins.

Conflict of interest

The authors declare no conflict of interest.

[1] a) H. Schönherr, T. Cernak, *Angew. Chem. Int. Ed.* **2013**, *52*, 12256–12267; *Angew. Chem.* **2013**, *125*, 12480–12492; b) E. J. Barreiro, A. E. Kümmerle, C. A. M. Fraga, *Chem. Rev.* **2011**, *111*, 5215–5246; c) J. N. Andexer, A. Rentmeister, *Nat. Chem.* **2020**, *12*, 791–792.

[2] I. J. W. McKean, P. A. Hoskisson, G. A. Burley, *ChemBioChem* **2020**, *21*, 2890–2897.

[3] a) S. Wu, R. Snajdrova, J. C. Moore, K. Baldenius, U. T. Bornscheuer, *Angew. Chem. Int. Ed.* **2020**, https://doi.org/10.1002/anie.202006648; *Angew. Chem.* **2020**, https://doi.org/10.1002/ange.202006648; b) B. Hauer, *ACS Catal.* **2020**, *10*, 8418–8427.

[4] a) A. W. Struck, M. L. Thompson, L. S. Wong, J. Micklefield, *ChemBioChem* **2012**, *13*, 2642–2655; b) L. Wessjohann, M. Dippe, M. Tengg, M. Gruber-Khadjawi, in *Cascade Biocatalysis* (Eds.: S. Riva, W. D. Fessner), Wiley-VCH, Weinheim, **2014**, pp. 393–426.

[5] a) T. D. Huber, B. R. Johnson, J. Zhang, J. S. Thorson, *Curr. Opin. Biotechnol.* **2016**, *42*, 189–197; b) C. Dalhoff, G. Lukinavičius, S. Klimašauskas, E. Weinhold, *Nat. Chem. Biol.* **2006**, *2*, 31–32; c) B. J. Law, M. R. Bennett, M. L. Thompson, C. Levy, S. A. Shepherd, D. Leys, J. Micklefield, *Angew. Chem. Int. Ed.* **2016**, *55*, 2683–2687; *Angew. Chem.* **2016**, *128*, 2733–2737; d) B. J. Law, A.-W. Struck, M. R. Bennett, B. Wilkinson, J. Micklefield, *Chem. Sci.* **2015**, *6*, 2885–2892.

[6] A. J. Herbert, S. A. Shepherd, V. A. Cronin, M. R. Bennett, R. Sung, J. Micklefield, *Angew. Chem. Int. Ed.* **2020**, *59*, 14950–14956; *Angew. Chem.* **2020**, *132*, 15060–15066.

[7] a) J. Kim, H. Xiao, J. B. Bonanno, C. Kalyanaraman, S. Brown, X. Tang, N. F. Al-Obaidi, Y. Patskovsky, P. C. Babbitt, M. P. Jacobson, Y.-S. Lee, S. C. Almo, *Nature* **2013**, *498*, 123–126; b) I. J. W. McKean, J. C. Sadler, A. Cuetos, A. Frese, L. D. Humphreys, G. Grogan, P. A. Hoskisson, G. A. Burley, *Angew. Chem. Int. Ed.* **2019**, *58*, 17583–17588; *Angew. Chem.* **2019**, *131*, 17747–17752.

[8] a) R. Borchardt, Y. S. Wu, *J. Med. Chem.* **1976**, *19*, 1099–1103; b) S. Khani-Oskouee, J. P. Jones, R. W. Woodard, *Biochem. Biophys. Res. Commun.* **1984**, *121*, 181–187.

[9] a) S. Singh, J. Zhang, T. D. Huber, M. Sunkara, K. Hurley, R. D. Goff, G. Wang, W. Zhang, C. Liu, J. Rohr, *Angew. Chem. Int. Ed.* **2014**, *53*, 3965–3969; *Angew. Chem.* **2014**, *126*, 4046–4050; b) F. Wang, S. Singh, J. Zhang, T. D. Huber, K. E. Helmich, M. Sunkara, K. A. Hurley, R. D. Goff, C. A. Bingman, A. J. Morris, *FEBS J.* **2014**, *281*, 4224–4239; c) M. Thomsen, S. B. Vogensen, J. Buchardt, M. D. Burkart, R. P. Clausen, *Org. Biomol. Chem.* **2013**, *11*, 7606–7610.

[10] C. Liao, F. P. Seebeck, *Nat. Catal.* **2019**, *2*, 696–701.

[11] a) Y. Nagatoshi, T. Nakamura, *Plant Biotechnol.* **2007**, *24*, 503–506; b) N. Itoh, H. Toda, M. Matsuda, T. Negishi, T. Taniguchi, N. Ohsawa, *BMC Plant Biol.* **2009**, *9*, 116.

[12] J. W. Schmidberger, A. B. James, R. Edwards, J. H. Naismith, D. O'Hagan, *Angew. Chem. Int. Ed.* **2010**, *49*, 3646–3648; *Angew. Chem.* **2010**, *122*, 3728–3730.

[13] A. S. Aslan-Üzel, A. Beier, D. Kovář, C. Cziegler, S. K. Padhi, E. D. Schuiten, M. Dörr, D. Böttcher, F. Hollmann, F. Rudroff, M. D. Mihovilovic, T. Buryška, J. Damborský, Z. Prokop, C. P. S. Badenhorst, U. T. Bornscheuer, *ChemCatChem* **2020**, *12*, 2032–2039.

[14] P. M. Bozeman, D. B. Learn, E. L. Thomas, *J. Immunol. Methods* **1990**, *126*, 125–133.

[15] Q. Tang, Y. M. Vianney, K. Weisz, C. W. Grathwol, A. Link, U. T. Bornscheuer, I. V. Pavlidis, *ChemCatChem* **2020**, *12*, 3721–3727.

[16] P. T. Männistö, S. Kaakkola, *Pharmacol. Rev.* **1999**, *51*, 593–628.

[17] C. Liao, F. P. Seebeck, *Angew. Chem. Int. Ed.* **2020**, *59*, 7184–7187; *Angew. Chem.* **2020**, *132*, 7251–7254.

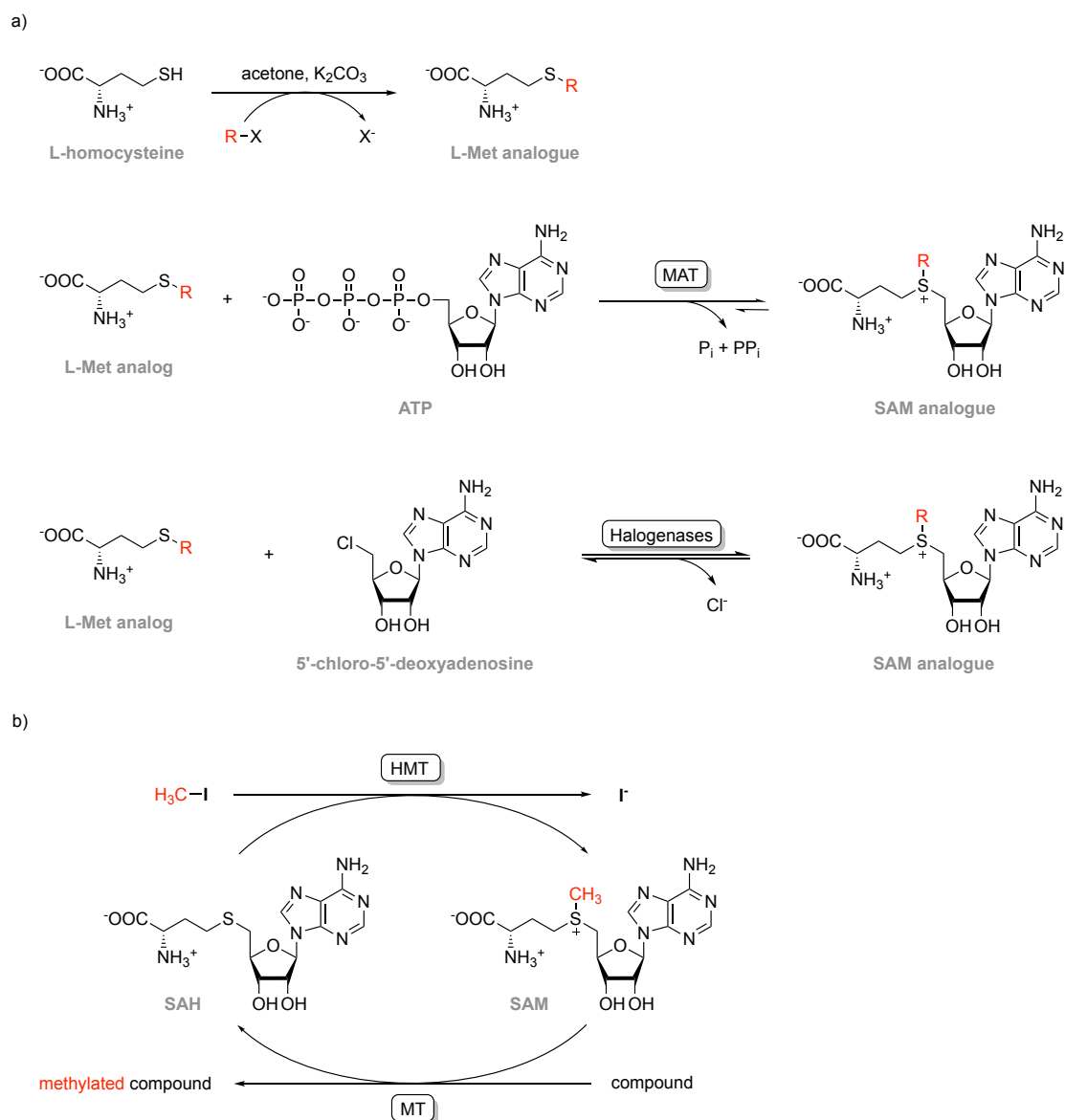[18] L. You, F. Arnold, *Protein Eng. Des. Sel.* **1996**, *9*, 77–83.

Supporting Information

Directed Evolution of a Halide Methyltransferase Enables Biocatalytic Synthesis of Diverse SAM Analogs

M. Eng. Qingyun Tang, Christoph W. Grathwol, Aşkın S. Aslan-Üzel, Shuke Wu, Andreas Link, Ioannis V. Pavlidis,* Christoffel P. S. Badenhorst,* and Uwe T. Bornscheuer*

anie_202013871_sm_miscellaneous_information.pdf

Table of Contents

Figure S1. Enzymatic synthesis of SAM analogues. a) L-methionine analogues can be chemically synthesized from L-homocysteine and alkyl halides. Methionine adenosyltransferases (MATs) or halogenases can then use ATP or 5'-chloro-5'-deoxyadenosine, respectively, to convert the L-methionine analogues to the corresponding SAM analogues. **b)** Halide methyltransferases (HMTs) can convert *S*-adenosyl homocysteine (SAH) to *S*-adenosyl methionine (SAM) in a single step, using methyl iodide as alkyl donor. This enables a simple regeneration of SAM from SAH, facilitating methyltransferase (MT)-catalyzed methylation of acceptor compounds using only catalytic amounts of SAH.[1] The purpose of our study was to expand the substrate specificity of an HMT to allow SAM analogues to be synthesized.

Additional results

Initial characterization of different HMTs

Table S1. Specific activities (μmol/min/mg) of HMTs towards different alkyl iodides for synthesizing the corresponding SAM analogues.[a]

Enzyme	⌐I	⌐⌐I	⌐⌐⌐I	⌐⌐⌐⌐I
Arabidopsis thaliana HMT	5942 ± 204	2.69 ± 0.15	<0.01	<0.01
Chloracidobacterium thermophilum HMT	132 ± 9	0.11 ± 0.07	<0.01	<0.01
Raphanus sativus HMT	8457 ± 361	1.18 ± 0.11	<0.01	<0.01

[a] Reaction times were 10 min for methyl iodide and ethyl iodide, and 4h for propyl iodide and butyl iodide. Negative controls performed without enzyme were subtracted. Activities were determined using the iodide assay described in Figure S2 and Method S3. Data reported are the means and standard deviations calculated from three independent measurements.
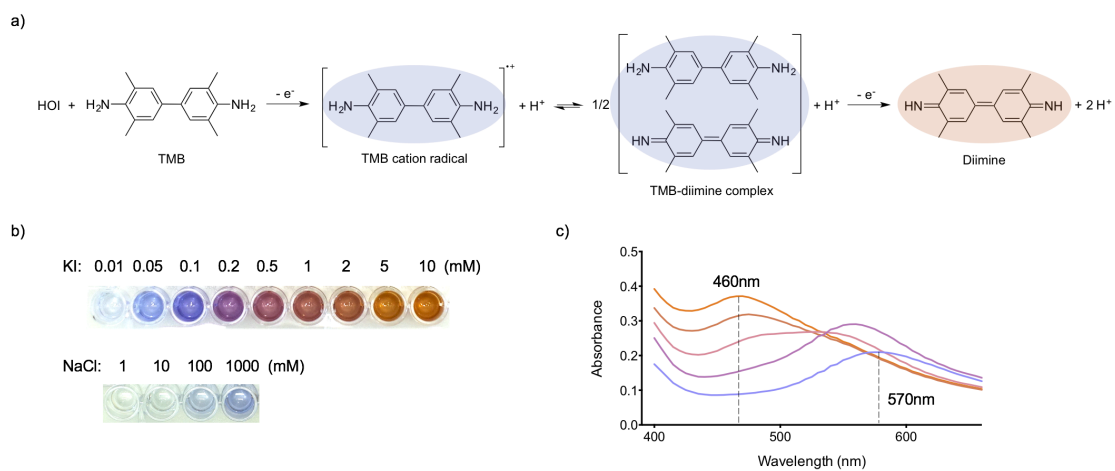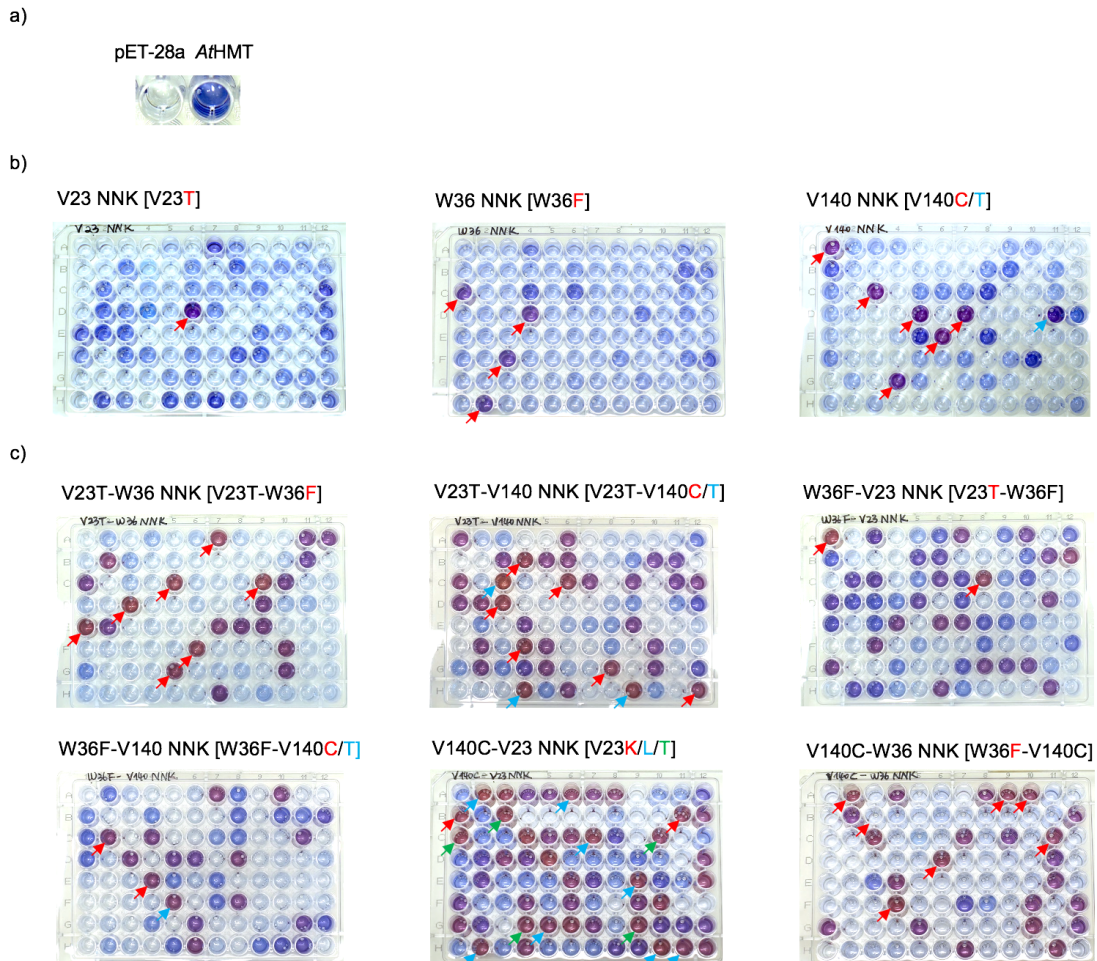
Figure S2. The iodide assay used for screening HMT libraries. a) Oxidation of TMB by HOI. **b)** Color of reaction mixtures of of KI and NaCl (1 µl; various concentrations) with the iodide assay solution (50 µl) after 30min. **c)** Absorbance spectra of the TMB oxidation products.

Comparison of the iodide assay to HPLC

a)

$$Y = 0.0109 \cdot X + 0.000157$$
$$R^2 = 0.9932$$

b)

$$Y = 7.218e{+}006 \cdot X + 135965$$
$$R^2 = 0.9966$$

c)
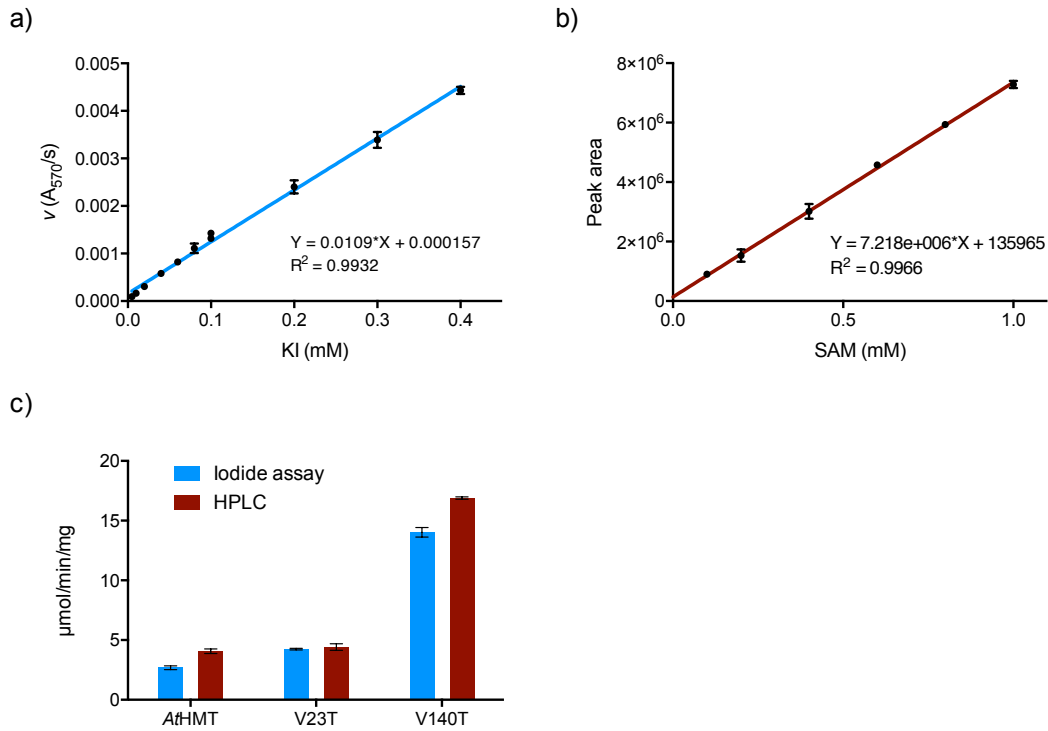
Figure S3. Comparison of the iodide assay to HPLC. a) Calibration curve for the iodide assay. The concentration of potassium iodide (KI) is plotted against the initial velocity of the increase in absorbance at 570 nm. **b)** Calibration curve for HPLC analysis of SAM. The concentration of SAM is plotted against the peak area determined by HPLC. **c)** Comparison of the ethyltransferase specific activities of wild-type *At*HMT and the V23T and V140T variants determined using the iodide assay and HPLC. Data plotted are the means and standard deviations calculated from three independent measurements.

Screening for improved ethyltransferase activity

a)

pET-28a *At*HMT

b)

V23 NNK [V23T]   W36 NNK [W36F]   V140 NNK [V140C/T]

c)

V23T-W36 NNK [V23T-W36F]   V23T-V140 NNK [V23T-V140C/T]   W36F-V23 NNK [V23T-W36F]

W36F-V140 NNK [W36F-V140C/T]   V140C-V23 NNK [V23K/L/T]   V140C-W36 NNK [W36F-V140C]

Figure S4. Examples of the iodide screening assay. Libraries were screened as described in Methods S2 and S3.

a) The empty pET-28a vector and the wild-type *At*HMT as negative and positive controls, respectively. **b)** Screening of

NNK libraries at positions V23, W36, and V140. Hits are indicated with colored arrows and named in square brackets.

c) Screening of libraries at secondary positions. Hits are indicated with colored arrows and named in square brackets.

7

Table S2. Libraries screened and hits obtained.

Library	Hits
P20 NNK	-
V23 NNK	V23T
L27 NNK	-
W36 NNK	W36F
W47 NNK	-
Y139 NNK	-
V140 NNK	V140C/T
C143 NNK	-
Y172 NNK	-
R214 NNK	-
V23T-W36 NNK	V23T-W36F
V23T-V140 NNK	V23T-V140C/T
W36F-V23 NNK	V23T-W36F
W36F-V140 NNK	W36F-V140C/T
V140C-V23 NNK	V23K/L/T-V140C
V140C-W36 NNK	W36F-V140C

Characterization of the hits

Table S3. Specific activities (µmol/min/mg) of the wild-type *At*HMT and mutants against ethyl iodide for synthesizing SAE.

Enzyme variant	Specific activity[a]
*At*HMT	2.69 ± 0.15
V23T	4.24 ± 0.07
W36F	2.80 ± 0.35
V140C	9.74 ± 0.85
V140T	14.02 ± 0.40
V23T-W36F	3.81 ± 0.23
V23T-V140C	10.28 ± 0.94
V23T-V140T	8.77 ± 1.39
V23K-V140C	6.23 ± 0.34
V23L-V140C	9.00 ± 0.62
W36F-V140C	8.21 ± 0.73
W36F-V140T	1.95 ± 0.15
V23T-W36F-V140C	6.74 ± 0.11
V23K-W36F-V140C	0.42 ± 0.03
V23L-W36F-V140C	1.59 ± 0.27
V23T-W36F-V140T	1.32 ± 0.19

[a] Data reported are the means and standard deviations calculated from three independent measurements. The data are visualized in Figure 2 of the main paper.

Michaelis-Menten kinetics for methyltransferase activity

a) *At*HMT

$k_{cat} = 364.4 \pm 12.7$ min^{-1}
$K_m = 1.43 \pm 0.19$ mM

b) V140T

$k_{cat} = 182.4 \pm 11.7$ min^{-1}
$K_m = 4.70 \pm 0.76$ mM
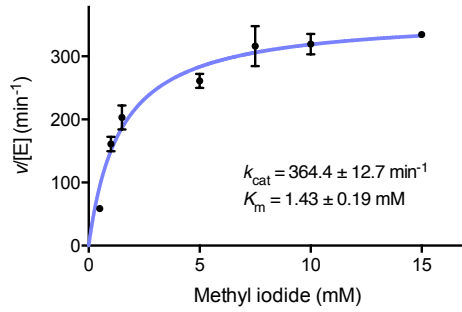
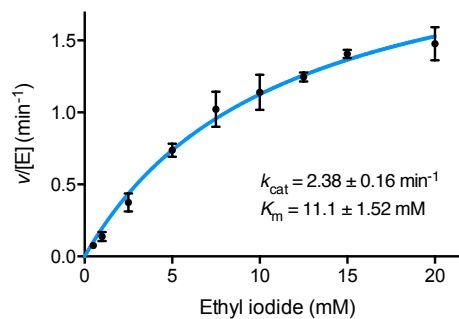**Figure S5. Michaelis-Menten kinetics for methyltransferase activity. a)** Wild-type *At*HMT and **b)** the V140T variant. Activities at various methyl iodide concentrations were determined as described in Method S4.2. Data plotted are the means and standard deviations calculated from three independent measurements.
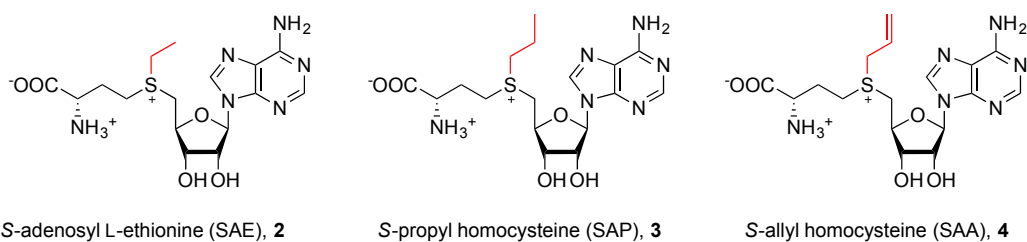
Michaelis-Menten kinetics for ethyltransferase activity
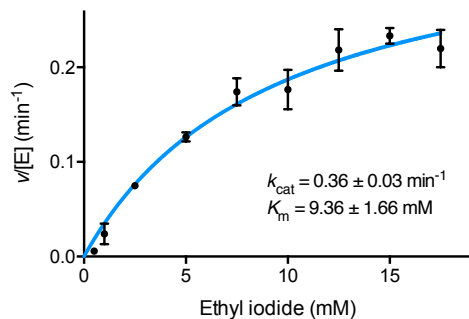
a) *At*HMT

b) V140T

Figure S6. Michaelis-Menten kinetics for ethyltransferase activity. a) Wild-type *At*HMT and **b)** the V140T variant. Activities at various ethyl iodide concentrations were determined as described in Method S4.2. Data plotted are the means and standard deviations calculated from three independent measurements.
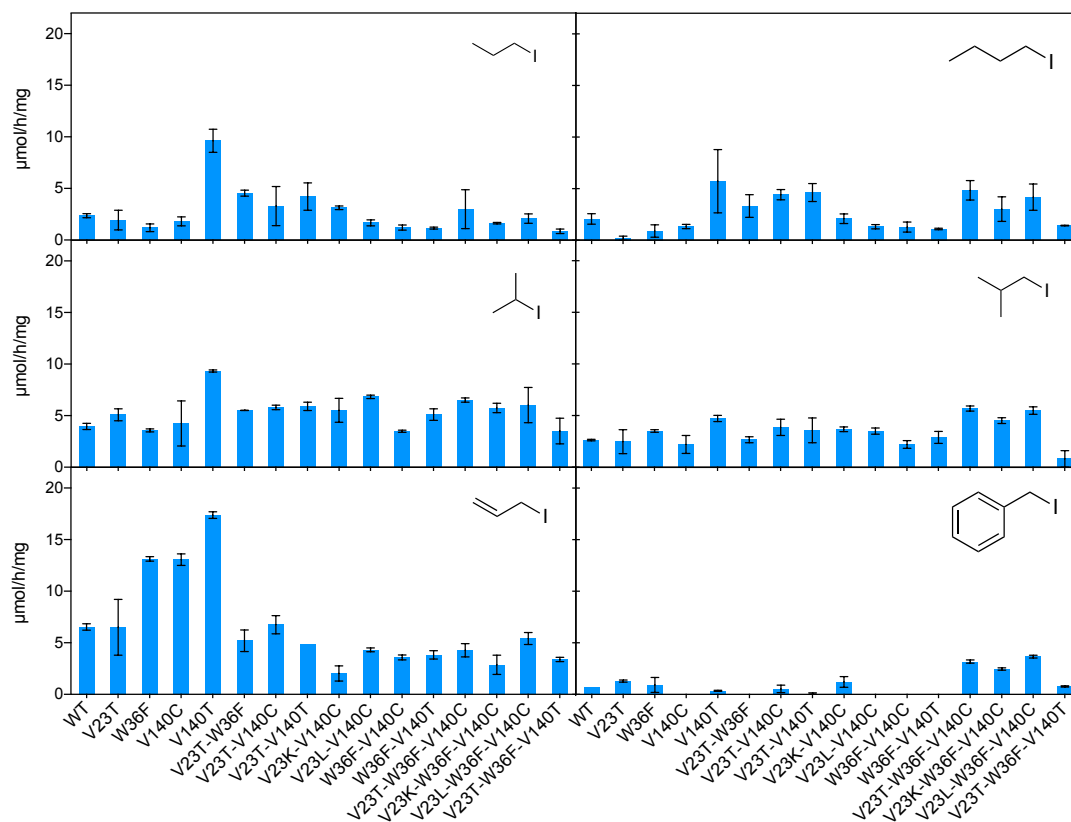
Figure S7. Specific activities (µmol/h/mg) of *At*HMT and several variants towards various alkyl iodides for synthesizing the corresponding SAM analogues. Specific activities were determined as described in Methods S3.3 and S4.2. Data plotted are the means and standard deviations calculated from two independent measurements.

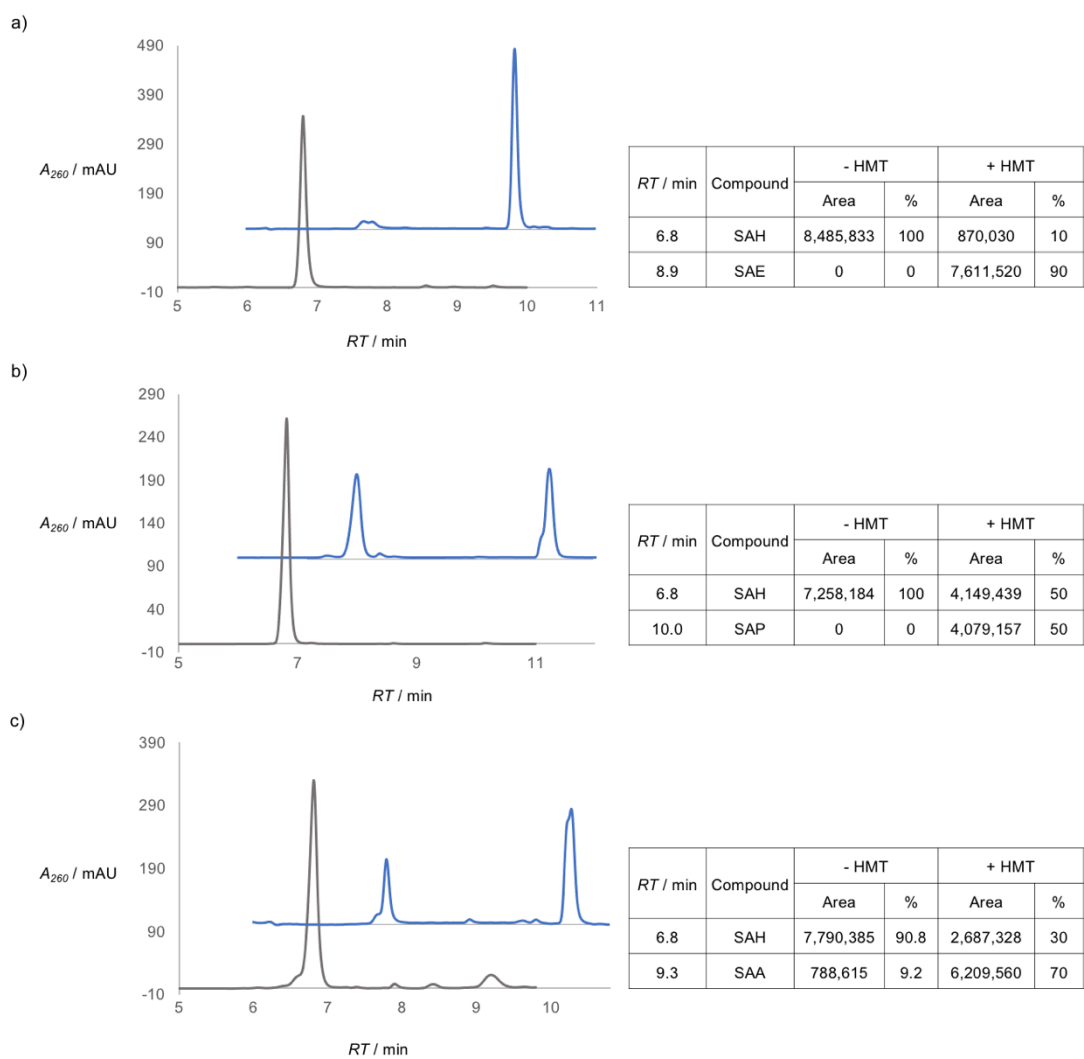# Preparative scale synthesis of SAM analogues and product analysis

Table S4. Scaled-up production of SAM analogues.

SAH	Alkyl iodide	V140T *At*HMT	Product	Reaction time	Conversion[a]
15 mg (10 mM)	80 mM EtI	5.0 mg/ml (182 µM)	SAE (**2**)	14h	90%
15 mg (10 mM)	80 mM PrI	7.7 mg/ml (273 µM)	SAP (**3**)	24h	50%
15 mg (10 mM)	80 mM Allyl-I	5.0 mg/ml (182 µM)	SAA (**4**)	14h	70%

[a]Determined by HPLC (Figure S9).

S-adenosyl L-ethionine (SAE), **2**   *S*-propyl homocysteine (SAP), **3**   *S*-allyl homocysteine (SAA), **4**

Figure S8. Structures of the SAM analogues. The iodoalkane-derived alkyl groups are shown in red.

a)

RT / min	Compound	- HMT		+ HMT	
		Area	%	Area	%
6.8	SAH	8,485,833	100	870,030	10
8.9	SAE	0	0	7,611,520	90

b)

RT / min	Compound	- HMT		+ HMT	
		Area	%	Area	%
6.8	SAH	7,258,184	100	4,149,439	50
10.0	SAP	0	0	4,079,157	50

c)

RT / min	Compound	- HMT		+ HMT	
		Area	%	Area	%
6.8	SAH	7,790,385	90.8	2,687,328	30
9.3	SAA	788,615	9.2	6,209,560	70

Figure S9. HPLC chromatograms of **a)** SAE (**2**), **b)** SAP (**3**), and **c)** SAA (**4**) syntheses on preparative scale. Structures of the SAM analogues are shown in Figure S8.

Preparative scale bioalkylation cascade reactions and product analysis

Table S5. Scaled-up production of alkylated products.

Substrate	Alkyl iodide	SAH	MT	HMT[a]	Product	Reaction time	Conversion [b]	Isolated yield	Number of cycles
5, 20 mg (10 mM)	EtI 80 mM	100 µM	IeOMT, 12 mg/ml (292 µM)	10 mg/ml (364 µM)	**6**	72h	41%	18%	41
7, 10 mg (10 mM)	PrI 80 mM	100 µM	COMT, 12 mg/ml (292 µM)	10 mg/ml (364 µM)	[c]	48h	5%	-	-
7, 10 mg (10 mM)	Allyl-I 80 mM	100 µM	COMT, 12 mg/ml (292 µM)	8 mg/ml (291 µM)	**8**	24h	48%	16%	48

[a] *At*HMT-V140T was used for all reactions. [b] Conversions were determined using HPLC (Figure S11). [c] Because of the low conversion, this product was not purified and identified by NMR spectroscopy.

luteolin, **5**   3'-O-ethylluteolin, **6**   3,4-dihydroxybenzaldehyde, **7**   4-allyloxy-3-hydroxybenzaldehyde, **8**
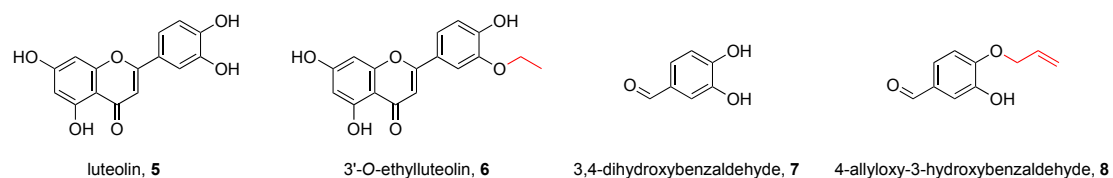
Figure S10. Structures of substrates and alkylated products. The iodoalkane-derived alkyl groups are shown in red.

a)

RT / min	Compound	- HMT/MT		+ HMT/MT	
		Area	%	Area	%
4.7	**5**	3,234,855	100	1,518,019	55
9.2	-	0	0	138,223	5
11.7	**6**	0	0	1,171,139	41

b)

RT / min	Compound	- HMT/MT		+ HMT/MT	
		Area	%	Area	%
2.0	**7**	8,606,287	100	11,284,861	95
5.6	-	0	0	527,122	5

c)

RT / min	Compound	- HMT/MT		+ HMT/MT	
		Area	%	Area	%
2.0	**7**	7,991,024	90.2	4,205,968	50
4.9	**8**	8,722,81	9.8	3,983,425	48
7.3	-	0	0	18,4431	2
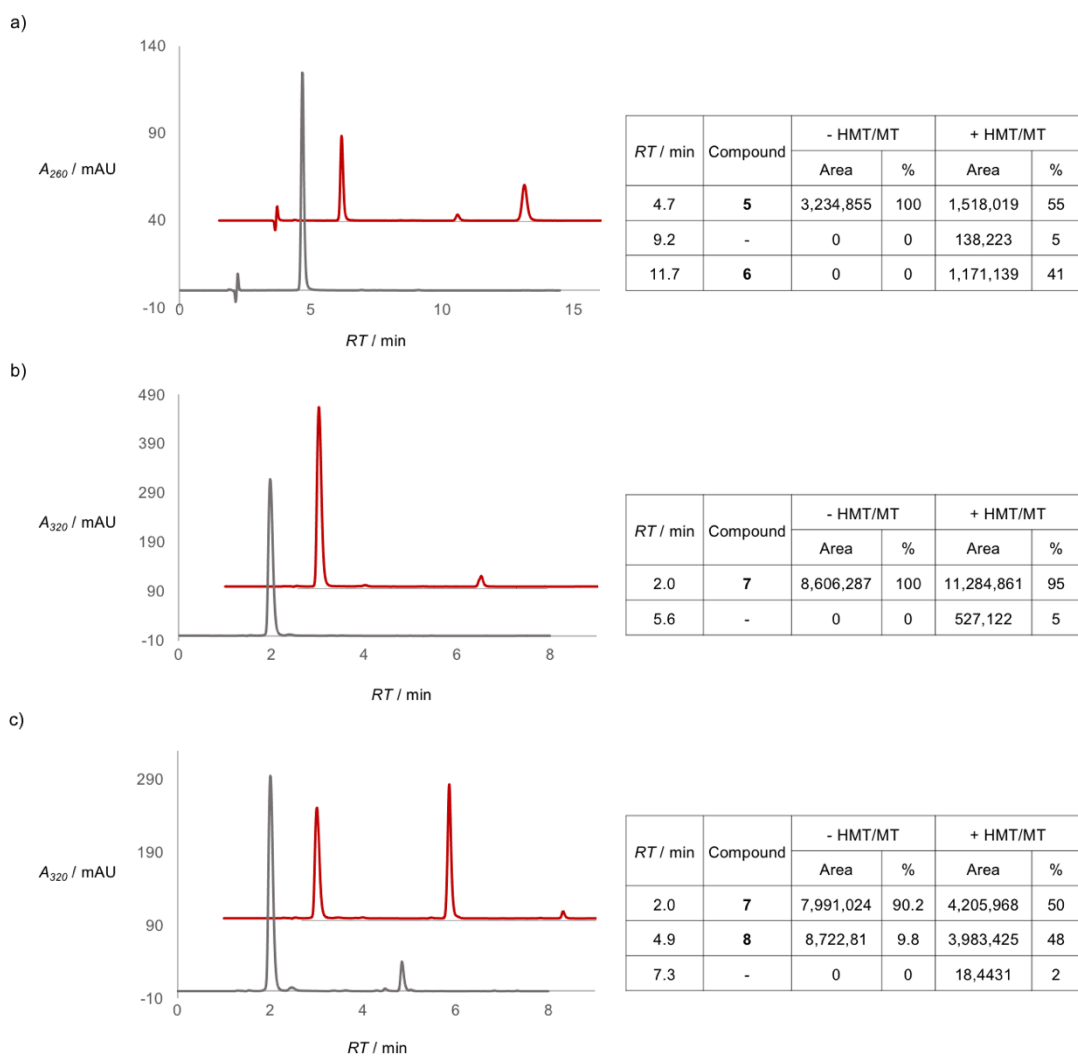
Figure S11. HPLC chromatograms of **a)** luteolin (**5**) and the product 3′-*O*-ethylluteolin (**6**), **b)** 3,4-dihydroxylbenzaldehyde (**7**) and the propylated product, and **c)** 3,4-dihydroxylbenzaldehyde (**7**) and the product 4-allyloxy-3-hydroxy-benzaldehyde (**8**).

3′-O-ethylluteolin (6). ^{1}H NMR, H,H-COSY (400 MHz, DMSO-d_6): δ (ppm) = 12.96 (s, 1H, OH), 7.55 (d, J = 1.8 Hz, 1H, 6′-H), 7.53 (s, 1H, 2′-H), 6.94 (d, J = 8.9 Hz, 1H, 5′-H), 6.84 (s, 1H, 3-H), 6.47 (s, 1H, 8-H), 6.17 (s, 1H, 6-H), 4.15 (q, J = 7.0 Hz, 1H, 7′-H), 1.37 (t, J = 7.0 Hz, 1H, 8′-H); ^{13}C NMR, DEPT-135, HSQC, HMBC (101 MHz, DMSO-d_6): δ (ppm) = 181.7 (4), 165.1 (7), 163.6 (2), 161.4 (5), 157.4 (9), 151.2 (4′), 147.2 (3′), 121.4 (1′), 120.3 (6′), 115.9 (5′), 111.4 (2′), 103.4 (10), 103.1 (3), 99.0 (6), 94.2 (8), 64.2 (7′), 14.7 (8′).
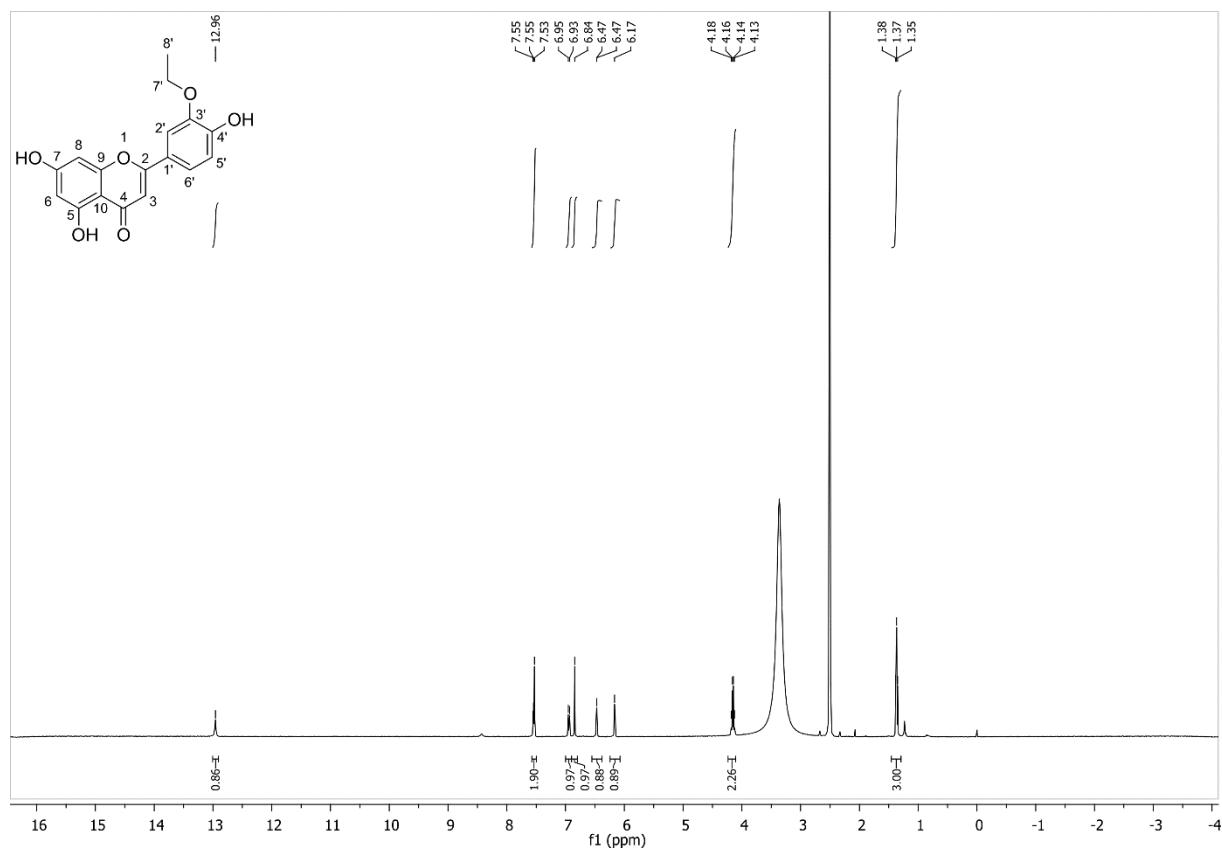
Figure S12. ^{1}H NMR spectrum of 3′-O-ethylluteolin (**6**).
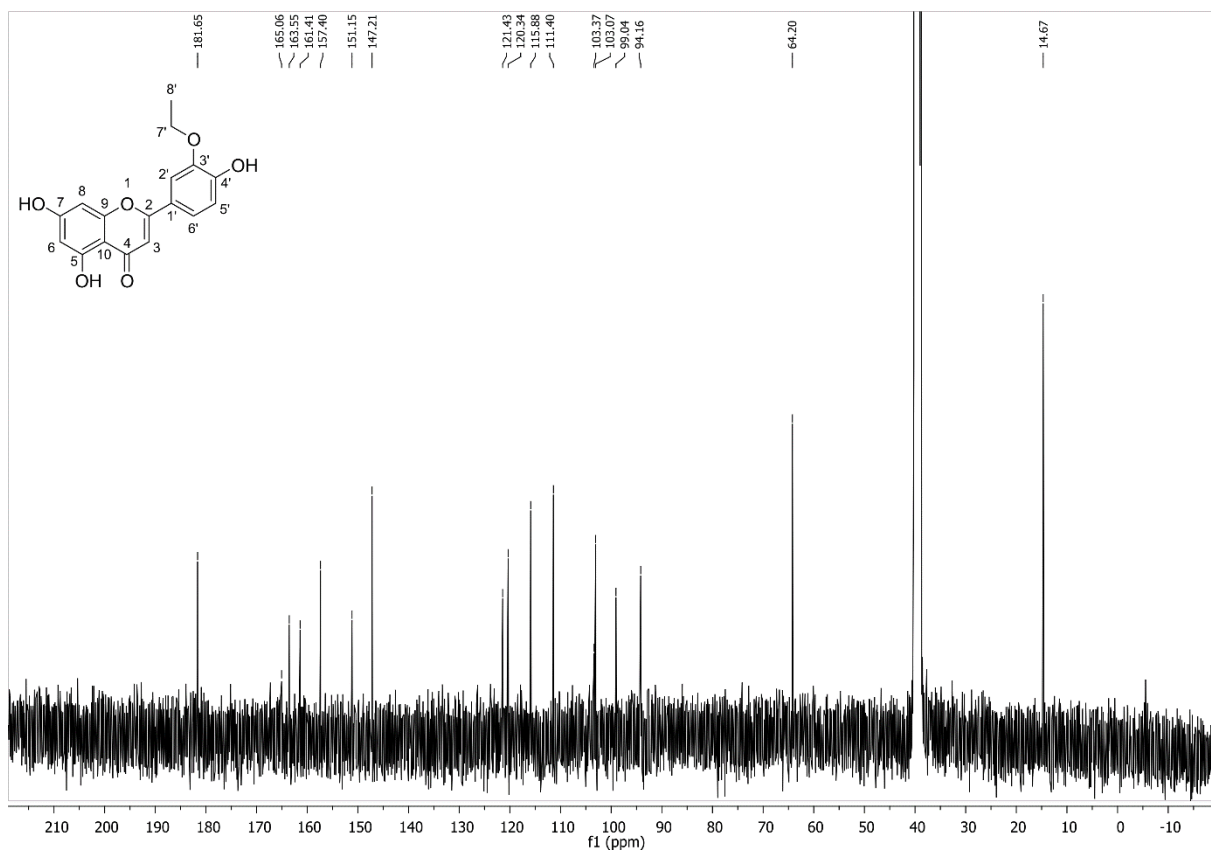
Figure S13. ^{13}C NMR spectrum of 3′-*O*-ethylluteolin (**6**).
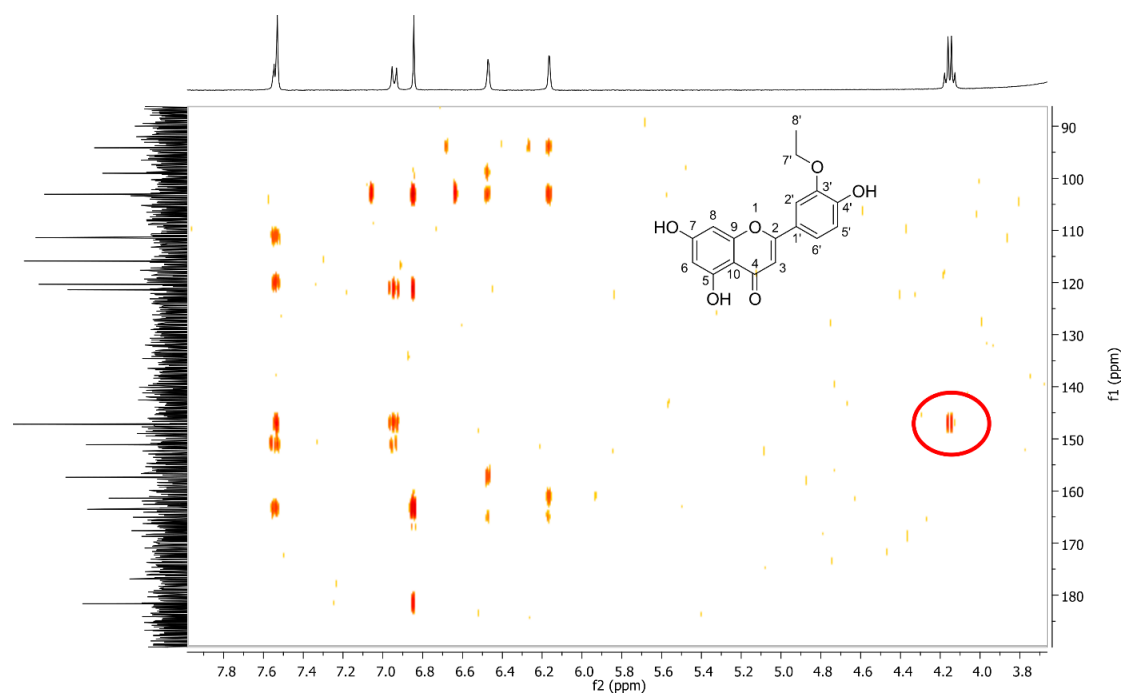
Figure S14. HMBC spectrum of 3′-*O*-ethylluteolin (**6**). The signal that defines the regioisomer is highlighted.

4-allyloxy-3-hydroxybenzaldehyde (8). ^{1}H NMR, H,H-COSY (400 MHz, DMSO-d_6): δ (ppm) = 9.77 (s, 1H, 7-H), 7.37 (dd, J = 8.3, 1.8 Hz, 1H, 6-H), 7.28 (d, J = 1.8 Hz, 1H, 2-H), 7.11 (d, J = 8.3 Hz, 1H, 5-H), 6.12–6.00 (m, 1H, 9-H), 5.44 (dd, J = 17.3, 1.6 Hz, 1H, 10-H$_a$), 5.29 (dd, J = 10.5, 1.5 Hz, 1H, 10-H$_b$), 4.68 (d, J = 5.2 Hz, 2H, 8-H); ^{13}C NMR, DEPT-135, HSQC, HMBC (101 MHz, DMSO-d_6): δ (ppm) =191.5 (7), 152.2 (4), 147.3 (3), 133.3 (9), 130.0 (1), 124.1 (6), 117.9 (10), 113.9 (2), 113.0 (5), 68.9 (8).
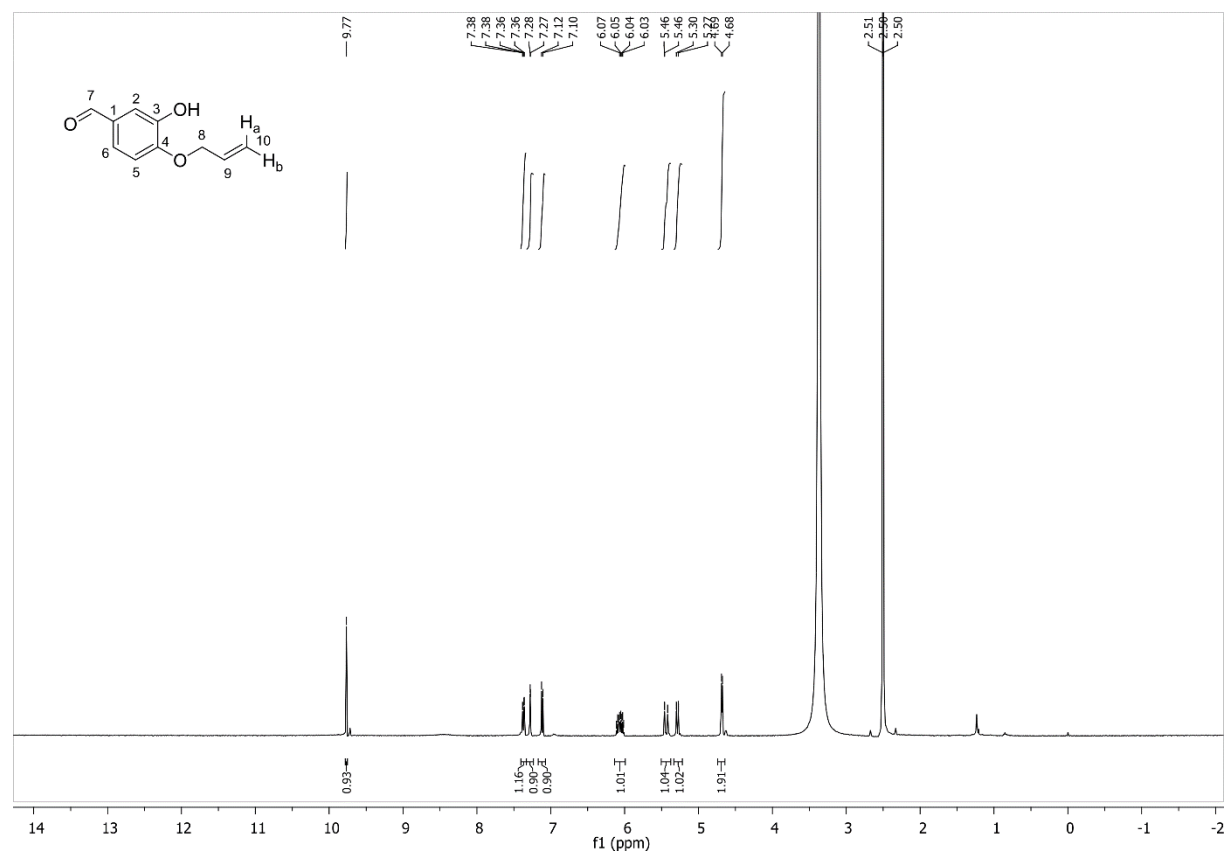
Figure S15. ^{1}H NMR spectrum of 4-allyloxy-3-hydroxybenzaldehyde (**8**).
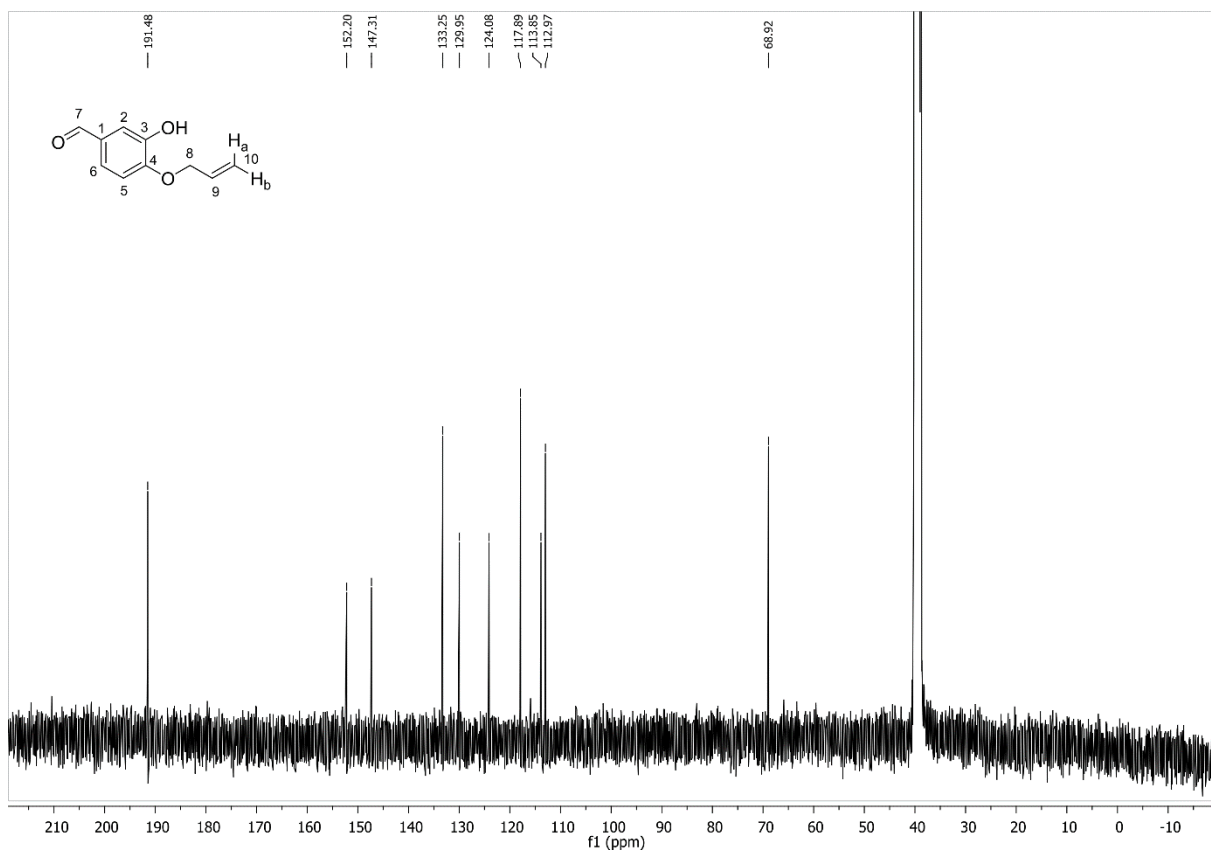
**Figure S16.** ^{13}C NMR spectrum of 4-allyloxy-3-hydroxybenzaldehyde (**8**).
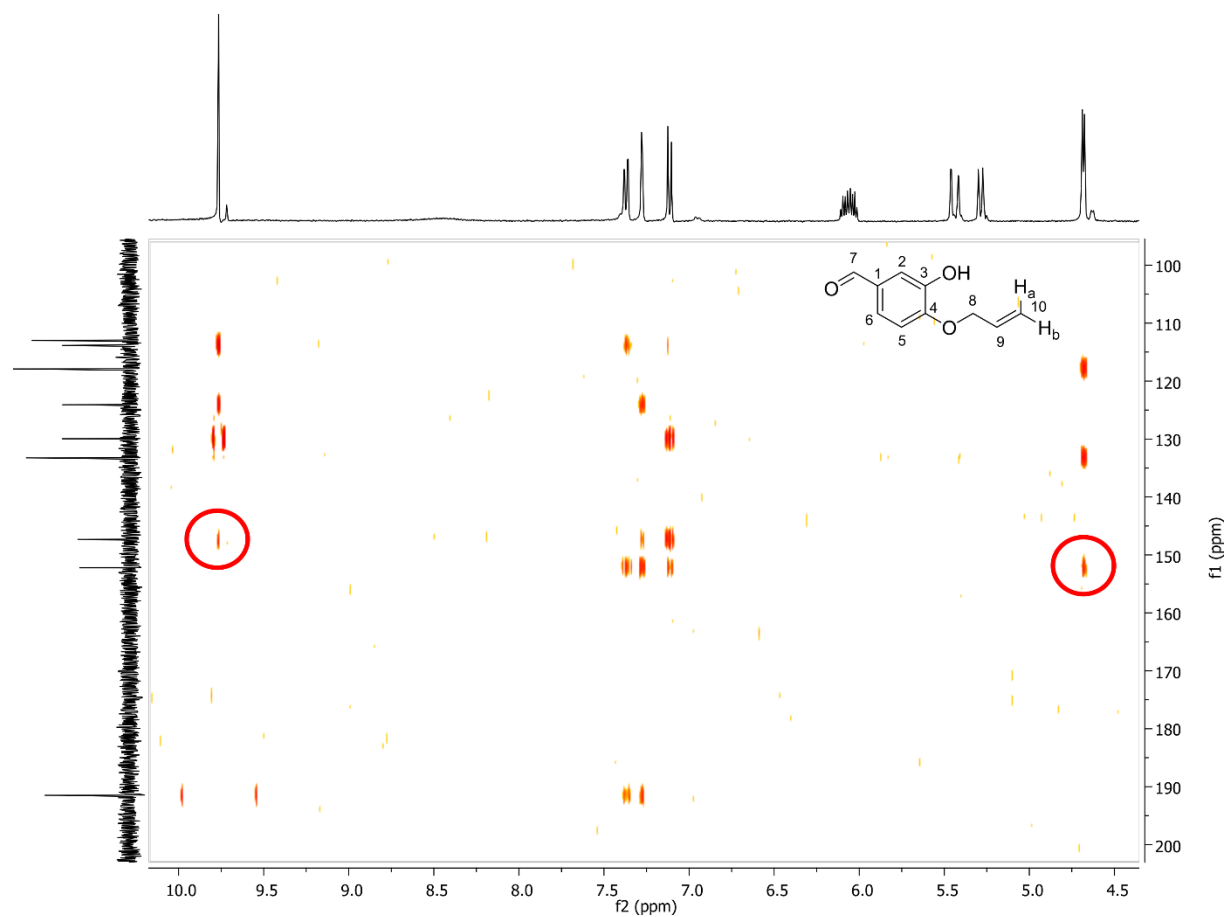
Figure S17. HMBC spectrum of 4-allyloxy-3-hydroxybenzaldehyde (**8**). Signals that define the regioisomer are highlighted.

S-propyl homocysteine (3). ^{1}H NMR, H,H-COSY (400 MHz, D$_2$O): δ (ppm) =8.36 (s, 2H, 2-H, 8-H), 6,07 (d, J = 3.7 Hz, 1H, 1′-H), 4.82–4.70 (m, 1H, 2′-H), 4.51 (pseudo-t, J = 5.8 Hz, 1H, 3′-H), 4.48–4.39 (m, 1H, 4′-H), 3.99–3.89 (m, 1H, H$_\alpha$), 3.88–3.79 (m, 2H, 5′-H), 3.62–3.38 (m, 2H, H$_\gamma$), 3.36–3.26 (m, 2H, 1″-H), 2.39–2.20 (m, 2H, H$_\beta$), 1.78–1.64 (m, 2H, 2″-H), 0.94–0.80 (m, 3H, 3″-H); ^{13}C NMR, DEPT-135, HSQC, HMBC (101 MHz, D$_2$O): δ (ppm) = 170.7 (COOH), 150.0 (6), 148.0 (4), 144.6 (2), 143.6 + 143.5 (8), 119.3 (5), 90.0 (1′), 78.6 + 78.5 (4′), 73.1 + 73.0 (2′), 72.8 (3′), 51.5 (5′), 42.7 + 42.0 (1″), 41.7 (α), 36.3 +36.2 (γ), 25.1 + 25.0 (β), 17.5 + 17.3 (2″), 11.9 (3″).
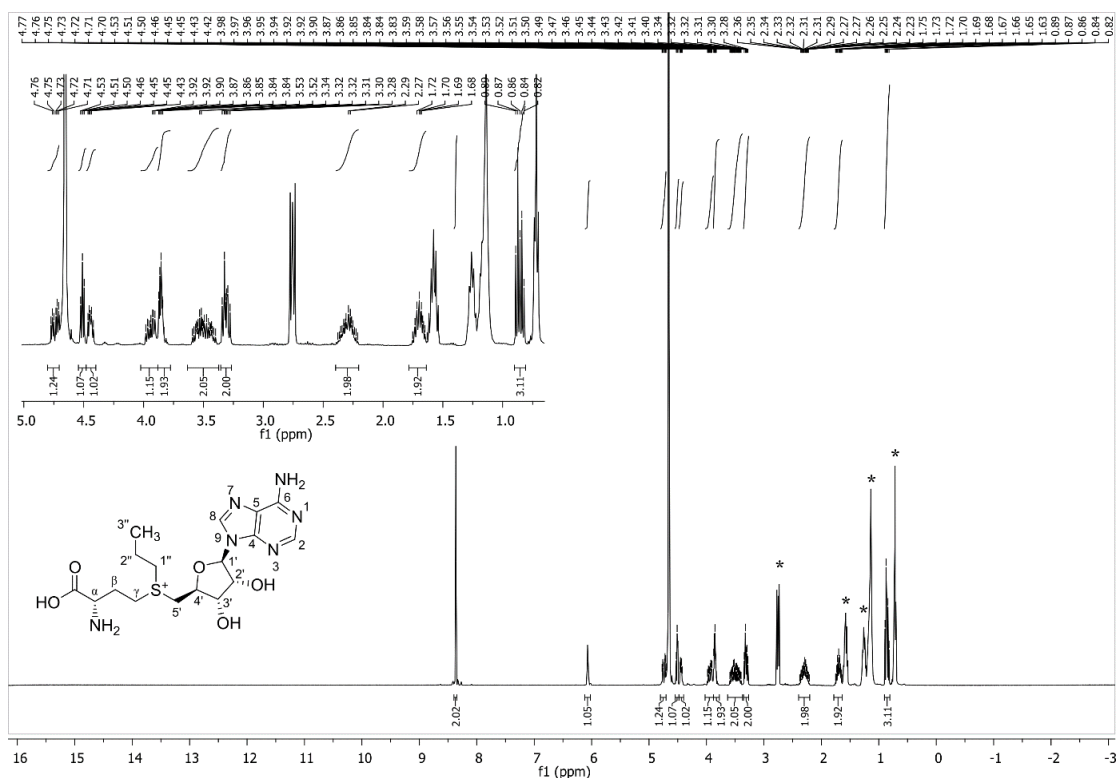
Figure S18. ^{1}H NMR spectrum of SAP (**3**) in the form of heptane sulfonate salt. Signals derived from heptane sulfonate are assigned asterisks.
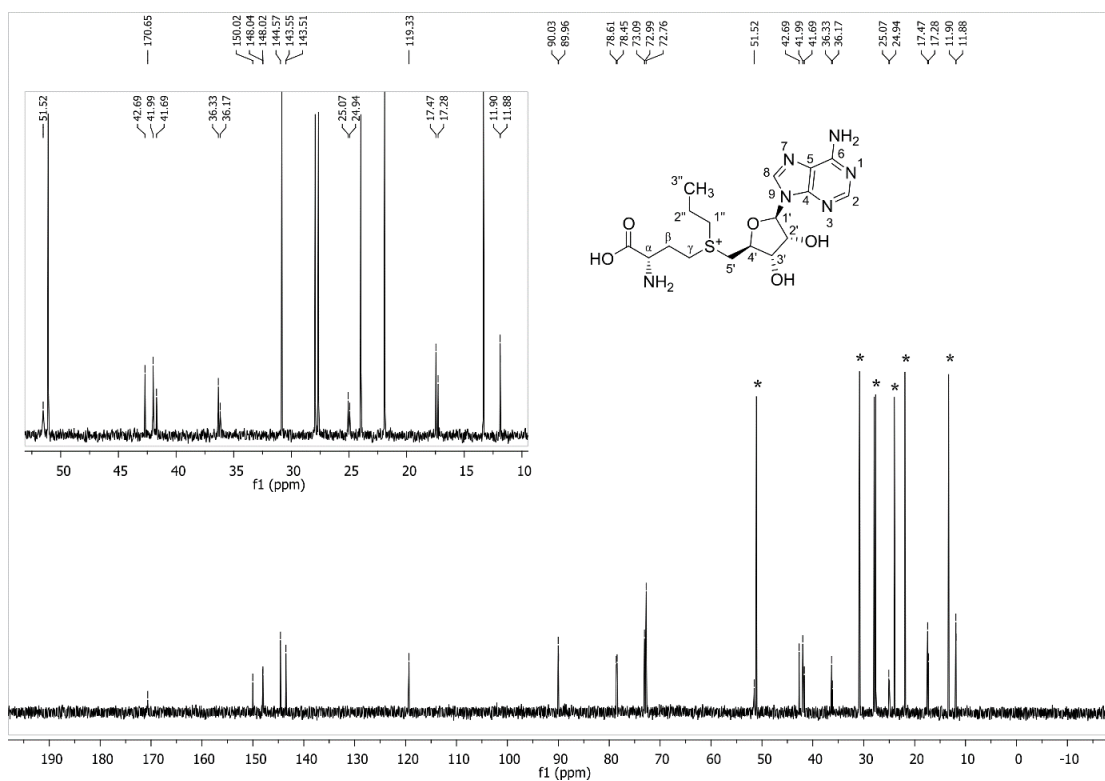
Figure S19. ^{13}C NMR spectrum of SAP (**3**) in the form of heptane sulfonate salt. Signals derived from heptane sulfonate are assigned asterisks.

Experimental procedures

Materials

Chemical reagents were purchased from Sigma Aldrich (Darmstadt, Germany), Alfa Aesar (Kandel, Germany), Fluorochem (Hadfield, UK), and Merck (Darmstadt, Germany), and were at least 95% pure. Reagents and solvents used for HPLC were purchased from Carl Roth GmbH (Karlsruhe, Germany) and VWR (Darmstadt, Germany), and were of HPLC grade. Molecular biology enzymes and kits were purchased from New England Biolabs (Ipswich, MA, US) unless otherwise stated. Oligonucleotides were purchased from Thermo Fisher Scientific (Hennigsdorf, Germany).

Method S1 Genes, protein expression, and mutagenesis

Method S1.1 Expression vectors

DNA sequences encoding the soluble form of *Homo sapiens* catechol *O*-methyltransferase (COMT) and the halide methyltransferases from *Arabidopsis thaliana* (*At*HMT), *Chloracidobacterium thermophilum* (*Ct*HMT), and *Raphanus sativus* (*Rs*HMT) were codon-optimized for expression in *Escherichia coli*. Genes were synthesized and cloned into pET-28a(+) or pET-28b(+) by BioCat GmbH (Heidelberg, Germany). The resulting constructs encoded either N- or C-terminal His$_6$-tags (amino acid sequences for all proteins are given in the section *Protein sequences*). The empty pET-28a(+) vector (Novagen®) was purchased from Merck (Darmstadt, Germany). The C-terminally His$_6$-tagged T133M-Y326L mutant of *Clarkia breweri* isoeugenol *O*-methyltransferase (IeOMT) in pET-21b(+) was described in our previous work.[2] All HMTs and MTs, as well as the empty pET-28a vector, were transformed into chemically competent *E. coli* BL21 Δmtn (DE3) cells (Method S1.3) using a standard heat-shock protocol.

Method S1.2 Site-saturation mutagenesis

Saturation mutagenesis was performed using mutagenesis PCR. Degenerate (NNK) oligonucleotides (Table S6) and the wild-type *At*HMT-encoding plasmid (DNA sequence is given below Table S6) were used. PCRs contained 1 x *Pfu* buffer, 2% DMSO, 200 nM each of forward and reverse primers, dNTPs (250 μM each), 0.2 ng/μl of template DNA (pET-28a-*At*HMT), and 0.05 U/μl of *Pfu* polymerase (Roboklon GmbH, Berlin, Germany). After initial denaturation for 1 min at 95 °C, 18 cycles of denaturation at 95 °C for 30 s, annealing at

60 °C for 1 min, and extension at 68 °C for 6.5 min were performed, followed by a final extension at 68 °C for 13 min. PCR products were treated twice by DpnI digestion to ensure complete removal of the template DNA.[3] The DpnI-digested PCR mixtures (5 µl) were used for heat shock transformation of chemically competent *E. coli* BL21 Δmtn (DE3) cells. Transformations were plated on LB Miller agar containing 50 µg/ml kanamycin and incubated overnight at 37 °C.

Table S6. Sequences of the degenerate (NNK) oligonucleotides used for site-saturation mutagenesis.

Primer	Sequence (5′-3′)
P20 NNK forward	ggcggtaatgtgattccgacc**nnk**gaagaagttgcaacctttctg
P20 NNK reverse	cagaaaggttgcaacttcttc**mnn**ggtcggaatcacattaccgcc
V23 NNK forward	atgtgattccgacccccggaagaa**nnk**gcaacctttctgcataaaac
V23 NNK reverse	gttttatgcagaaaggttgc**mnn**ttcttccggggtcggaatcacat
L27 NNK forward	cccggaagaagttgcaaccttt**nnk**cataaaaccgttgaagaaggcg
L27 NNK reverse	cgccttcttcaacggttttatg**mnn**aaaggttgcaacttcttccggg
W36 NNK forward	gcataaaaccgttgaagaaggcggt**nnk**gaaaaatgctgggaagaagaaatta
W36 NNK reverse	taatttcttcttcccagcattttc**mnn**accgccttcttcaacggttttatgc
W47 NNK forward	ggaagaagaaattaccccg**nnk**gatcagggtcgtgcaaccc
W47 NNK reverse	gggttgcacgaccctgatc**mnn**cggggtaatttcttcttcc
Y139 NNK forward	gtccgaccgaactgtttgatctgatttttgat**nnk**gtgttcttctgtgcaat
Y139 NNK reverse	attgcacagaagaacac**mnn**atcaaaaatcagatcaaacagttcggtcggac
V140 NNK forward	cgtccgaccgaactgtttgatctgatttttgattat**nnk**ttcttctgtgcaatcgaac
V140 NNK reverse	gttcgattgcacagaagaa**mnn**ataatcaaaaatcagatcaaacagttcggtcggacg
C143 NNK forward	ctgattttttgattatgtgttcttc**nnk**gcaatcgaaccggaaatgcgccc
C143 NNK reverse	gggcgcatttccggttcgattgc**mnn**gaagaacacataatcaaaaatcag
Y172 NNK forward	ggcgaactgattaccctgatg**nnk**ccgattaccgatcatgttggt
Y172 NNK reverse	accaacatgatcggtaatcgg**mnn**catcagggtaatcagttcgcc
R214 NNK forward	aatccgcatgccattccgacc**nnk**aaaggtaaagaaaaactgggc
R214 NNK reverse	gcccagtttttctttacctttt**mnn**ggtcggaatggcatgcggatt

DNA sequence of the wild-type *At*HMT template used for mutagenesis

```
ATGGGCAGTAGTCATCATCATCATCACCATAGTAGTGGTCTGGTGCCGCGCGGCAGTCATATGGCCGAAGAACAGCAGAATAGTGA
TCAGAGTAATGGCGGTAATGTGATTCCGACCCCGGAAGAAGTTGCAACCTTTCTGCATAAAACCGTTGAAGAAGGCGGTTGGGAAA
AATGCTGGGAAGAAGAAATTACCCCGTGGGATCAGGGTCGTGCAACCCCGCTGATTGTGCATCTGGTGGATACCAGCAGCCTGCCG
CTGGGTCGCGCCTTAGTGCCGGGTTGCGGTGGTGGTCATGATGTGGTTGCCATGGCCAGTCCGGAACGCTTTGTTGTTGGTCTGGA
TATTAGTGAAAGCGCCCTGGCAAAAGCCAATGAAACCTATGGTAGCAGTCCGAAAGCAGAATATTTTAGCTTTGTTAAAGAGGATG
TGTTCACCTGGCGTCCGACCGAACTGTTTGATCTGATTTTTGATTATGTGTTCTTCTGTGCAATCGAACCGGAAATGCGCCCGGCA
TGGGCAAAAAGTATGTATGAACTGCTGAAACCGGATGGCGAACTGATTACCCTGATGTATCCGATTACCGATCATGTTGGTGGCCC
GCCGTATAAAGTTGATGTTAGCACCTTTGAAGAAGTGCTGGTGCCGATTGGCTTTAAAGCCGTGAGCGTGGAAGAAAATCCGCATG
CCATTCCGACCCGTAAAGGTAAAGAAAAACTGGGCCGCTGGAAAAAGATTAATTAA
```

Method S1.3 Construction of the MTA/SAH nucleosidase knockout *E. coli* BL21(DE3) strain

As reported by Liao and Seebeck, the 5′-methylthioadenosine/*S*-adenosyl homocysteine (MTA/SAH) nucleosidase from *E. coli* BL21(DE3) can degrade the substrate SAH to adenine and the enzyme cannot be completely removed by immobilized metal-affinity chromatography and size exclusion chromatography.[1] Thus, an MTA/SAH nucleosidase knockout strain of *E. coli* BL21(DE3), named *E. coli* BL21 Δmtn (DE3), was constructed for the expression of HMTs and MTs. The *E. coli* BL21 Δmtn (DE3) strain was constructed using a previously developed two-plasmid-based CRISPR-Cas9 system.[4] The two key plasmids pCas and pTargetF (plasmid IDs #62225 and #62226) were purchased from Addgene (Watertown, MA, US). The pTarget-dmtn plasmid was first constructed by engineering the flanking sequence of the *mtn* gene by assembly of three DNA fragments, the mtn-Up fragment and the mtn-Down fragment, amplified from the *E. coli* genome using the corresponding primers (Table S7), and the pTarget fragment amplified from the pTargetF plasmid using the corresponding primers (Table S7). The three fragments were assembled using the SLiCE cloning method.[5] Next, the gRNA was introduced with the dmtn-gRNA primers using the Q5 Site-Directed Mutagenesis Kit from New England Biolabs. The resulting pTarget-dmtn plasmid was sequenced to confirm introduction of the gRNA and the flanking sequence of the *mtn* gene. The *E. coli* BL21 Δmtn (DE3) strain was constructed by using pCas and pTarget-dmtn according to a previously developed procedure,[4] briefly outlined here. 1) Transform the pCas plasmid into chemically competent *E. coli* BL21(DE3) cells following a standard heat-shock protocol. 2) Prepare electrocompetent cells of *E. coli* BL21(DE3) pCas with the arabinose-inducible λ-Red system. 3) Introduce the pTarget-dmtn plasmid by electroporation. 4) Select on LB agar plates containing 50 µg/ml kanamycin (for pCas) and 50 µg/ml streptomycin (for pTarget-dmtn). 5) Perform colony PCR for genotyping the colonies. 6) Sequentially cure the plasmid pTarget-dmtn (by adding 1 mM IPTG) and pCas (by culturing at 37°C for 12 h). The resulting *E. coli* BL21 Δmtn (DE3) strain was confirmed by PCR amplification and sequencing the genomic DNA flanking the *mtn* gene. Chemically competent *E. coli* BL21 Δmtn (DE3) cells were prepared using the Hanahan method.[6]

Table S7. Oligonucleotides used for construction of the MTA/SAH nucleosidase knockout strain *E. coli* BL21 Δmtn (DE3).

Primer	Sequence (5′-3′)
mtn-Up-forward	gagtcgacctgcagaagcttgaaaccagtcatttatcgctgc
mtn-Up-reverse	cagtgacttaagatttactcgcgataagccc
mtn-Down-forward	gagtaaatcttaagtcactgttcagggcgct
mtn-Down-reverse	gagctgcacatgaactcgaggcggattaatgccgaattgcag
pTarget-forward	ctcgagttcatgtgcagctc
pTarget-reverse	aagcttctgcaggtcgactc
dmtn-gRNA-forward	tctgttggtcgttttagagctagaaatagcaagtt
dmtn-gRNA-reverse	accgttgatgactagtattatacctaggactgagc

Method S2 High-throughput screening

Method S2.1 Expression of variants in the NNK libraries

For each NNK library, 96 colonies were picked into 96-well microtiter plates filled with 150 µl of lysogeny broth (LB) containing 50 µg/ml kanamycin and grown overnight at 37°C, shaking at 700 rpm in an MTP shaker (Edmund Bühler GmbH, Bodelshausen, Germany). Terrific Broth (1 ml in 96-deepwell plates) containing 50 µg/ml kanamycin and 100 µM of isopropyl β-D-1-thiogalactopyranoside (IPTG) were inoculated with 50 µl of the pre-cultures and incubated at 30°C (900 rpm) for 24 h. The remaining pre-cultures were mixed with glycerol to a final concentration of 40% (v/v) and stored at -80 °C. The cells grown in deep-well plates were collected by centrifugation at 4,000 g for 20 min (4 °C). Cell pellets were resuspended in 250 µl of lysis buffer (1 mg/ml lysozyme, 10 µg/ml DNase, and 0.1 × BugBuster ® Protein Extraction Reagent (Novagen®) in 50 mM sodium phosphate (pH 7.5) by shaking at 700 rpm (30°C) for 2 h. Cell debris was removed by centrifugation at 4,000 g for 20 min (4°C) and the resulting supernatant was used for HMT assays.

Method S2.2 HMT assay and general screening workflow

Libraries were screened in 96-well microtiter plates. Reactions contained 1 mM SAH (1 µl of a 100 mM stock in DMSO), 5 mM ethyl iodide (5 µl of a 100 mM stock in DMSO) and 94 µl of crude cell lysate, for a final assay volume of 100 µl, and were incubated at room temperature for 30 min. Ethyl iodide stocks in DMSO were freshly prepared for each experiment to avoid hydrolysis. After the 30 min incubation, 1 µl of each reaction mixture was added to 50 µl of iodide assay reagent (Method S3.2) in transparent 96-well polystyrene microtiter plates. The plates were shaken immediately and then incubated at room temperature for 30 min. The production of iodide by the HMT reaction was detected by observing the color of the assay solution (Figures S2 and S4). Hits showing higher activities than the wild-type AtHMT were cultivated from the respective glycerol stocks. Plasmid DNA was isolated (innuPREP Plasmid Mini Kit 2.0, Analytic Jena, Jena, Germany) and sent for Sanger sequencing (Eurofins Genomics Germany GmbH, Ebersberg, Germany).

Method S3   The iodide assay for high-throughput screening

Method S3.1   Assay principle

We developed a sensitive and convenient iodide assay based on our published HOX assay for halides.[7] The assay described here has the advantage of being very sensitive to iodide but relatively insensitive to chloride, making it ideal for screening using chloride-containing crude lysates. Iodide was first converted to hypoiodous acid (HOI) by a vanadium-dependent chloroperoxidase from *Curvularia inaequalis* (*Ci*VCPO, Method S3.4). The *Ci*VCPO uses hydrogen peroxide to oxidize iodide to HOI under mild conditions. The bleach HOI is detected by oxidation of the chromogen 3,3′,5,5′-tetramethylbenzidine (TMB).[8] TMB is first converted to the one-electron oxidized cation radical which is in equilibrium with the TMB-diimine complex (blue, λ_{max} = 570 nm) and finally to the two-electron oxidized diimine product (orange, λ_{max} = 460 nm) (Figure S2a).[8a] In the assay setup we used, iodide concentrations from 5 µM to 10 mM could be detected. For lower concentrations, the assay turns blue and for higher concentrations the mixture turns purple, copper, or orange (Figure S2b) and the maximum absorption shifts from 570 to 460 nm (Figure S2c). We used only 1 µl of sample in a 50 µl iodide assay so much less than 5 µM of iodide could be detected by adjusting the sample volume. Importantly, the iodide assay is not sensitive to millimolar concentrations of chloride (Figure S2b).

Method S3.2   Preparation of the assay reagent

An assay solution containing 1.37 U/ml CiVCPO, 25 mM TMB, 2 mM hydrogen peroxide, and 1 mM sodium orthovanadate in 20 mM phosphate (pH 6.0) was used for preliminary experiments. However, to simplify the assay setup for high-throughput screening, we used a commercially available premixed solution of TMB and H_2O_2 in a mildly acidic buffer (Liquid Substrate System for Membranes; Sigma Aldrich; Catalog Number T0565). The concentrations of the individual components in this proprietary mixture are unknown but the commercial solution was suitable for our purposes. We prepared our iodide assay reagent by adding recombinantly expressed CiVCPO, to a final concentration of 1.37 U/ml, to the Liquid Substrate System. Expression and purification of the CiVCPO is described in Method S3.4.

Method S3.3   Iodide quantification and calibration curves

For the quantification of iodide, the iodide assay was carried out in transparent 384-well polystyrene microtiter plates at room temperature. Absorbance measurements were performed using a Tecan Infinite M200PRO3 Plate Reader (Tecan, Männedorf, Switzerland). Samples or standards (1 µl) were added to iodide assay reagent (50 µl) and the initial rate of change in absorbance at 570 nm determined. The calibration curve for iodide was determined using potassium iodide (KI) solutions ranging from 5 µM to 400 µM. All reactions were performed in triplicate. The rate of change in absorbance at 570 nm (v (A_{570nm}/s)) showed a linear relationship to iodide concentration from 5 µM to 400 µM (Figure S3a).

Method S3.4   Expression, purification, and activity determination of CiVCPO

The CiVCPO was expressed from the pBAD-VCPO plasmid in $E.$ $coli$ TOP10 as previously described.[7] In short, a culture of 10 ml of LB medium containing 100 µg/ml ampicillin was inoculated with a single colony from an LB agar plate containing 100 µg/ml ampicillin and incubated overnight at 37°C. The overnight culture was used to inoculate a 1 L culture of LB medium containing 100 µg/ml ampicillin. The culture was incubated at 37°C (180 rpm) until OD_{600} reached ~0.5. Expression was then induced by the addition of 0.02% L-arabinose, followed by incubation at 25°C for 24 h. Cells were harvested by centrifugation at 4500 g for 30 min (4°C) and resuspended in ice-cold lysis buffer (1 ml per gram of cells). The lysis buffer was composed of 50 mM Tris-H_2SO_4 (pH 8.1) containing 2 mg/ml lysozyme, 1 mg/ml DNaseI, and cOmplete™ protease inhibitor cocktail (Roche, Mannheim, Germany). The resuspended cells were lysed by sonication on ice (three cycles of 5 min; 50% pulse and 50% power) using a SONOPULS HD 2070 (BANDELIN electronic GmbH & Co. KG, Berlin, Germany). The crude

lysate was clarified at 4 °C by centrifugation at 10,000 g for 1 h. The clarified lysate was transferred to a new tube, thoroughly mixed with an equal volume of isopropanol, and incubated on ice for 20 min. Precipitated proteins were removed by centrifugation at 10,000 g for 30 min (4°C). The CiVCPO was then purified by chromatography on a 5 ml DEAE Sephacel column equilibrated with 50 mM Tris-H_2SO_4 (pH 8.1). The column was washed with 25 ml of 50 mM Tris-H_2SO_4 (pH 8.1) and then with 25 ml of the same buffer containing 100 mM sodium chloride. Protein was eluted from the column using 50 mM Tris-H_2SO_4 (pH 8.1) containing 1 M sodium chloride. The eluate (30 ml) was dialyzed three times (twice for 4 h and then overnight) against 4.5 L of 50 mM sodium phosphate (pH 8.0) supplemented with 100 µM sodium orthovanadate. The orthovanadate converts the apoenzyme to the active vanadium-bound holoenzyme. Protein concentration was determined by measuring absorbance at 280 nm using a NanoDrop™ (Thermo Fisher, Hennigsdorf, Germany). The specific activity of CiVCPO was determined using the monochlorodimedone assay. Bromination of monochlorodimedone results in a decrease in absorbance at 290 nm ($\Delta\varepsilon$ = 20,000 $M^{-1}cm^{-1}$). Reactions (600 µl) containing 42 µM monochlorodimedone, 100 µM bromide, 8.8 mM H_2O_2, and 1 mM orthovanadate in 20 mM phosphate (pH 6.0) were initiated by addition of 10 µl (4.5 µg) of purified CiVCPO. The decrease in absorbance at 290 nm was monitored using a 1 cm-pathlength quartz cuvette and a JASCO V-550 spectrophotometer (JASCO, Easton, MD, US). Specific activity was expressed as µmol monochlorodimedone brominated/min/(mg CiVCPO).

Method S4   Characterization of purified HMTs

Method S4.1   Expression and purification of HMTs and MTs

All HMTs and MTs were transformed into chemically competent *E. coli* BL21 Δmtn (DE3) cells for protein expression. Pre-cultures (5 ml LB medium containing 50 µg/ml kanamycin for pET-28 plasmids or 100 µg/ml ampicillin for pET-21 plasmids) were inoculated with single colonies and incubated overnight at 37°C. For the *At*HMT mutants, pre-cultures were inoculated from glycerol stocks of the NNK libraries (2% (v/v)). The cells were grown in Terrific Broth (TB) containing antibiotics (50 µg/ml kanamycin for pET-28 plasmids or 100 µg/ml ampicillin for pET-21 plasmids) at 37°C for about 2 h until OD_{600} reached ~0.7. Cultures were then cooled to 20°C before IPTG was added to a final concentration of 200 µM. Cultures were then incubated at 20°C for another 20 h. Cells were harvested by centrifugation at 10,000 g at 4°C for 10 min. Cell pellets were resuspended in lysis buffer (20 mM sodium phosphate, 500 mM sodium chloride, and 20 mM imidazole, pH 7.5) and lysed by sonication on ice (as described in Method S3.4). Lysates were clarified by centrifugation at 10,000 g and 4°C for 1 h. The His_6-tagged recombinant proteins were purified by immobilized metal-affinity chromatography as follows. The clarified lysates were loaded onto 2 ml of Roti®garose-His/Co Beads (Carl Roth, Karlsruhe,

Germany) equilibrated with lysis buffer. After incubation on ice for 15 min, the flow-through was discarded and weakly bound proteins were removed by washing the resin with lysis buffer. The target proteins were then eluted with elution buffer (20 mM sodium phosphate, 500 mM sodium chloride, and 200 mM imidazole, pH 7.5) and desalted using PD-10 desalting columns (GE healthcare, Amersham, UK) equilibrated with storage buffer. Storage buffer was chloride-free 50 mM sodium phosphate (pH 7.5) prepared by dissolving 0.041 mol Na_2HPO_4 and 0.0094 mol NaH_2PO_4 in 1 L of Milli-Q water. Further desalting was achieved by dilution into 40 volumes of storage buffer and concentration using 10 kDa MWCO Vivaspin™ protein concentrators (Sartorius AG, Germany, Göttingen). Protein concentrations were determined by measuring absorbance at 280 nm using a NanoDrop™ (Thermo Fisher, Hennigsdorf, Germany). Extinction coefficients (280 nm) for all proteins are given in the section *Protein sequences*. Protein purities were analyzed by SDS-PAGE.[9]

Method S4.2   HMT specific activity and kinetic measurements

HMTs and mutants were assayed using various haloalkanes. Reaction mixtures (50 μl) contained 1 mM SAH, 10 mM haloalkane, and the purified enzyme in 50 mM chloride-free sodium phosphate (pH 7.5). Reactions were incubated in an Eppendorf Thermomixer at 25°C and 1000 rpm. All haloalkane stocks in DMSO were freshly prepared before the reaction to avoid hydrolysis. Purified enzymes were added to a final concentration of 0.15 mg/ml (5.46 μM wild-type or mutant *At*HMT) for methyl iodide assays. For other iodoalkanes, 2 mg/ml protein (72.85 μM wild-type or mutant *At*HMT) was used. The reaction time was 10 min for methyl and ethyl iodide and 4 h (Table S1) or 16 h (Figure S7) for all other iodoalkanes. After incubation, 1 μl of reaction mixture was added to 50 μl iodide assay reagent and the iodide concentration determined as described in Method S3.3. Kinetic measurements were carried out at various concentrations of alkyl iodide while keeping the SAH concentration constant at 1 mM. Initial velocities (amount of iodide produced per minute) were fit to the Michaelis-Menten model using GraphPad Prism 6.0 (GraphPad Software Inc., San Diego, CA, US) to determine the k_{cat} and K_m values. All reactions were measured in triplicate. Rates of HMT-dependent iodide formation were corrected for autohydrolysis, which was determined for each iodoalkane at each concentration employed.

Method S4.3   HPLC analysis of MT and HMT reactions

Reactions using the purified HMTs were performed as described for the iodide assay (Method S2.2) and quenched by adding an equal volume of acetonitrile. Samples were vortexed vigorously and then centrifuged at 17,000 *g* for 10 min to remove precipitated protein.

The supernatants were transferred to HPLC sample vial inserts prior to HPLC analysis. Analyses were performed on a Hitachi Elite LaChrom system equipped with a Kinetex EVO C18 (4.6 x 250 mm column, 5 µm particle size) reversed-phase column (Phenomenex, Torrance, CA, US). For separation of SAH, SAM, and SAM analogues, the mobile phases were (A) 10 mM sodium phosphate containing 5 mM sodium 1-heptanesulfonate (pH 3.5), and (B) acetonitrile. A gradient from 5% to 20% (B) over 10 min was followed by a gradient from 20% to 5% B over 1 min, then kept constant for 4 min. The flow rate was constant at 1 ml/min. SAH, SAM, and SAM analogues were detected at 260 nm and quantified using a SAM (J&K Chemical Ltd., China, Beijing) standard and calibration curve (Figure S3b).[10] Previously reported methods were adapted for separations of luteolin (**5**) and its ethylated product (**6**), and 3,4-dihydroxybenzaldehyde (**7**) and its propylated or allylated (**8**) products.[2, 11] A Kinetex EVO C18 (4.6 x 250 mm column, 5 µm particle size) reversed-phase column was used. The mobile phases were 0.1% acetic acid (A) and acetonitrile (B). A gradient elution from 20% to 60% (B) over 5 min, was followed by a gradient from 60% to 20% (B) over 1 min, and then kept constant for 4 min.

Method S5   Preparative-scale HMT reactions

Method S5.1   Preparative-scale production of SAM analogues and alkylated compounds

Scaled-up production of ethylated luteolin (**6**), and propylated or allylated DHBA (**8**) was performed as described in Table S5. In short, 10 or 20 mg of the MT substrate, 80 mM alkyl iodide, 100 µM SAH, MT, and HMT were shaken (1200 rpm) in HPLC vials at 25°C. Subsequently, the reaction mixtures were extracted three times with an equal volume of ethyl acetate. Ethyl acetate was then removed by rotary evaporation and the residues dissolved in hexadeuterodimethyl sulfoxide (DMSO-d_6) for NMR spectroscopy. The alkylated products were analyzed by HPLC (Method S4.3) and then purified by preparative HPLC (Method S5.2).

Preparative-scale production of SAE, SAP, and SAA started with 15 mg SAH (10 mM), 80 mM of the corresponding alkyl iodide, and purified V140T-*At*HMT (Table S4). Reactions (3.9 ml of each) were performed in HPLC vials at 25°C, with shaking at 1200 rpm, for 14 to 24 h. Subsequently, the reaction mixtures were extracted three times with an equal volume of diethyl ether to remove the remaining alkyl iodides and to precipitate protein. The supernatants containing the SAM analogues were analyzed by analytical HPLC (Method S4.3) and then purified by preparative HPLC (Method S5.2).

Method S5.2   Preparative HPLC purification

Prior to purification, SAM analogues and alkylated compounds were analyzed using an analytical LiChrospher® 100 RP-18 (5 µm) LiChroCART® (250×4 mm, Merck) column. The preparative column was a LiChrospher® 100 RP-18 (5 µm) Hibar® RT (250×25 mm, Merck). Both HPLCs were equipped with Shimadzu devices CBM-20A, LC-20A P, SIL-20A, FRC-10A, and an SPD 20A UV/Vis detector. The mobile phases and gradient elution methods were the same as for measurements on the Kinetex EVO C18 column (Method S4.3) and the flow rate was 32 ml/min. Product fractions were collected and dried using a rotary evaporator, lyophilized, and dissolved in DMSO-d_6 (for **6** and **8**) or D_2O (for **3**) for NMR spectroscopy (Method S5.3).

Method S5.3   NMR spectroscopy

NMR spectra were recorded on a Bruker Avance III instrument (^1H NMR: 400 MHz, ^{13}C NMR: 100.6 MHz). Chemical shifts were referenced to tetramethylsilane (TMS) as internal standard in DMSO-d_6 and reported in parts per million (ppm). For D_2O as solvent, chemical shifts were referenced to residual H_2O in the samples. Signals are described using the abbreviations: br = broad, s = singlet, d = duplet, t = triplet, q = quartet, m = multiplet, and combinations thereof.

Protein sequences

Molecular weights and extinction coefficients, used for quantification of protein concentration by measurements of absorbance at 280 nm, were calculated using the ExPASy ProtParam tool.

Arabidopsis thaliana halide methyltransferase

- NCBI Reference Sequence: NP_001318420.1
- Extinction coefficient (280 nm): 42065 $M^{-1}cm^{-1}$
- Molecular weight: 27.452 kDa

MGSSHHHHHHSSGLVPRGSHMAEEQQNSDQSNGGNVIPTPEEVATFLHKTVEEGGWEKCWEEEITPWDQ
GRATPLIVHLVDTSSLPLGRALVPGCGGGHDVVAMASPERFVVGLDISESALAKANETYGSSPKAEYFS
FVKEDVFTWRPTELFDLIFDYVFFCAIEPEMRPAWAKSMYELLKPDGELITLMYPITDHVGGPPYKVDV
STFEEVLVPIGFKAVSVEENPHAIPTRKGKEKLGRWKKIN

Chloracidobacterium thermophilum halide methyltransferase

- GenBank: AEP12557.1
- Extinction coefficient (280 nm): 65680 $M^{-1}cm^{-1}$
- Molecular weight: 24.617 kDa

MGHHHHHHHAENLYFQGSGLGMDADTASFWEEKYRADLTAWDRGGVSPALEHWLAEGALKPGRILIPGCG
YGHEVLALARRGFEVWGLDIALTPVRRLQEKLAQAGLTAHVVEGDVRTWQPEQPFDAVYEQTCLCALSP
EDWPRYEAQLCRWLRPGGRLFALWMQTDRPGGPPYHCGLEAMRVLFALERWRWVEPPQRTVPHPTGFFE
YAAILERLV

Raphanus sativus halide methyltransferase

- GenBank: BAH84870.1
- Extinction coefficient (280 nm): 46535 $M^{-1}cm^{-1}$
- Molecular weight: 26.530 kDa

MAEGQQNSGNSNGENIIPPEDVAKFLPKTVEEGGWEKCWEDGVTPWDQGRATPLVVHLVESSSLPLGRA
LVPGCGGGHDVVAMASPERYVVGLDISESALEKAAETYASSPKAKYFTFVKEDFFTWRPSELFDLIFDY
VVFCAIEPEMRAAWAKTMYELLKPDGELITLMYPITDHDGGPPYKVAVSTYEDVLVPVGFKAVSIEENP
YSIATRKGKEKLGRWKKINKLAAALEHHHHHH

Homo sapiens catechol *O*-methyltransferase (soluble-form)

- NCBI Reference Sequence: NP_009294.1
- Extinction coefficient (280 nm): 23295 M^{-1}cm^{-1}
- Molecular weight: 24.449 kDa

```
MGDTKEQRILNHVLQHAEPGNAQSVLEAIDTYCEQKEWAMNVGDKKGKIVDAVIQEHQPSVLLELGAYC
GYSAVRMARLLSPGARLITIEINPDCAAITQRMVDFAGVKDKVTLVVGASQDIIPQLKKKYDVDTLDMV
FLDHWKDRYLPDTLLLEECGLLRKGTVLLADNVICPGAPDFLAHVRGSSCFECTHYQSFLEYREVVDGL
EKAIYKGPGSEAGPGSSGHHHHHH
```

Curvularia inaequalis vanadium-dependent chloroperoxidase

- Described by Hasan et al.[12]
- Extinction coefficient (280 nm): 93280 M^{-1}cm^{-1}
- Molecular weight: 70.105 kDa

```
MKKLLFAIPLVVPFYSHSTMASHMGSVTPIPLPKIDEPEEYNTNYILFWNHVGLELNRVTHTVGGPLTG
PPLSARALGMLHLAIHDAYFSICPPTDFTTFLSPDTENAAYRLPSPNGANDARQAVAGAALKMLSSLYM
KPVEQPNPNPGANISDNAYAQLGLVLDRSVLEAPGGVDRESASFMFGEDVADVFFALLNDPRGASQEGY
HPTPGRYKFDDEPTHPVVLIPVDPNNPNGPKMPFRQYHAPFYGKTTKRFATQSEHFVADPPGLRSNADE
TAEYDDAVRVAIAMGGAQALNSTKRSPWQTAQGLYWAYDGSNLIGTPPRFYNQIVRRIAVTYKKEEDLA
NSEVNNADFARLFALVDVACADAGIFSWKEKWEFEFWRPLSGVRDDGRPDHGDPFWLTLGAPATNTNDI
PFKDPFPAYPSGHATFGGAVFQMVRRYYNGRVGTWKDDEPDNIAIDMMISEELNGVNRDLRQPYDPTAP
IEDQPGIVRTRIVRHFDSAWELMFENAISRIFLGVHWRFDAAAARDILIPTTTKDVYAVDNNGATVFQN
VEDIRYTTRGTREDPEGLFPIGGVPLGIEIADEIFNNGLKPTPPEIQPMPQETPVQKPVGQQPVKGMWE
EEQAPVVKEAP
```

References

[1]   C. Liao, F. P. Seebeck, *Nat. Catal.* **2019**, *2*, 696-701.

[2]   Q. Tang, Y. M. Vianney, K. Weisz, C. W. Grathwol, A. Link, U. T. Bornscheuer, I. V. Pavlidis, *ChemCatChem* **2020**, *12*, 3721–3727.

[3]   M. T. Reetz, J. D. Carballeira, *Nat. Protoc.* **2007**, *2*, 891.

[4]   Y. Jiang, B. Chen, C. Duan, B. Sun, J. Yang, S. Yang, *Appl. Environ. Microbiol.* **2015**, *81*, 2506-2514.

[5]   Y. Zhang, U. Werling, W. Edelmann, *Nucleic Acids Res.* **2012**, *40*, e55-e55.

[6]   R. Green, E. J. Rogers, *Methods Enzymol.* **2013**, *529*, 329.

[7]   A. S. Aslan-Üzel, A. Beier, D. Kovář, C. Cziegler, S. K. Padhi, E. D. Schuiten, M. Dörr, D. Böttcher, F. Hollmann, F. Rudroff, M. D. Mihovilovic, T. Buryška, J. Damborský, Z. Prokop, C. P. S. Badenhorst, U. T. Bornscheuer, *ChemCatChem* **2020**, *12*, 1-9.

[8]   a) P. D. Josephy, T. Eling, R. P. Mason, *J. Biol. Chem.* **1982**, *257*, 3669-3675; b) P. M. Bozeman, D. B. Learn, E. L. Thomas, *J. Immunol. Methods* **1990**, *126*, 125-133.

[9]   U. K. Laemmli, *Nature* **1970**, *227*, 680-685.

[10]   B. J. Law, A.-W. Struck, M. R. Bennett, B. Wilkinson, J. Micklefield, *Chem. Sci.* **2015**, *6*, 2885-2892.

[11]   Q. Tang, U. T. Bornscheuer, I. V. Pavlidis, *ChemCatChem* **2019**, *11*, 3227-3233.

[12]   Z. Hasan, R. Renirie, R. Kerkman, H. J. Ruijssenaars, A. F. Hartog, R. Wever, *J. Biol. Chem.* **2006**, *281*, 9738-9744.

Affirmation

Hiermit erkläre ich, dass diese Arbeit bisher von mir weder an der Mathematisch-Naturwissenschaftlichen Fakultät der Universität Greifswald noch einer anderen wissenschaftlichen Einrichtung zum Zwecke der Promotion eingereicht wurde.

Ferner erkläre ich, dass ich diese Arbeit selbstständig verfasst und keine anderen als die darin angegebenen Hilfsmittel und Hilfen benutzt und keine Textabschnitte eines Dritten ohne Kennzeichnung übernommen habe.

Qingyun Tang

Curriculum Vitae

PERSONAL INFORMATION

Name, family name:	Qingyun Tang
Birthday:	01.10.1990
Birthplace:	Zhuhai, China
Nationality:	Chinese
Email:	lornatown@hotmail.com

EDUCATION

2016.09-present: **Ph.D. candidate**

Dept. of Biotechnology & Enzyme Catalysis, Institute of Biochemistry, University of Greifswald, Germany

2017.01-2017.07:   **Visiting student**

Dept. of Biochemistry, Institute of Biology, University of Kassel, Germany

2012.09-2016.06:   **M. Eng. in Food Science**

College of Light Industry and Food Sciences, South China University of Technology, China

2008.09-2012.06:   **B. Eng. in Food Quality and Safety**

College of Light Industry and Food Sciences, South China University of Technology, China

List of publications

- Q. Tang, U. T. Bornscheuer*, I. V. Pavlidis*, Specific residues expand the substrate scope and enhance the regioselectivity of a plant *O*-methyltransferase, *ChemCatChem* **2019**, *11*, 3227.

- Q. Tang, Y. M. Vianney, K. Weisz, C. W. Grathwol, A. Link, U. T. Bornscheuer*, I. V. Pavlidis*, Influence of substrate binding residues on the substrate scope and regioselectivity of a plant *O*-methyltransferase against flavonoids, *ChemCatChem* **2020**, *12*, 3721.

- Q. Tang, C. W. Grathwol, A. S. Aslan-Üzel, S. Wu, A. Link, I. V. Pavlidis*, C. P. S. Badenhorst*, U. T. Bornscheuer*, Directed evolution of a halide methyltransferase enables biocatalytic synthesis of diverse SAM analogs, *Angew. Chem., Int. Ed.* **2021**, *60*, 1524; *Angew. Chem.* **2021**, *133*, 1547.

Acknowledgements

I also want to express my heartfelt thankfulness to the international and the Chinese community in Greifswald. My dear Ani, Goran, Gunnar, Jan, Kasia, Mori, Nikkito, Oleg, Olga A., Pau, Sahar, Wardito, and William, we have had so much fun at all those movie nights, Sushi nights, Christmas parties, Laser tag, Chinese buffets and barbecues. Many thanks to my Chinese friends Baozhu, Chen Fan, Fukun, Guokai, Haitao, Haomin, Junfeng, Liu Fan, Meimei, Mu Yu, Rong Piao, Yangjing, Yiming, and Zhang Chong. We can talk in Mandarin and enjoy the Chinese cuisine, which makes me feel at home. And I will not forget the great times I spent with Chonggan, Jiayu and Jiqing, from the German course in China to our meeting in Germany. All in all, there are a lot of great memories in Greifswald, the IBZ, the market place and the Christmas market, Fritz and Kontor, the harbour and fischbrötchen, Eldena, Hiddensee, Rügen, Szczecin, etc., which I will cherish for the rest of my life.

I would like to acknowledge the Chinese Scholarship Council (CSC) for financial support over the last four years which enabled me to pursue my PhD degree in Germany (File No.: 201606150073). And Prof. Yonghua Wang, who has offered me massive support during my master study and with my PhD application, will never be forgotten.

Last but not least, I would like to dedicate my deepest gratitude to my parents, my grandparents and all the family members for your unconditional love, trust and support. Home is where the heart is.

Qingyun Tang (Lorna)

14.12.2020 Greifswald