

# **Biosynthesis of oxygenated monoterpenes in *Thyme, Melaleuca, and Eucalyptus* species**

## **Dissertation**

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

der

Naturwissenschaftlichen Fakultät I – Biowissenschaften –

der Martin-Luther-Universität  
Halle-Wittenberg

vorgelegt

von Frau Sandra Krause

geb. am 21.09.1984 in Karl-Marx-Stadt, jetzt Chemnitz



1. Gutachter: Prof. Dr. Jörg Degenhardt

2. Gutachter: Prof. Dr. Alain Tissier

3. Gutachter: Prof. William Foley

Tag der öffentlichen Verteidigung: 31.03.2016



---

## Table of Contents

|  |           |
|--|-----------|
| <b>List of Figures</b> .....   | <b>i</b>  |
| <b>List of Tables</b> .....  | <b>iv</b> |
| <b>Abbreviations</b> .....   | <b>v</b>  |
| <b>General Introduction</b> .....  | <b>1</b>  |
| I. Essential Oils: Valuable products of plant secondary metabolism .....   | 1         |
| II. Thyme ( <i>Thymus vulgaris</i> ) – a medicinal plant with long tradition .....   | 2         |
| III. <i>Melaleuca</i> and <i>Eucalyptus</i> species are important resources for the Australian essential oil industry .....      | 3         |
| IV. The chemistry and biosynthesis of terpenes.....  | 4         |
| V. Terpenes and terpenoids are characterized by a great chemical diversity .....   | 7         |
| VI. Aims of the study .....  | 8         |
| <b>Chapter I</b> .....   | <b>9</b>  |
| <b>Stereochemical characterization of monoterpene synthases producing cyclic oxygenated monoterpenes in essential oil plants</b> |           |
| <b>1 Introduction</b> .....  | <b>10</b> |
| 1.1 Monoterpene synthases are able to form an array of structural diverse compounds .....  | 10        |
| 1.2 Stereochemical aspects of terpene synthase reactions .....   | 12        |
| 1.3 Structure-function relationships in monoterpene synthases.....   | 13        |
| <b>2 Materials and Methods</b> .....   | <b>15</b> |
| 2.1 Plant material.....  | 15        |
| 2.2 Microbiological methods.....   | 15        |
| 2.2.1 Bacteria strains.....  | 15        |
| 2.2.2 Cultivation of <i>E. coli</i> .....  | 15        |
| 2.2.3 Transformation of <i>E. coli</i> .....   | 16        |
| 2.3 Nucleic acid techniques.....   | 16        |
| 2.3.1 RNA extraction from leaf material and cDNA/RACE cDNA synthesis .....   | 16        |
| 2.3.2 Amplification of DNA .....   | 16        |
| 2.3.3 Transcript quantification by quantitative real-time PCR.....   | 17        |
| 2.3.4 Electrophoresis of DNA and RNA .....   | 18        |
| 2.3.5 Isolation of DNA fragments from agarose gels .....   | 18        |
| 2.3.6 Restriction of DNA .....   | 18        |
| 2.3.7 Cloning of DNA fragments.....  | 19        |

|          |   |           |
|----------|---|-----------|
| 2.3.8    | Multiplication of plasmid DNA .....   | 19        |
| 2.3.9    | Sequencing of DNA .....   | 19        |
| 2.3.10   | Analysis of DNA and protein sequences .....   | 20        |
| 2.3.11   | Quantification of DNA and RNA .....   | 20        |
| 2.4      | Isolation of terpene synthases from <i>Thymus</i> and <i>Melaleuca</i> species .....  | 20        |
| 2.5      | <i>In vitro</i> mutagenesis of terpene synthase open reading frames.....  | 22        |
| 2.6      | Dendrogram analysis.....  | 22        |
| 2.7      | Terpene synthase overexpression in <i>E. coli</i> and protein extraction .....  | 22        |
| 2.8      | Purification of terpene synthases.....  | 23        |
| 2.9      | Enzyme assays.....  | 23        |
| 2.10     | Terpene extraction from thyme leaves .....  | 24        |
| 2.11     | GC-MS analysis of leaf terpenes and terpene synthase assay products .....   | 24        |
| 2.12     | Stereospecific synthesis of LPP .....   | 25        |
| 2.13     | Determination of $K_m$ values of terpene synthases.....   | 25        |
| 2.14     | Modeling of the terpene synthase active site and docking studies.....   | 26        |
| <b>3</b> | <b>Results .....</b>  | <b>27</b> |
| 3.1      | Isolation and characterization of two sabinene hydrate synthases from thyme ( <i>Thymus vulgaris</i> ).....   | 27        |
| 3.1.1    | The multiproduct monoterpene synthases TvTPS6 and TvTPS7 provide most of the monoterpene spectrum of the <i>Thymus vulgaris</i> ( <i>E</i> )-sabinene hydrate chemotype ..... | 27        |
| 3.1.2    | TvTPS6 and TvTPS7 produce monoterpenes with opposite configuration .....  | 29        |
| 3.1.3    | The configuration of the monoterpene products is determined by the configuration of the LPP intermediate in the TvTPS6 and TvTPS7 active centers .....                        | 31        |
| 3.1.4    | Mutagenesis studies revealed an amino acid responsible for the different stereospecificity of TvTPS6 and TvTPS7.....  | 33        |
| 3.1.5    | Modeling of the active sites of TvTPS6 and TvTPS7 and substrate docking studies .....   | 34        |
| 3.1.6    | The pathway to ( <i>E</i> )- and ( <i>Z</i> )-sabinene hydrate involves water quenching ..  | 35        |
| 3.1.7    | Mutagenesis studies of the 1,8-cineole synthase of <i>Salvia fruticosa</i> show no function of Asn-338 in stereocontrol .....   | 36        |
| 3.2      | Isolation and characterization of an $\alpha$ -terpineol synthase from thyme ( <i>Thymus vulgaris</i> ) .....   | 37        |
| 3.2.1    | The multiproduct terpene synthase TvTPS5 produces most of the monoterpenes found in the $\alpha$ -terpineol chemotype of <i>Thymus vulgaris</i> .....                         | 37        |

---

|   |   |           |
|---|---|-----------|
| 3.2.2   | In TvTPS5, the configuration of the monoterpene products is influenced by the same amino acid position that was critical in TvTPS6 and TvTPS7.....  | 39        |
| 3.2.3   | Water capture takes place at different steps of the pathway of TvTPS5 and TvTPS7 .....  | 40        |
| 3.3   | Isolation of monoterpene synthase genes from <i>Melaleuca linariifolia</i> and <i>Melaleuca trichostachya</i> .....   | 41        |
| <b>4</b>  | <b>Discussion.....</b>  | <b>44</b> |
| 4.1   | The stereospecific reaction mechanism of the sabinene hydrate synthases TvTPS6 and TvTPS7 from <i>Thymus vulgaris</i> .....   | 44        |
| 4.1.1   | TvTPS6 and TvTPS7 are most likely responsible for the characteristic monoterpene composition of the ( <i>E</i> )-sabinene hydrate chemotype of <i>Thymus vulgaris</i> .....   | 44        |
| 4.1.2   | The opposite initial binding conformation of GPP leads to the opposite stereospecificity of TvTPS6 and TvTPS7.....  | 45        |
| 4.1.3   | One amino acid in the active site controls most of the stereospecificity of TvTPS6 and TvTPS7 .....   | 47        |
| 4.1.4   | Modeling studies suggest a steric influence of Asn-350 (TvTPS6) and Ile-346 (TvTPS7) on the binding conformation of GPP .....   | 48        |
| 4.2   | Characterization and stereochemical aspects of the reaction mechanism of the $\alpha$ -terpineol synthase TvTPS5 .....  | 49        |
| 4.2.1   | TvTPS5 may be the main monoterpene synthase of the $\alpha$ -terpineol chemotype of <i>Thymus vulgaris</i> .....  | 49        |
| 4.2.2   | TvTPS5 and TvTPS7 show a close phylogenetic relationship and a similar stereospecificity .....  | 50        |
| 4.2.3   | The two amino acids difference between the TvTPS5 and TvTPS7 active sites may affect water binding .....  | 51        |
| 4.3   | The monoterpene synthase genes isolated from <i>Melaleuca trichostachya</i> and <i>Melaleuca linariifolia</i> show a close relationship to 1,8-cineole and sabinene hydrate synthases from other <i>Myrtaceae</i> ..... | 52        |
| <b>Chapter II</b>   | <b>.....</b>  | <b>55</b> |
| <b>Monoterpene modification by cytochrome P450 monooxygenases in essential oil plants</b> |   |           |
| <b>1</b>  | <b>Introduction.....</b>  | <b>56</b> |
| 1.1   | Cytochrome P450 monooxygenases and their role in plant oxidative metabolism .....   | 56        |
| 1.2   | Nomenclature and gene organization of P450 monooxygenases.....  | 56        |
| 1.3   | Overall cytochrome P450 architecture and structure-function relationships.....  | 57        |
| 1.4   | Reaction mechanism of P450 enzymes: The catalytic cycle.....  | 58        |

|          |   |           |
|----------|---|-----------|
| 1.5      | P450 enzymes participate in thymol and carvacrol formation in thyme and oregano .....   | 59        |
| <b>2</b> | <b>Materials and Methods.....</b>   | <b>62</b> |
| 2.1      | Plant material.....   | 62        |
| 2.2      | Isolation of cytochrome P450 genes from <i>Eucalyptus</i> species .....   | 62        |
| 2.3      | <i>In vitro</i> mutagenesis of P450 monooxygenase open reading frames .....   | 62        |
| 2.4      | Yeast expression system.....  | 63        |
| 2.4.1    | Yeast strain.....   | 63        |
| 2.4.2    | Media for yeast cultivation, transformation, and protein overexpression ...   | 63        |
| 2.4.3    | Yeast transformation .....  | 63        |
| 2.4.4    | Cytochrome P450 monooxygenase overexpression and microsomes extraction .....  | 64        |
| 2.5      | P450 enzyme assay.....  | 65        |
| 2.6      | Limonene feeding of thyme cuttings.....   | 65        |
| 2.7      | <i>Thymus vulgaris</i> transcriptome sequencing .....   | 66        |
| 2.8      | Modeling of CYP71D179 .....   | 66        |
| <b>3</b> | <b>Results .....</b>  | <b>67</b> |
| 3.1      | The formation of the phenolic monoterpenes thymol and carvacrol from a monoterpene precursor – one cytochrome P450 reaction or a multiple step pathway?.....          | 67        |
| 3.1.1    | CYP71D179 and CYP71D181 form <i>p</i> -cymene from $\gamma$ -terpinene .....  | 67        |
| 3.1.2    | Additional cytochrome P450 enzymes might be involved in the formation of thymol and carvacrol.....  | 69        |
| 3.1.3    | A dehydrogenase might catalyze a crucial step in the thymol and carvacrol formation.....  | 70        |
| 3.1.4    | Limonene as an alternative monoterpene precursor of thymol and carvacrol . .....  | 71        |
| 3.2      | The ( <i>R</i> )-limonene hydroxylation patterns of CYP71D179 and CYP71D181 are presumably dependent on specific amino acids in the substrate recognition sites ..... | 77        |
| 3.3      | An approach to P450 monooxygenases involved in the metabolism of monoterpenes in <i>Eucalyptus</i> species .....  | 80        |
| <b>4</b> | <b>Discussion.....</b>  | <b>83</b> |
| 4.1      | Different pathways for the formation of the phenolic monoterpenes thymol and carvacrol are conceivable .....  | 83        |
| 4.1.1    | CYP71D179 and CYP71D181 are active cytochrome P450 monooxygenases, but their explicit role in thymol and carvacrol formation remains unclear .....                    | 83        |



---

|       |   |            |
|-------|---|------------|
| 4.1.2 | The transformation from $\gamma$ -terpinene into thymol and carvacrol may require the activity of multiple enzymes .....              | 84         |
| 4.1.3 | Limonene as a potential origin of thymol and carvacrol .....  | 86         |
| 4.2   | The oxygenation site of cyclic monoterpenes is determined by sterical constraints in the P450 active center .....                     | 88         |
| 4.2.1 | Alteration of amino acids in SRS 4-6 did not affect the ( <i>R</i> )-limonene hydroxylation patterns of CYP71D179 and CYP71D181 ..... | 88         |
| 4.2.2 | The biological function of CYP71D178-182 might be the C-6 hydroxylation of either ( <i>S</i> )-limonene or $\gamma$ -terpinene .....  | 89         |
| 4.3   | The <i>Eucalyptus</i> cytochrome P450s identified in this study belong to the large and diverse CYP71 family .....                    | 91         |
|       | <b>Summary.....</b>   | <b>93</b>  |
|       | <b>Zusammenfassung.....</b>   | <b>95</b>  |
|       | <b>References.....</b>  | <b>97</b>  |
|       | <b>Supplementary Material Chapter I.....</b>  | <b>105</b> |
|       | <b>Supplementary Material Chapter II .....</b>  | <b>113</b> |



## List of Figures

|         |   |    |
|---------|---|----|
| Fig. 1  | Topview on a glandular trichome of a <i>Thymus vulgaris</i> leaf.....   | 2  |
| Fig. 2  | The major components of the essential oils from <i>Thymus vulgaris</i> chemotypes found in Southern France.....   | 3  |
| Fig. 3  | Sub-epidermal shizo-lysigenous secretory cavities in a <i>Eucalyptus globulus</i> leaf.....   | 4  |
| Fig. 4  | Overview of the terpene biosynthesis in plants.....   | 6  |
| Fig. 5  | Carbocationic reaction mechanism of monoterpene synthases.....  | 11 |
| Fig. 6  | The binding conformation of GPP determines the stereochemical configuration of LPP, which is converted to the respective enantiomers of the $\alpha$ -terpinyl cation intermediate.....   | 13 |
| Fig. 7  | Scheme of the universal monoterpene synthase sequence motifs and their respective functions.....  | 14 |
| Fig. 8  | The monoterpene spectrum of the ( <i>E</i> )-sabinene hydrate chemotype (U-type) of <i>Thymus vulgaris</i> . (A) The terpene synthases TvTPS6 and TvTPS7 produce monoterpenes of the ( <i>E</i> )-sabinene hydrate chemotype of <i>Thymus vulgaris</i> . (B)..... | 28 |
| Fig. 9  | The monoterpene synthase genes <i>Tvtps6</i> and <i>Tvtps7</i> are both expressed in leaves of <i>Thymus vulgaris</i> .....   | 29 |
| Fig. 10 | TvTPS6 and TvTPS7 produce monoterpenes with opposite stereochemical configuration. A-D.....   | 30 |
| Fig. 11 | The configuration of the monoterpene products of TvTPS6 and TvTPS7 is determined by the configuration of the LPP intermediate.....  | 31 |
| Fig. 12 | The configuration of monoterpene minor products of TvTPS6 and TvTPS7 is determined by the configuration of the LPP intermediate. A-B.....   | 32 |
| Fig. 13 | A critical amino acid is located seven amino acids upstream of the DDxxD motif in TvTPS6, TvTPS7, and SfCinS1.....  | 33 |
| Fig. 14 | Site-directed mutagenesis of a critical amino acid residue of TvTPS6 and TvTPS7 inverts the configuration of the sabinene hydrate products.....   | 34 |
| Fig. 15 | Model of the TvTPS6 and TvTPS7 active sites.....  | 35 |
| Fig. 16 | Alignment of the C-terminal domain of TvTPS6 and TvTPS7.....  | 36 |
| Fig. 17 | The amino acid residue crucial for the stereospecificity of TvTPS6 and TvTPS7 does not affect the enantiomeric composition of a 1,8-cineole synthase of <i>Salvia fruticosa</i> (SfCinS1).....  | 36 |
| Fig. 18 | Dendrogram analysis by Maximum Likelihood method with <i>Tvtps5</i> , <i>Tvtps6</i> , <i>Tvtps7</i> , and monoterpene synthase genes from other <i>Lamiaceae</i> species.....   | 38 |
| Fig. 19 | The monoterpene spectrum of the $\alpha$ -terpineol chemotype (A-type) of <i>Thymus vulgaris</i> . (A) The terpene synthase TvTPS5 produces monoterpenes of the $\alpha$ -terpineol chemotype of <i>Thymus vulgaris</i> . (B).....                                | 39 |

|   |    |
|---|----|
| Fig. 20 Site-directed mutagenesis of a critical amino acid residue of TvTPS5 changes the enantiomeric ratio of the $\alpha$ -terpineol product.....   | 40 |
| Fig. 21 Model of the TvTPS5 and TvTPS7 active sites.....  | 41 |
| Fig. 22 Dendrogram analysis by Maximum Likelihood method with <i>Mttps1</i> , <i>Mttps2</i> , <i>Mttps3</i> , <i>Mltps1</i> , and other monoterpene synthase genes from <i>Melaleuca</i> and <i>Eucalyptus</i> species.....                       | 43 |
| Fig. 23 The products of TvTPS6 and TvTPS7 constitute the monoterpene spectrum of the ( <i>E</i> )-sabinene hydrate chemotype (U-type) of <i>Thymus vulgaris</i> .....   | 45 |
| Fig. 24 Putative reaction mechanism of TvTPS6 and TvTPS7.....   | 46 |
| Fig. 25 Scheme of the universal P450 sequence motifs and their respective functions.....  | 58 |
| Fig. 26 The catalytic cycle of P450 enzymes.....  | 59 |
| Fig. 27 Originally suggested pathway from $\gamma$ -terpinene to thymol and carvacrol postulated by Pouloule and Croteau (Pouloule and Croteau, 1978).....  | 60 |
| Fig. 28 Pathway from $\gamma$ -terpinene to thymol and carvacrol suggested by C. Crocoll (Crocoll, 2011).<br>.....  | 61 |
| Fig. 29 <i>In vitro</i> product spectrum of CYP71D181. (A) <i>S. 67 P</i> -cymene is formed from $\gamma$ -terpinene by CYP71D181. (B).....   | 68 |
| Fig. 30 <i>In vitro</i> product spectra of CYP71D179 and CYP71D181 at different incubation temperatures.....  | 69 |
| Fig. 31 Comparison of the first steps of the pathway from (-)-limonene to menthol in <i>Mentha x piperita</i> (A) (Croteau et al., 2005) and the postulated pathway from $\gamma$ -terpinene to thymol in <i>Thymus vulgaris</i> (B).....         | 70 |
| Fig. 32 Amino acid sequence alignment of the putative short-chain dehydrogenase from <i>Thymus vulgaris</i> (ISPD-like) and the (-)-( <i>trans</i> )-isopiperitenol dehydrogenase from <i>Mentha x piperita</i> (ISPD) (Ringer et al., 2005)..... | 71 |
| Fig. 33 Dendrogram analysis of short-chain dehydrogenases by Maximum Likelihood method...71   | 71 |
| Fig. 34 <i>In vitro</i> product spectrum of CYP71D179 and CYP71D181.....  | 72 |
| Fig. 35 Alternative hypothetic pathway leading to thymol (A) and carvacrol (B) with limonene as origin.....   | 73 |
| Fig. 36 Dendrogram analysis by Maximum Likelihood method of <i>Lamiaceae</i> cytochrome P450 enzymes that hydroxylate limonene as their main substrate.....   | 74 |
| Fig. 37 Limonene was incorporated into thyme leaves by evaporation in closed glass vessels for 24h.....   | 75 |
| Fig. 38 Monoterpene spectrum of essential oils extracted from stems of different <i>Thymus vulgaris</i> chemotypes after feeding with ( <i>S</i> )-limonene and ( <i>R</i> )-limonene for 24 h.....   | 76 |
| Fig. 39 Amino acid sequence alignment of the C-terminal region of CYP71D179 and CYP71D181. (A) Model of CYP71D179. (B).....   | 78 |
| Fig. 40 Site-directed mutagenesis of amino acid residues of CYP71D179 and CYP71D181 in substrate recognition sites 4-6.....   | 79 |

---

|  |     |
|--|-----|
| Fig. 41 Dendrogram analysis by Maximum Likelihood method of the <i>Eucalyptus</i> cytochrome P450s <i>CYP71AP17</i> , <i>CYP71AP18</i> , <i>CYP71AP19</i> , <i>CYP71AH18</i> , <i>CYP71CJ1</i> , and further P450s of the CYP71 family from various species..... | 82  |
| Fig. 42 Amino acid sequence alignment of the SRS 5 region of CYP71D enzymes isolated from thyme and oregano.....   | 90  |
| Fig. 43 Model of the CYP71D179 active site.....  | 90  |
| Fig. S1 Amino acid alignment of TvTPS6 and TvTPS7.....   | 108 |
| Fig. S2 Amino acid sequence of TvTPS5.....   | 109 |
| Fig. S3 Model of the TvTPS5 and TvTPS7 active sites.....   | 109 |
| Fig. S4 Amino acid alignment of MtTPS1-TPS3 and MtTPS1.....  | 110 |
| Fig. S5 Amino acid alignment of the identified (partial) P450 enzyme sequences CYP71D_u1-u3 and CYP71D179.....   | 115 |
| Fig. S6 Comparison of the recorded mass spectra of the isopiperitenone GC standard and substances “a” and “b” (recorded in the limonene feeding experiment with <i>Thymus vulgaris</i> , Results II, section 3.1.4).....   | 116 |
| Fig. S7 Proposed pathway from limonene to piperitone.....  | 116 |
| Fig. S8 Amino acid alignment of the <i>Eucalyptus</i> cytochrome P450 enzymes CYP71AP17, CYP71AP18, CYP71AP19, CYP71AH18, and CYP71CJ1.....  | 117 |

**List of Tables**

|          |  |     |
|----------|--|-----|
| Table 1  | Classification of terpenes.....  | 5   |
| Table 2  | The monoterpene composition of the essential oil of the ( <i>E</i> )-sabinene hydrate chemotype of <i>Thymus vulgaris</i> in comparison to the product blends formed by TvTPS6 and TvTPS7..... | 29  |
| Table 3  | Monoterpene formation from the ( <i>3R</i> )-LPP substrate by TvTPS6 and TvTPS7...   | 33  |
| Table 4  | Monoterpene composition of the <i>Melaleuca trichostachya</i> and <i>Melaleuca linariifolia</i> samples collected and used for RNA extraction.....   | 42  |
| Table 5  | Monoterpene composition of the <i>Eucalyptus dives</i> , <i>Eucalyptus piperita</i> , and <i>Eucalyptus elata</i> samples used for RNA extraction.....   | 80  |
| Table S1 | Primers used in this study - Chapter I.....  | 111 |
| Table S2 | Primers used in this study - Chapter II.....   | 118 |
| Table S3 | Amino acid sequence identity of CYP71D_u1-u3 and other P450 enzymes from the CYP71D subfamily.....   | 119 |

---

## Abbreviations

|         |  |
|---------|--|
| aa      | amino acid                                   |
| A-type  | $\alpha$ -terpineol chemotype                |
| ANU     | Australian National University               |
| bp      | base pair                                    |
| BLAST   | Basic Local Alignment Search Tool            |
| CNRS    | Centre national de la recherche scientifique |
| CoA     | coenzyme A                                   |
| CPR     | cytochrome P450 reductase                    |
| C-type  | carvacrol chemotype                          |
| DMAPP   | dimethylallyl diphosphate                    |
| DNA     | deoxyribonucleic acid                        |
| dNTP    | deoxynucleoside triphosphate                 |
| DTT     | dithiothreitol                               |
| E. coli | <i>Escherichia coli</i>                      |
| EDTA    | ethylenediaminetetraacetic acid              |
| ER      | endoplasmatic reticulum                      |
| FID     | flame ionization detector                    |
| FPP     | farnesyl diphosphate                         |
| GC      | gas chromatograph                            |
| GC-MS   | gas chromatography - mass spectrometry       |
| GPP     | geranyl diphosphate                          |
| GGPP    | geranyl geranyl diphosphate                  |
| G-type  | geraniol chemotype                           |
| IPP     | isopentenyl diphosphate                      |
| kb      | kilobase pairs                               |
| LB      | Luria-Bertani (medium)                       |
| LPP     | linalyl diphosphate                          |

## Abbreviations

---

|                      |   |
|----------------------|---|
| L-type               | linalool chemotype  |
| MEP                  | methylerythritol phosphate                                |
| MLU                  | Martin-Luther-University                                  |
| MPI                  | Max Planck Institute                                      |
| MS                   | mass spectrometer   |
| MVA                  | mevalonic acid  |
| NADPH                | Nicotinamide adenine dinucleotide phosphate hydrogen      |
| NCBI                 | National Center for Biotechnology Information             |
| OD                   | optical density   |
| ORF                  | open reading frame  |
| PCR                  | polymerase chain reaction                                 |
| PEG                  | polyethylene glycol                                       |
| PMSF                 | phenylmethylsulfonyl fluoride                             |
| qRT PCR              | quantitative real-time PCR                                |
| RACE                 | rapid amplification of cDNA ends                          |
| RNA                  | ribonucleic acid  |
| Rpm                  | rounds per minute   |
| SC                   | selection (medium)  |
| <i>S. cerevisiae</i> | <i>Saccharomyces cerevisiae</i>                           |
| SDS-PAGE             | sodium dodecyl sulfate polyacrylamide gel electrophoresis |
| SDR                  | short-chain dehydrogenase/reductase                       |
| SOC                  | Super Optimal broth with Catabolite repression            |
| SPME                 | solid phase microextraction                               |
| SRS                  | substrate recognition site                                |
| TAE                  | Tris-acetate-EDTA buffer                                  |
| TIC                  | total ion chromatogram                                    |
| TLC                  | thin-layer chromatography                                 |
| TPS                  | terpene synthase  |



|        |  |
|--------|--|
| Tris   | tris(hydroxymethyl)aminomethane                          |
| T-type | thymol chemotype   |
| Tv     | <i>Thymus vulgaris</i>                                   |
| U      | units  |
| U-type | ( <i>E</i> )-sabinene hydrate chemotype                  |
| UV     | ultra violet   |
| YPGA   | yeast-bacto peptone-glucose-adenine hemisulfate (medium) |



## General Introduction

### I. Essential Oils: Valuable products of plant secondary metabolism

Essential oils are plant products of great commercial value for pharmaceuticals, cosmetics, and food products (Burt, 2004; Baser and Buchbauer, 2009). They are used in many medications for their antibacterial, antifungal, and antiviral effects. (Guynot et al., 2005; Schnitzler et al., 2007; Lang and Buchbauer, 2012). Due to their pleasant fragrance they are added in body hygiene products such as deodorants and shower gels. In the food industry, pure essential oils are applied as aromatic additive, and dried herbs with essential oils are commonly used as spices.

Essential oils consist of mostly lipophilic compounds with a molecular weight lower than 500 Da and can comprise over 100 different compounds that are usually of terpenoid, phenylpropanoid, or fatty acid origin (Nakatsu et al., 2000; Baser and Demirci, 2007). Among these compounds, terpenoids are predominant (Baser and Buchbauer, 2009). The amount of the respective constituents of the essential oil can vary due to ecological influences or genetic variation (Raut and Karuppayil, 2014). In plants, essential oils fulfill various functions such as antiherbivore defense (Bryant et al., 1991; Gershenzon and Croteau, 1991), attraction of pollinators (Ayasse et al., 2000), or adaptation to varying climatic conditions (Amiot et al., 2005). Because of their cell toxic properties, essential oils are often stored in separated plant compartments. This can be oil or resin ducts found in the needles of conifers, or the glandular trichomes typical for *Lamiaceae* species (Gershenzon, 1998; Stahl-Biskup, 2002; Baser and Demirci, 2007).

Plants from about 60 different families are known to produce essential oils. Among these, essential oils from the *Alliaceae*, *Apiaceae*, *Asteraceae*, *Lamiaceae*, *Myrtaceae*, and *Rutaceae* belong to the most important (Raut and Karuppayil, 2014). Some of the highest valued essential oils worldwide are orange oil from *Citrus sinensis*, eucalyptus oil from *Eucalyptus globulus*, and lemon oil from *Citrus limon*. Furthermore, corn mint oil from *Mentha arvensis*, peppermint oil from *Mentha x piperita*, and other *Lamiaceae* herb oils, such as thyme oil from *Thymus vulgaris*, are important (Lawrence and Tucker, 2002; Brud, 2010). The oils of *Thyme*, *Melaleuca*, and *Eucalyptus* species are especially appreciated for their health benefits and used in medicinal applications. In this study, aspects of the essential oil formation in these plants are investigated.

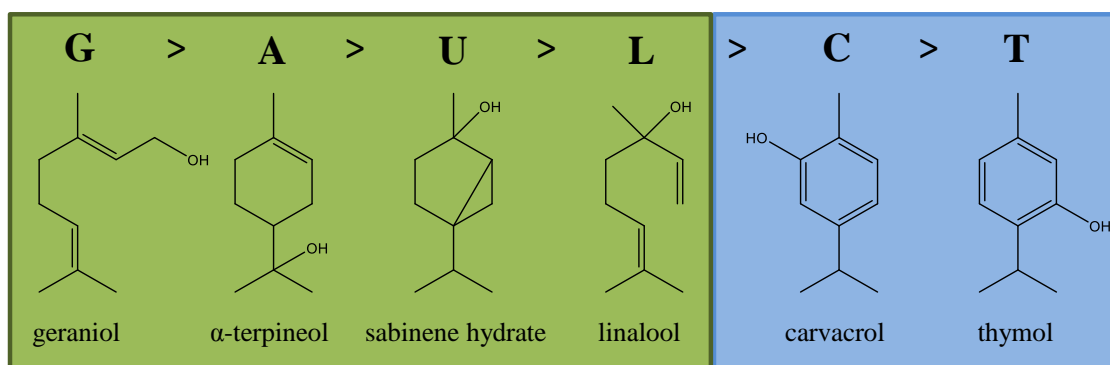
## II. Thyme (*Thymus vulgaris*) – a medicinal plant with long tradition

The genus *Thymus* comprises about 215 different species and is one of the most important genera of the *Lamiaceae* (Morales, 2002). These evergreen dwarf shrubs are endemic to Central and South Europe, the Balkan countries, and the Caucasus. In addition, they are cultivated in Central Europe, East Africa, India, Turkey, Morocco, and North America (Czygan and Hiller, 2002). The most prominent species is *Thymus vulgaris* because its essential oil is a natural resource for a wide range of monoterpenes. Due to their antibacterial and spasmolytic activities, these monoterpenes are used in many pharmaceuticals, especially in medications for the respiratory system. In addition, thyme is widely used as a spice and as a preservative in the food industry (Zarzuelo and Crespo, 2002). The essential oil is stored in glandular trichomes found on the surface of the leaves (Fig. 1). The trichomes consist of 12 secretory cells in which the essential oil components are produced. These cells are covered by a cuticle. For storage, the compounds are secreted into the sub-cuticular space (Stahl-Biskup, 2002).



**Fig. 1 Topview on a glandular trichome of a *Thymus vulgaris* leaf.** One trichome consists of 12 secretory cells. The essential oil is stored in the sub-cuticular space. Figure taken from: Mikroskopischer Farbatlas pflanzlicher Drogen, Bettina Rahfeld, Spektrum Akademischer Verlag, 2011.

The composition of the essential oil varies strongly between thyme plants. Natural populations of thyme often consist of several chemotypes, which are morphologically identical sub-populations with a distinct composition of their essential oils. These chemotypes are characterized by their major monoterpene alcohols. In Southern France, six chemotypes of *Thymus vulgaris* are distinguished:  $\alpha$ -terpineol (A-type), carvacrol (C-type), geraniol (G-type), linalool (L-type), thymol (T-type), and (*E*)-sabinene hydrate (U-type), and their formation underlies an epistatic control (Fig. 2) (Vernet et al., 1986; Schimmel, 2014). Commercial thyme is usually dominated by thymol and carvacrol, which are the most valuable monoterpenes for pharmaceutical applications and as a spice. Since these two phenolic monoterpenes are comparatively rare in other plant genera, *Thymus* oils are the most important resource (Stahl-Biskup, 2002).



**Fig. 2** The major components of the essential oils from *Thymus vulgaris* chemotypes found in Southern France. The chemotypes are shown in their epistatic order. The non-phenolic chemotypes are shaded in green, the phenolic chemotypes are shaded in blue.

### III. *Melaleuca* and *Eucalyptus* species are important resources for the Australian essential oil industry

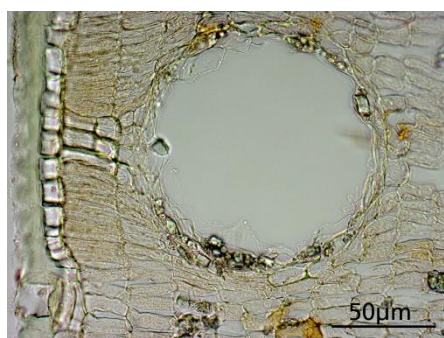
Both the genera *Melaleuca* and *Eucalyptus* belong to the *Myrtaceae* family and are indigenous to Australia and some islands of South-East Asia and the Pacific (Southwell, 1999; Coppen, 2002a). Especially eucalyptus is furthermore dispersed and cultivated in many additional areas of the world including South Africa, Brazil, or China (Turnbull and Booth, 2002). The genera are large and diverse, with *Melaleuca* comprising about 300 species (Brophy et al., 2013a) and *Eucalyptus* even consisting of over 800 species (Coppen, 2002a). However, only few of these species are of commercial interest.

For *Melaleuca* oil extraction, first *Melaleuca cajuputi* (cajuput oil) and *Melaleuca quinquenervia* (niaouli oil) were exploited due to their high content of 1,8-cineole. The basis of the Australian tea tree oil industry then became *Melaleuca alternifolia* (tea tree oil), sometimes including *Melaleuca linariifolia* and *Melaleuca dissitiflora*. This oil is especially appreciated for its high content of terpinen-4-ol (Brophy et al., 2013b), which has antimicrobial and antifungal activities (Mondello et al., 2006). The Australian standard for tea tree essential oil demands a high content of terpinen-4-ol (30-48%) along with lower amounts of 1,8-cineole (<15%) (Padovan et al., 2013). In order to produce oil of these specifications, chemotypes rich in terpinen-4-ol have to be cultivated. To select such chemotypes in breeding processes without an essential oil analysis, molecular markers for *Melaleuca* might be a useful tool. The development of such markers requires a detailed knowledge of terpene biosynthesis and its regulatory mechanisms on the molecular level. For example, the identification of the enzymes responsible for the terpene formation might allow the development of reliable markers. Young seedlings can thus be screened for the

respective genes to determine the plant's chemotype at an early stage (Brophy et al., 2013b).

The *Eucalyptus* species of most commercial significance for essential oil production are *Eucalyptus globulus*, *Eucalyptus exserta*, *Eucalyptus polybractea* and, *Eucalyptus smithii* (Brooker, 2002; Copen, 2002b). The Australian standards for eucalyptus oil demand a high content of 1,8-cineole (70-85%) (Padovan et al., 2013), which is present in these species. But piperitone and citronellal are also commercially valuable eucalyptus oil constituents. These compounds are found in large amounts in *Eucalyptus citriodora* and *Eucalyptus dives*, respectively (Davis, 2002). 1,8-Cineole is a mucolytic agent and therefore used in many pharmaceuticals for respiratory infection treatments (Worth et al., 2009). In the past, piperitone was used for menthol production, but nowadays it is applied for flavor and fragrance purposes due to its peppermint-like aroma (Brooker, 2002; Davis, 2002). The third major compound found in eucalyptus essential oils, citronellal, is a very effective and commonly used insect repellent (Kim et al., 2005).

In both *Eucalyptus* and *Melaleuca* leaves, the essential oil is stored in sub-epidermal shizo-lysigenous secretory cavities (Fig. 3) (List, 1995; Turner, 1999a).



**Fig. 3 Sub-epidermal shizo-lysigenous secretory cavity in a *Eucalyptus globulus* leaf.** The surrounding cell layers produce and secrete the essential oil in the internal cavity. Figure taken from: Mikroskopischer Farbatlas pflanzlicher Drogen, Bettina Rahfeld, Spektrum Akademischer Verlag, 2011.

#### IV. The chemistry and biosynthesis of terpenes

Terpenes form the largest group of plant organic secondary compounds (Gershenzon and Croteau, 1991). Generally, terpenes are of lipophilic character as they are built from a C<sub>5</sub>-hydrocarbon, also called isoprene unit. The classification of terpenes is determined by the number of isoprene units they are built from (Table 1). The simplest of all terpenes are the hemiterpenes, composed of only one C-5 unit. C-10 compounds are called monoterpenes, consisting of two isoprene units. Accordingly, sesquiterpenes (C-15) consist of three isoprene units and diterpenes (C-20) are built from four isoprene units.

High-molecular terpenes consist of C-30 backbones (triterpenes) or C-40 backbones (tetraterpenes). Polyterpenes have a chain length of  $(C-5)_n$ , ( $n > 8$ ).

**Table 1 Classification of terpenes.**

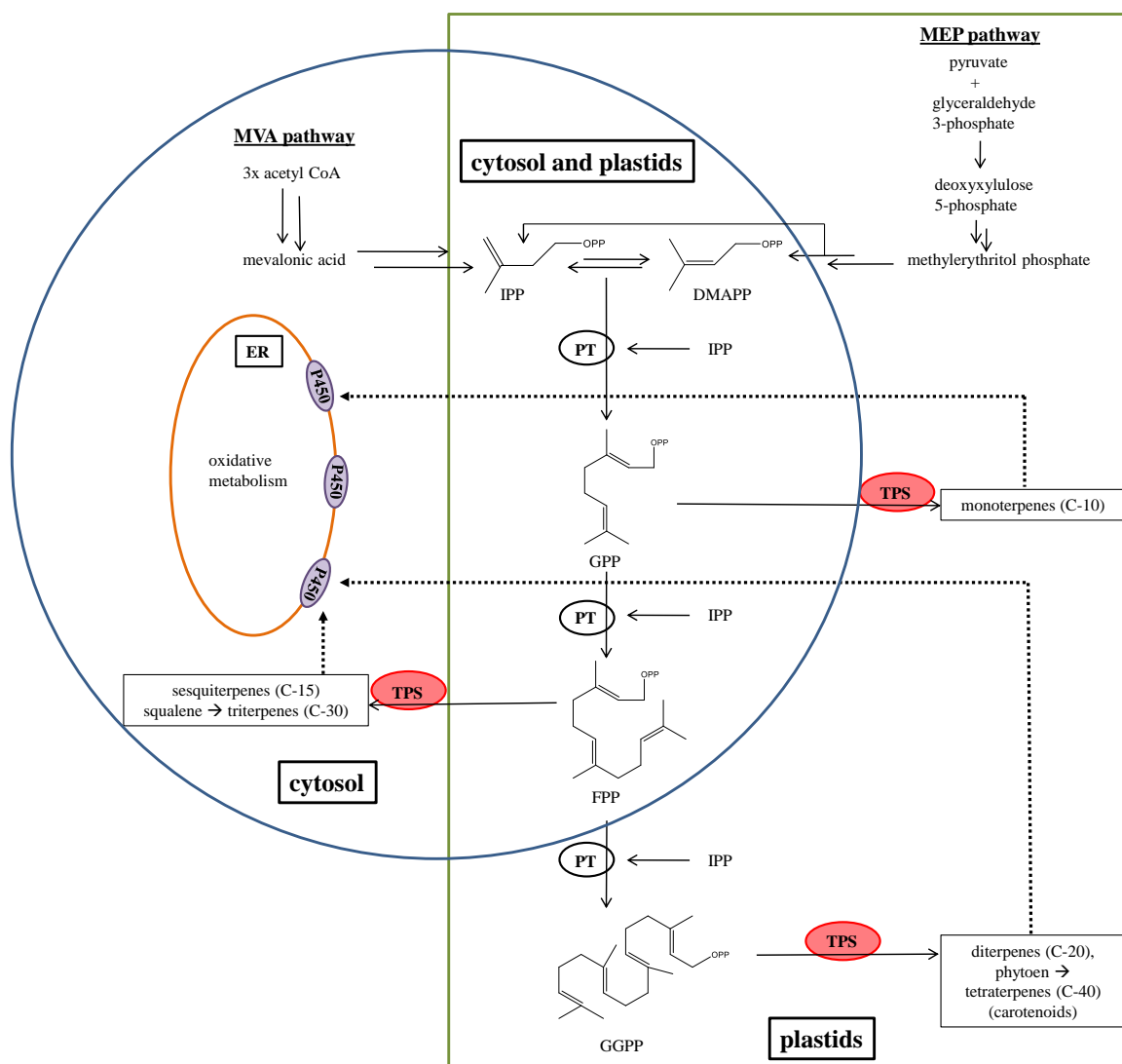
| Type           | C-5 Isoprene units  | Examples                                       |
|----------------|---------------------|--|
| Hemiterpenes   | 1 (C-5)             | isoprene                                       |
| Monoterpenes   | 2 (C-10)            | sabinene, $\alpha$ -pinene                     |
| Sesquiterpenes | 3 (C-15)            | $\beta$ -bisabolene,<br>$\beta$ -caryophyllene |
| Diterpenes     | 4 (C-20)            | casbene  |
| Triterpenes    | 6 (C-30)            | squalene                                       |
| Tetraterpenes  | 8 (C-40)            | $\beta$ -carotene                              |
| Polyterpenes   | $(C-5)_n$ , $n > 8$ | natural rubber                                 |

In plant cells, the biosynthesis of the central building blocks isopentenyl diphosphate (IPP) and its isomer dimethylallyl diphosphate (DMAPP) is ensued via two different pathways, which are located in the cytosol and the plastids, respectively. However, also an exchange of intermediates between the compartments has been observed (Eisenreich et al., 1998) (Fig. 4).

In the cytosol, the mevalonic acid pathway (MVA) starts with the fusion of three molecules acetyl CoA, resulting in the C-6-intermediate hydroxymethylglutaryl CoA (HMG-CoA), that is subsequently reduced to mevalonic acid. Following phosphorylation, decarboxylation, and dehydration reactions lead to the basic isoprene unit IPP (McGarvey and Croteau, 1995). The MVA pathway proceeds with the biosynthesis of sesqui- and triterpenes.

On the other hand, the plastidic 2-C-methyl-D-erythritol 4-phosphate pathway (MEP) is initiated by the formation of 1-deoxy-D-xylulose 5-phosphate from pyruvate and D-glyceraldehyde 3-phosphate. Via the intermediate 2-C-methyl-D-erythritol 4-phosphate, IPP and DMAPP are formed in equal amounts (Eisenreich et al., 1998; Lichtenthaler,

1999). The MEP pathway in the plastids proceeds with the biosynthesis of mono-, di-, and tetraterpenes.



**Fig. 4 Overview of the terpene biosynthesis in plants.** Shown is the intracellular compartmentation of the mevalonic acid pathway (MVA) in the cytosol and the 2-C-methyl-D-erythritol 4-phosphate pathway (MEP) in the plastids. Both pathways result in the formation of isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP), respectively. The terpenes are derived from these isoprene units, which are first condensated by prenyltransferases (PT). The resulting prenyl diphosphates are metabolized to the terpenes by the respective terpene synthases (TPS). Terpenes can be modified by P450 enzymes, which are located at the cytosolic surface of the endoplasmic reticulum (ER). Figure modified from (Trapp and Croteau, 2001) and (Staniek et al., 2013).

In both compartments, the activated isoprene unit IPP and its isomer DMAPP combine in a head-to-tail condensation to form the prenyl diphosphates, catalyzed by prenyltransferases. The electrophilic allylic CH<sub>2</sub> group (tail) of DMAPP hereby reacts with the nucleophilic methyl group (head) of IPP. This can happen multiple times, within each step the carbon skeleton is enlarged with C-5 units. The first product is the C-10 compound geranyl diphosphate (GPP), the precursor of all monoterpenes. A second condensation with



IPP leads to farnesyl diphosphate (FPP). This C-15 molecule is the precursor for sesquiterpenes. Consequentially, the addition of another molecule IPP results in the C-20 compound geranylgeranyl diphosphate (GGPP), which is metabolized to diterpenes. High-molecular compounds emerge after a tail-to-tail condensation of two molecules FPP from squalene (C-30 compound) or after a tail-to-tail condensation of two molecules GGPP from phytoen (C-40 compound) (Gershenzon, 1998).

The enzymes finally processing the prenyl diphosphates to terpenes are called terpene synthases. Each of these enzymes is named after the terpene class they produce, for example enzymes synthesizing monoterpenes from GPP are called monoterpene synthases and so on.

## **V. Terpenes and terpenoids are characterized by a great chemical diversity**

Although all terpenes are derived from a single isoprene precursor, they show an extraordinary structural diversity. Over 25,000 diverse terpene structures were found in plants (Connolly and Hill, 1991). The basic carbon skeleton of the terpene structures is formed by terpene synthases, the enzymes which metabolize the central prenyl diphosphates GPP, FPP, and GGPP. After formation of the basic carbon structure, additional modifications can follow. For example, various oxygen containing functional groups are found in terpenes (Gershenzon and Croteau, 1991). Such modifications can be mediated by cytochrome P450 monooxygenases (Fig. 4) or dehydrogenases (Ringer et al., 2005; Weitzel and Simonsen, 2013). Modified terpenes are often called terpenoids (Zwenger and Basu, 2008).

Regio- and stereochemical aspects also contribute strongly to the chemical diversity of terpenes. Terpene forming and modifying enzymes often exhibit rather strict stereospecificities, resulting in defined ratios of product stereoisomers, for example, in essential oils. The resulting monoterpene enantiomers can differ from each other in their function in plant-insect interactions (Phillips et al., 2003). The enantiomeric ratio of terpene compounds can also be used as a quality index to verify the authenticity, origin, and purity of particular commercial essential oils (Kreis and Mosandl, 1994; König et al., 1997). Furthermore, only small differences in the stereoisomeric composition can completely change the sensation of the respective essential oils. For example, in spearmint and caraway, (*S*)-limonene and (*R*)-limonene, respectively, are hydroxylized by two distinct limonene-6-hydroxylases to yield either (-)-(*trans*)-carveol or (+)-(*trans*)-carveol

(Bouwmeester et al., 1998; Lupien et al., 1999). The resulting (-)-carvon in spearmint and (+)-carvon in caraway can be distinguished by different smells, each of them determining the characteristic features of the respective plant's essential oil (Sell, 2009).

## **VI. Aims of the study**

Due to their pharmaceutical and commercial value, the biosynthesis of terpenes in medicinal plants has been the subject of intense research within the last decades. The focus of the present study was the investigation of the biosynthesis of cyclic oxygenated monoterpenes in *Thymus*, *Melaleuca*, and *Eucalyptus* species. Most cyclic monoterpenes bear one or more stereocenters and the stereochemical configuration of monoterpenes can affect their properties significantly. Also, oxygenation reactions underlie regioselective mechanisms.

Two enzyme classes are predominantly involved in the formation and modification of monoterpenes from a prenyl diphosphate precursor: terpene synthases and cytochrome P450 enzymes. Genes encoding these enzyme classes will be isolated from *Thymus*, *Melaleuca*, and *Eucalyptus* species, and biochemically characterized *in vitro* after heterologous expression. The characterization of the catalytic properties of the enzymes will have a particular emphasis on the stereochemical mechanism and control of regiospecific oxygenations. These investigations will help to understand how the formation of oxygenated monoterpenes is enzymatically controlled. The understanding of the biochemistry of this biosynthetic pathway can improve the options for plant metabolic engineering and breeding processes to increase the yield of valuable terpenoids.

# Chapter I<sup>1</sup>

## **Stereochemical characterization of monoterpene synthases producing cyclic oxygenated monoterpenes in essential oil plants**

---

<sup>1</sup> Parts of this chapter are published in:

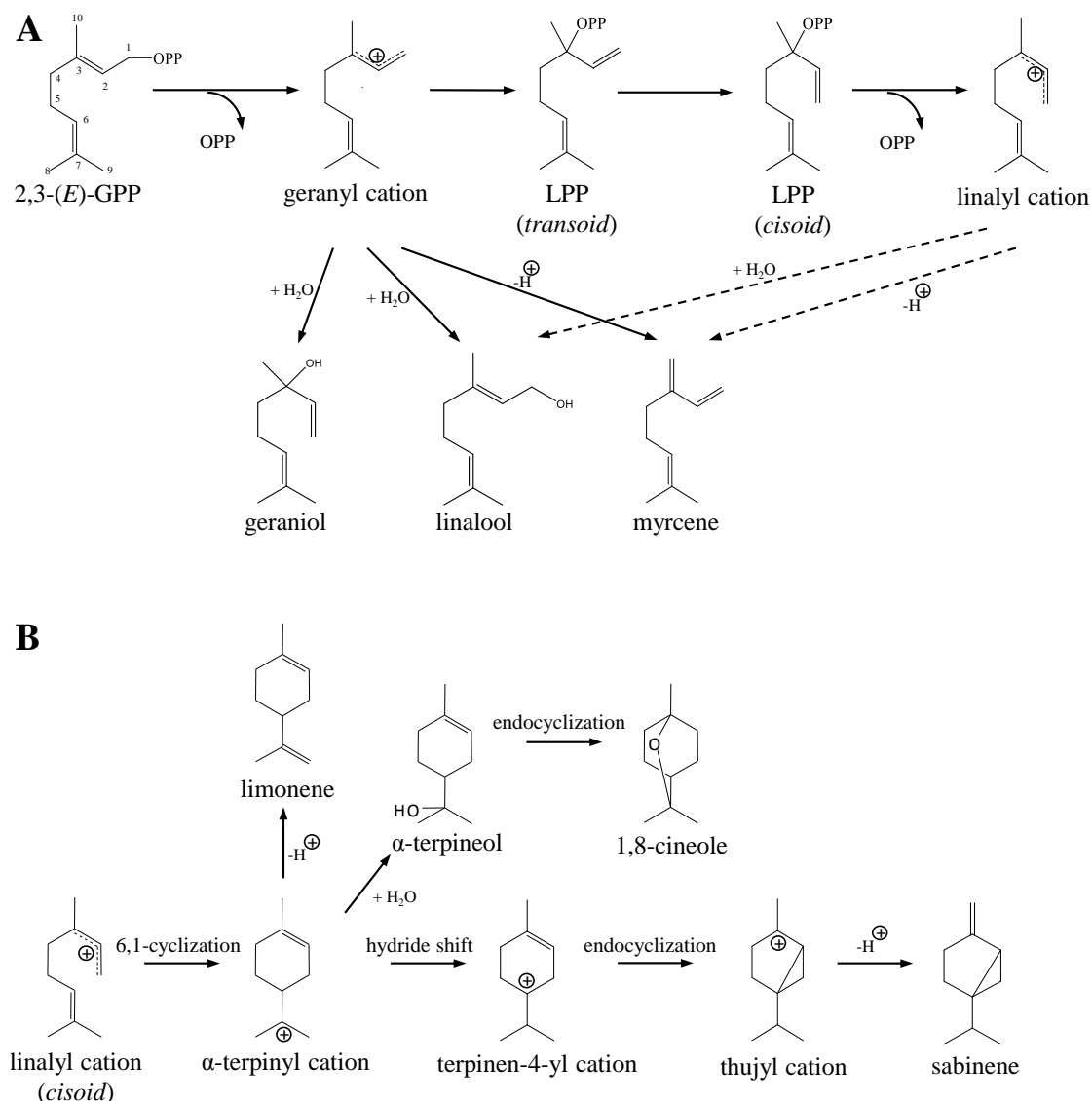
**Krause, S.T., Köllner T.G., Asbach, J., and Degenhardt, J.** (2012). Stereochemical mechanism of two sabinene hydrate synthases forming antipodal monoterpenes in thyme (*Thymus vulgaris*). Archives of Biochemistry and Biophysics **529**, 112-121.

## 1 Introduction

### 1.1 Monoterpene synthases are able to form an array of structural diverse compounds

The structural range of terpenes is formed by the enzyme class of terpene synthases, which convert the prenyl diphosphates to the basic terpene olefins or alcohols (Davis and Croteau, 2000). Most of the products contain a cyclic carbon structure, which is why the respective enzymes are also called “cyclases” (Croteau et al., 2000). A factor contributing to the terpene structure diversity is the unique feature of terpene synthases to form multiple products from one substrate. One terpene synthase can generate complex terpene blends with over 50 compounds (Steele et al., 1998). The terpene synthase reaction mechanism that realizes this variety of structures has been in the focus of intensive research in the last years (Degenhardt et al., 2009). The reaction pathway to the monoterpenes starts with the ionization of the 2,3-(*E*)-GPP substrate in the presence of divalent cations, mostly magnesium ( $Mg^{2+}$ ). A trinuclear magnesium cluster causes the elimination of the diphosphate moiety (Christianson, 2006), which leads to the geranyl cation, a carbocationic intermediate that is highly reactive and undergoes a series of further reactions. Acyclic monoterpenes arise directly after the ionization through water binding or deprotonation (Fig. 5 A). For example, water capture leads to geraniol or linalool, while deprotonation results in myrcene.

The (*E*)-configuration of GPP at the 2,3-double bond prevents the formation of cyclic products. Therefore, cyclization reactions require a rearrangement of GPP to linalyl diphosphate (LPP), enabled by the transfer of the diphosphate moiety to C-3. This allows the rotation of the new-risen 2,3-single bond and the conformation change of LPP from the *transoid* to the *cisoid* form. The following 6,1-cyclization of the *cisoid* LPP is initiated by another ionization leading to the linalyl cation. The cyclization results in the  $\alpha$ -terpinyl cation, the central intermediate and origin of all cyclic monoterpenes (Davis and Croteau, 2000). The array of diverse cyclic monoterpenes is then formed after hydride shifts and rearrangements. The reactions are terminated by deprotonation, additional endocyclizations, or water capture which leads to the formation of monoterpene alcohols (Davis and Croteau, 2000) (Fig. 5 B).



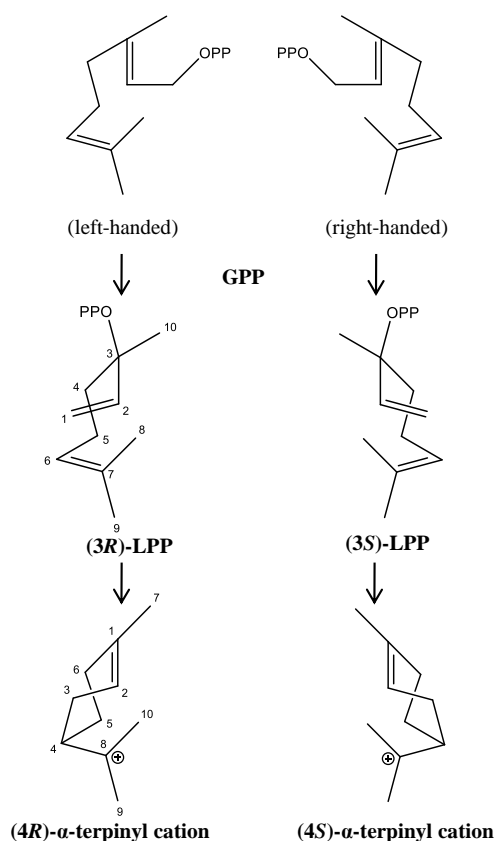
**Fig. 5 Carbocationic reaction mechanism of monoterpene synthases.** The formation of acyclic terpenes starts with the ionization of 2,3-(*E*)-geranyl diphosphate (GPP). Proton loss or water capture lead to the acyclic monoterpenes. Alternatively, these reactions can occur after formation of the linalyl cation. (A) The formation of cyclic monoterpenes requires the conformation change of linalyl diphosphate (LPP) from the *transoid* to the *cisoid* conformation. After ionization to the linalyl cation, a 6,1-cyclization results in the  $\alpha$ -terpinyl cation, the central precursor of all cyclic monoterpenes. Hydride shifts, endocyclizations, proton loss, and water capture lead to the cyclic terpene products. (B) Figure modified from (Degenhardt et al., 2009).

Terpene synthases reacting via this carbocationic intermediates are called class I terpene synthases on the basis of their reaction mechanism. Other terpene synthases are dependent on the protonation of GGPP, which subsequently initiates the cyclization of the substrate. A representative of these class II terpene synthases is the copalyl diphosphate synthase (Yamaguchi, 2008). However, class II terpene synthases are rare in the plant terpene synthase family (Chen et al., 2011).

## 1.2 Stereochemical aspects of terpene synthase reactions

Most monoterpene synthases have a stereospecific reaction mechanism and convert the achiral GPP substrate to compounds that contain one or more stereocenters. Studies on a fenchol synthase from fennel (Satterwhite et al., 1985), bornyl diphosphate synthases from tansy and sage (Croteau et al., 1986), or  $\alpha$ -pinene synthases of loblolly pine (Phillips et al., 2003) suggested that the stereospecificity of the enzyme reaction is defined in the early steps of the pathway by the initial folding of GPP in the active site pocket. The right-handed folding of GPP in the active site leads to the formation of (3*S*)-LPP and the left-handed folding leads to (3*R*)-LPP, the first chiral intermediates in the pathway. Finally, ionization and 6,1-cyclization of both enantiomers give the central cyclic intermediate, the  $\alpha$ -terpinyl cation, in the respective conformations (Fig. 6). At this stage, the configuration at C-4 is determined and fixed throughout the whole reaction to the end product. Thus, the GPP binding orientation in the monoterpene active center is crucial for the stereochemical configuration of the products.

Although a large number of plant terpene synthases has been identified to date, little is known about the structure-function relationships in the active center of these enzymes and their impact on product specificity and stereoselectivity (Degenhardt et al., 2009). In monoterpene synthases, site-directed mutagenesis and domain swapping have been utilized to identify structural elements that determine the product specificity (Peters and Croteau, 2003; Hyatt and Croteau, 2005; Kampranis et al., 2007). However, none of these studies reported on structural elements of monoterpene synthases which influence the stereochemical configuration of monoterpenes.



**Fig. 6** The binding conformation of GPP determines the stereochemical configuration of LPP, which is converted to the respective enantiomers of the  $\alpha$ -terpinyl cation intermediate.<sup>2</sup>

### 1.3 Structure-function relationships in monoterpene synthases

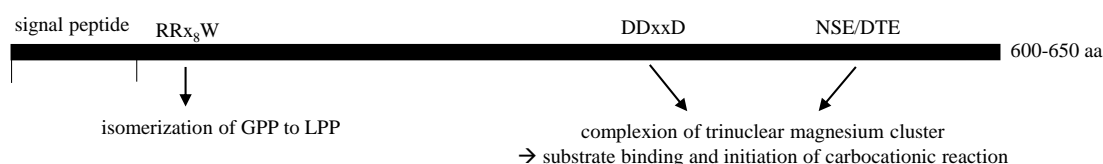
Monoterpene synthases show average lengths of 600-650 amino acids. Since the biosynthesis of monoterpenes is located in the plastids, they contain an N-terminal signal peptide which directs the enzymes to the plastids after expression (Turner et al., 1999b). These signal peptides are rich in serine and threonine residues, whereas acidic amino acids are rare. There have been, however, no typical sequence motifs identified (Bohlmann et al., 1997). To date, hundreds of terpene synthases have been cloned and characterized. These studies revealed some structural elements with defined general functions (Degenhardt et al., 2009).

Solved crystal structures of mono-, sesqui-, and diterpene synthases show a very similar overall fold in two structural domains. They can occur in monomeric or dimeric forms (Alonso et al., 1992; Whittington et al., 2002) and consist of  $\alpha$ -helices and short connecting loops (Starks et al., 1997; Whittington et al., 2002; Kampranis et al., 2007). The C-terminal

<sup>2</sup> The numbering of the carbon atoms changes after the ring formation. For reasons of comprehensibility, the numbering of the carbon atoms of the  $\alpha$ -terpinyl cation is maintained throughout this study for all cyclic monoterpenes.

domain bears the active center, a hydrophobic pocket built by six helices and covered by two loops on the protein surface (Degenhardt et al., 2009). The N-terminal domain resembles the structure of some glycosylhydrolases. Its function, however, remains unclear (Dudareva et al., 1996). It is proposed that this domain functions as a scaffold, stabilizing the correct fold of the C-terminal domain (Köllner et al., 2004). The N-terminus is furthermore assumed to cover the active center after substrate binding to shield it from the surrounding aqueous milieu (Starks et al., 1997).

Some sequence motifs highly conserved among terpene synthases and their catalytic functions have been identified (Fig. 7) (Starks et al., 1997; Degenhardt et al., 2009). The aspartate-rich DDxxD motif is located at the entrance of the catalytic pocket and is involved in the binding of magnesium ions. Thus, this motif is essential for the correct substrate binding and the formation of the carbocation intermediate through interaction with the diphosphate group of the substrate (Cane et al., 1996). Another magnesium-binding motif has been located at the opposite site of the catalytic center. Because of its consensus sequence (L,V)(V,L,A)-(N,D)D(L,I,V)X(S,T)XXXE, it is designated as NSE/DTE motif. It seems to have evolved from another aspartate-rich motif, which is conserved in prenyltransferases (Whittington et al., 2002). Together, the DDxxD and the NSE/DTE motif bind three magnesium ions which interact with the diphosphate moiety of the substrate. This causes the initial substrate ionization and rearrangement of the diphosphate group (Christianson, 2006). Another motif, predominantly found in monoterpene synthases, is the RR<sub>x</sub>W motif, which is located about 60 amino acids from the N-terminus. This motif is important for the isomerization of the GPP substrate to the LPP intermediate. Deletion studies showed that monoterpene synthases lacking this motif do not accept GPP as substrate anymore, but are able to metabolize LPP (Bohlmann et al., 1998; Williams et al., 1998). Hence, this motif seems to be crucial for the formation of cyclic monoterpenes.



**Fig. 7 Scheme of the universal monoterpene synthase sequence motifs and their respective functions.**



## 2 Materials and Methods

### 2.1 Plant material

The chemotypes of thyme (*Thymus vulgaris*) were originally collected in Southern France at CNRS, Montpellier, France (Thompson et al., 2003). Plants were grown in the greenhouse under following conditions: temperature day (13 h light) 20-22 °C, temperature night 18-20 °C, humidity 55 %, luminosity approximately 320  $\mu\text{mol}$  photosynthetically active radiation.

*Salvia fruticosa* plants were purchased from a local garden center.

Leaves of *Melaleuca linariifolia* and *Melaleuca trichostachya* were collected in Bulahdelah at the East Coast of New South Wales, Australia.

### 2.2 Microbiological methods

#### 2.2.1 Bacteria strains

For the multiplication of plasmid DNA and heterologous expression of monoterpene synthases, the *Escherichia coli* (*E. coli*) strain TOP10 (Invitrogen, Carlsbad, USA) was used.

Genotype: F<sup>-</sup> *mcrA*  $\Delta$ (*mrr-hsdRMS-mcrBC*)  $\phi$ 80*lacZ* $\Delta$ M15  $\Delta$ *lacX74* *recA1* *araD139*  $\Delta$ (*araleu*) 7697 *galU* *galK* *rpsL* (Str<sup>R</sup>) *endA1* *nupG*.

#### 2.2.2 Cultivation of *E. coli*

*E. coli* cells were cultivated in liquid Luria-Bertani (LB) medium (20 g LB-Medium (Roth, Karlsruhe, Germany) in 1 l H<sub>2</sub>O) at 37 °C and 220 rpm shaking (“CERTOMAT IS”, Sartorius, Göttingen, Germany) or on LB agar plates (25 g LB-Agar (Roth, Karlsruhe, Germany) in 1 l H<sub>2</sub>O) at 37 °C in an incubator (“Heraeus, Function Line, Type T6”, Thermo Scientific, Waltham, MA, USA).

For the selection of transformed *E. coli*, the media were supplied with 50  $\mu\text{g/ml}$  kanamycin or 100  $\mu\text{g/ml}$  ampicillin (Roth, Karlsruhe, Germany).

For the conservation of transformed *E. coli* cultures, 500  $\mu\text{l}$  culture were supplied with 150  $\mu\text{l}$  glycerol and stored at -80 °C.

### **2.2.3 Transformation of *E. coli***

Chemically competent *E. coli* TOP10 cells (Invitrogen, Carlsbad, USA) were transformed with the heat shock method. Therefore, 1-10  $\mu$ l (50-500 ng) of the ligation mix or plasmid were mixed with 20-100  $\mu$ l of *E. coli* TOP10 cells. After incubation on ice for 15 min, the cells were placed in a water bath at 42 °C for 45 s and immediately cooled on ice for 2 min. Then, 200  $\mu$ l SOC medium (Invitrogen, Carlsbad, USA) were added and the bacteria were incubated at 37 °C and 220 rpm shaking for 1 h. 50-200  $\mu$ l of the transformation mix were plated on LB agar plates with the respective selection medium.

## **2.3 Nucleic acid techniques**

### **2.3.1 RNA extraction from leaf material and cDNA/RACE cDNA synthesis**

From *Thymus vulgaris*, total RNA was extracted from 100 mg homogenized young leaves of the respective chemotypes, pooled from five plants. The RNA was extracted with the “RNeasy Plant Mini Kit” (Qiagen, Hilden, Germany) according to the manufacturer’s instructions. To remove residual genomic DNA, the RNA was treated with RQ1 RNase-free DNase (Promega GmbH, Mannheim, Germany).

For cDNA synthesis, the “Fermentas First Strand cDNA Kit” (Fermentas, St. Leon-Rot, Germany) was used according to the manufacturer’s instructions.

For the completion of partial open reading frames (ORF), 3’- and 5’-RACE cDNA libraries were synthesized with specific adaptor sequences. The libraries were constructed with the “SMARTer RACE cDNA Amplification Kit” according to the manufacturer’s instructions (Clontech Laboratories Inc., Mountain View, CA, USA).

From *Melaleuca* species, total RNA was extracted from 50 mg homogenized young leaves from one plant. The RNA was extracted with the “Spectrum Plant Total RNA Kit” (Sigma-Aldrich, St. Louis, MO, USA) according to the manufacturer’s instructions.

### **2.3.2 Amplification of DNA**

The amplification of DNA was conducted with the polymerase chain reaction (PCR). Various Taq polymerases were used: the Go Taq Polymerase (5 U/ $\mu$ l; Promega Corporation, Madison, WI, USA) for standard DNA amplifications, the Advantage 2 Polymerase Mix (Clontech Laboratories Inc., Mountain View, CA, USA) for RACE PCRs

and the Pfu DNA Polymerase (2.5 U/ $\mu$ l; Stratagene, La Jolla, CA, USA) as a proof-reading polymerase for *in vitro* mutagenesis reactions.

The general composition of the reaction was as follows:

|                        |   |
|------------------------|---|
| 0.5-4 $\mu$ l          | DNA template (5-200 ng)                       |
| 2.5-10 $\mu$ l         | 5-10x polymerase buffer                       |
| 0.25-1 $\mu$ l         | DNA polymerase                                |
| 0.5-2 $\mu$ l          | deoxynucleoside triphosphates (dNTPs) (10 mM) |
| 0.5-2 $\mu$ l          | oligonucleotide (primer) forward (10 $\mu$ M) |
| 0.5-2 $\mu$ l          | oligonucleotide (primer) reverse (10 $\mu$ M) |
| $\Sigma$ 25-50 $\mu$ l | H <sub>2</sub> O                              |

The reaction was conducted in a thermocycler. A “Thermocycler MWG-Biotech primus 96 plus” (MWG-Biotech, Ebersbach, Germany) or a “peqSTAR Universal 96” (PeqLab, Erlangen, Germany) was programmed as follows:

|                  |            |   |
|------------------|------------|---|
| 95°C             | 2 min      | Denaturation of DNA template                      |
| ┌ 95°C           | 30 s       | Denaturation of DNA template                      |
| 20-35x   50-70°C | 30 s       | Primer binding (Annealing)                        |
| └ 68-72°C        | 1.5-10 min | Synthesis of complementary DNA strand (Extension) |
| 68-72°C          | 5 min      | Completion of DNA synthesis                       |

Detailed compositions and temperature protocols for the isolation of terpene synthase genes or cytochrome P450 monooxygenase genes are given in section 2.4 (Chapter I) and section 2.2 (Chapter II).

### 2.3.3 Transcript quantification by quantitative real-time PCR

Transcript quantification was performed with the “CFX96 Real Time System” (BioRAD, Munich, Germany). The components of the quantitative real-time PCR (qRT PCR) were: 10  $\mu$ l Maxima SYBR Green/ROX qPCR Master Mix (Fermentas, St. Leon-Rot, Germany), 0.5  $\mu$ l Primer 1 and 0.5  $\mu$ l Primer 2 (Table S1), 5  $\mu$ l cDNA template (1:5 diluted) and 4  $\mu$ l PCR-grade water. Controls included non-template controls (water template). PCR thermocycles were run as follows: 10 min at 95 °C, 40 cycles of 30 s at 95 °C, 30 s at 62 °C and 40 s at 72 °C. Fluorescence was determined after each extension phase. After each run,

a melting curve analysis from 60 °C to 95 °C was performed. The amplification products were cloned and sequenced to validate the primer specificity. The amplification plots were analyzed with the “BioRAD CFX manager” to receive Ct values. For relative qRT PCR, 18S ribosomal RNA was employed as housekeeping gene. Relative quantification of the gene copy number in each cDNA sample was conducted using a standard curve. The standard curve was generated with cDNA containing the respective genes, therefore, a dilution series from 3 to 1/27-fold was made.

### **2.3.4 Electrophoresis of DNA and RNA**

PCR products or linearized plasmids were separated according to their fragment size in 1.5 % agarose gels (1.5 % agarose (w/v), (Roth, Karlsruhe, Germany); 1.27 µM ethidium bromide; 0.5x TAE buffer: 20 mM Tris base, 9.5 mM acetic acid, 0.5 mM EDTA, with HCl pH 7.6 - 7.8) through electrophoresis in 0.5x TAE buffer as electrophoresis buffer. Therefore, the DNA was mixed with 1/5 volume of 5x loading buffer (50 % glycerin (v/v), 0.05 % bromphenol blue (w/v), 100 mM EDTA) and placed in a gel slot. For the determination of the fragment size, a 1 kb DNA marker was used (Roth, Karlsruhe, Germany). In general, the duration of electrophoresis in the migration chamber “Mupid-One” (Eurogentec, Cologne, Germany) at 100 V was 20 min.

For the electrophoresis of RNA, the same procedure was conducted, but the chamber was cleaned thoroughly with detergents in advance, in order to avoid RNA degradation during the electrophoresis through RNase contamination.

### **2.3.5 Isolation of DNA fragments from agarose gels**

DNA bands were cut from an agarose gel with a scalpel under UV light (260 nm). For the extraction and purification of DNA from the agarose gel, the “QIA quick Gel Extraction Kit” (QIAGEN, Hilden, Germany) or the “Nucleo Spin Extract Kit” (Macherey-Nagel, Düren, Germany) was used according to the manufacturer’s instructions. The DNA was eluted in 10 mM Tris-HCl buffer (pH 8.5).

### **2.3.6 Restriction of DNA**

The restriction of DNA fragments or plasmids was conducted using specific restriction enzymes from New England Biolabs (Ipswich, MA, USA) or Fermentas (Thermo Fisher

Scientific Inc., Waltham, MA, USA). The reaction mix contained restriction enzyme (1-10 U), 10x enzyme buffer and DNA (1-10 mg) in a total volume of 10-30  $\mu$ l. The restriction mix was incubated at the optimum temperature for the respective restriction enzyme (37-50 °C) for 1-3 h.

### **2.3.7 Cloning of DNA fragments**

In general, DNA fragments were amplified by PCR with primers containing specific restriction sites. The PCR products were then digested with the respective restriction enzymes to obtain sticky ends (section 2.3.6) and ligated into plasmids which were linearized by digestion with the same restriction enzymes to produce compatible ends.

For the ligation, a T4 DNA ligase (Thermo Fisher Scientific Inc., Waltham, MA, USA) and the supplied 10x ligase buffer with ATP was used. The ligation mix contained vector and DNA fragment in a 1:5 ratio and 1 U ligase. It was incubated at 10-37 °C for 1-12 h.

The pCR4-TOPO sequencing vector (Invitrogen, Carlsbad, CA, USA) was used without any enzymes. Here, the Topoisomerase I, covalently bound to the vector, worked as ligase for fragments with a 3'-A overhang.

### **2.3.8 Multiplication of plasmid DNA**

Plasmids carrying the desired DNA fragment were transformed into *E. coli* cells as described in section 2.2.3.

5 ml LB medium were inoculated with a positive transformed *E. coli* clone and grown overnight at 37 °C and 220 rpm shaking.

Plasmid DNA was isolated from *E. coli* cells by using the “NucleoSpin Plasmid Kit” (Macherey-Nagel, Düren, Germany) according to the manufacturer’s instructions. The plasmid DNA was eluted and stored in 5 mM Tris-HCl buffer.

### **2.3.9 Sequencing of DNA**

Plasmid DNA samples (50-100 ng/ $\mu$ l) were sent to Eurofins MWG Operon (Martinsried, Germany) for sequencing by the cycle sequencing technology, a modification of the chain termination method after Sanger. The reaction was analyzed on an “ABI 3730XL” sequencer (Applied Biosystems, Waltham, MA, USA).

### 2.3.10 Analysis of DNA and protein sequences

DNA and protein sequences were analyzed with the software “DNA Star Lasergene” (DNASTAR Inc., Madison, USA) and the software “BioEdit” (Hall, T.A. 2004; Ibis Biosciences, Carlsbad, CA, USA). For signal peptide prediction, the following databases were used: ChloroP, SignalP, and TargetP of the CBS Prediction Server (<http://www.cbs.dtu.dk/services/>). BLAST searches of isolated sequences were performed via the NCBI sequence database (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>).

### 2.3.11 Quantification of DNA and RNA

DNA and RNA were quantified using the “NanoQuant infinite M200” (Tecan, Männedorf, Switzerland) by using the „NanoQuant Plate“ (a quartz sample plate). It measures the absorbance of the sample at 260 nm.

## 2.4 Isolation of terpene synthases from *Thymus* and *Melaleuca* species

In general: Terpene synthases isolated in this study are named as follows: Genes begin with the abbreviation of the species the genes were isolated from (e.g. *Tv* for *Thymus vulgaris*), followed by *tps* for *terpene synthase*. The gene numbers are assigned in order of isolation. The corresponding enzymes are designated with capital letters.

### *Thymus vulgaris*:

From the (*E*)-sabinene hydrate chemotype, 5' and 3'- RACE libraries were generated. For the first PCRs, degenerate primers based on sequences of terpene synthase genes from other *Lamiaceae* were used to obtain partial 3'- sequences. The components of the PCRs were: 0.8 µl Advantage Taq DNA Polymerase Mix (5 U/µl), 4 µl Advantage Taq PCR buffer, 1 µl dNTPs (10 mM each), 1 µl universal primer mix, 1 µl degenerate forward primer (10 µM) (Table S1), 3 µl 3'-RACE cDNA, and PCR grade water added to a final volume of 40 µl. The PCRs were conducted with an initial denaturation at 95 °C for 2 min, 30-35 cycles of denaturation at 95 °C for 30 s, annealing ranging from 48 to 60 °C for 30 s, extension at 68 °C for 2 min, and a final step at 68 °C for 5 min. The PCR fragments were cloned into the pCR4-TOPO vector (TOPO TA cloning kit for sequencing, Invitrogen, Carlsbad, CA, USA) and subsequently sequenced. The obtained sequences were compared by BLAST searches via the NCBI sequence database and showed similarity to monoterpene synthases from other plants. The sequence fragments were used to design primers for the

isolation of the 5'- ends of the full length genes. This time, the components of the PCRs were: 1  $\mu$ l Advantage Taq DNA Polymerase Mix (5 U/ $\mu$ l), 5  $\mu$ l Advantage Taq PCR buffer, 1  $\mu$ l dNTPs (10 mM each), 5  $\mu$ l universal primer mix and 1  $\mu$ l gene specific reverse primer (10  $\mu$ M) (Table S1), 2.5  $\mu$ l 5'-RACE cDNA and PCR grade water added to a final volume of 50  $\mu$ l. PCR thermocycles were run as follows: initial denaturation at 95 °C for 2 min, 30-35 cycles of denaturation at 94 °C for 30 s, annealing ranging from 64 to 68 °C for 30 s, extension at 72 °C for 1.5 min, and a final step at 72 °C for 5 min. PCR fragments were cloned into the pCR4-TOPO vector and sequenced. All 5'- end and 3'- end sequences were assembled with the "SeqMan" program (Lasergene DNASTar V5.05, Madison, WI, USA). This assembly revealed two complete open reading frames called *Tvtps6* and *Tvtps7*.

From the  $\alpha$ -terpineol chemotype, cDNA was synthesized. A primer pair, originally designed for subcloning *Tvtps6* and *Tvtps7* into the pASK-IBA37+ expression vector (IBAGmbH, Göttingen, Germany), was used in a PCR (Table S1). The components and conditions were: 0.125  $\mu$ l Go Taq DNA Polymerase, 5  $\mu$ l Go Taq buffer, 0.5  $\mu$ l dNTPs (10 mM each), 1  $\mu$ l forward primer (10  $\mu$ M) and 1  $\mu$ l reverse primer (10  $\mu$ M) (Table S1), 1  $\mu$ l cDNA, and PCR grade water added to a final volume of 25  $\mu$ l. PCR thermocycles were run as follows: initial denaturation at 95 °C for 2 min, 35 cycles of denaturation at 95 °C for 30 s, annealing at 53 °C for 30 s, extension at 72 °C for 1.5 min, and a final step at 72 °C for 5 min. The amplified fragment was cloned into the pASK-IBA37+ vector and sequenced. The obtained sequence was called *Tvtps5*.

*Melaleuca trichostachya* and *Melaleuca linariifolia*:

From both species, 5'-RACE cDNA libraries were generated. Primer pairs designed for subcloning monoterpene synthases from *Melaleuca alternifolia* into the pASK-IBA37+ expression vector (Table S1) were used in PCRs. The components of the PCRs were: 1  $\mu$ l Advantage Taq DNA Polymerase Mix (5 U/ $\mu$ l), 5  $\mu$ l Advantage Taq PCR buffer, 1  $\mu$ l dNTPs (10 mM each), 2  $\mu$ l forward primer (10  $\mu$ M) and 2  $\mu$ l reverse primer (10  $\mu$ M), 4  $\mu$ l 5'- RACE cDNA and PCR grade water added to a final volume of 50  $\mu$ l. PCR thermocycles were run as follows: initial denaturation at 95 °C for 2 min, 35 cycles of denaturation at 95 °C for 30 s, annealing at 53 °C for 30 s, extension at 72 °C for 1.5 min, and a final step at 72 °C for 5 min. The PCR fragments were cloned into the pCR4-TOPO vector and sequenced. The obtained sequences were called *Mttps1*, *Mttps2*, *Mttps3*, and *Mltps1*.

## 2.5 *In vitro* mutagenesis of terpene synthase open reading frames

For site-directed mutagenesis, the “Stratagene QuickChange method” (Stratagene, La Jolla, CA, USA) was used. The PCR based mutagenesis protocol was performed with the ORFs of *Tvtps5*, *Tvtps6*, and *Tvtps7* cloned into the expression vector pASK-IBA37+. Primers containing the point mutations were designed to bind at opposite strands of the vector (Table S1). The PCR resulted in the amplification of the complete vector containing the desired mutation. Therefore, the PCR elongation phase was prolonged to up to 10 min (in 18 cycles). After vector amplification, the PCR product was digested with the restriction enzyme *DpnI* (10 U) for 1 h to ensure the destruction of the methylated parental vector used as template. The expression vector containing the mutated ORF was subsequently transformed into *E. coli* TOP10 cells. The constructs were sequenced before expression to confirm positive mutagenesis.

## 2.6 Dendrogram analysis

DNA sequence alignments (Muscle codon alignment) and dendrogram analyses with the Maximum Likelihood method were performed using the program “MEGA” version 6 (Tamura et al., 2013).

The analyses were based on the General Time Reversible model (Nei and Kumar, 2000), which was the best fitting model according to “MEGA”.

## 2.7 Terpene synthase overexpression in *E. coli* and protein extraction

The ORFs of the terpene synthases with 5'-signal peptide truncations were cloned into the bacterial expression vector pASK-IBA37+ (IBAGmbH, Göttingen, Germany). This vector contains a 6x His-tag and a *tet*-promotor. The genes were amplified with primers created with the program “Primer D'Signer” (IBA GmbH, Göttingen, Germany) (Table S1) with the Go Taq DNA Polymerase. The amplification products were digested and cloned into the pASK-IBA37+ expression vector according to the manufacturer's instructions. The expression constructs were verified by sequencing and transformed into *E. coli* TOP10 cells (Invitrogen, Carlsbad, CA, USA). A starter culture of 5 ml LB medium with 100 µg/ml ampicillin was grown overnight at 37 °C. 3 ml of the starter culture were used to inoculate 100 ml of LB medium with 100 µg/ml ampicillin and the bacteria were grown at 37 °C and 220 rpm to an OD of 0.6. The terpene synthase expression was induced by the addition of anhydrotetracycline (final concentration: 200 µg/l). The cultures were shaken for 20 h at



18 °C. The cells were harvested by centrifugation for 10 min at 5,000 x g and 4 °C. The pellet was resuspended in 3 ml extraction buffer (50 mM Tris-HCl pH 7.5, 10 % glycerol, 5 mM MgCl<sub>2</sub>, 5 mM DTT, 5 mM sodium ascorbate pH 7.0, 0.5 mM PMSF). After disruption by sonication (3 x 30 s at 50 % power, “Branson Sonifier 250”, Dietzenbach, Germany), the cell fragments were removed by centrifugation for 20 min at 20,800 x g and 4 °C. The crude protein extract was transferred into assay buffer (10 mM Tris-HCl pH 7.5, 10 % glycerol, 1 mM DTT) by using 10DG columns (Bio-Rad Laboratories, Hercules, CA, USA) according to the manufacturer’s instructions.

## 2.8 Purification of terpene synthases

For enzyme purification, 4 ml crude protein extract was mixed with 5 ml of Profinity IMAC Ni-charged resin (Bio-Rad Laboratories, Hercules, CA, USA). After incubation for 1 h at 4 °C and shaking at 170 rpm, the resin was transferred to a Poly-Prep Chromatography column (Bio-Rad Laboratories, Hercules, CA, USA). After two washing steps with washing buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 20 mM imidazole pH 8.0, 10 % glycerol), the enzyme was eluted with elution buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 250 mM imidazole pH 8.0, 10 % glycerol). The enzyme concentration was determined according to a method described by Gill and von Hippel (Gill and Von Hippel, 1989).

## 2.9 Enzyme assays

Terpene synthase activity assays were performed with 30 µl crude enzyme extract and 70 µl reaction mix (10 mM Tris-HCl pH 7.5, 10 % glycerol, 1 mM DTT, 60 µM GPP (Echelon Research Laboratories, Salt Lake City, USA), and 10 mM MgCl<sub>2</sub>). Enzyme products were collected by a solid phase microextraction fiber (SPME) (100 µm polydimethylsiloxane; Supelco, Belafonte, PA, USA). The fiber was exposed in the headspace of the assay mixture for 45 min at 35 °C in a water bath.

For quantitative analyses, 400 µl of crude enzyme extract and 600 µl reaction mix (10 mM Tris-HCl pH 7.5, 10 % glycerol, 1 mM DTT, 60 µM GPP (Echelon Research Laboratories, Salt Lake City, USA), 10 mM MgCl<sub>2</sub>) were mixed and overlaid with 300 µl hexane, supplied with 10 µg/ml nonyl acetate as internal standard. The assay was incubated at 37 °C for 3 h or 20 min for kinetic analyses, respectively. Terpene products were transferred into the organic solvent by shaking intensively for 2 min. The mixture was frozen in liquid

nitrogen, and the hexane was extracted with a glass pipette. The hexane extract was subsequently used for analysis.

### **2.10 Terpene extraction from thyme leaves**

Young leaves from three plants of the (*E*)-sabinene hydrate chemotype or the  $\alpha$ -terpineol chemotype of *Thymus vulgaris* were combined and ground to a fine powder with mortar and pestle. The powder (50-100 mg) was soaked in 600  $\mu$ l hexane (supplied with 10  $\mu$ g/ml nonyl acetate as internal standard) and incubated for 1 h at room temperature. Alternatively, a polydimethylsiloxane SPME fiber (Supelco, Belafonte, PA, USA) was exposed to the leaf volatiles for 5 s.

### **2.11 GC-MS analysis of leaf terpenes and terpene synthase assay products**

The terpenes were identified with a gas chromatograph (GC) (GC-2010, Shimadzu, Duisburg, Germany) coupled to a mass spectrometer (MS) or flame ionization detector (FID) (GCMS-QP 2010 Plus, Shimadzu, Duisburg, Germany). 1  $\mu$ l of hexane extract was injected at 220 °C injector temperature. Alternatively, a SPME fiber was introduced into the injector. All volatiles were separated on an EC-5 column (5 % phenyl)- 95 % methylpolysiloxane) (Grace, Deerfield, IL, USA) and identified by comparison to authentic standards or with the Shimadzu software “GCMS Postrun Analysis“ with the mass spectra libraries “Wiley8” (Hewlett&Packard, Palo Alto, USA) and “Adams” (Adams, 2007).

For qualitative GC-MS analyses, the following GC program was used: 50 °C for 2 min, first ramp 7 °C/min to 150 °C, second ramp 100 °C/min to 300 °C, final 2 min hold. GC-MS carrier gas: hydrogen (1 ml/min).

For quantitative FID analyses, the following temperature program was used: 40 °C for 3 min, first ramp 6 °C/min to 280 °C, second ramp 100 °C/min to 300 °C, final 2 min hold. For the analysis of chiral compounds, the Rt-bDex sm column (2,3-di-O-methyl-6-O-tert-butyl dimethylsilyl beta cyclodextrin added into 14 % cyanopropylphenyl/86 % dimethyl polysiloxane) (Restek, Bad Homburg, Germany) was used with the following conditions: GC-program: 50 °C for 1 min, first ramp 2 °C/min to 170 °C, second ramp 100 °C/min to 220 °C, final 2 min hold. GC-MS carrier gas: hydrogen (1 ml/min). Alternative GC program optimized for chiral separations: 40 °C for 1 min, first ramp 1 °C/min to 120 °C, second ramp 100 °C/min to 220 °C, final 2 min hold, column flow: 2 ml/min.

## 2.12 Stereospecific synthesis of LPP

(3*S*)- and (3*R*)-LPP were synthesized according to a method described by Keller and Thompson (Keller and Thompson, 1993).

0.2 mmol (3*R*)-linalool (Sigma-Aldrich Co., St. Louis, MO, USA) or (3*S*)-linalool were resolved in 500 µl trichloroacetonitrile (Sigma-Aldrich, St. Louis, MO, USA). Due to the lack of commercial available (3*S*)-linalool, coriander oil (SAFC Supply solutions, St. Louis, MO, USA) was used for phosphorylation. This oil consists mostly of (3*S*)-linalool, with a 10:1 enantiomeric ratio (Özek et al., 2010).

Fresh TEAP solution was prepared. Therefore, 2.5 ml concentrated phosphoric acid was mixed with 9.5 ml acetonitrile (Roth, Karlsruhe, Germany) (solution A) and 5.5 ml triethyl amine was mixed with 5.0 ml acetonitrile (solution B). The final TEAP solution was made by mixing 0.91 ml of solution A with 1.5 ml of solution B. 500 µl of this TEAP solution was added to the dissolved linalool and stirred for 5 min at 37 °C until it was a clear yellow solution. 500 µl TEAP solution were added and stirred for 5 min at 37 °C for two further times. The final volume was 2 ml. This reaction mixture was applied on a preparative thin-layer chromatography (TLC) plate in a thin line (Analtech, Inc., Newark, DE, USA; silica gel 60A; layer thickness 1000 µm; 20x20 cm). The TLC was performed overnight, the mobile phase consisted of 120 ml isopropanol, 60 ml ammonia, and 20ml water. The migration distance of LPP was supposed to be 6-9 cm, so this band was scraped off and extracted two times with 10 mM ammonium hydrogen carbonate solution. The extracts were centrifuged and the solid components were discarded. The fresh LPP solution was used immediately for terpene synthase assays.

## 2.13 Determination of $K_m$ values of terpene synthases

30 µl crude enzyme extract was incubated with 5 mM  $MgCl_2$  and 5 µM  $^3H$ -GPP between 5-30 min to determine the linear phase of the reactions. For the determination of the substrate  $K_m$  values, the enzymes were incubated with 5 mM magnesium and  $^3H$ -GPP in a range of 1-30 µM. All assays were overlaid with 1 ml pentane and incubated at 30 °C for 7.5 or 10 min, depending on the linear phase of TvTPS6 and TvTPS7. The assays were stopped by shaking at 1.400 rpm for 2 min to partition terpene volatiles in the solvent phase. 500 µl pentane were mixed with 2 ml of scintillation cocktail (RotiSzint2200, Roth, Karlsruhe, Germany), and counts per minute were measured in a scintillator (“LS 6500”,

Beckman Coulter Inc., Krefeld, Germany). All assays were performed in triplicate. The  $K_m$  values were determined by using the Lineweaver-Burke method.

#### **2.14 Modeling of the terpene synthase active site and docking studies**

Models of the three-dimensional structure of TvTPS5, TvTPS6, and TvTPS7 were generated using the Swiss Model Server (<http://www.expasy.org>) (Schwede et al., 2003; Arnold et al., 2006). For modeling, the respective amino acid sequences were fitted to the template structure of the bornyl diphosphate synthase (PDB code: 1n1zA) (Whittington et al., 2002). For docking studies, hydrogen atoms were added to the structures of TvTPS6 and TvTPS7 by using the program “AutoDock Tools” (Sanner, 1999). Energy-minimized ligand structures were generated with the software “ChemDraw” and “Chem3D” (CambridgeSoft, Cambridge, USA). Docking of GPP into the model of the TvTPS6 and TvTPS7 active site cavity was performed with the program “AutoDock Vina” (Trott and Olson, 2010). The resulting models were visualized with the program “PyMOL” (DeLano, 2002).

### 3 Results

#### 3.1 Isolation and characterization of two sabinene hydrate synthases from thyme (*Thymus vulgaris*)

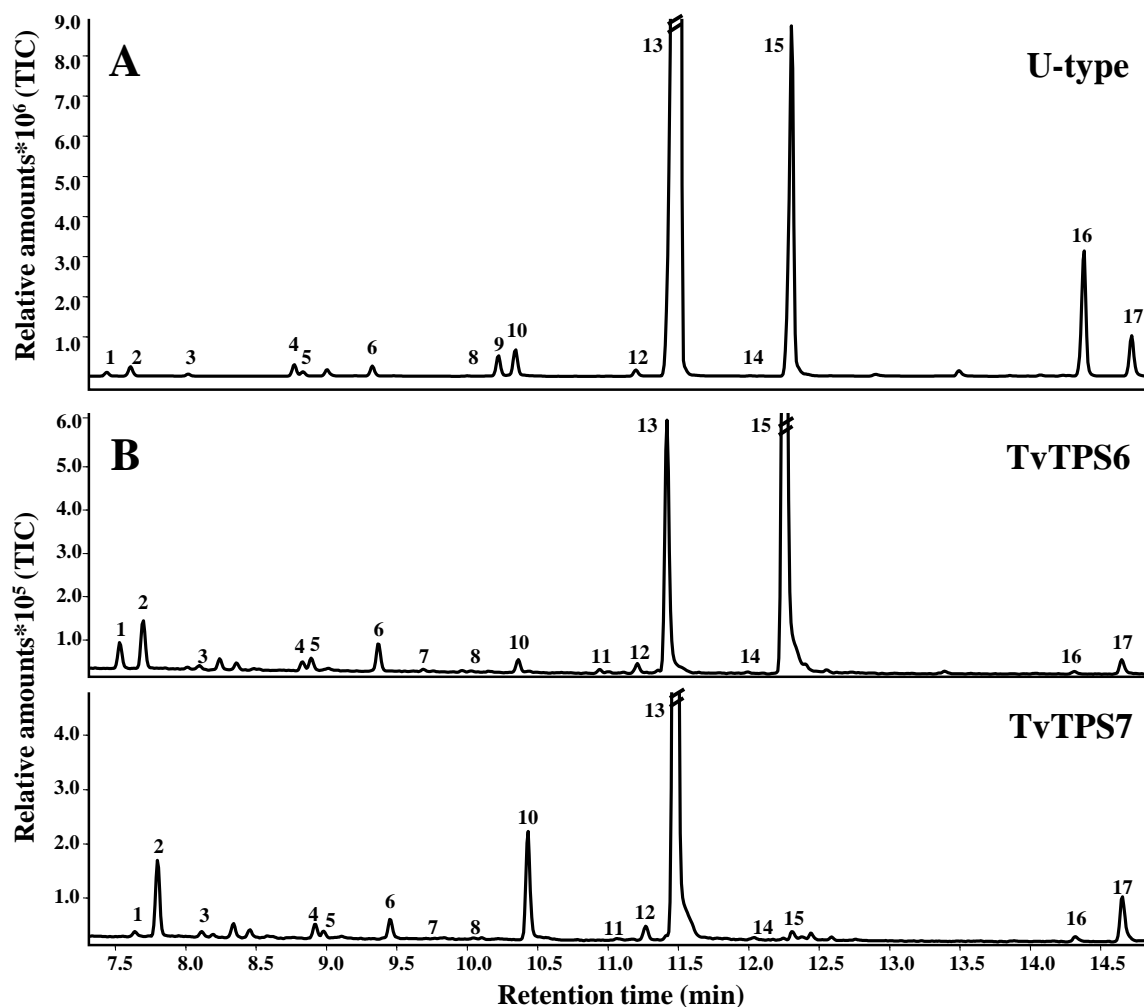
##### 3.1.1 The multiproduct monoterpene synthases TvTPS6 and TvTPS7 provide most of the monoterpene spectrum of the *Thymus vulgaris* (*E*)-sabinene hydrate chemotype

To identify monoterpene synthase genes of the (*E*)-sabinene hydrate chemotype of *Thymus vulgaris*, 5'- and 3'-RACE cDNA libraries from this plant were generated. 3'-RACE PCRs were conducted with degenerate forward primers based on previously identified monoterpene synthase genes of the *Lamiaceae* (Table S1). This screen revealed two fragments of about 1000 bp in length. These partial sequences were isolated by J. Asbach (MPI for Chemical Ecology, Jena, Germany) and provided for this study. The sequences were extended by 5'-RACE PCRs to obtain the complete ORF. The final ORFs of the putative monoterpene synthases comprised 1812 bp and 1794 bp, respectively. They shared a sequence identity of 89 % on DNA level and were designated as *Tvtps6* and *Tvtps7*.

The deduced amino acid sequences TvTPS6 and TvTPS7 contained the magnesium-binding DDxxD motif in the C-terminal moiety. Furthermore, the RRx<sub>8</sub>W motif, that is typical for monoterpene synthases, was found 60 amino acids from the N-terminus. According to the prediction by the ChloroP database, the sequences contained an N-terminal signal peptide of 47 amino acids (Fig. S1). Since plastids are the location of monoterpene biosynthesis, *Tvtps6* and *Tvtps7* were assumed to encode monoterpene synthases. This assumption was supported by their high similarity to monoterpene synthases from other plants, especially from the *Lamiaceae*. *Origanum vulgare* cultivar d0601 *terpene synthase 1*, for example, showed a DNA sequence identity of 91 % to *Tvtps6* and 90 % to *Tvtps7*. Furthermore, *Thymus caespititius* isolate C2-01  $\alpha$ -terpineol synthase (*Tctps5*) showed a DNA sequence identity of 90 % to *Tvtps6* and *Tvtps7*. The respective putative enzymes TvTPS6 and TvTPS7 were heterologously expressed in *E. coli* after truncation of the first 43 amino acids that may encode the signal peptide. With GPP as substrate, both enzymes produced 16 different monoterpenes. The product spectra of TvTPS6 and TvTPS7 differed in their major product, sabinene hydrate. TPS6 converted GPP into both (*E*)- and (*Z*)-sabinene hydrate<sup>3</sup>, while TPS7 formed mostly the (*E*)-isomer

<sup>3</sup> The (*E*)/(*Z*)-notation of sabinene hydrate (instead of *trans* and *cis*) was chosen according to the reference literature: (Larkov et al., 2005).

and only a very small quantity of the (*Z*)-isomer (Fig. 8 B, Table 2). The most abundant minor products were  $\alpha$ -pinene, myrcene, limonene, and  $\alpha$ -terpineol (Table 2). The product spectra of both enzymes were almost identical with the monoterpene blend of thyme plants of the (*E*)-sabinene hydrate chemotype (Fig. 8 A, Table 2). Only the *p*-cymene present in the essential oil of the chemotype was not formed by these enzymes.

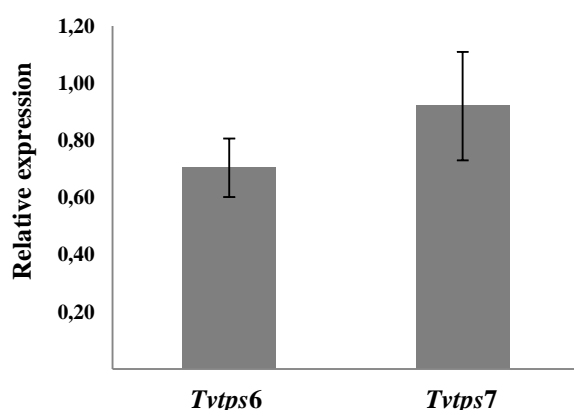


**Fig. 8** The monoterpene spectrum of the (*E*)-sabinene hydrate chemotype (U-type) of *Thymus vulgaris*. (A) Terpenes were extracted with hexane and analyzed by GC-MS. The terpene synthases TvTPS6 and TvTPS7 produce monoterpenes of the (*E*)-sabinene hydrate chemotype of *Thymus vulgaris*. (B) The enzymes were expressed in *E. coli*, partially purified, and incubated with the substrate GPP. The resulting monoterpene products were analyzed by GC-MS. The numbers given are: 1:  $\alpha$ -thujene, 2:  $\alpha$ -pinene, 3: camphene, 4: sabinene, 5:  $\beta$ -pinene, 6: myrcene, 7:  $\alpha$ -phellandrene, 8:  $\alpha$ -terpinene, 9: *p*-cymene, 10: limonene, 11:  $\beta$ -ocimene, 12:  $\gamma$ -terpinene, 13: (*E*)-sabinene hydrate, 14: terpinolene, 15: (*Z*)-sabinene hydrate, 16: terpinen-4-ol, 17:  $\alpha$ -terpineol.

|    |                               | U-type [%] | TvTPS6 [%] | TvTPS7 [%] |
|----|-------------------------------|------------|------------|------------|
| 1  | $\alpha$ -Thujene             | 0,64       | 1,83       | 0,56       |
| 2  | $\alpha$ -Pinene              | 0,95       | 3,23       | 3,64       |
| 3  | Camphene                      | 0,23       | 0,32       | 0,23       |
| 4  | Sabinene                      | 1,50       | 1,09       | 0,99       |
| 5  | $\beta$ -Pinene               | 0,48       | 1,32       | 0,58       |
| 6  | Myrcene                       | 1,41       | 2,60       | 1,21       |
| 7  | $\alpha$ -Phellandrene        | --         | 0,43       | 0,34       |
| 8  | $\alpha$ -Terpinene           | 0,65       | 0,82       | 0,68       |
| 9  | <i>p</i> -Cymene              | 2,00       | --         | --         |
| 10 | Limonene                      | 3,31       | 1,97       | 6,28       |
| 11 | $\beta$ -Ocimene              | --         | 0,34       | n.e.i.     |
| 12 | $\gamma$ -Terpinene           | 1,65       | 1,82       | 1,37       |
| 13 | ( <i>E</i> )-Sabinene hydrate | 58,01      | 20,91      | 77,82      |
| 14 | Terpinolene                   | 0,33       | 0,46       | 0,38       |
| 15 | ( <i>Z</i> )-Sabinene hydrate | 17,21      | 60,37      | 1,43       |
| 16 | Terpinen-4-ol                 | 8,03       | 0,82       | 0,63       |
| 17 | $\alpha$ -Terpineol           | 3,58       | 1,68       | 3,84       |

**Table 2** The monoterpene composition of the essential oil of the (*E*)-sabinene hydrate chemotype of *Thymus vulgaris* in comparison to the product blends formed by TvTPS6 and TvTPS7. The amounts of monoterpenes were determined by GC-FID analysis after hexane extraction and are given as a percentage of the total monoterpene content. Nonylacetaate (10 $\mu$ g/ml) was used as internal standard.

Gene expression analysis by real-time quantitative PCR revealed that both *Tvtps6* and *Tvtps7* were expressed at similar levels in the leaves of the (*E*)-sabinene hydrate chemotype (Fig. 9). The  $K_m$  values of the enzymes, however, displayed significant differences. For TvTPS6, the  $K_m$  value for GPP was 33.5  $\mu$ M, which is high in comparison to those of other monoterpene synthases. In contrast, the  $K_m$  value of TvTPS7 was determined as 6.1  $\mu$ M, which is typical for monoterpene synthases (Wise and Croteau, 1999).

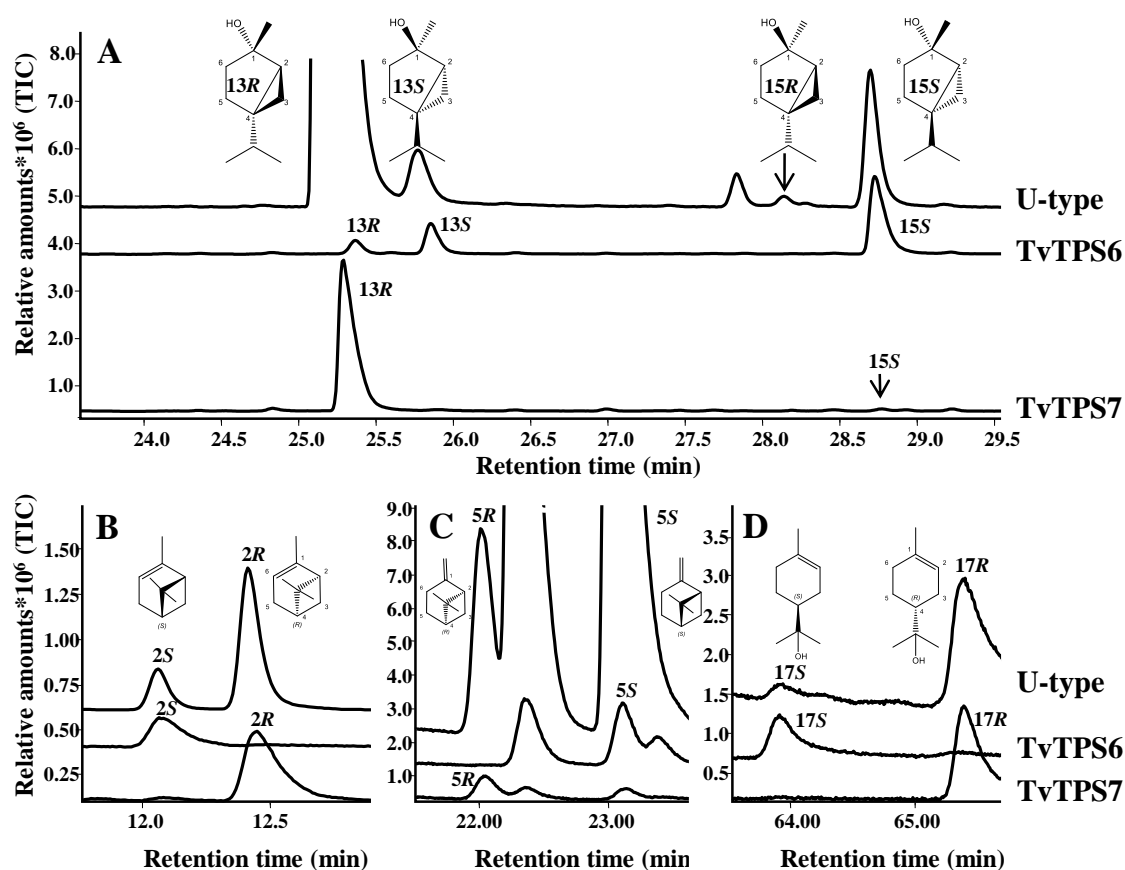


**Fig. 9** The monoterpene synthase genes *Tvtps6* and *Tvtps7* are both expressed in leaves of *Thymus vulgaris*. Transcript concentrations were determined relative to control genes by real-time quantitative PCR.

### 3.1.2 TvTPS6 and TvTPS7 produce monoterpenes with opposite configuration

(*E*)- and (*Z*)-sabinene hydrate, the main products of TvTPS6 and TvTPS7, have chiral centers at C-1, C-2, and C-4 (Fig. 10 A). To compare the enantiomeric compositions of the sabinene hydrate found in the essential oil of the (*E*)-sabinene hydrate chemotype with the ones formed by TvTPS6 and TvTPS7, the plant's essential oil and TvTPS6/TvTPS7 products were analyzed on a chiral-phase column. Plants of the (*E*)-sabinene hydrate

chemotype contained all four stereoisomers of (*E*)- and (*Z*)-sabinene hydrate, with (1*S*,2*S*,4*R*)-(*E*)-sabinene hydrate dominating the essential oil (Fig. 10 A). After heterologous expression, the terpene synthase TvTPS6 formed (1*S*,2*R*,4*S*)-(*Z*)-sabinene hydrate as the major product and minor amounts of both enantiomers of (*E*)-sabinene hydrate. Conversely, TvTPS7 only produced (1*S*,2*S*,4*R*)-(*E*)-sabinene hydrate, along with traces of (1*S*,2*R*,4*S*)-(*Z*)-sabinene hydrate. The (1*R*,2*S*,4*R*)-isomer of (*Z*)-sabinene hydrate was not detected in the product profile of both enzymes. The opposite configuration of TvTPS6 and TvTPS7 products was also observed among all their minor chiral monoterpene products. For example, while TvTPS6 produced (*S*)- $\alpha$ -pinene, (*S*)- $\beta$ -pinene, and (*S*)- $\alpha$ -terpineol, TvTPS7 formed (*R*)- $\alpha$ -pinene, (*R*)- $\beta$ -pinene, and (*R*)- $\alpha$ -terpineol (Fig. 10 B, C, D). The essential oil of the (*E*)-sabinene hydrate chemotype contained both enantiomers of the respective monoterpenes, indicating that both, TvTPS6 and TvTPS7, contribute to the blend.

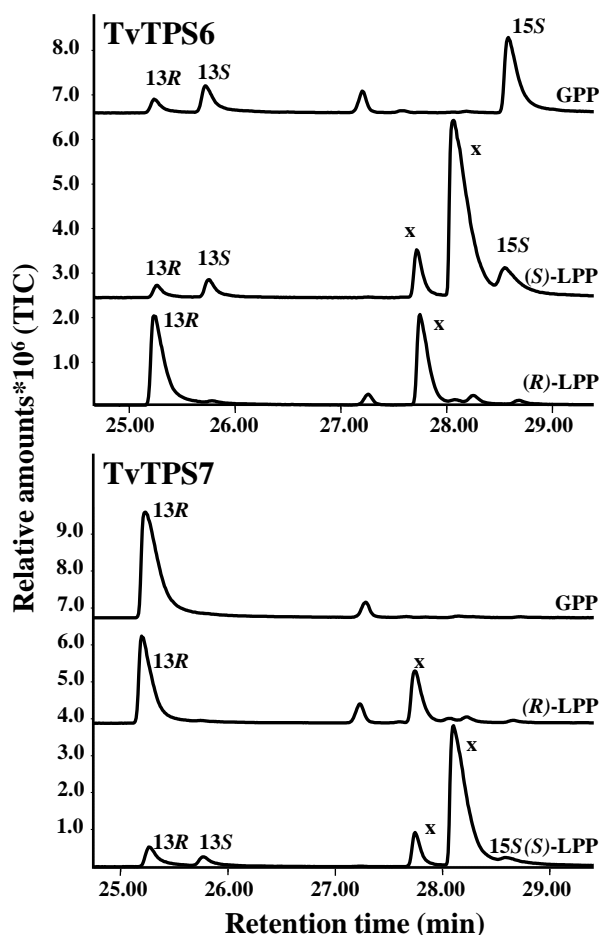


**Fig. 10 TvTPS6 and TvTPS7 produce monoterpenes with opposite stereochemical configuration.** The enantiomeric composition of sabinene hydrate (A),  $\alpha$ -pinene (B),  $\beta$ -pinene (C) and  $\alpha$ -terpineol (D) of TvTPS6 and TvTPS7 products from GPP and the (*E*)-sabinene hydrate chemotype (U-type) was determined by enantioselective separation and identification by chiral-phase GC-MS. The numbers given are: 13*R*: (1*S*,2*S*,4*R*)-(*E*)-sabinene hydrate, 13*S*: (1*R*,2*R*,4*S*)-(*E*)-sabinene hydrate, 15*R*: (1*R*,2*S*,4*R*)-(*Z*)-sabinene hydrate, 15*S*: (1*S*,2*R*,4*S*)-(*Z*)-sabinene hydrate, 2*S*: (*S*)- $\alpha$ -pinene, 2*R*: (*R*)- $\alpha$ -pinene, 5*R*: (*R*)- $\beta$ -pinene, 5*S*: (*S*)- $\beta$ -pinene, 17*S*: (*S*)- $\alpha$ -terpineol, 17*R*: (*R*)- $\alpha$ -terpineol. The *R* and *S* designations after the numbers refer to the configuration at C-4.



### 3.1.3 The configuration of the monoterpene products is determined by the configuration of the LPP intermediate in the TvTPS6 and TvTPS7 active centers

In order to understand how the opposite stereospecificity of TvTPS6 and TvTPS7 is determined, the step at which the crucial difference in the reaction mechanism emerged, was examined. On the pathway from GPP to the monoterpenes, LPP is the earliest chiral reaction intermediate (Introduction I, Fig. 6) and has previously been utilized to study the reaction mechanism of monoterpene synthases (Croteau et al., 1986). To test the stereoselective preference of TvTPS6 and TvTPS7 at this step of the reaction, enantiopure (3*R*)-LPP and (3*S*)-LPP were synthesized and used as substrates. TvTPS6 and TvTPS7 accepted both enantiomers of LPP and converted them into the corresponding enantiomers of the monoterpenes (Fig. 11).

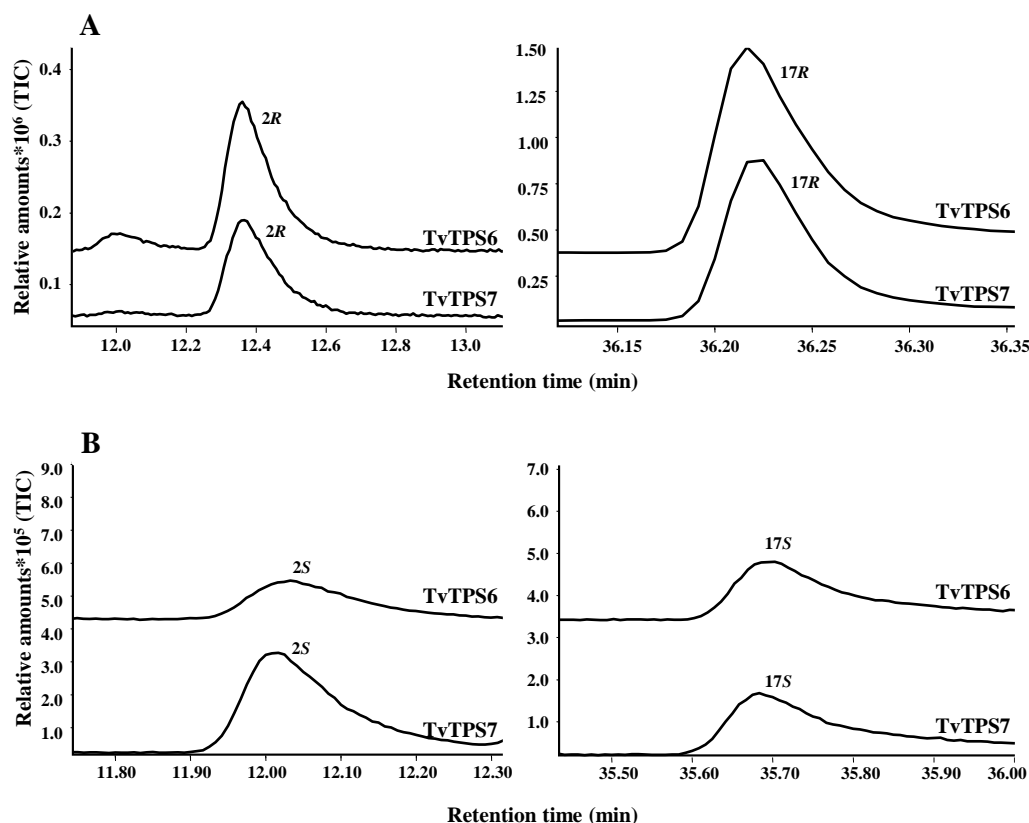


**Fig. 11** The configuration of the monoterpene products of TvTPS6 and TvTPS7 is determined by the configuration of the LPP intermediate. The enantiomeric composition of TvTPS6 and TvTPS7 products after incubation with GPP, (3*R*)-LPP, and (3*S*)-LPP was determined by GC-MS analysis on a chiral-phase column.

The numbers given are: 13*R*: (1*S*,2*S*,4*R*)-(*E*)-sabinene hydrate, 13*S*: (1*R*,2*R*,4*S*)-(*E*)-sabinene hydrate, 15*S*: (1*S*,2*R*,4*S*)-(*Z*)-sabinene hydrate. The *R* and *S* designations after the numbers refer to the configuration at C-4. The linalool detected in these traces is a substrate artefact (x).

When (3*R*)-LPP was offered, both enzymes converted it exclusively into (1*S*,2*S*,4*R*)-(*E*)-sabinene hydrate. Conversely, the incubation with the (3*S*)-LPP substrate resulted in (1*R*,2*R*,4*S*)-(*E*)-sabinene hydrate and (1*S*,2*R*,4*S*)-(*Z*)-sabinene hydrate. The relatively high

abundance of (1*S*,2*S*,4*R*)-(*E*)-sabinene hydrate after incubation with (3*S*)-LPP was most likely due to the fact that (3*S*)-LPP was not completely enantiopure, and contained approximately 10 % (3*R*)-LPP (see Materials and Methods I, section 2.12). As with the GPP substrate, the stereoisomer (1*R*,2*S*,4*R*)-(*Z*)-sabinene hydrate was not formed by the enzymes. The configuration of the minor chiral monoterpene products also followed the chirality of the respective LPP substrate (Fig. 12).



**Fig. 12** The configuration of monoterpene minor products of TvTPS6 and TvTPS7 is determined by the configuration of the LPP intermediate. The enantiomeric composition of TvTPS6 and TvTPS7 products after incubation with (3*R*)-LPP (A) and (3*S*)-LPP (B) was determined by GC-MS analysis on a chiral-phase column. The numbers given are: 2*S*: (*S*)- $\alpha$ -pinene, 2*R*: (*R*)- $\alpha$ -pinene, 17*S*: (*S*)- $\alpha$ -terpineol, 17*R*: (*R*)- $\alpha$ -terpineol. The *R* and *S* designations after the numbers refer to the configuration at C-4.

To investigate the kinetic preferences of TvTPS6 and TvTPS7 for the reaction intermediate LPP, enzyme assays with the (3*R*)-LPP substrate were performed and the resulting product concentrations were determined. Therefore, the enzymes were purified, the enzyme concentrations were equalized, and the enzymes were incubated with the substrate for the same time interval. TvTPS7 produced twice the concentration of the most abundant monoterpenes (*E*)-sabinene hydrate,  $\alpha$ -terpineol, and limonene than TvTPS6, indicating a higher turnover of (3*R*)-LPP by TvTPS7 (Table 3).

|                      | TvTPS6 [ $\mu\text{g/ml}$ ] | TvTPS7 [ $\mu\text{g/ml}$ ] | <b>Table 3 Monoterpene formation from the (3R)-LPP substrate by TvTPS6 and TvTPS7.</b> The amounts of monoterpenes were determined by GC-FID after hexane extraction. Nonylacetate (10 $\mu\text{g/ml}$ ) was used as internal standard. |
|----------------------|-----------------------------|-----------------------------|--|
| (E)-Sabinene hydrate | 2,26                        | 5,22                        |  |
| $\alpha$ -Terpineol  | 1,28                        | 2,46                        |  |
| Limonene             | 0,27                        | 0,6                         |  |

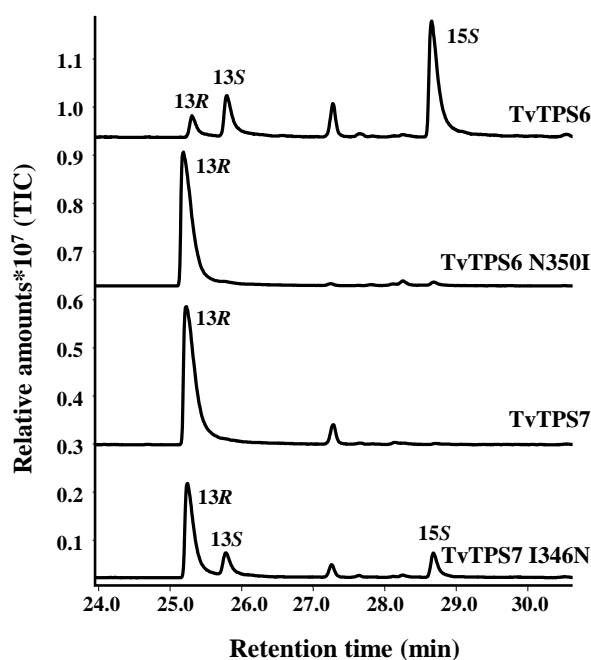
### 3.1.4 Mutagenesis studies revealed an amino acid responsible for the different stereospecificity of TvTPS6 and TvTPS7

The high amino acid identity between TvTPS6 and TvTPS7 (85 %) provided an opportunity to identify the residue(s) responsible for the opposite stereospecificity by sequence comparison and site-directed mutagenesis. The amino acid Asn-350 in TvTPS6 corresponds to Ile-346 in TvTPS7. Both residues were located seven amino acids upstream of the DDxxD motif. In a closely related 1,8-cineole synthase (SfCinS1) from *Salvia fruticosa*, the corresponding Asn-338 residue was demonstrated to be important for the product specificity of the enzyme (Kampranis et al., 2007) (Fig. 13).

|                 |                     |              |       |            |            |
|-----------------|---------------------|--------------|-------|------------|------------|
| <b>Tv TPS6</b>  | AAV <b>N</b> ILITAI | DDVYD        | VYGTL | DELRLFTDVI | RRWDTQSIDQ |
| <b>Tv TPS7</b>  | AAI <b>I</b> ILITAI | DDVYD        | VYGTL | DELQLFTDVI | RRWDTQSIDQ |
| <b>Sf CinS1</b> | TKI <b>N</b> ALVTTI | DDVFD        | IYGTL | EELQLFTTAI | QRWDIESMKQ |
|                 |                     | <b>DDxxD</b> |       |            |            |

**Fig. 13 A critical amino acid is located seven amino acids upstream of the DDxxD motif in TvTPS6, TvTPS7, and SfCinS1.** Asn-350 (TvTPS6) and Asn-338 (SfCinS1) are highlighted in purple, Ile-346 (TvTPS7) is highlighted in yellow.

To test whether this residue can determine the stereospecificity of TvTPS6 and TvTPS7, Asn-350 of TvTPS6 was altered to Ile, the corresponding amino acid in TvTPS7, by site-directed mutagenesis of the ORF. After overexpression and incubation with the GPP substrate, the N350I mutant of TvTPS6 showed an altered stereospecificity and produced only the (1*S*,2*S*,4*R*)-enantiomer of (*E*)-sabinene hydrate, like TvTPS7. Conversely, the I346N mutant of TvTPS7 displayed a stereospecificity similar to that of TvTPS6. Although (1*S*,2*S*,4*R*)-(*E*)-sabinene hydrate was still the main product of this mutant, TvTPS7 I346N also produced the (4*S*)-enantiomers of (*E*)-sabinene hydrate and (*Z*)-sabinene hydrate (Fig. 14).

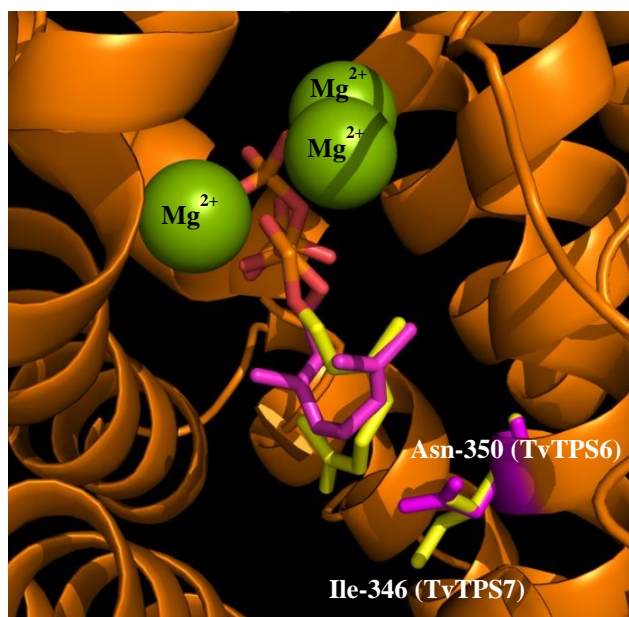


**Fig. 14 Site-directed mutagenesis of a critical amino acid residue of TvTPS6 and TvTPS7 inverts the configuration of the sabinene hydrate products.** The enantiomeric composition of the products of TvTPS6, TvTPS7, and their respective mutants after incubation with GPP was determined by GC-MS analysis on a chiral-phase column. The numbers given are: 13R: (1*S*,2*S*,4*R*)-(E)-sabinene hydrate, 13S: (1*R*,2*R*,4*S*)-(E)-sabinene hydrate, 15S: (1*S*,2*R*,4*S*)-(Z)-sabinene hydrate. The *R* and *S* designations after the numbers refer to the configuration at C-4.

### 3.1.5 Modeling of the active sites of TvTPS6 and TvTPS7 and substrate docking studies

To illustrate the mechanism by which one amino acid, Asn-350 in TvTPS6 or Ile-346 in TvTPS7, can affect the stereospecificity of the reaction mechanism of the enzymes, a model of the TvTPS6 and TvTPS7 active sites was created by threading of the sequences onto the structure of the bornyl diphosphate synthase from *Salvia officinalis* (Whittington et al., 2002). The model revealed that the surface of both active sites is formed by 31 amino acid residues, of which 5 residues differed between TvTPS6 and TvTPS7 (Fig. S1).

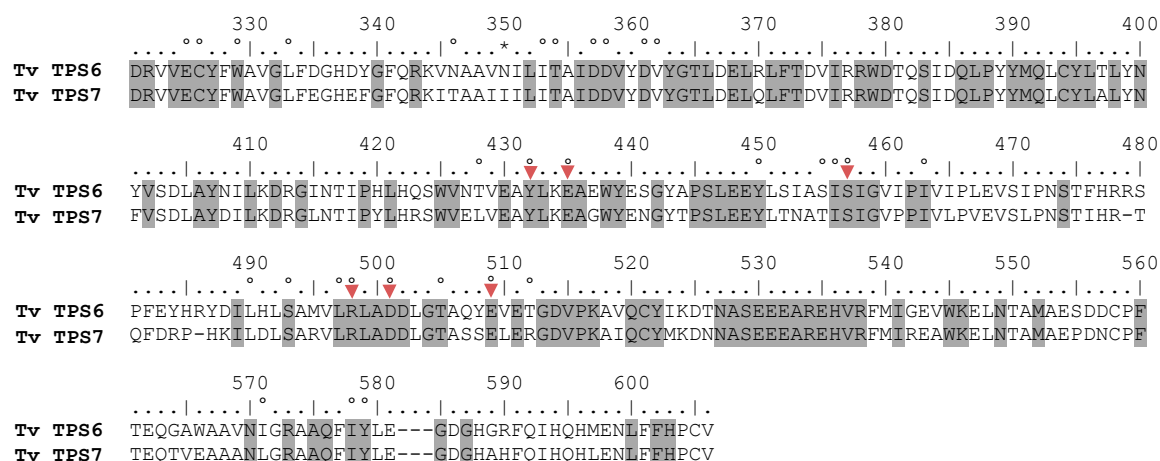
Furthermore, the model suggested that Asn-350 in TvTPS6 and Ile-346 in TvTPS7 were located at the bottom of the active site cavity, in the middle of an  $\alpha$ -helix that was designated as helix D in the bornyl diphosphate synthase. The same helix accommodates also the magnesium-binding DDxxD motif at the C-terminal end. The GPP substrate was docked into the active site pocket (Fig. 15). In TvTPS6, the side chain of Asn-350 may force the GPP molecule to form an upward kink while Ile-346 in TvTPS7 provides more space.



**Fig. 15 Model of the TvTPS6 and TvTPS7 active sites.** Both Asn-350 of TvTPS6 (shown in purple) and Ile-346 of TvTPS7 (shown in yellow) might affect the conformation of the GPP substrate that is shown in the corresponding color.

### 3.1.6 The pathway to (*E*)- and (*Z*)-sabinene hydrate involves water quenching

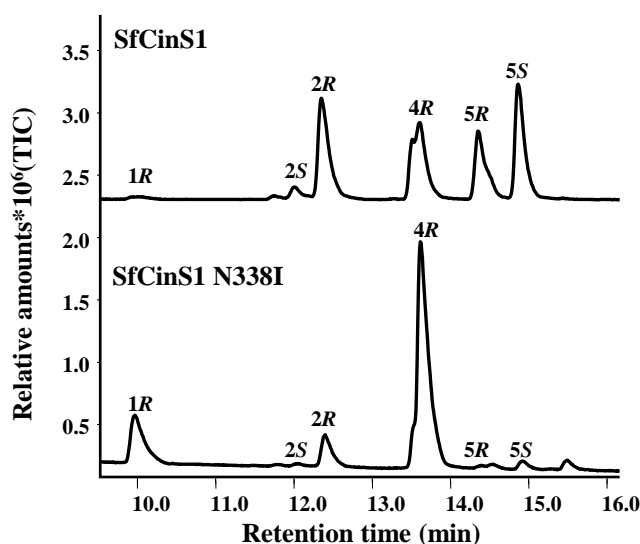
The main products of TvTPS6 and TvTPS7, (*E*)- and (*Z*)-sabinene hydrate, are both hydroxylated compounds. Water is supposed to provide the hydroxyl group that quenches the sabinyl cation, the direct carbocationic precursor of sabinene hydrate (Hallahan and Croteau, 1988). In order to find crucial amino acids that could bind and activate water molecules during the reaction process, all amino acids with polar or hydrophilic properties and direct contact to the active site pocket were investigated. To find the amino acids involved in water binding, these amino acids were mutated to nonpolar amino acids via site-directed mutagenesis. The following amino acid residues were exchanged: Tyr-432, Glu-435, Ser-457, Arg-498, Asp-501, and Glu-509 (numbers refer to TvTPS6) (Fig. 16). The positions were replaced with Leu, Val, or Phe, depending on their residue size. However, all mutations rendered the enzymes essentially inactive (data not shown). Thus, a deduction on the specific catalytic function of the exchanged amino acids was not possible.



**Fig. 16 Alignment of the C-terminal domain of TvTPS6 and TvTPS7.** Identical amino acids are shaded in grey. The amino acids forming the active site (according to the 3D model, Results I, section 3.1.5) are denoted with a circle. The amino acids investigated for their putative role in water binding are marked with an arrow. The amino acid position crucial for the stereospecificity of the enzymes is marked with an asterisk.

### 3.1.7 Mutagenesis studies of the 1,8-cineole synthase of *Salvia fruticosa* show no function of Asn-338 in stereocontrol

The Asn-338 residue has a crucial function within the 1,8-cineole synthase from *Salvia fruticosa* (SfCinS1). In this enzyme, a mutation of Asn-338 to Ile abolished the hydroxylation of the  $\alpha$ -terpinyl intermediate and thereby the subsequent formation of  $\alpha$ -terpineol and 1,8-cineole (Kampranis et al., 2007). To test whether this amino acid also has an effect on the stereocontrol of the reaction mechanism as in TvTPS6 and TvTPS7, the gene encoding for SfCinS1 was isolated from *Salvia fruticosa* and the N338I mutation of the corresponding protein was generated. The mutant did not show the opposite stereospecificity, only a different ratio of the formed products. No changes in the configuration of the chiral monoterpene products were observed (Fig. 17), indicating that in this case the Asn-338 residue only activates water for the hydroxylation.

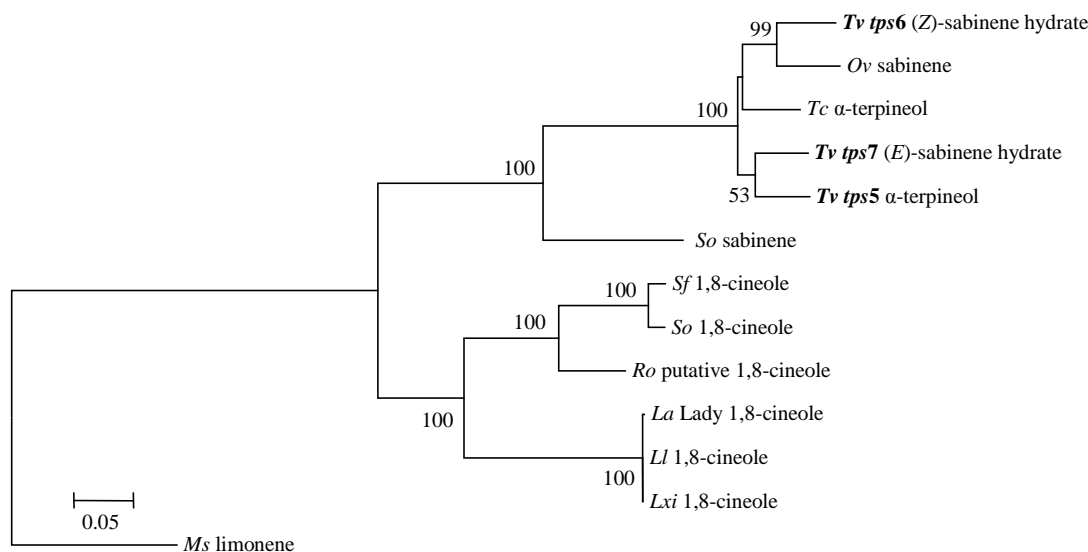


**Fig. 17 The amino acid residue crucial for the stereospecificity of TvTPS6 and TvTPS7 does not affect the enantiomeric composition of a 1,8-cineole synthase of *Salvia fruticosa* (SfCinS1).** Site-directed mutagenesis of Asn-338 to Ile (SfCinS1 N338I) does not alter the stereospecificity of the enzyme. The numbers are: 1R: (*R*)- $\alpha$ -thujene, 2S: (*S*)- $\alpha$ -pinene, 2R: (*R*)- $\alpha$ -pinene, 4R: (*R*)-sabinene, 5R: (*R*)- $\beta$ -pinene, 5S: (*S*)- $\beta$ -pinene. The *R* and *S* designations after the numbers refer to the configuration at C-4.

### 3.2 Isolation and characterization of an $\alpha$ -terpineol synthase from thyme (*Thymus vulgaris*)

#### 3.2.1 The multiproduct terpene synthase TvTPS5 produces most of the monoterpenes found in the $\alpha$ -terpineol chemotype of *Thymus vulgaris*

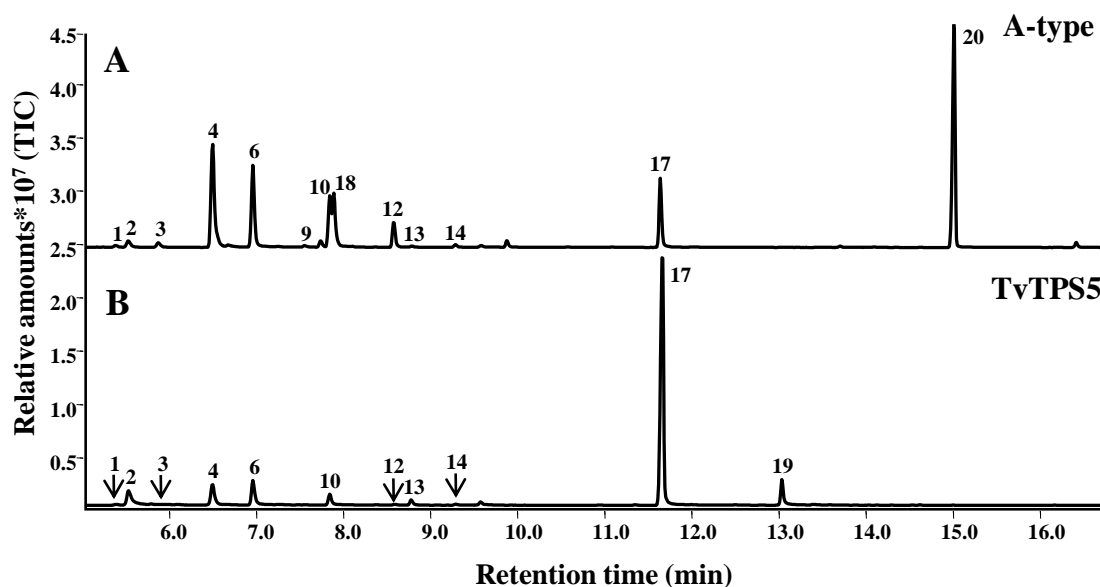
The genes encoding for TvTPS6 and TvTPS7 were expressed in the  $\alpha$ -terpineol chemotype of *Thymus vulgaris*, according to Northern Blot results (J. Asbach, MPI for Chemical Ecology Jena, personal communication). Therefore, I used the primers designed for subcloning of these genes into the pASK-IBA 37+ expression vector for sequence isolation from cDNA of the  $\alpha$ -terpineol chemotype of thyme. This PCR resulted in one fragment of 1500 bp length which significantly differed from *Tvtps6* and *Tvtps7*, but still shared a high sequence similarity, especially with *Tvtps7* (82.5 % with TvTPS6 and 90 % with TvTPS7 on deduced amino acid level, respectively). Since the fragment, designated as *Tvtps5*, was amplified with the primers for subcloning into the expression vector, the sequence lacked the N-terminal signal peptide. The sequence was completed to the 5'-end by J. Schimmel (Schimmel, 2014). The deduced amino acid sequence was predicted to contain a signal peptide of 47 amino acids in length (Chloro P database). It showed the RRx<sub>8</sub>W motif 60 amino acids from the N-terminus, the C-terminal moiety contained the magnesium-binding DDxxD motif (Fig. S2). A dendrogram analysis confirmed the close relationship of *Tvtps5* to *Tvtps6* and *Tvtps7* (Fig. 18). Interestingly, among the *Lamiaceae*,  $\alpha$ -terpineol synthases seemed to be more closely related to sabinene hydrate synthases and sabinene synthases than to 1,8-cineole synthases, despite the fact that  $\alpha$ -terpineol is the direct precursor on the pathway to 1,8-cineole formation. 1,8-cineole is supposed to result from endocyclization of  $\alpha$ -terpineol after protonation (Kampranis et al., 2007; Fähnrich et al., 2011).



**Fig. 18 Dendrogram analysis by Maximum Likelihood method with *Tvtps5*, *Tvtps6*, *Tvtps7*, and monoterpene synthase genes from other *Lamiaceae* species.** Each branch is designated with the abbreviation of the species and the main product of the corresponding enzyme. *Tvtps5*, *Tvtps6*, and *Tvtps7* are highlighted. Bootstrap values greater than 50 are given as a percentage of 1000 replicates. The branch lengths are measured in the number of substitutions per site. The tree was rooted on the limonene synthase from *Mentha spicata*. GenBank IDs in parentheses: *La Lavandula angustifolia*: 1,8-cineole synthase (JN701461); *Ll Lavandula latifolia*: 1,8-cineole synthase (JN701460); *Lxi Lavandula x intermedia*: 1,8-cineole synthase (JN701459); *Ms Mentha spicata*: (+)-limonene synthase (L13459); *Ov Origanum vulgare*: TPS1-d06-01 sabinene synthase (GU385980); *Ro Rosmarinus officinalis*: putative 1,8-cineole synthase (JX050194); *So Salvia officinalis*: 1,8-cineole synthase (AF051899), (+)-sabinene synthase (AF051901); *Sf Salvia fruticosa*: 1,8-cineole synthase (DQ785793); *Tc Thymus ceasptitius*: C2-01 TPS5 (KC181101); *Tv Thymus vulgaris*: TPS5  $\alpha$ -terpineole synthase (KC461937), TPS6 (Z)-sabinene hydrate synthase (JX946357), TPS7 (E)-sabinene hydrate syntase (JX946358).

After heterologous expression in *E. coli*, the enzyme activity was tested with GPP as substrate. TvTPS5 formed 12 different monoterpenes with  $\alpha$ -terpineol as main product (Fig. 19 B). Sabinene, myrcene, and limonene were the most abundant side products. Compared to the  $\alpha$ -terpineol chemotype, TvTPS5 covered most of the monoterpenes found in the plant (Fig. 19 A). The monoterpenes not produced by TvTPS5 but found in the chemotype were *p*-cymene, 1,8-cineole, camphor, borneol, bornyl acetate, and  $\alpha$ -terpineol acetate.



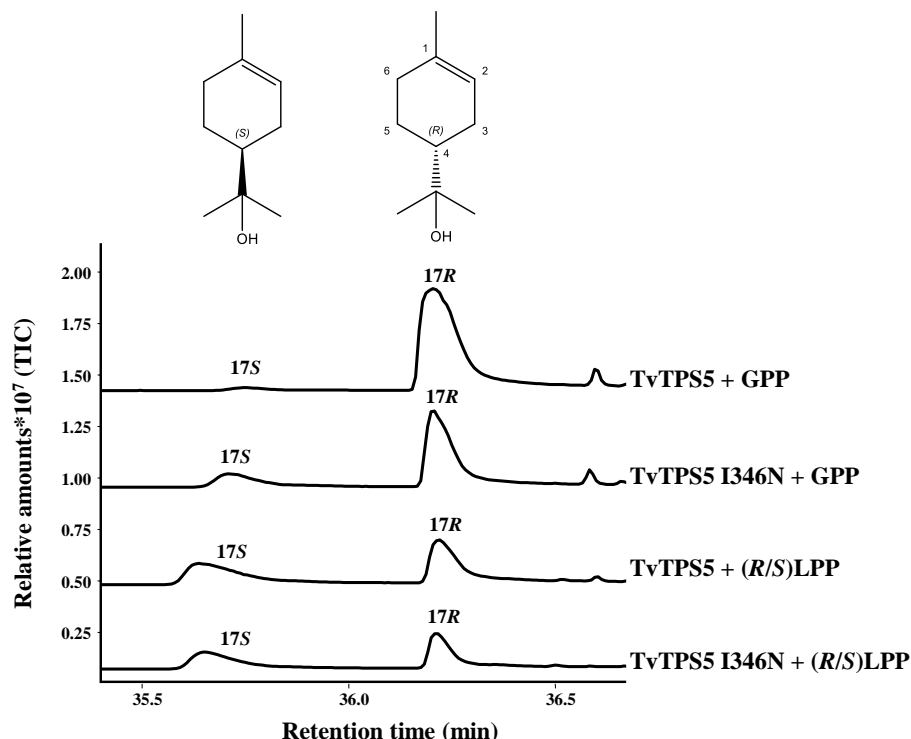


**Fig. 19** The monoterpene spectrum of the  $\alpha$ -terpineol chemotype (A-type) of *Thymus vulgaris*. (A) Terpenes were extracted with hexane and analyzed by GC-MS. The terpene synthase TvTPS5 produces monoterpenes of the  $\alpha$ -terpineol chemotype of *Thymus vulgaris*. (B) TvTPS5 was heterologously expressed in *E. coli*, extracted, and incubated with the substrate GPP. The resulting monoterpene products were identified by GC-MS. The numbers given are: 1:  $\alpha$ -thujene, 2:  $\alpha$ -pinene, 3: camphene, 4: sabinene, 6: myrcene, 9: *p*-cymene, 10: limonene, 18: 1,8-cineole, 12:  $\gamma$ -terpinene, 13: (*E*)-sabinene hydrate, 14: terpinolene, 17:  $\alpha$ -terpineol, 19: nerol, 20:  $\alpha$ -terpineol acetate.

### 3.2.2 In TvTPS5, the configuration of the monoterpene products is influenced by the same amino acid position that was critical in TvTPS6 and TvTPS7

TvTPS5 shared a high sequence identity with both TvTPS6 and TvTPS7, the main terpene synthases from the (*E*)-sabinene hydrate chemotype, but the identity between TvTPS5 and TvTPS7 was remarkably higher (82.5 % with TvTPS6 vs. 90 % with TvTPS7, on amino acid level). Furthermore, the analysis of the main product of TvTPS5 on a chiral-phase column showed that the enzyme displayed the same stereospecificity as TvTPS7: Mainly the *R*-enantiomer of  $\alpha$ -terpineol was formed from GPP and only a very little amount of the *S*-enantiomer was present (Fig. 20). The same was true for the two most abundant chiral side products, sabinene and limonene (data not shown). At the amino acid position responsible for the stereospecificity in TvTPS6 and TvTPS7, there was an Ile found in TvTPS5 (Ile-346), just as in TvTPS7 (Fig. S2). Also, protein structure modeling showed that the steric position of Ile-346 in TvTPS5 resembles that of Ile-346 in TvTPS7 (Fig. S3). Therefore, the effect of a change of this amino acid was examined. Ile-346 was altered to Asn, the amino acid found in TvTPS6 at the corresponding position. The product spectrum of the resulting TPS5 I346N mutant with the GPP substrate showed a slight increase of the *S*-enantiomer of the  $\alpha$ -terpineol product, but (*R*)- $\alpha$ -terpineol was still predominant. When TvTPS5 was fed with a racemic mixture of (*R*)- and (*S*)-LPP, the relative amounts of

(*S*)- and (*R*)- $\alpha$ -terpineol were similar to those formed by the mutant. When the TvTPS5 I346N mutant was fed by the racemate of LPP, nearly equal amounts of both  $\alpha$ -terpineol enantiomers were formed (Fig. 20).

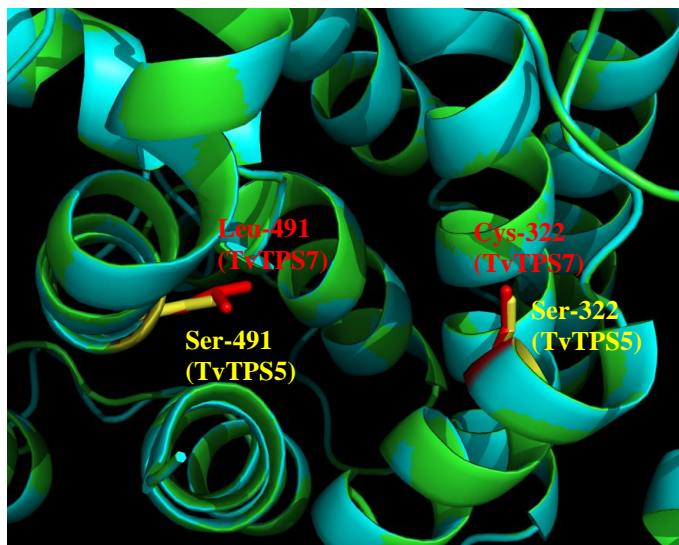


**Fig. 20 Site-directed mutagenesis of a critical amino acid residue of TvTPS5 changes the enantiomeric ratio of the  $\alpha$ -terpineol product.** GC-traces show the enantiomeric composition of the main product of TvTPS5 and the mutant TvTPS5 I364N with GPP and with (*R/S*)-LPP, respectively. The numbers given are: 17*S*: (*S*)- $\alpha$ -terpineol, 17*R*: (*R*)- $\alpha$ -terpineol. The *R* and *S* designations after the numbers refer to the configuration at C-4.

### 3.2.3 Water capture takes place at different steps of the pathway of TvTPS5 and TvTPS7

$\alpha$ -Terpineol and (*E*)-sabinene hydrate, the main products of TvTPS5 and TvTPS7, respectively, are both hydroxylated products. Originated from the first cyclic intermediate, the  $\alpha$ -terpinyl cation, water quenching of the carbocationic reaction mechanism takes place at different steps of the pathway. Considering the high amino acid sequence identity between TvTPS5 and TvTPS7, a model of TvTPS5 was created and compared to that from TvTPS7. The amino acid residues reaching into the active site cavities of these two enzymes were determined in order to find the differences that might cause the altered site of water quenching. According to the model, there were only two amino acid differences between TvTPS5 and TvTPS7: At the positions corresponding to Cys-322 and Leu-491 in TvTPS7, serine was found in TvTPS5 (Ser-322 and Ser-491). Cys-322 (TvTPS7) and the

corresponding Ser-322 (TvTPS5) were positioned at the C-terminal end of  $\alpha$ -helix C and reached towards helix D, where also the DDxxD motif is located. Leu-491 (TvTPS7) and the corresponding Ser-491 (TvTPS5) were placed in the middle of  $\alpha$ -helix H at the opposite site of the active site cavity (Fig. 21). All helix-designations refer to those of the bornyl diphosphate synthase from *Salvia officinalis*, which was used as modeling template structure (Whittington et al., 2002).



**Fig. 21 Model of the TvTPS5 and TvTPS7 active sites.** Two amino acid residues reaching into the active site cavity differed between both enzymes: Cys-322 (red) and Leu-491 (red) in TvTPS7 and Ser-322 and Ser-491 (yellow) in TvTPS5.

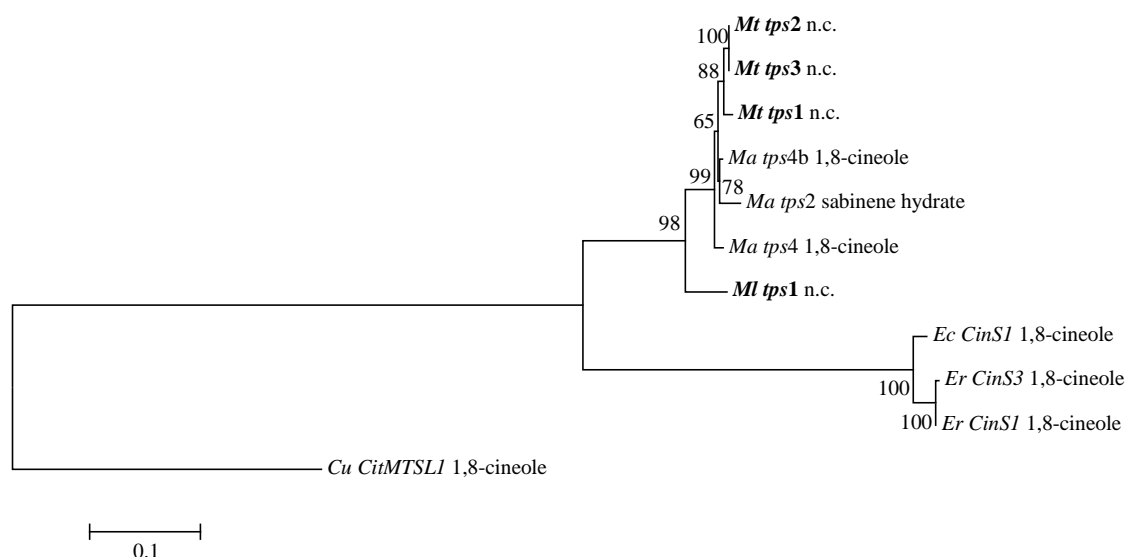
### 3.3 Isolation of monoterpene synthase genes from *Melaleuca linariifolia* and *Melaleuca trichostachya*

In order to study further plant species that produce pharmaceutically valuable oxygenated monoterpenes, various *Melaleuca* species were investigated. Samples of *Melaleuca linariifolia* and *Melaleuca trichostachya* were collected at the East Coast of Australia in New South Wales and used for RNA extraction and sequence isolation. GC-MS analysis of the monoterpene fraction extracted from the plant material revealed that *Melaleuca trichostachya* belonged to the 1,8-cineole chemotype. *Melaleuca linariifolia* belonged to the  $\alpha$ -pinene chemotype, but contained considerable amounts of 1,8-cineole, too. Along with the major monoterpenes, minor amounts of a range of other monoterpenes were found in the essential oil of the leaves (Table 4).

|    |                                  | <i>M. trichostachya</i> | <i>M. linariifolia</i> |
|----|----------------------------------|-------------------------|------------------------|
| 1  | $\alpha$ -Thujene                | 0,028                   | 0                      |
| 2  | $\alpha$ -Pinene                 | 0,212                   | 3,580                  |
| 3  | Sabinene                         | 0,229                   | 0,054                  |
| 4  | $\beta$ -Myrcene                 | 0,076                   | 0,061                  |
| 5  | $\beta$ -Pinene                  | 0,128                   | 0,169                  |
| 6  | $\alpha$ -Terpinene              | 0                       | 0,024                  |
| 7  | Limonene                         | 0,247                   | 0,232                  |
| 8  | $\beta$ -Phellandrene            | 0,044                   | 0,047                  |
| 9  | <i>p</i> -Cymene                 | 0,064                   | 0,054                  |
| 10 | 1,8-Cineole                      | 2,213                   | 1,896                  |
| 11 | $\gamma$ -Terpinene              | 0,180                   | 0                      |
| 12 | ( <i>Z</i> )- $\beta$ -Terpineol | 0,197                   | 0                      |
| 13 | Terpinolene                      | 0,009                   | 0                      |
| 14 | Terpinen-4-ol                    | 0,316                   | 0                      |
| 15 | $\alpha$ -Terpineol              | 0,426                   | 0,216                  |
| 18 | $\alpha$ -Terpinylacetate        | 0                       | 0,314                  |

**Table 4 Monoterpene composition of the *Melaleuca trichostachya* and *Melaleuca linariifolia* samples collected and used for RNA extraction.** The amounts were determined by GC-MS (A. Padovan, ANU Canberra, Australia) and are given in mg/g fresh weight.

Four ORFs of putative monoterpene synthases were amplified: *Mttps1*, *Mttps2*, and *Mttps3* from *Melaleuca trichostachya* and *Mltps1* from *Melaleuca linariifolia*. The deduced amino acid sequences lacked the N-terminal signal peptide since the genes were amplified with primers for subcloning the terpene synthases in the pASK-IBA37+ expression vector (previously used for 1,8-cineole synthases from *Melaleuca alternifolia*; A. Keszei, ANU Canberra, Australia). Nevertheless, they showed the typical monoterpene synthase sequence motifs like the RRx<sub>8</sub>W motif and the important magnesium-binding DDxxD motif (Fig. S4). Furthermore, they all showed a close relationship and a very high sequence identity (between 91-98 % on DNA level) to other 1,8-cineole synthases and sabinene hydrate synthases which were isolated from *Melaleuca alternifolia* and different *Eucalyptus* species (Fig. 22). These enzymes were characterized previously (A. Keszei and V. Lam, ANU Canberra, Australia, unpublished results). After heterologous expression in *E. coli*, the enzymes showed no activity when fed with GPP as substrate. However, due to the low quality of the sequencing reactions, the correctness of the ORFs of the terpene synthases in the expression vector pASK-IBA37+ could not be validated.



**Fig. 22 Dendrogram analysis by Maximum Likelihood method with *Mttps1*, *Mttps2*, *Mttps3*, *Mltps1*, and other monoterpene synthase genes from *Melaleuca* and *Eucalyptus* species.** Each branch is designated with the abbreviation of the species, the gene name, and the main product of the corresponding enzyme. N.c: not characterized. The branch lengths are measured in the number of substitutions per site. Bootstrap values are given as a percentage of 1000 replicates. The tree was rooted on the 1,8-cineole synthase *CitMTSL1* from *Citrus unshiu* (GenBank ID: AB110638). *Ma* *Melaleuca alternifolia*, *Ml* *Melaleuca linariifolia*, *Mt* *Melaleuca trichostachya*, *Ec* *Eucalyptus cladocalyx*, *Er* *Eucalyptus radiata*.

## 4 Discussion

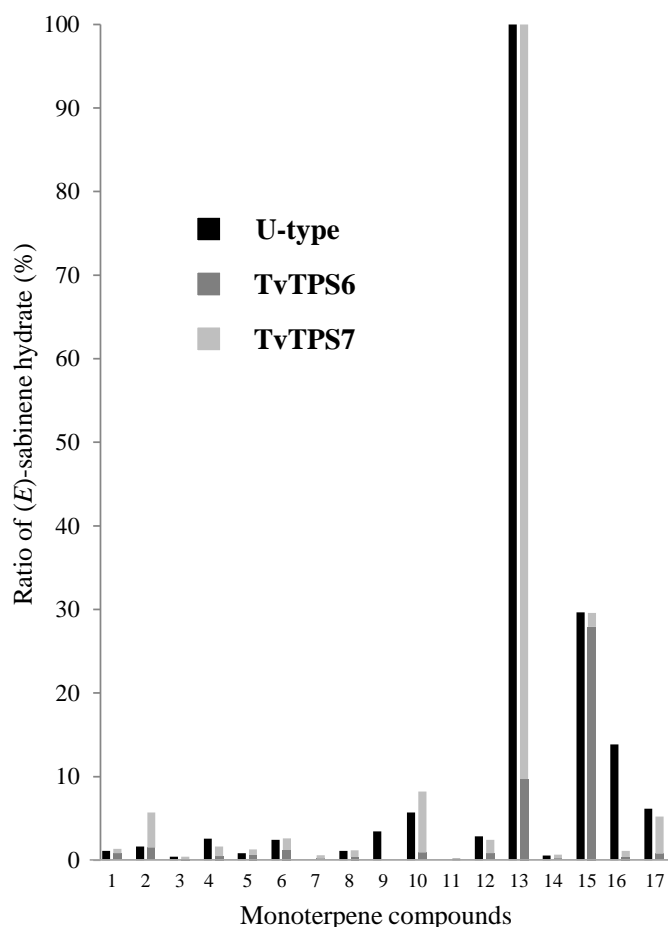
### 4.1 The stereospecific reaction mechanism of the sabinene hydrate synthases TvTPS6 and TvTPS7 from *Thymus vulgaris*

#### 4.1.1 TvTPS6 and TvTPS7 are most likely responsible for the characteristic monoterpene composition of the (*E*)-sabinene hydrate chemotype of *Thymus vulgaris*

The essential oil of the (*E*)-sabinene hydrate chemotype of *Thymus vulgaris* is dominated by high concentrations of both (*E*)- and (*Z*)-sabinene hydrate (Fig. 8 A, Table 2). The two *Thymus vulgaris* terpene synthases TvTPS6 and TvTPS7 produce not only sabinene hydrate but minor amounts of sixteen additional monoterpenes. Since TvTPS7 is expected to be more active than TvTPS6 due to its lower  $K_m$  value ( $K_m$  6.1  $\mu\text{M}$  and  $K_m$  33.5  $\mu\text{M}$  for GPP, respectively), the dominance of (1*S*,2*S*,4*R*)-(*E*)-sabinene hydrate, the main product of TvTPS7, in the essential oil seems reasonable. A quantitative model adjusting the product concentrations of both terpene synthases to fit the ratio of (*E*)- to (*Z*)-sabinene hydrate found in the essential oil demonstrates that the two enzymes are sufficient to produce nearly the complete monoterpene blend of the chemotype (Fig. 23).

Only one monoterpene compound, *p*-cymene, appears to be an additional constituent of the oil. *p*-Cymene can be derived from  $\gamma$ -terpinene, either by spontaneous conversion of the cyclohexene ring during plant growth under the influence of UV light and oxygen or during extraction using the SPME method (Sefidkon et al., 1999; Zabarar and Wyllie, 2002). Also, the action of a cytochrome P450 monooxygenase can result in *p*-cymene formation (Keszei et al., 2008). The concentration of terpinen-4-ol is higher in the essential oil than in the product spectra of TvTPS6 and TvTPS7. Most likely, terpinen-4-ol is formed non-enzymatically by rearrangement of sabinene hydrate, which is produced in large amounts by both TvTPS6 and TvTPS7. In *Melaleuca* species for example, the sabinene hydrate concentration in the essential oils was reported to decrease during leaf aging while the concentration of terpinen-4-ol increased in the same period of time (Cornwell et al., 1999). In addition to the monoterpenes, low concentrations of sesquiterpenes including (*E*)- $\beta$ -caryophyllene, germacrene D and nerolidol were identified in *Thymus* species. These sesquiterpenes were found in most chemotypes of thyme and are most likely not affected by the mechanism of chemotype formation (Stahl-Biskup, 2002). The characterization of sesquiterpene synthases in the closely related oregano (*Origanum vulgare*) suggests that

two or three sesquiterpene synthases are sufficient to produce the complete sesquiterpene blend of the (*E*)-sabinene hydrate chemotype in thyme (Crocoll et al., 2010).

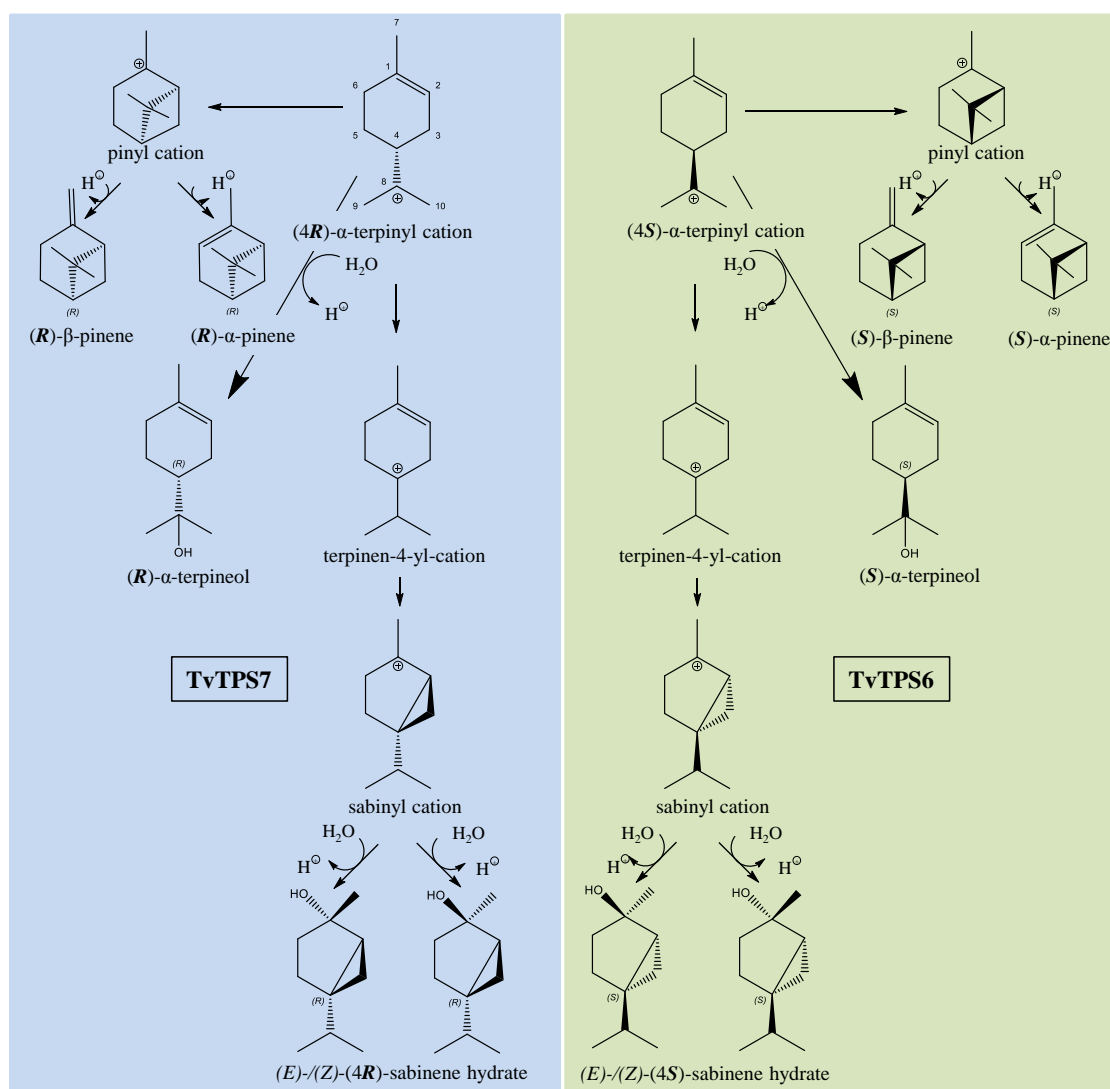


**Fig. 23** The products of TvTPS6 and TvTPS7 constitute the monoterpene spectrum of the (*E*)-sabinene hydrate chemotype (U-type) of *Thymus vulgaris*. The amount of (*E*)-sabinene hydrate was set as 100% in both the chemotype and the sum of the blend produced by TvTPS6 and TvTPS7. The ratio between (*E*)- and (*Z*)-sabinene hydrate in the chemotype was set as the same as in the sum of TvTPS6 and TvTPS7 products. Therefore, the monoterpene amounts produced by TvTPS7 were multiplied with the factor 2.4. The relative amounts of the monoterpenes were calculated as a percentage of (*E*)-sabinene hydrate. The numbers given are: 1:  $\alpha$ -thujene, 2:  $\alpha$ -pinene, 3: camphene, 4: sabinene, 5:  $\beta$ -pinene, 6: myrcene, 7:  $\alpha$ -phellandrene, 8:  $\alpha$ -terpinene, 9: *p*-cymene, 10: limonene, 11:  $\beta$ -ocimene, 12:  $\gamma$ -terpinene, 13: (*E*)-sabinene hydrate, 14: terpinolene, 15: (*Z*)-sabinene hydrate, 16: terpinen-4-ol, 17:  $\alpha$ -terpineol.

#### 4.1.2 The opposite initial binding conformation of GPP leads to the opposite stereospecificity of TvTPS6 and TvTPS7

In terpene synthases, the succession of unstable carbocationic intermediates is responsible for the formation of multiple products (Davis and Croteau, 2000). Previous biochemical studies on the sabinene hydrate synthase activities in sweet majoram (*Majorana hortensis*) suggested that (*E*)- and (*Z*)-sabinene hydrate are not formed via sabinene or  $\alpha$ -thujene intermediates (Hallahan and Croteau, 1988). The characterization of TvTPS6 and TvTPS7

from *Thymus vulgaris* suggests that the reaction proceeds over the  $\alpha$ -terpinyl cation, the terpinen-4-yl cation, and the sabinyl cation, which is quenched by a water molecule to give the sabinene hydrate stereoisomers (Fig. 24). A striking feature of TvTPS6 and TvTPS7 is the opposite configuration of their chiral products at the carbon atom corresponding to C-4 in the  $\alpha$ -terpinyl cation. TvTPS6 forms primarily monoterpenes of the *S*-configuration series, while TvTPS7 forms monoterpenes of the *R*-configuration series (Figs. 10, 24). The presence of two terpene synthases with opposing stereospecificity has been observed in several other plants, including (-)- $\alpha$ -pinene synthase and (+)- $\alpha$ -pinene synthase of loblolly pine (Phillips et al., 2003), (+)-germacrene D synthase and (-)-germacrene D synthase from goldenrod (Schmidt et al., 1999), and sesquiterpene synthases TPS4 and TPS5 from maize (Köllner et al., 2004).



**Fig. 24 Putative reaction mechanism of TvTPS6 and TvTPS7.** The monoterpene products of TvTPS6 and TvTPS7 are derived from the  $\alpha$ -terpinyl cation intermediates by cyclizations, hydride shifts, and termination by proton loss or quenching with water. TvTPS6 forms primarily monoterpenes of the *S*-configuration series, while TvTPS7 forms monoterpenes of the *R*-configuration series.



To identify the step of the reaction mechanism that introduces the stereospecificity, (3*R*)- and (3*S*)-LPP were utilized as substrates. These likely pathway intermediates contain a stereocenter at C-3 and 6,1-cyclization of both enantiomers gives the central cyclic intermediate, the  $\alpha$ -terpinyl cation, in the respective configurations (Introduction I, Fig. 6). As it was shown previously for other monoterpene synthases, the configuration of the LPP intermediate determined the configuration of the TvTPS6 and TvTPS7 chiral products (Figs. 11, 12). The ability of monoterpene synthases to control the stereochemistry indicates that the conversion of GPP to LPP is stereospecific and most likely influenced by the initial helical fold of the native substrate. According to this theory, the left-handed folding facilitates the formation of (3*R*)-LPP and the right-handed folding yields (3*S*)-LPP (Croteau et al., 1986). This suggests that GPP is bound in the active center of TvTPS7 in the left-handed folding, whereas in TvTPS6 it is bound preferentially in the right-handed folding. However, the enzymes are not capable of excluding the unfavored LPP enantiomers (Figs. 11, 12). This may be due to the fact that both configurations of the intermediate show similar hydrophobic properties, especially in the *anti-endo* conformation (Satterwhite et al., 1985). However, for the (-)-bornyl diphosphate cyclase from tansy, a loss of function was reported when fed with the unnatural LPP enantiomer (Croteau et al., 1986). Additionally, the main product sabinene hydrate is formed by TvTPS7 almost exclusively in the (*E*)-configuration, while TvTPS6 formed both, (*E*)- and (*Z*)-sabinene hydrate. Thus, also the way water can approach the sabinyl cation intermediate seems to be influenced by the configuration of the molecules.

#### **4.1.3 One amino acid in the active site controls most of the stereospecificity of TvTPS6 and TvTPS7**

The 85 % sequence similarity between TvTPS6 and TvTPS7 enabled the identification of the structures in the active site that determine the stereospecificity of the enzymes. A similar approach was successful with two terpene synthases in maize (Köllner et al., 2004), while other pairs of stereospecific terpene synthases including (-)- $\alpha$ -pinene synthase and (+)- $\alpha$ -pinene synthase of loblolly pine share a sequence identity of only 66 % and are therefore not suitable for structure-function analyses (Phillips et al., 2003).

In TvTPS6 and TvTPS7, a single amino acid located seven positions upstream the magnesium-binding DDxxD motif, Asn-350 in TvTPS6 and Ile-346 in TvTPS7, was shown to determine the stereospecificity of the enzymes (Fig. 14). The corresponding Asn-338 residue in a 1,8-cineole synthase of *Salvia fruticosa* (SfCinS1) was demonstrated to be

involved in the binding of a water molecule that is used for the hydroxylation of intermediates (Kampranis et al., 2007). However, the formation of hydroxylated terpenes is not abolished by an Ile at the respective position in TvTPS7, suggesting that this residue is not directly involved in water binding and activation in TvTPS7. On the other hand, Asn-338 of SfCinS1 was shown to have no effect on the configuration of the chiral enzyme products (Fig. 17). This finding shows that amino acid residues located at the same site in terpene synthase active centers can have different catalytic functions.

To identify the amino acid residue(s) responsible for the binding of water in TvTPS6 and TvTPS7, site-directed mutagenesis on six amino acid residues with polar or hydrophilic properties was conducted (Fig. 16). The mutations of Tyr-432, Glu-435, Ser-457, Arg-498, Asp-501, and Glu-509 (numbers refer to TPS6) resulted in inactive enzymes, suggesting that these residues are important for protein stability or correct folding of the active site. Due to the loss of overall activity, it was not possible to determine whether these amino acid residues participate in water interaction.

#### **4.1.4 Modeling studies suggest a steric influence of Asn-350 (TvTPS6) and Ile-346 (TvTPS7) on the binding conformation of GPP**

The model of the TvTPS6 and TvTPS7 reaction centers (Fig. 15) illustrates the likely binding conformations of the GPP substrate. According to this model, the steric interaction between the amino acid residue and the carbon moiety of the GPP appears to be crucial. The Asn-350 residue (TvTPS6) reached further into the active site cavity than the Ile-346 residue (TvTPS7) and might lead to the right-handed folding of the GPP substrate. The position of the GPP molecule in TvTPS6 would then result in the formation of the (3*S*)-LPP intermediate and determine the configuration of the reaction products. With an Ile residue in this position, like in TvTPS7, the cavity might be bigger and the substrate could be more flexible. The left-handed conformation of the GPP might be energetically favored in the larger cavity and thereby preferred in TvTPS7. Also, the right-handed conformation of the substrate might be destabilized by specific amino acid residues that change the steric or charge interactions in the active site of TvTPS7. Thus, the two conformations illustrated in Fig. 15 could represent the right- and left-handed helical fold of GPP that determines the stereospecific reaction.

The mutation studies showed a complete reversal of stereospecificity when Asn-350 was substituted by Ile in TvTPS6 (Fig. 14). This might be explained by an increased space in the active center that allows for an energetically favored conformation of GPP or a steric

clash of Ile with GPP in TvTPS6. In TvTPS7, the opposite substitution (Ile-346 to Asn) did not completely invert the product spectrum (Fig. 14). This indicates that an Asn at this position in TvTPS7 causes the formation of equal amounts of both enantiomers of LPP. There are no obvious amino acid differences nearby to explain these relative effects since they are similar to those of TvTPS6 in size and chemical properties. Most likely, this effect is due to conformation changes of the enzyme structure that are caused by distant amino acid residues. Such effects of distant residues on the active site conformation have already been shown for the 5-*epi*-aristolochene synthase from *Nicotiana tabacum* and the premnaspirodiene synthase from *Hyoscyamus muticus*. According to modeling studies, the amino acids in direct contact with the substrate were almost identical despite a differing product outcome. Hence, it was presumed that amino acids surrounding the active site influenced the interaction of those residues reaching directly in the active site cavity (Greenhagen et al., 2006).

The position of Asn-350 (TvTPS6) and Ile-346 (TvTPS7) supports the observation in a study of Köllner *et al.* (2004), who located the crucial amino acids for stereospecificity in the terpene synthases TPS4 and TPS5 of maize also at the bottom of the active site. Also, Schwab *et al.* (Schwab et al., 2001) compared the active site openings of terpene synthases which generate antipodal configurations of LPP and recognized no significant differences. Thus, the mechanistic differences were assumed to reflect in structural differences deeper in the active site pocket.

Interestingly, two stereospecific (*R*)-limonene synthases from *Schizonepeta tenuifolia* and *Agastache rugosa* contain an Ile residue seven amino acids upstream the DDxxD motif, while several (*S*)-limonene synthases of *Mentha longifolia*, *Mentha spicata*, *Perilla frutescens*, and *Perilla citriodora* contain an Asn residue at this position (Maruyama et al., 2001; Maruyama et al., 2002). This may indicate a general function of this amino acid position in the stereocontrol of monoterpene synthases.

## **4.2 Characterization and stereochemical aspects of the reaction mechanism of the $\alpha$ -terpineol synthase TvTPS5**

### **4.2.1 TvTPS5 may be the main monoterpene synthase of the $\alpha$ -terpineol chemotype of *Thymus vulgaris***

$\alpha$ -Terpineol and its derivate  $\alpha$ -terpineol acetate are the main components of the *Thymus vulgaris*  $\alpha$ -terpineol chemotype's essential oil. TvTPS5, a terpene synthase isolated from this chemotype, produces mainly  $\alpha$ -terpineol and next to this 11 minor monoterpene

products which can all be found in the plant (Fig. 19). The chemotype's monoterpene blend consists of six additional compounds which are not formed by TvTPS5. These are *p*-cymene, 1,8-cineole, camphor, borneol, bornyl acetate, and  $\alpha$ -terpineol acetate. Most of them are presumably the products of rearrangements or metabolization of the other monoterpenes produced by TvTPS5. As in the (*E*)-sabinene hydrate chemotype, *p*-cymene might be the result of spontaneous oxidation of  $\gamma$ -terpinene or the product of a P450 monooxygenase. 1,8-Cineole could be derived from  $\alpha$ -terpineol by spontaneous endocyclization. However, terpene synthases producing 1,8-cineole are already known from other plant species (see Results I, section 3.3) and thus another 1,8-cineole synthase could be responsible for the 1,8-cineole formation in thyme, too. Finally,  $\alpha$ -terpineol acetate is most likely formed from  $\alpha$ -terpineol and acetyl CoA by the action of an acyltransferase. Only borneol might be the product of an additional monoterpene synthase. Bornyl acetate and camphor could both be formed by further metabolization of borneol.

Gene expression analysis by real-time quantitative PCR showed that *Tvtps5* is expressed in high levels in the leaves of the  $\alpha$ -terpineol chemotype (Schimmel, 2014). This result supports the assumption that TvTPS5 is the main monoterpene synthase responsible for the formation of most of the monoterpenes found in the essential oil of the  $\alpha$ -terpineol chemotype.

#### **4.2.2 TvTPS5 and TvTPS7 show a close phylogenetic relationship and a similar stereospecificity**

TvTPS5 showed a high amino acid sequence similarity to TvTPS7 from the (*E*)-sabinene hydrate chemotype. An Ile found in the TvTPS5 sequence (Ile-346) at the amino acid position responsible for the stereospecificity of TvTPS6 and TvTPS7 seems to control the reaction mechanism in TvTPS5 in the same way as in TvTPS7, where also an Ile was found at this position. TvTPS5 preferably formed the *R*-enantiomers of the chiral monoterpene products, but when Ile-346 was altered into Asn, the amount of the *S*-enantiomers increased slightly (Fig. 20). The same effect was observed with TvTPS7 and its mutant TvTPS7 I364N. Since according to modeling studies Ile-346 in TvTPS5 had the same steric orientation in the active site center as in TvTPS7 (Fig. S3), the change of stereospecificity in TvTPS5 could be explained in the same way. The steric modification of the active site must have changed the binding conformation of GPP, so that consequentially the formation of the (*S*)-LPP intermediate increased. Feeding of the racemate of LPP to the native TvTPS5 resulted in almost the same enantiomeric pattern of the main product  $\alpha$ -terpineol

as with the TvTPS5 I364N mutant in the presence of the GPP substrate (Fig. 20). This finding supports the theory that the binding conformation of GPP was changed by the mutation of Ile-346 to Asn, leading to differing amounts of (*R*)- and (*S*)-LPP during the reaction pathway. However, the stereospecificity was not reversed completely.

Dendrogram analyses show that *Tvtps5* may have evolved from *Tvtps7* (Fig. 18). The Maximum-Likelihood dendrogram shows that *Tvtps5*, *Tvtps6*, and *Tvtps7* cluster with sabinene synthases and  $\alpha$ -terpineol synthases from other *Lamiaceae* species. Another group of genes encoding for 1,8-cineole synthases from *Lamiaceae* species cluster in a second group which is more distantly related. Presumably, the catalytic function to form 1,8-cineole from  $\alpha$ -terpineol through an additional protonation has evolved independently from the  $\alpha$ -terpineol synthase activity in *Lamiaceae*.

#### **4.2.3 The two amino acids difference between the TvTPS5 and TvTPS7 active sites may affect water binding**

Both the main products from TvTPS5 and TvTPS7 are hydroxylated at different steps of the carbocationic reaction pathway. In TvTPS5, an immediate water quenching occurs at C-8 of the  $\alpha$ -terpinyl cation as main reaction to form  $\alpha$ -terpineol. In TvTPS6 and TvTPS7, water quenching takes place predominantly at C-1 after two additional steps, a hydride shift to C-4 (terpinen-4-yl cation) and the subsequent endocyclization to the sabinyl cation (Fig. 24). Sequence comparison and structure modeling of the closely related  $\alpha$ -terpineol synthase TvTPS5 and the (*E*)-sabinene hydrate synthase TvTPS7 revealed only two differing amino acids whose residues reach into the active center. Those are Cys-322 and Leu-491 in TvTPS7, while a Ser was found in TvTPS5 at both corresponding positions (Fig. 21). The serines in TvTPS5 carry a hydrophilic hydroxyl group in their residues, which may be able to interact with water through a hydrogen bridge bond. As a consequence, water could be bound and activated at the two sites of the active site pocket entrance, ready to quench the carbocationic reaction mechanism when the  $\alpha$ -terpinyl cation is formed. This results in the formation of  $\alpha$ -terpineol. Cys-322 in TvTPS7 contains a thiol group, but the dipole of this side chain is generally too weak to undergo a hydrogen bridge bond. Leu-491 on the other hand is a hydrophobic amino acid and not involved in water binding. Only, due to its longer side chain, it could influence the binding conformation of GPP in TvTPS7 so that water attack is impeded sterically. In TvTPS7 (and TvTPS6), the amino acids involved in water binding could not be identified. It is conceivable that they are located at different positions than in TvTPS5, causing the hydroxylation of the

intermediates at different sites. These assumptions need still to be verified by site-directed mutagenesis studies.

In addition, more distant regions of the enzyme could affect binding and access of water to the active site pocket. The N-terminus of the enzymes, for example, is positioned over the entrance of the active pocket after substrate binding (Starks et al., 1997) and thus could prevent water access to a different extent in both enzymes.

#### **4.3 The monoterpene synthase genes isolated from *Melaleuca trichostachya* and *Melaleuca linariifolia* show a close relationship to 1,8-cineole and sabinene hydrate synthases from other *Myrtaceae***

In this study, four monoterpene synthase genes from *Melaleuca linariifolia* and *Melaleuca trichostachya* were isolated. These genes are highly identical to other 1,8-cineole synthases and sabinene hydrate synthases from other *Myrtaceae*, but their function is yet to be determined. In a dendrogram analysis, the sequences cluster with 1,8-cineole and sabinene hydrate synthases from *Melaleuca alternifolia* (Fig. 22), so that a close relationship can be assumed.

The sabinene hydrate synthase MaTPS2 from *Melaleuca alternifolia* is highly expressed in the terpinen-4-ol chemotype (A. Keszei, ANU Canberra, Australia, unpublished results). It is proposed that sabinene hydrate rearranges non-enzymatically to terpinen-4-ol (Cornwell et al., 1995). Presumably, the same fact is true for the terpinen-4-ol chemotypes of other *Melaleuca* species. In *Myrtaceae*, this is the case to a much larger extent than in *Thymus* species, for example (Fig. 8, Table 2). Most likely, this is caused by different storage of the essential oil. In thyme, terpenes are secreted in subcuticular storage cavities of glandular trichomes (Fig. 1) (Stahl-Biskup, 2002), where they are protected from hydrolyzing influences. In *Melaleuca* and *Eucalyptus* species, these secondary compounds are stored in subepidermal schizogenous or lysigenous secretory cavities in the leaf (Fig. 3) (List, 1995), where they could be attacked by hydrolyzing or oxidizing components of the lysogenic plant cells. Also, the duration of storage may be longer in sub-epidermal oil cavities.

Another interesting fact is that the ratio of (*Z*)- and (*E*)- sabinene hydrate differs between *Melaleuca linariifolia* and *Melaleuca alternifolia* flush leaves. In *Melaleuca alternifolia*, the (*Z*)-isomer is dominant over the (*E*)-isomer in a 7.1:1 ratio. In *Melaleuca linariifolia*, the (*E*)-isomer is present in a higher amount, the ratio is 0.65:1 (Southwell and Stiff, 1990). In order to study the mechanism on how the formation of the respective stereoisomers is influenced, more sabinene hydrate synthases from these *Melaleuca* species have to be

isolated. Structure modeling and sequence comparison to other enzymes like the sabinene hydrate synthases TvTPS6 and TvTPS7 from thyme could reveal important features of the enzymes that change substrate or water binding, resulting in these different (*Z*)/(*E*)-ratios of sabinene hydrate.





## **Chapter II**

**Monoterpene modification by cytochrome P450  
monooxygenases in essential oil plants**

## 1 Introduction

### 1.1 Cytochrome P450 monooxygenases and their role in plant oxidative metabolism

In plants, the biosynthesis of many secondary metabolites requires oxygenation reactions, especially hydroxylations of unactivated hydrocarbons (Hamdane et al., 2008). Oxygenating enzymes can generally be divided in dioxygenases and monooxygenases, depending on how many oxygen atoms are incorporated into the substrate. Upon the reaction of monooxygenases, the second atom of molecular oxygen is reduced to water. Among the family of monooxygenases, cytochrome P450 monooxygenases are prevalent (Cochrane and Vederas, 2014). These enzymes are found in all life forms and form one of the the largest protein family in plants (Podust and Sherman, 2012). For example, 244 cytochrome P450 genes were found in the *Arabidopsis thaliana* genome (Bak et al., 2011). Plant P450 enzymes are membrane-bound hemoproteins, most of them are anchored to the endoplasmatic reticulum (ER) on the cytosolic site via a hydrophobic N-terminal anchor segment (Bolwell et al., 1994; Halkier, 1996). Heme b is located as prosthetic group in the active site. It is anchored at its iron center through a conserved cysteine residue of the surrounding apoprotein (Hasemann et al., 1995). In general, the activity of P450 enzymes is dependent on electron-donating flavoproteins. They provide the reduction equivalents from a cofactor, which are needed during the catalytic cycle of P450 enzymes. In plants, P450 enzymes bound to the ER require a cytochrome P450 reductase (CPR) which contains flavin adenine dinucleotide and flavin mononucleotide (Halkier, 1996, Bak et al., 2011), and NADPH serves as cofactor. As the CPR is membrane-bound as well, these enzymes are closely associated with P450s at the surface of the ER. The details of their interaction, however, are still under investigation (Jensen and Møller, 2010).

### 1.2 Nomenclature and gene organization of P450 monooxygenases

The nomenclature of P450 enzymes is based on phylogenetic grouping and the amino acid sequences of the proteins (Nelson, 2006). All cytochrome P450 names begin with the “CYP” prefix, followed by an Arabic numeral which designates the P450 family. Among single families, the amino acid sequence identity is at least 40 % or higher. Each family is divided in further subfamilies, which share a sequence identity of at least 55 %. They are denoted with capital letters. The last Arabic numeral stands for the respective isoform. A central cytochrome P450 nomenclature committee (David Nelson, dnelson@uthsc.edu) is

responsible for the naming. Usually, isoforms are named in chronological order of submission (Bak et al., 2011). Plant P450 enzymes range from CYP families 71 to 91 and 701 and above. In order to find a higher-order classification of the rapidly growing number of identified P450 enzymes, clans were defined, each of them consisting of the P450 families which group together in phylogenetic analyses. They are named after the member with the lowest family number. 14 plant P450 clans were identified, with CYP71 being the largest clan (Nelson et al., 2004).

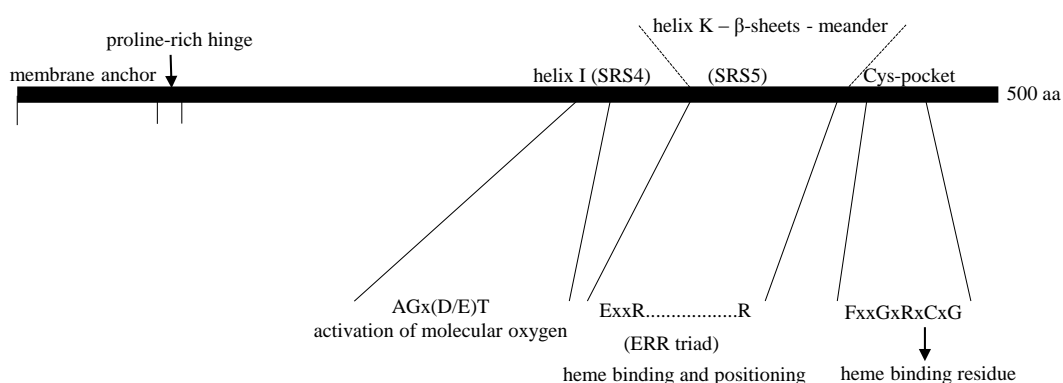
### 1.3 Overall cytochrome P450 architecture and structure-function relationships

Sequence identities between P450 enzymes can be as low as 20 %, but the three-dimensional structures solved so far suggest that the overall fold is conserved (Poulos and Johnson, 2005; Rupasinghe and Schuler, 2006). Eleven  $\alpha$ -helices (A-K) and four  $\beta$ -sheets (1-4) are common structural elements. They contain the active site and carry the overall core structure (Graham and Peterson, 1999). To date, no crystal structure of plant P450 enzymes has been solved. The X-ray diffraction analysis is hampered by the membrane-anchorage of plant P450 enzymes. The synthesis of functional soluble plant P450 derivatives, however, did not lead to enhanced results. Therefore, only models based on available crystallographic templates can be used for structural analysis of plant P450s, even though the sequence similarities of plant P450s and the respective templates are often very low. This circumstance makes it difficult to deduce valid assertions from the model alone (Rupasinghe and Schuler, 2006).

Average P450 primary sequences comprise about 500 amino acids. Between enzymes, usually low sequence similarities are found, the differences can include over 80 % of the sequence. Within these non-synonymous segments, some defined regions are thought to determine the substrate specificity of the single enzymes. Six substrate recognition sites (SRS) were originally described for the family CYP2, which differ significantly between enzymes but had crucial functions in substrate recognition and binding (Gotoh, 1992). As it turned out, these SRS can be assigned to the other CYP families as well. However, not only primary sequences but secondary structures need to be taken into account for correct SRS assignment (Rupasinghe and Schuler, 2006).

Nevertheless, there are few highly conserved sequence motifs which are crucial for the general activity of all P450 enzymes (Fig. 25). There is the proline-rich membrane hinge, which separates the N-terminal hydrophobic anchor segment from the cytosolic domain.

Additionally, this hinge seems to be responsible for the correct incorporation of heme in the active center (Halkier, 1996). Furthermore, there is the AGx(D/E)T motif in helix I (SRS 4), in which especially the threonine residue was found to be important for oxygen protonation. This protonation leads to the activation of the bound molecular oxygen during the catalytic cycle (Atkins and Sligar, 1988; Podust and Sherman, 2012). In addition, there are the ExxR motif in helix K just in front of SRS 5, and the FxxGxRxCxG motif containing the heme binding cysteine residue. This cysteine is the only absolutely conserved residue in all known P450 enzymes, it functions as the heme iron proximal ligand through a thiolate bond (Podust and Sherman, 2012).

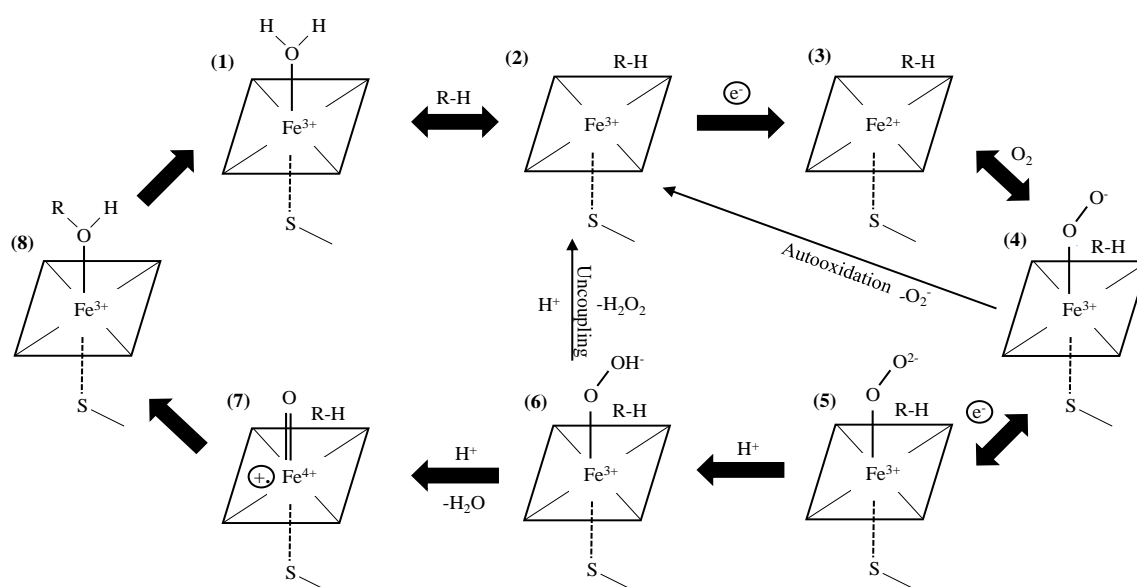


**Fig. 25 Scheme of the universal P450 enzyme sequence motifs and their respective functions.** Modified from (Bak et al., 2011).

#### 1.4 Reaction mechanism of P450 enzymes: The catalytic cycle

Despite the diverse reactions catalyzed by P450 enzymes, the general reaction mechanism is thought to be similar. The conventional catalytic cycle (Fig. 26) starts with the resting state (1), in which the iron is in the ferric state and a water molecule is bound in the active center (low spin state). Upon substrate binding, the water molecule is displaced (2), resulting in an increase of the redox potential of the heme iron due to the iron shift from a hexa-coordinated (low spin state) to a penta-coordinated state (high spin state). This initiates the redox partner to transfer one electron from NADPH to the P450 enzyme, yielding the reduced ferrous state (3). Molecular oxygen is subsequently bound, leading to the ferrous dioxygen complex (4). A second electron transfer results in the peroxo intermediate (5) and the subsequent proton transfer from water gives the hydroperoxo intermediate (Compound 0) (6). The following heterolytic O-O cleavage is achieved by a second protonation and release of water and yields the highly reactive oxyferryl

intermediate (Compound I) (7). This intermediate is believed to transfer one oxygen atom to the bound substrate via abstraction of a hydrogen atom of the substrate and radical recombination (8). The last step is the release of the oxygenated product and the return to the initial resting state. With the exception of the formation of the activated Compound I, all other steps are radical processes (Halkier, 1996). The ferrous dioxygen complex (4) can also undergo autooxidation to the ferric P450 and release of a superoxide anion. Furthermore, uncoupling reactions can occur, such as the release of hydrogen peroxide from the hydroperoxo intermediate after protonation (6) (Halkier, 1996; Hamdane et al., 2008; De Montellano, 2010; Podust and Sherman, 2012; Cochrane and Vederas, 2014).

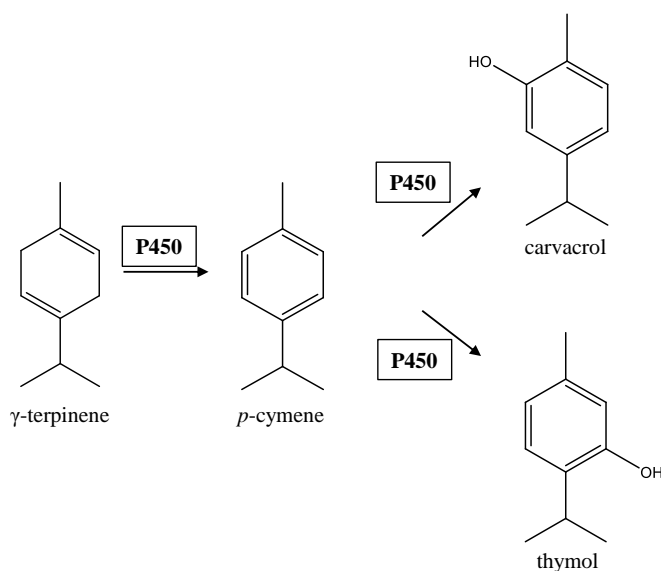


**Fig. 26 The catalytic cycle of P450 enzymes.** The heme is represented by black boxes. It is anchored at its iron center (Fe) through a conserved cysteine residue (S) of the surrounding apoprotein. R-H: hydrocarbon substrate; R-OH: hydroxylated product. The reaction is explained in the section above (1.4). Figure modified from (Hamdane et al., 2008; Cochrane and Vederas, 2014).

### 1.5 P450 enzymes participate in thymol and carvacrol formation in thyme and oregano

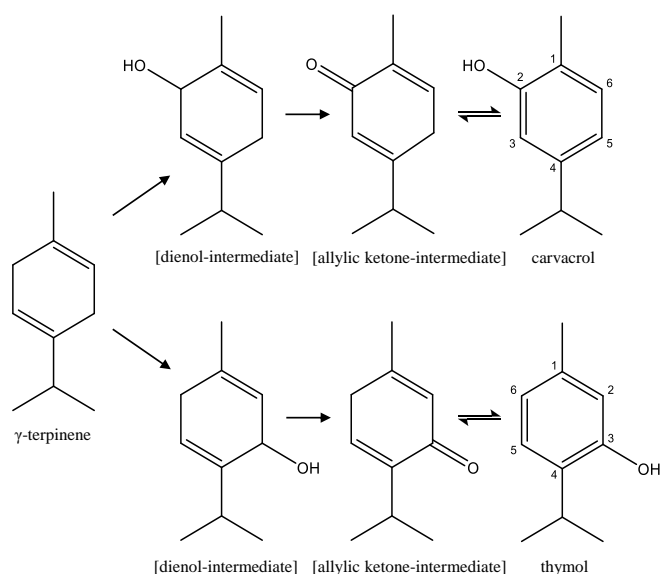
The biosynthesis of the phenolic monoterpenes is particular in that they are aromatic compounds and bear a phenolic scaffold. Thus, their formation involves not only terpene synthases but other metabolizing enzymes able to form aromatic rings, like cytochrome P450 enzymes. In previous studies,  $\gamma$ -terpinene was suggested to be the precursor of thymol and carvacrol. An intermediate of this pathway was thought to be *p*-cymene (Fig. 27).  $^3\text{H}$ -labelling studies showed that  $\gamma$ -terpinene was converted into thymol and carvacrol after incubation with young thyme leaves (Poulose and Croteau, 1978). This study also proved that it is not just a spontaneous rearrangement of  $\gamma$ -terpinene, but an enzymatic reaction.

Furthermore, in plant species which contain either thymol or carvacrol in their essential oil, there are always considerable amounts of  $\gamma$ -terpinene and *p*-cymene found. This lead to the conclusion that these monoterpenes are part of one biosynthetic pathway (Stahl-Biskup, 2002).



**Fig. 27** Originally suggested pathway from  $\gamma$ -terpinene to thymol and carvacrol postulated by Poulou and Croteau (Poulou and Croteau, 1978).  $\gamma$ -Terpinene is metabolized to carvacrol and thymol by P450 enzymes via *p*-cymene as intermediate.

The formation of thymol and carvacrol in *Thyme* and *Oregano* species was also in the focus of a study by C. Crocoll (Crocoll, 2011). He identified eleven cytochrome P450 gene sequences. The deduced amino acid sequences were summarized to five cytochrome P450 enzymes of the CYP71D subfamily: CYP71D178-182. Since the expression of the corresponding genes correlated with thymol and carvacrol presence *in planta*, these enzymes were supposed to be responsible for their formation. After heterologous expression of CYP71D178, CYP71D180, and CYP71D181 in *Saccharomyces cerevisiae* (*S. cerevisiae*), mainly *p*-cymene formation next to trace amounts of thymol and carvacrol was detected. However, limonene was accepted as substrate as well, leading to the either C-3 or C-6 hydroxylated allylic alcohols isopiperitenol and carveol. Based on this result, a new pathway from  $\gamma$ -terpinene to thymol and carvacrol was postulated (Fig. 28). Therein, not *p*-cymene was assumed to be the intermediate but a dienol compound corresponding to the allylic alcohols formed from limonene, which is then oxidized to the respective ketones able to undergo a spontaneous aromatization (keto-enol-tautomerism) to thymol and carvacrol. Thus, *p*-cymene was assumed to be not an intermediate of the pathway, as previously postulated.



**Fig. 28 Pathway from  $\gamma$ -terpinene to thymol and carvacrol suggested by C. Crocoll (Crocoll, 2011).** The model suggests two oxidations steps that are catalyzed by a single cytochrome P450 enzyme. First,  $\gamma$ -terpinene is hydroxylated at either C-3 or C-6 (C-2). In a second step, the dienol-intermediates are oxidized to the respective ketones, which can undergo a keto-enol-tautomerism to form carvacrol or thymol.

Furthermore, C. Crocoll could show that after overexpressing CYP71D178 and CYP71D180 in *Arabidopsis thaliana* and feeding with  $\gamma$ -terpinene, carvacrol was produced in the transgenic plants and stored as glycosides. In *Arabidopsis* plants overexpressing CYP71D178, also traces of glycosidically bound thymol were found (Crocoll, 2011). To date, however, definite proof of the biosynthetic roles of these P450 enzymes has not been found.

## 2 Materials and Methods

For microbiological methods and nucleic acid techniques see Materials and Methods I, section 2.2 and 2.3.

### 2.1 Plant material

Thyme chemotypes (*Thymus vulgaris*) were collected in Southern France at CNRS, Montpellier, France (Thompson et al., 2003), implanted in outdoor fields at the university campus in Halle (Saale), Germany, and grown under natural conditions.

Leaves from eucalyptus (*Eucalyptus dives*, *Eucalyptus piperita*, and *Eucalyptus elata*) were collected at the campus of the Australian National University, Canberra, Australia.

### 2.2 Isolation of cytochrome P450 genes from *Eucalyptus* species

An *Eucalyptus grandis* proteome BLAST search on the Phytozome database (<http://www.phytozome.net/search.php>) was performed using the protein sequence of the limonene-C3-hydroxylase CYP71D13 (Lupien et al., 1999). The result contained about 20 sequences, which showed a sequence identity of 55 % or more and thus probably belonged to the same subfamily of cytochrome P450 enzymes. Some of these sequences were used to design primers for the sequence isolation of cytochrome P450 monooxygenases from RACE cDNAs of *Eucalyptus dives*, *Eucalyptus piperita*, and *Eucalyptus elata*. The components of the PCRs were: 0.5 µl Advantage Taq DNA Polymerase Mix (5 U/µl), 2.5 µl Advantage Taq PCR buffer, 0.5 µl dNTPs (10 mM each), 1 µl forward primer (10 µM) and 1 µl reverse primer (10 µM) (Table S2), 2 µl 5'-RACE cDNA and PCR grade water added to a final volume of 25 µl. PCR thermocycles were run as follows: initial denaturation at 95 °C for 2 min, 35 cycles of denaturation at 95 °C for 30 s, annealing at 53 °C for 30 s, extension at 72 °C for 1.5 min, and a final step at 72 °C for 5 min. The obtained PCR fragments were cloned into the pCR4-TOPO vector and subsequently sequenced. The sequences were called *CYP71AP17* (*Eucalyptus dives*), *CYP71AP18* (*Eucalyptus piperita*), *CYP71AP19* (*Eucalyptus elata*), *CYP71AH18* (*Eucalyptus piperita*), and *CYP71CJ1* (*Eucalyptus elata*) (D. Nelson, cytochrome P450 nomenclature committee).

### 2.3 *In vitro* mutagenesis of P450 monooxygenase open reading frames

Site-directed mutagenesis was performed according to Materials and Methods I, section 2.5. The PCR based mutagenesis protocol was performed with the ORFs of *CYP71D179*



and *CYP71D181* cloned into the expression vector pESC-LEU2d (Ro et al., 2008). Previously, the ORFs were optimized for yeast codon usage and synthesized by Genart (Regensburg, Germany).

## 2.4 Yeast expression system

### 2.4.1 Yeast strain

For the heterologous expression of cytochrome P450 enzymes, the engineered *Saccharomyces cerevisiae* strain WAT11 was used. Its genome contains the gene encoding for ATR1, the *Arabidopsis thaliana* NADPH-P450 reductase (Pompon et al., 1996).

### 2.4.2 Media for yeast cultivation, transformation, and protein overexpression

|                |   |
|----------------|---|
| YPGA (Glc):    | 10 g yeast extract (AppliChem GmbH, Darmstadt, Germany), 20 g Bacto Peptone (Beckton, Dickinson; Le Point de Claix, France), 74 mg adenine hemisulfate (Sigma-Aldrich, St. Louis, MO, USA), 100 ml glucose solution (200 g/l), and H <sub>2</sub> O added to 1 l  |
| YPGA (Gal):    | 10 g yeast extract, 20 g Bacto Peptone, 74 mg adenine hemisulfate, 100 ml galactose solution (200 g/l), and H <sub>2</sub> O added to 1 l   |
| YPGA agar:     | YPGA (Glc) medium components, 20 g agar-agar (Roth, Karlsruhe, Germany), and H <sub>2</sub> O added to 1 l  |
| SC-Leu medium: | 6.7 g yeast nitrogen base (without amino acids, with sulfate) (Sigma-Aldrich, St. Louis, MO, USA),<br>100 mg of each adenine hemisulfate, arginine, cysteine, lysine, threonine, tryptophan, uracil, and 50 mg of each aspartic acid, histidine, isoleucine, methionine, phenylalanine, proline, serine, tyrosine, and valine (all amino acids from Roth, Karlsruhe, Germany)<br>100 ml glucose solution (200 g/l), and H <sub>2</sub> O added to 1 l |
| SC-Leu agar:   | SC-Leu medium components, 20 g agar-agar (Roth, Karlsruhe, Germany), and H <sub>2</sub> O added to 1 l  |

### 2.4.3 Yeast transformation

The transformation of *S. cerevisiae* strain WAT11 was conducted according to a method originally described by Gietz and Woods (Gietz and Woods, 2002). 30 ml YPGA (Glc) medium were inoculated with a WAT11 clone and grown overnight at 180 rpm and 28 °C. 50 ml of YPGA (Glc) was inoculated with this culture to an OD(600) of 0.4. After 4 h

incubation at 180 rpm and 28 °C, the culture was centrifuged for 5 min at 4,000 x g and 4 °C. The yeast cells were washed in 30 ml H<sub>2</sub>O, centrifuged under the same conditions as before and washed in 1 ml H<sub>2</sub>O. After centrifugation at 17,900 x g for 30 s, the cells were resuspended in 1 ml H<sub>2</sub>O and 100 µl aliquots of this yeast cell suspension were used for each transformation.

Following transformation mix was prepared: 240 µl PEG 3500 (50 % w/v), 36 µl 1 M lithium acetate, and 50 µl boiled salmon sperm carrier DNA (2 mg/ml). 34 µl H<sub>2</sub>O with 0.5-1.0 µg plasmid were added. 100 µl fresh cell suspension were briefly centrifuged, mixed with the transformation mix, and incubated in a water bath at 42 °C for 40 min. The transformed cells were briefly centrifuged and resuspended in 1 ml H<sub>2</sub>O. 300 µl of the cell suspension were plated on a SC-Leu agar plate and incubated at 28 °C for 3-4 days.

#### **2.4.4 Cytochrome P450 monooxygenase overexpression and microsome extraction**

The isolated cytochrome P450 genes were cloned into the pESC-LEU2d yeast expression vector (Ro et al., 2008) by sticky end cloning (for primers used, see Table S2) with the appropriate restriction enzymes. The vector contains the auxotrophic selection marker gene LEU2d, which allows maintaining the expression vector in *S. cerevisiae* cells on media without leucin. LEU2d is a modification from the selection marker LEU2 with a shortened length of the promoter sequence, resulting in higher copy numbers in *S. cerevisiae* in order to compensate the low LEU2 expression (Ro et al., 2008). The vector carries the GAL1 and GAL10 yeast promoters, which are repressible with glucose.

To prepare the expression culture, 30 ml SC-Leu medium were inoculated with a WAT11 clone transformed with the desired pESC-Leu2d plasmid, and the cells were grown overnight at 180 rpm and 28 °C. 100 ml YPGA (Glc) medium were inoculated with one unit OD(600) of the overnight culture, and the cells were grown under the same conditions for 30 h. After centrifugation (5,000 x g, 18 °C, 5 min) the enzyme expression was induced by resuspension of the cells in 100 ml YPGA (Gal) medium and the expression culture was shaken for 18 h (160 rpm, 25 °C).

During the following steps of yeast microsome extraction, the yeast cells were constantly cooled on ice or in the machines at 4 °C, to avoid enzyme degradation. The expression culture was centrifuged (7,500 x g, 10 min) and the cells were resuspended in 30 ml TEK buffer (50mM Tris-HCl pH 7.5, 1mM EDTA, 100 mM KCl). After centrifugation (7,500 x g, 10 min), the cells were resuspended in 2 ml TES buffer (50 mM Tris-HCl pH

7.5, 1 mM EDTA, 600 mM sorbitol, 10 g/l Bovin serum fraction V protein and 1.5 mM  $\beta$ -mercaptoethanol). For cell disruption, glass beads (0.45-0.5 mm diameter, Sigma-Aldrich, St. Louis, MO, USA) were added until they reached the upper level of the culture, then the cultures were shaken by hand 5 times for 1 min, with cooling on ice for 1 min in between, respectively. 5 ml TES buffer were added, the disrupted cell culture was gently shaken and the supernatant was carefully removed and stored in a new falcon tube. This step was repeated three times, the last time all of the liquid was removed from the glass beads. After centrifugation (7,500 x g, 10 min), the supernatant was transferred into an ultracentrifuge tube and the cell lysate including the microsomes was centrifuged in an ultracentrifuge (100,000 x g, 90 min). The supernatant was carefully removed and the microsome pellet was gently washed with 5 ml TES buffer, then with 2.5 ml TEG buffer (50 mM Tris-HCl pH 7.5, 1 mM EDTA, 30 % glycerol). The microsome pellet was transferred into a 2 ml glass homogenizer and thoroughly homogenized in 2 ml TEG buffer. The microsomes were stored at -80 °C or immediately used for enzyme assays.

## 2.5 P450 enzyme assay

For standard cytochrome P450 enzyme activity assays, the following mixture was prepared in a 1.5 ml GC vial: 225  $\mu$ l sodium phosphate buffer (75 mM) pH 7.4, 50  $\mu$ l H<sub>2</sub>O, 1.2  $\mu$ l monoterpene substrate (25 mM in DMSO), 20  $\mu$ l microsomes, and 3  $\mu$ l NADPH (100 mM) to start the reaction. The vial was placed in a water bath at 30 °C for 1 h, with a SPME fibre (100  $\mu$ m polydimethylsiloxane; Supelco, Belafonte, PA, USA) positioned in the headspace of the assay for the collection of volatile monoterpene products. The enzyme products were analyzed according to Methods and Materials I, section 2.11.

For the determination of temperature dependent reactions, the incubation temperature was increased from 16 °C to 34 °C in 6 °C steps.

## 2.6 Limonene feeding of thyme cuttings

Cuttings of each *Thymus vulgaris* chemotype were placed in small beakers with water in two big glass vessels. In each vessel, either 100  $\mu$ l (*R*)-limonene or 100  $\mu$ l (*S*)-limonene (Sigma-Aldrich (Fluka), St. Louis, MO, USA) were added in a small beaker next to the cuttings. Limonene evaporates at room temperature and due to its lipophilic character it is able to penetrate through plant cuticulas and cell walls. The thyme cytochrome P450 enzymes were expected to remain their activity for at least 24 h after harvesting of the

stems. After 24 h, the cuttings of each chemotype were removed from the vessel, ground in liquid nitrogen, and stored at -80 °C. As a negative control, cuttings of each chemotype were removed before the experiment and processed the same way as the limonene treated cuttings.

15 mg ground leaf powder of each sample were extracted with 200 µl hexane for one hour at room temperature in a 1.5 ml GC vial. The extracted terpenes were analyzed according to Materials and Methods I, section 2.11.

### **2.7 *Thymus vulgaris* transcriptome sequencing**

The sequencing of total RNA isolated from the thymol chemotype and the carvacrol chemotype of *Thymus vulgaris* was conducted at the Max Planck-genome-centre Cologne, Germany. A HiSeq2500 Illumina sequencing with 75.000.000 reads per run was requested.

### **2.8 Modeling of CYP71D179**

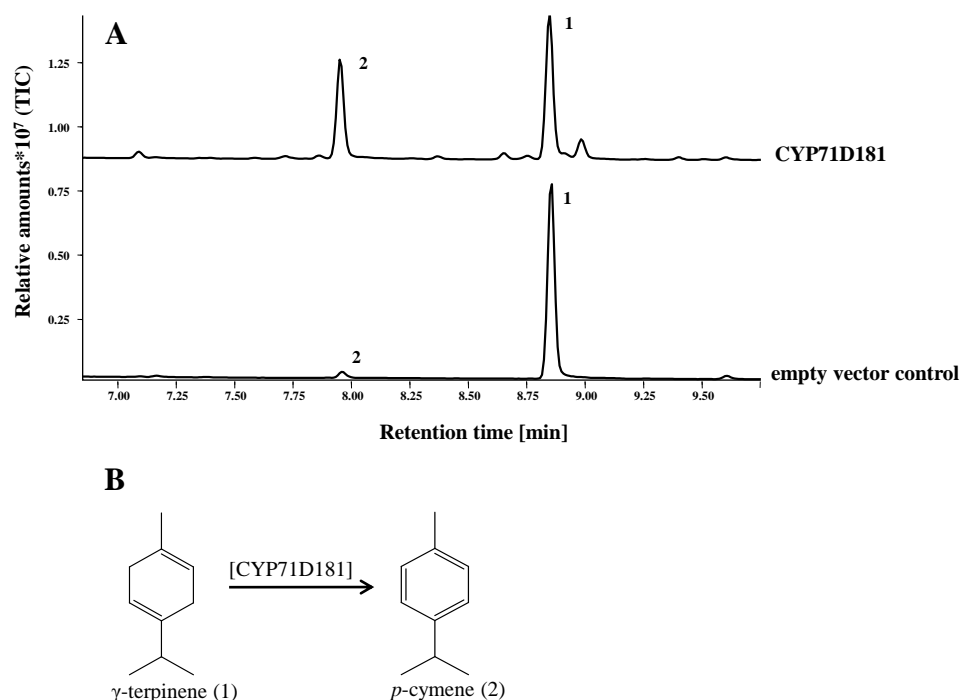
A model of the three-dimensional structure of CYP71D179 was generated using the Swiss Model Server (<http://www.expasy.org>) (Schwede et al., 2003; Arnold et al., 2006). For modeling, the amino acid sequence was fitted to the template structure of mammalian cytochrome P450 2B4 (PDB code: 1po5A) (Scott et al., 2003). The resulting model was visualized with the program PyMOL (DeLano, 2002).

### 3 Results

#### 3.1 The formation of the phenolic monoterpenes thymol and carvacrol from a monoterpene precursor – one cytochrome P450 reaction or a multiple step pathway?

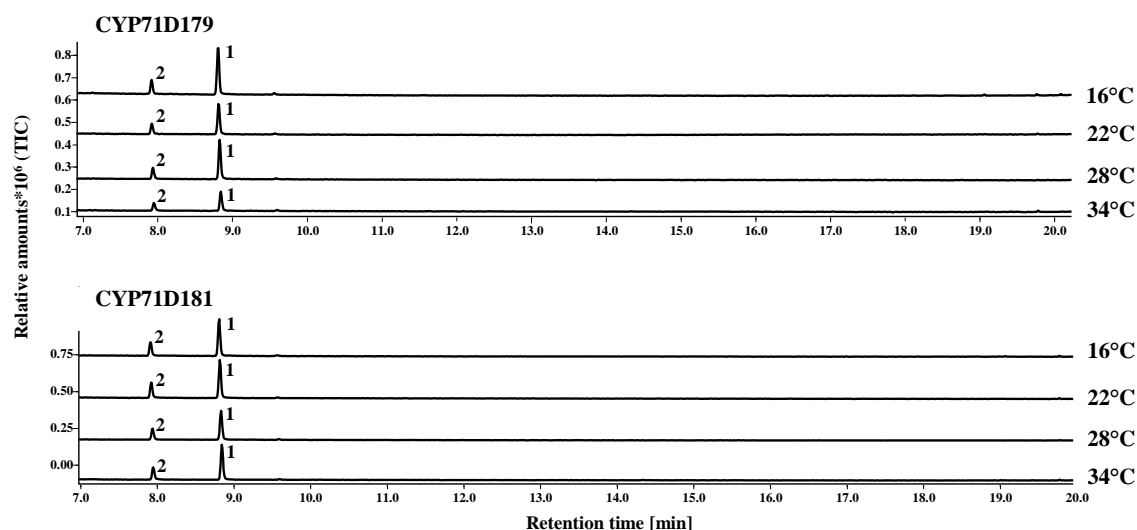
##### 3.1.1 CYP71D179 and CYP71D181 form *p*-cymene from $\gamma$ -terpinene

The phenolic monoterpenes thymol and carvacrol, which occur in large amounts in the essential oil of the thymol chemotype (T-type) and carvacrol chemotype (C-type) of *Thymus vulgaris*, were assumed to be formed from  $\gamma$ -terpinene by cytochrome P450 monooxygenases (see Introduction II, Fig. 28). The gene expression patterns of two recently characterized P450 monooxygenases, *CYP71D179* and *CYP71D181* isolated from *Origanum vulgare* and *Thymus vulgaris*, correlated well with the presence of thymol and carvacrol in the essential oils of these species. This result suggested that they might play a role in thymol and carvacrol formation (Crocoll, 2011). The aim of my study was to biochemically characterize the enzyme activities of CYP71D179 and CYP71D181 with the goal to confirm the activities that were proposed by C. Crocoll. For heterologous expression in *S. cerevisiae*, a yeast codon usage optimization was performed and the optimized sequences were synthesized by Genart (Regensburg, Germany) in the pESC-Leu2d yeast expression vector. These vectors were provided by C. Crocoll (MPI for Chemical Ecology, Jena, Germany) for further studies. When heterologously expressed in *S. cerevisiae* strain WAT11 and fed with  $\gamma$ -terpinene, the main product of both cytochrome P450s was *p*-cymene (Fig. 29, data for CYP71D179 not shown). Thus, it appeared that the enzymes performed only one oxidation step which led to the aromatic ring, but no hydroxylation was observed. In addition to *p*-cymene, C. Crocoll (2011) also reported traces of thymol and carvacrol as enzyme products. Unfortunately, this result could not be reproduced.



**Fig. 29** *In vitro* product spectrum of CYP71D181. (A) The enzyme was heterologously expressed in *S. cerevisiae* strain WAT11 and microsomes containing the recombinant protein were incubated with the substrate  $\gamma$ -terpinene. The resulting monoterpene products were extracted with hexane and identified by GC-MS. 1:  $\gamma$ -terpinene, 2:  $p$ -cymene.  $p$ -Cymene is formed from  $\gamma$ -terpinene by CYP71D181. (B)

Based on the assumption that the enzyme assay conditions needed to be optimized to facilitate proper hydroxylation and oxidation of  $\gamma$ -terpinene, several assay parameters were changed in order to investigate the effect on the product outcome. Varying pH-values had already been tested and the optimum pH value was determined to be pH 6.8 (Crocoll, 2011). In a study from Novak and coworkers (Novak et al., 2010) it was shown that thymol and carvacrol formation in two *Oregano* species was temperature dependent. While the ratio of thymol in the essential oil decreased with increasing temperatures, the impact on the amount of carvacrol was opposite. Therefore, assays with CYP71D179 and CYP71D181 were performed at different incubation temperatures to test whether the enzyme specificity was influenced by temperature. The temperature was increased from 16 °C to 34 °C in 6 °C steps. In both cases, the main enzyme product was  $p$ -cymene, neither thymol nor carvacrol were detected (Fig. 30).



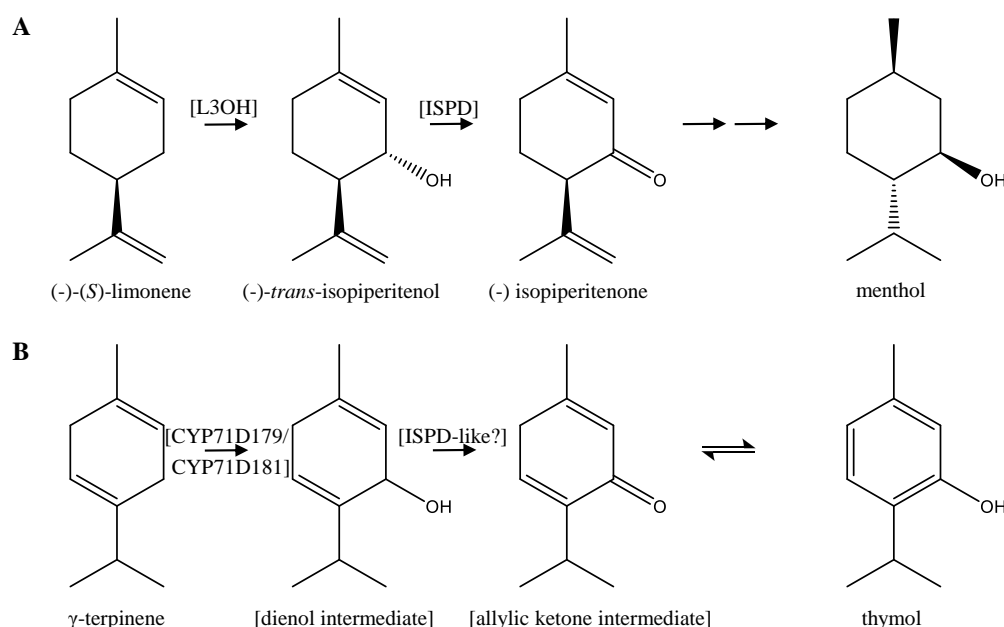
**Fig. 30** *In vitro* product spectra of CYP71D179 and CYP71D181 at different incubation temperatures. The enzymes were heterologously expressed in *S. cerevisiae* strain WAT11 and microsomes containing the recombinant protein were incubated with the substrate  $\gamma$ -terpinene. The resulting monoterpene products were extracted with hexane and identified by GC-MS. 1:  $\gamma$ -terpinene, 2: *p*-cymene.

### 3.1.2 Additional cytochrome P450 enzymes might be involved in the formation of thymol and carvacrol

Recently, high throughput sequencing of the transcriptomes of the C- and T-type of *Thymus vulgaris* was conducted (N. Arndt, MLU Halle-Wittenberg, Germany, unpublished results). A BLAST search of the transcriptomes using *CYP71D179* as template revealed one full length ORF and two partial P450 gene sequences (part of the N-terminus missing), which showed 66 % identity on DNA level to *CYP71D179*. The sequences were temporary called *CYP71D\_u1*, *CYP71D\_u2*, and *CYP71D\_u3* (Fig. S5). A BLAST search in the NCBI database with the deduced amino acid sequences showed the closest similarities with cytochrome P450s from the CYP71D subfamily. The sequence identities to this subfamily were higher than 55 % (Table S3). According to the P450 nomenclature, the identified genes presumably belong to the CYP71D subfamily as well (Nelson, 2006). Since this subfamily is associated with monoterpene metabolism in plants in general (Bak et al., 2011; Weitzel and Simonsen, 2013), the identified P450 enzymes may also play a role in the thymol and carvacrol formation.

### 3.1.3 A dehydrogenase might catalyze a crucial step in the thymol and carvacrol formation

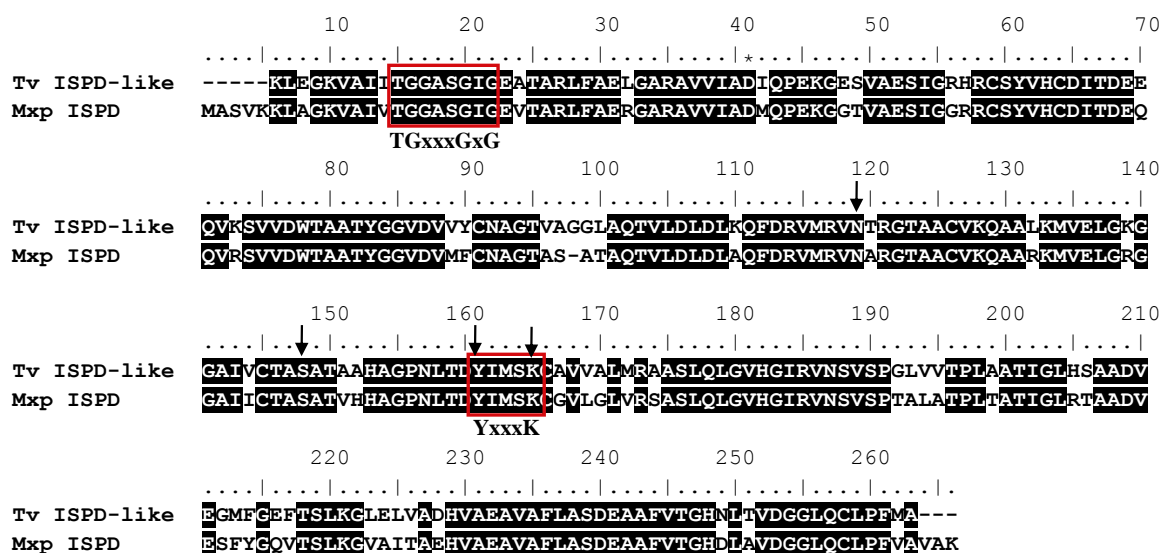
In mint (*Mentha x piperita*), the formation of the oxygenated monoterpene menthol is initiated by a P450-mediated hydroxylation of limonene at C-3 to form isopiperitenol, an allylic alcohol. A subsequent second oxidation is catalyzed by a short-chain dehydrogenase and leads to the respective ketone, before it is rearranged to menthol in following steps (Croteau et al., 2005). It is thus conceivable that in thyme, the postulated allylic alcohol (dienol intermediate) formed by a P450 enzyme from  $\gamma$ -terpinene (Introduction II, Fig. 28) (Crocoll, 2011) could be dehydrogenated by a dehydrogenase similar to the mint enzyme, leading to the ketone which can undergo the keto-enol-tautomerism to form thymol or carvacrol (Fig. 31).



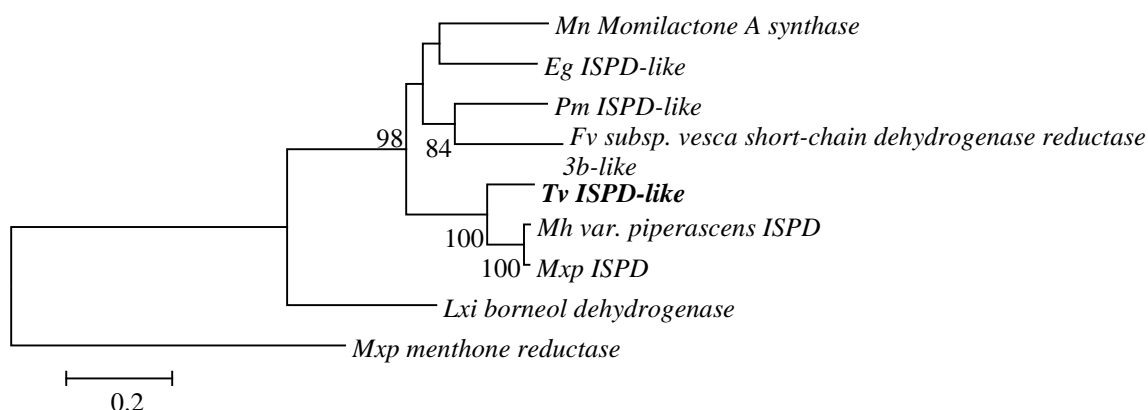
**Fig. 31** Comparison of the first steps of the pathway from (-)-limonene to menthol in *Mentha x piperita* (A) (Croteau et al., 2005) and the postulated pathway from  $\gamma$ -terpinene to thymol in *Thymus vulgaris* (B). Enzyme abbreviations in square brackets are: L3OH: (-)-limonene-3-hydroxylase (CYP71D13), ISPD: (-)-(*trans*)-isopiperitenol dehydrogenase.

The availability of transcriptomes from the C-type and T-type of *Thymus vulgaris* provided the opportunity to search for a homologue of the (-)-(*trans*)-isopiperitenol dehydrogenase from *Mentha x piperita*. In both transcriptomes, a gene with an overall sequence identity of 85 % on DNA level could be identified. An alignment with the deduced amino acid sequences is shown in Fig. 32. The dendrogram analysis in Fig. 33 displays a close relationship of this sequence to two isopiperitenol dehydrogenases from *Mentha* species.





**Fig. 32** Amino acid sequence alignment of the putative short-chain dehydrogenase from *Thymus vulgaris* (ISPD-like) and the (-)-(trans)-isopiperitenol dehydrogenase from *Mentha x piperita* (ISPD). Identical amino acids between both sequences are shaded in black. Some important short-chain dehydrogenase sequence motifs are designated: The NAD<sup>+</sup>-binding motif (TGxxxGxG) and the active site sequence motif (YxxxK). The amino acids belonging to the catalytic tetrad are marked with an arrow.

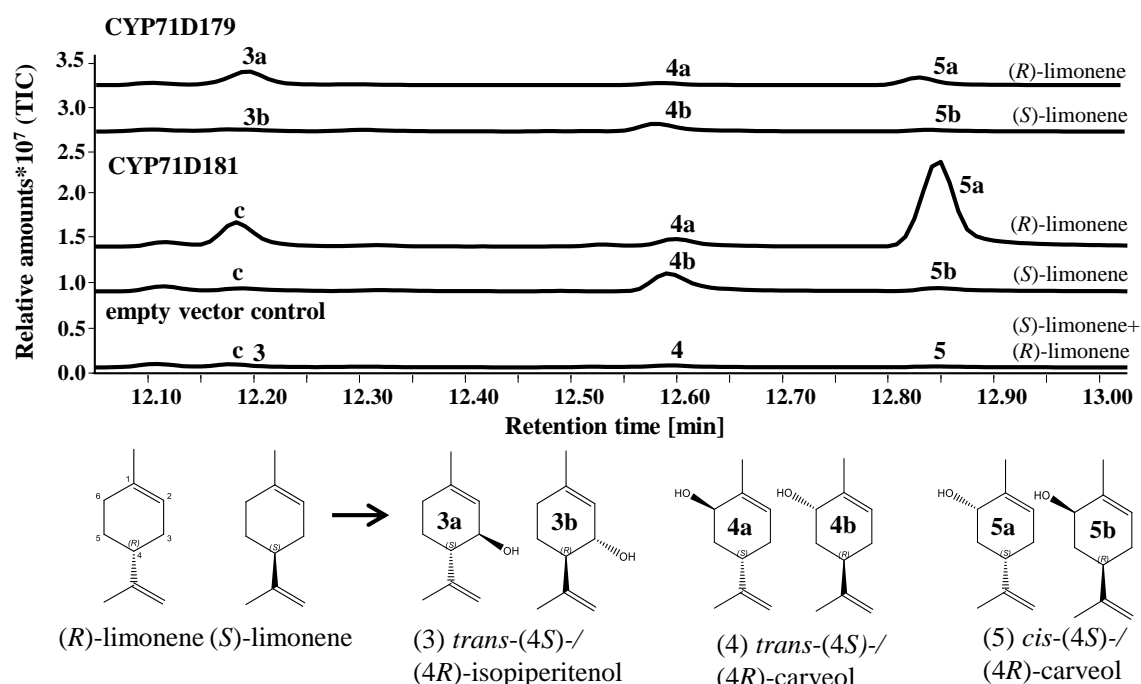


**Fig. 33** Dendrogram analysis of short-chain dehydrogenases by Maximum Likelihood method. Each branch is designated with the abbreviation of the species and the gene name. Bootstrap values greater than 50 are given as a percentage of 1000 replicates. The branch lengths are measured in the number of substitutions per site. The tree was rooted on *Mxp menthone reductase*. ISPD: (-)-(trans)-isopiperitenol dehydrogenase. GenBank IDs in parentheses: *Eg Eucalyptus grandis* ISPD-like (XM\_010038837); *Fv Fragaria vesca* subsp. *vesca* short-chain dehydrogenase reductase 3b-like (XM\_004288707); *Lxi Lavandula x intermedia* borneol dehydrogenase (JX972167); *Mh Mentha haplocalyx* var *piperascens* ISPD (EF426465); *Mn Morus notabilis* Momilactone A synthase (XM\_010114296); *Mxp Mentha x piperita* ISPD (AY641428), menthone reductase (AY288138); *Pm Prunus mume* ISPD-like (XM\_008241627).

### 3.1.4 Limonene as an alternative monoterpene precursor of thymol and carvacrol

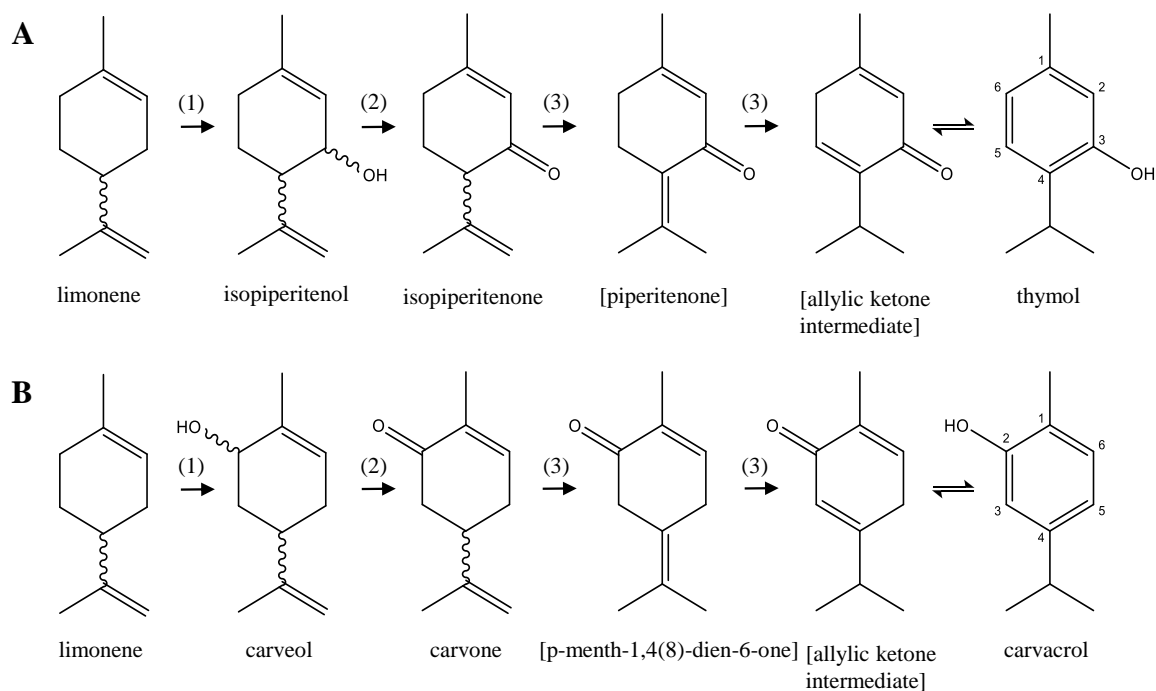
C. Crocoll showed that the cytochrome P450 enzymes CYP1D179 and CYP71D181 from thyme and oregano not only metabolized  $\gamma$ -terpinene, but also a few other structural related monoterpenes, including limonene (Crocoll, 2011). This finding was confirmed in the

present study. Fig. 34 shows the product spectra of recombinant CYP71D179 and CYP71D181 after the incubation with either (*S*)-limonene or (*R*)-limonene.



**Fig. 34** *In vitro* product spectrum of CYP71D179 and CYP71D181. The enzymes were heterologously expressed in *S. cerevisiae* strain WAT11 and microsomes containing the recombinant protein were incubated with the substrates (*S*)-limonene and (*R*)-limonene. The resulting monoterpene products were collected by SPME and identified by GC-MS. “c”: mass spectra similar to perillyl alcohol or menth-1-en-9-ol.

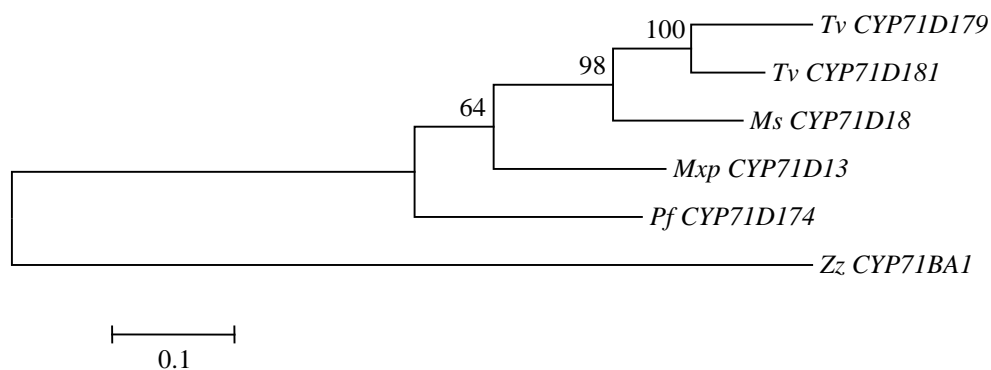
Using (*S*)-limonene as substrate, CYP71D179 and CYP71D181 formed predominantly *trans*-carveol. Additionally, CYP71D179 formed minor ratios of *trans*-isopiperitenol and *cis*-carveol, CYP71D181 formed minor ratios of *cis*-carveol and an unidentified product with a mass spectrum similar to perillyl alcohol or menth-1-en-9-ol. (*R*)-limonene was also accepted as substrate by both enzymes, however, the hydroxylation profile differed. CYP71D179 had a slight preference to perform the C-3 hydroxylation resulting in *trans*-isopiperitenol, next to minor ratios of *trans*- and *cis*-carveol. CYP71D181 formed mainly *cis*-carveol with the hydroxyl group at C-6, next to minor ratios of *trans*-carveol and the unidentified product similar to perillyl alcohol or menth-1-en-9-ol. The opposite hydroxylation patterns with (*R*)-limonene resembled those of thymol and carvacrol, which also carry a hydroxyl group at either C-3 or C-2 (corresponding to C-6 in carveol), respectively (Crocoll, 2011). Given this result, an alternative reaction pathway for the formation of thymol and carvacrol was postulated in this study, suggesting limonene as origin (Fig. 35).



**Fig. 35 Alternative hypothetical pathway leading to thymol (A) and carvacrol (B) with limonene as origin.** A P450 enzyme (1) could catalyze the hydroxylation of limonene, leading to either isopiperitenol or carveol. Following this, a dehydrogenase (2) could form the respective ketones. Subsequent isomerizations by an isomerase (3) could lead to the allylic ketones, which spontaneously rearrange to thymol or carvacrol.

Similar to the pathway of menthol biosynthesis in *Mentha x piperita* (Croteau et al., 2005), in which different enzymes catalyze subsequent steps, the reaction pathway leading to thymol and carvacrol could include several other enzymes with a cytochrome P450 to catalyze the first step, i.e. the hydroxylation of limonene. The products of the pathway, isopiperitenol or carveol, could be metabolized in a similar way as in mint. This reaction would involve a dehydrogenase to form the ketone, e.g. by the isopiperitenol dehydrogenase-like enzyme found in the *Thymus vulgaris* transcriptomes (Fig. 32). A subsequent isomerization by an isomerase, including two double bond shifts, would then result in an allylic ketone which is able to spontaneously undergo an isomerization to form either thymol or carvacrol.

A dendrogram analysis with several *Lamiaceae* cytochrome P450s that are able to metabolize limonene as their main substrate showed a close relationship between these enzymes and CYP71D179 as well as CYP71D181 (Fig. 36).

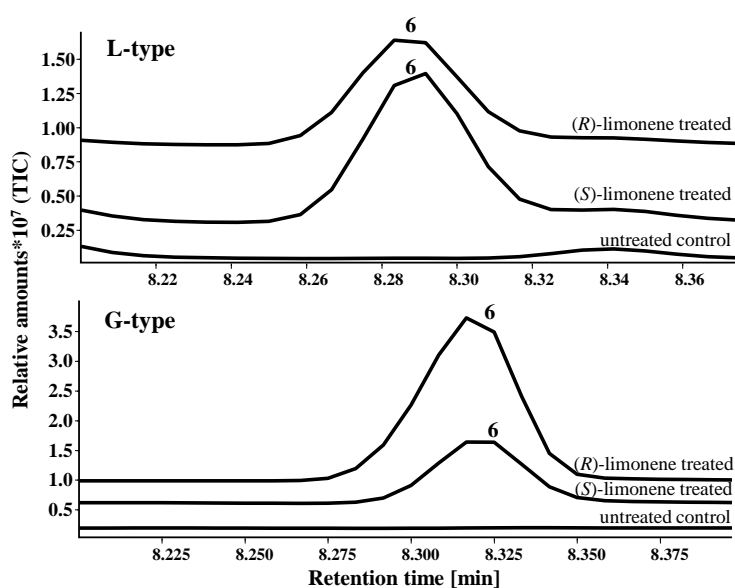


**Fig. 36 Dendrogram analysis by Maximum Likelihood method of *Lamiaceae* cytochrome P450 enzymes that hydroxylate limonene as their main substrate.** Each branch is designated with the abbreviation of the species and the gene name. Bootstrap values are given as a percentage of 1000 replicates. The branch lengths are measured in the number of substitutions per site. The tree was rooted on *CYP71BA1* ( $\alpha$ -humulene hydroxylase). GenBank IDs in parentheses: *Ms Mentha spicata*: *CYP71D18* (limonene-6-hydroxylase) (AF124815); *Mxp Mentha x piperita*: *CYP71D13* (limonene-3-hydroxylase) (AF124816); *Pf Perilla frutescens*: *CYP71D174* (limonene-7-hydroxylase) (GQ120438); *Tv Thymus vulgaris*: *CYP71D179*, *CYP71D181*; *Zz Zingiber zerumbet*: *CYP71BA1* (AB331234).

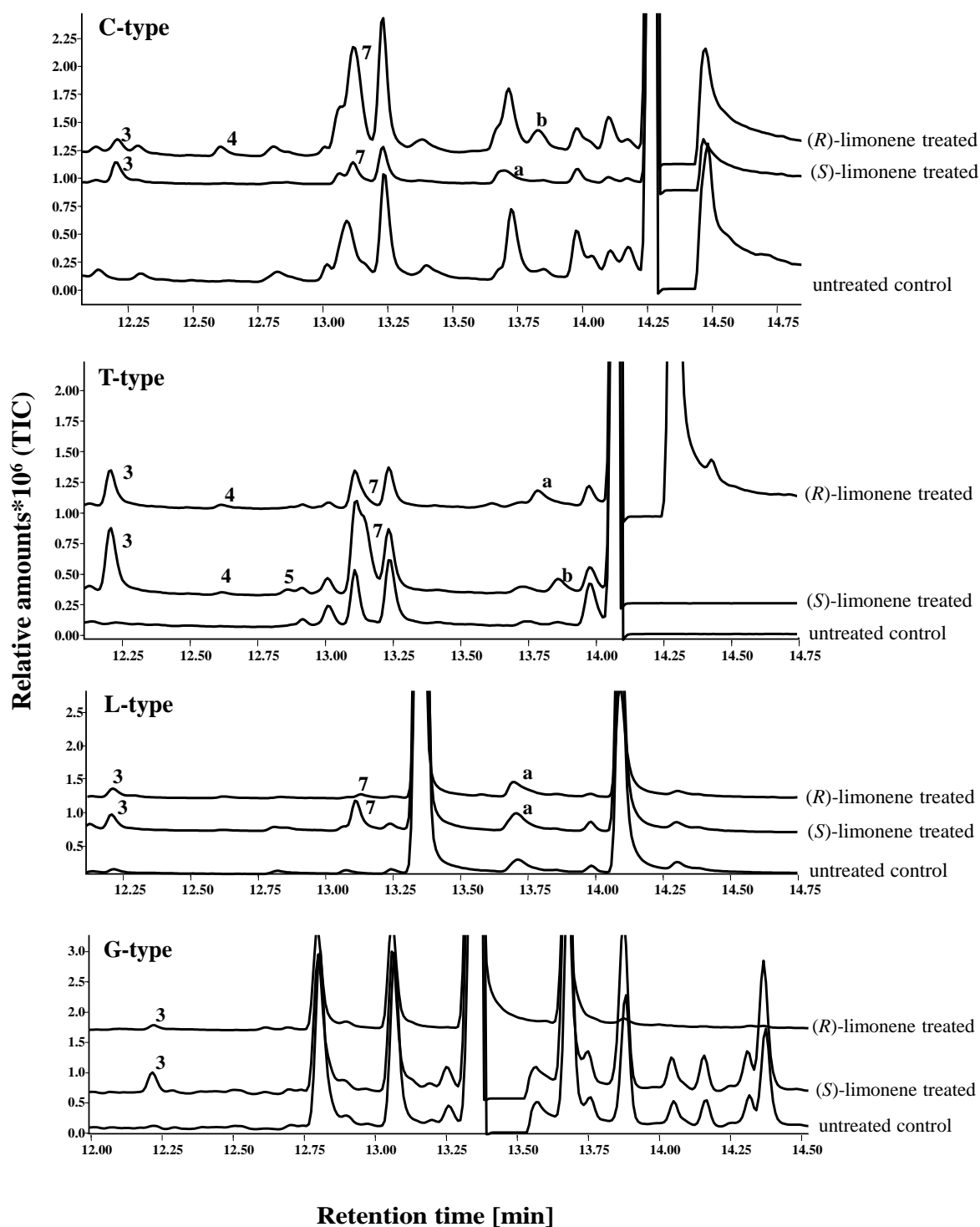
In order to acquire indications for this alternative pathway, limonene was fed to thyme stems. Cuttings of the phenolic chemotypes T and C were used for its high concentration of the aromatic monoterpene alcohols thymol and carvacrol. Cuttings of the L-type were chosen for their low content of thymol. Cuttings of the G-type were used as a control, because they lack P450-metabolized monoterpenes. The limonene was applied separately in both enantiomeric forms by evaporation in glass vessels for 24 h. I wanted to determine how an oversupply of limonene changed the monoterpene profile of the leaves and whether some of the intermediates of the alternative reaction pathway accumulate in the essential oil. After 24 h incubation, the cuttings were analyzed for their monoterpene composition (Figs. 37, 38). The GC traces of the terpenes extracted from the limonene-fed cuttings of the L-type and the G-type showed that limonene was incorporated into the leaves, while untreated cuttings contained no limonene at all (Fig. 37). It is also possible that some of the evaporated limonene stucked on the leaf surface. In the phenolic chemotypes, limonene was overlaid by *p*-cymene in the GC spectra and thus the ratio of the incorporated limonene could not be displayed.

After the limonene feeding, *trans*-isopiperitenol was detected in all investigated chemotypes. *Trans*-carveol was identified in the spectra of the C- and T-type. In the profile of the T-type, *cis*-carveol was also detected. Carvone was found in the T-, C- and L-type, and two additional peaks were present, designated as “a” and “b” (Fig. 38). The substances “a” and “b” could not be reliably identified with the GC libraries, but their mass spectra

showed great resemblance to the mass spectrum of the isopiperitenone GC standard (Fig. S6). In the G-type, investigated as control, only *trans*-isopiperitenol was detected. Thus, the intermediates of the first two hypothetical steps from the alternative pathway from limonene (Fig. 35) could be identified in the T-, C-, and L-types. The remaining intermediates were not observed. This is not surprising, since they might have been metabolized and rearranged in the very moment they were formed. Differences between the total amount of thymol and carvacrol in the limonene-fed stems and the negative controls could not be compared in the T- and C-type. Both compounds accumulate in large amounts in the essential oils of these chemotypes and a potential difference between limonene-treated and untreated plants would be too small to be discernible. In the L-type, the peak height of thymol was similar in all GC traces and no significant larger amount of thymol could be noticed in the limonene-treated stems. In all investigated chemotypes, there was no discernible difference between the metabolization of (*S*)-limonene and (*R*)-limonene.



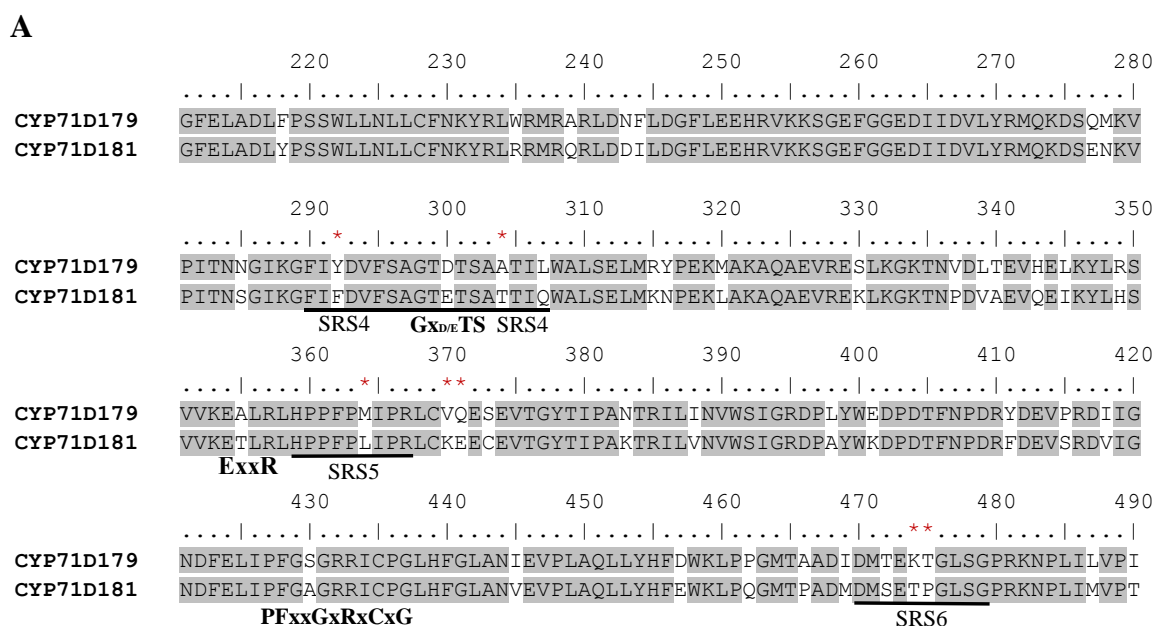
**Fig. 37** Limonene was incorporated into thyme leaves by evaporation in closed glass vessels for 24h. Terpenes were extracted with hexane and analyzed by GC-MS. 6: limonene.



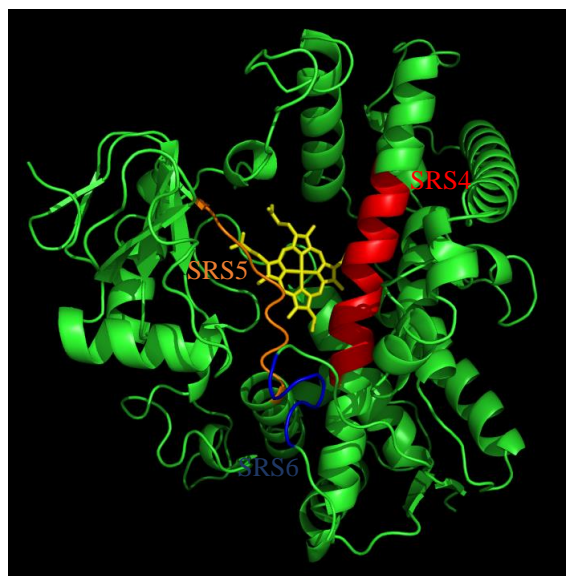
**Fig. 38** Monoterpene spectrum of essential oils extracted from stems of different *Thymus vulgaris* chemotypes after feeding with (*S*)-limonene and (*R*)-limonene for 24 h. Terpenes were extracted with hexane and analyzed by GC-MS. The numbers given are: 3: *trans*-isopiperitenol, 4: *trans*-carveol, 5: *cis*-caveol, 7: carvone, “a” and “b”: mass spectra similar to isopiperitenone (see Supplementary Material II, Fig. S6).

### 3.2 The (*R*)-limonene hydroxylation patterns of CYP71D179 and CYP71D181 are presumably dependent on specific amino acids in the substrate recognition sites

When fed with the (*R*)-limonene substrate, CYP71D179 and CYP71D181 showed different hydroxylation patterns (Crocoll, 2011). CYP71D179 had a slight preference for the C-3 hydroxylation to form *trans*-isopiperitenol. On the other hand, CYP71D181 was more stereospecific in producing almost only *cis*-carveol (Fig. 34). Thus, it appeared that this enzyme had a preference for the hydroxylation at C-6. In another study, one amino acid position had been proved to determine the regiospecificity of the mint P450 enzymes CYP71D13 and CYP71D18. In CYP71D18, the limonene hydroxylation site was altered from C-6 to C-3 after a F363I substitution in SRS 5 (Schalk and Croteau, 2000). Since all isolated cytochrome P450s from thyme and oregano showed the same amino acid at the respective position, namely Phe-362, this site was not supposed to be critical in these enzymes. However, two amino acids downstream in CYP71D179, a methionine (Met-364) was found instead of a leucine (Leu-364) in CYP71D181. This amino acid position may play an important role in substrate binding, because of its position in SRS 5 close to the heme center in the active site (Fig. 39 A, B) (Seifert and Pleiss, 2009). To test whether this amino acid position determines the regiospecificity of the two enzymes, site-directed mutagenesis was performed. The respective positions in CYP71D179 and CYP71D181 were interchanged. Both the exchange of Met-364 to Leu in CYP71D179 and the exchange of Leu-364 to Met in CYP71D181 did not lead to an altered product spectrum from the substrate (*R*)-limonene (Fig. 40). Absolute amounts of the recombinant mutants were not determined, therefore only the ratios between the peak heights were compared. The sequence comparison of CYP71D179 and CYP71D181 shown in Fig. 39 A revealed two other amino acid polymorphisms near position 364 at the C-terminal end of SRS 5. However, all mutations in these positions, CYP71D179 V370K, CYP71D179 Q371E, CYP71D179 V370K + Q371E, CYP71D181 K370V, CYP71D181 E371Q, and CYP71D181 K370V + E371Q appeared to have identical regiospecificity (Fig. 40). It was not possible to determine whether these mutations even inactivated the enzymes, since the product peak heights were low and the empty vector control showed that the yeast microsomes were also able to produce small amounts of *trans*-isopiperitenol and *trans*- and *cis*-carveol by itself.



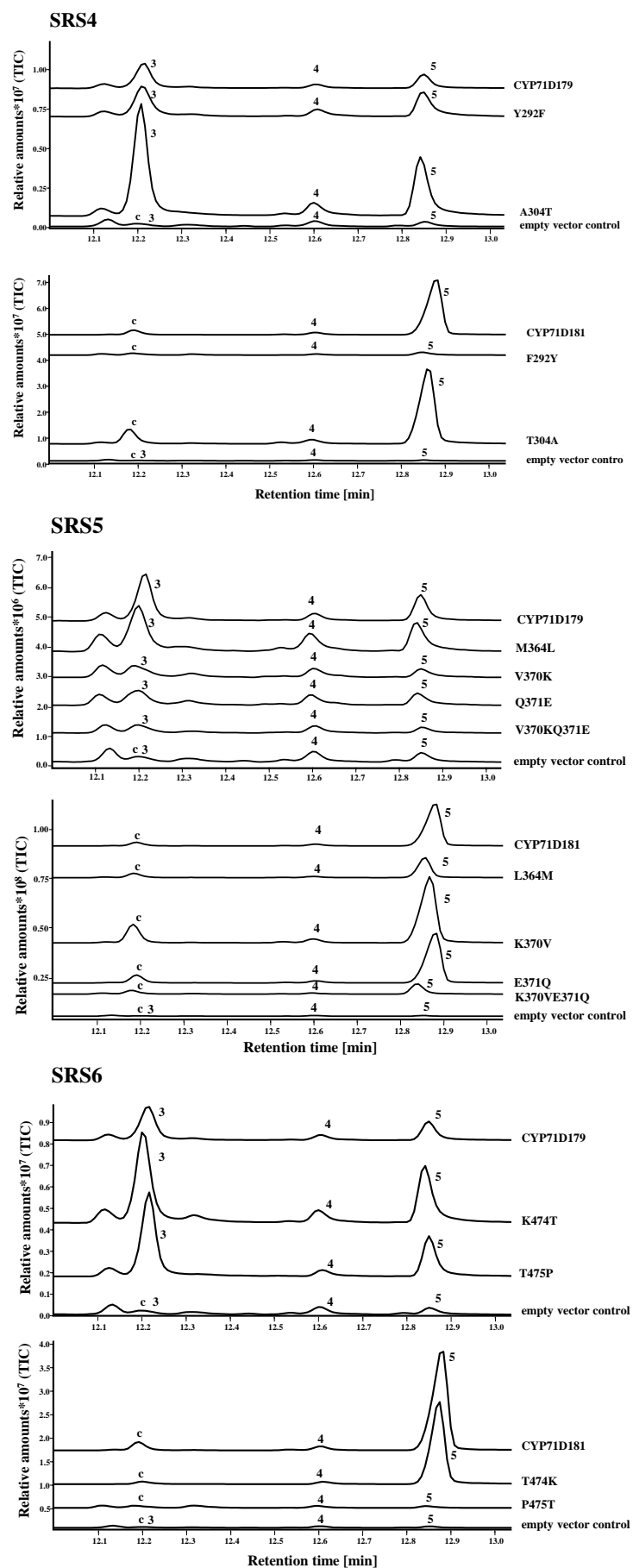
**B**



**Fig. 39 Amino acid sequence alignment of the C-terminal region of CYP71D179 and CYP71D181.** (A) The predicted SRS 4-6 are underlined, the amino acid positions investigated via site-directed mutagenesis studies are marked with an asterisk. Typical P450 sequence motifs are designated. **Model of CYP71D179.** (B) The predicted SRS regions are colored. SRS 4: red, SRS 5: orange, SRS 6: blue. The heme center is shown in yellow.

Two additional SRS were investigated, SRS 4 and SRS 6, because they had been reported to play a role in substrate binding and positioning in the active center, too (Schoch et al., 2003; Sawada et al., 2002; Seitz et al., 2007; Kahn et al., 2001). Fig. 39 B shows the location of SRS 4-6, surrounding the heme center of the P450 enzyme. Tyr-292, Ala-304, Lys-474, and Thr-475 from CYP71D179 were exchanged with the corresponding amino acids from CYP71D181: Phe-292, Thr-304, Thr-474, and Pro-475, and vice versa (Fig. 39 A). As already observed for the SRS 5 mutants, the mutations in SRS 4 and SRS 6 did not affect the hydroxylation pattern of (*R*)-limonene. Only the mutants CYP71D181 F292Y and CYP71D181 P475T appeared to be inactive as they showed no peaks that were significantly higher than the empty vector control (Fig. 40).





**Fig. 40 Site-directed mutagenesis of amino acid residues of CYP71D179 and CYP71D181 in substrate recognition sites 4-6.** The GC-traces show the product composition of both enzymes and their respective mutants from the substrate (*R*)-limonene. Terpenes were collected by SPME and analyzed by GC-MS. 3: *trans*-isopiperitenol, 4: *trans*-carveol, 5: *cis*-carveol, “c”: mass spectra similar to perillyl alcohol or menth-1-en-9-ol.

### 3.3 An approach to P450 monooxygenases involved in the metabolism of monoterpenes in *Eucalyptus* species

Piperitone is a commercially important monoterpene that is found in significant amounts in some *Eucalyptus* species. This was the basis for an *in silico* search of monoterpene-modifying cytochrome P450s in eucalyptus. The limonene-3-hydroxylase CYP71D13 was used as template for a BLAST search in the *Eucalyptus grandis* genome (<http://www.phytozome.net/>), since this enzyme catalyzes the first reaction of the pathway to form piperitone from limonene in mint (Fig. S7) (Diemer et al., 2001). The cytochrome P450 coding sequences with the highest identity were further used for sequence isolation in the *Eucalyptus* species of interest.

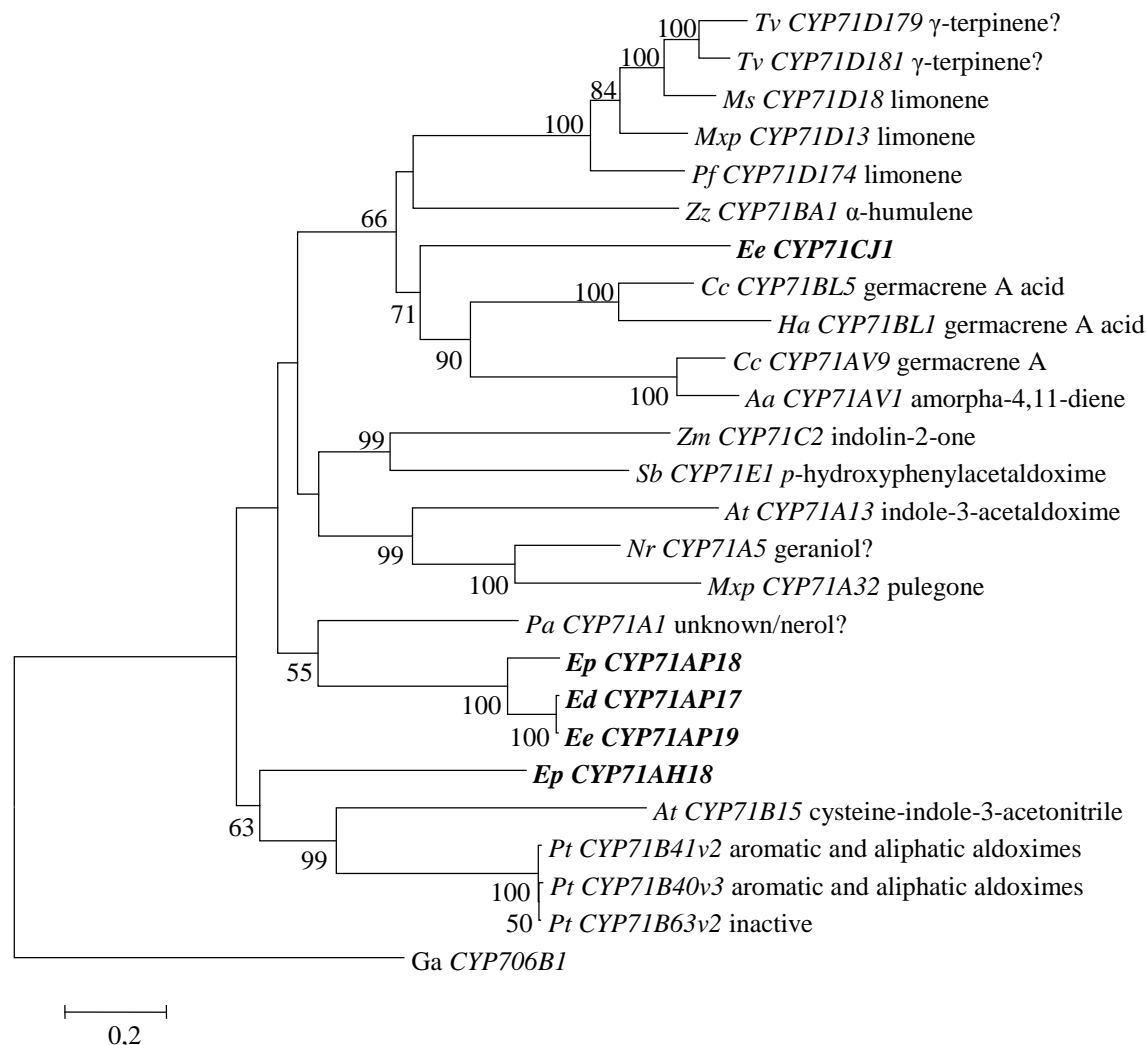
Leaves from *Eucalyptus dives*, *Eucalyptus piperita*, and *Eucalyptus elata* were collected in Canberra, Australia and used for RNA isolation, because they were reported to contain considerable amounts of piperitone in their respective essential oil (Coppen, 2003). Table 5 shows the monoterpene composition of the collected sample leaves.

**Table 5 Monoterpene composition of the *Eucalyptus dives*, *Eucalyptus piperita*, and *Eucalyptus elata* samples used for RNA extraction.** The amounts were determined by GC-MS (A. Padovan, ANU Canberra, Australia) and are given in mg/g fresh weight.

|    |                       | <i>E. piperita</i> | <i>E. dives</i> | <i>E. elata</i> |
|----|-----------------------|--------------------|-----------------|-----------------|
| 1  | $\alpha$ -Thujene     | 1,66               | 2,96            | 1,52            |
| 2  | $\alpha$ -Pinene      | 0,23               | 0,29            | 0,17            |
| 3  | Sabinene              | 0,10               | 0,22            | 0,12            |
| 4  | $\beta$ -Myrcene      | 0,53               | 0,90            | 0,48            |
| 5  | $\alpha$ -Terpinene   | 0,16               | 0,69            | 0,20            |
| 6  | Limonene              | 0,21               | 0,27            | 0,18            |
| 7  | $\beta$ -Phellandrene | 0,44               | 0,76            | 0,31            |
| 8  | 1,8-Cineole           | 0,20               | 0,20            | 0,18            |
| 9  | $\gamma$ -Terpinene   | 0,17               | 0,48            | 0,14            |
| 10 | Terpinolene           | 0,22               | 1,72            | 0,50            |
| 11 | Linalool              | 0,60               | 0,29            | 0,31            |
| 12 | p-Menth-2-en-1-ol     | 0,51               | 0,72            | 0,28            |
| 13 | Terpinen-4-ol         | 1,52               | 2,35            | 1,00            |
| 14 | (Z)-Piperitol         | 0                  | 0,22            | 0,09            |
| 15 | $\alpha$ -Terpineol   | 0,06               | 0,71            | 0,35            |
| 16 | (E)-Piperitol         | 0,42               | 0,53            | 0,20            |
| 17 | Thymol                | 0,11               | 0,07            | 0,05            |
| 18 | Sabinol               | 0,14               | 0,09            | 0,06            |
| 19 | Piperitone            | 2,93               | 0,29            | 0               |
| 20 | Carvacrol             | 0                  | 0,07            | 0               |
| 21 | exo-2-hydroxycineole  | 0                  | 0,16            | 0               |

Five complete ORFs encoding for P450 enzymes were isolated (Fig. S8) and designated as *CYP71AP17* (*Eucalyptus dives*), *CYP71AP18* (*Eucalyptus piperita*), *CYP71AP19* (*Eucalyptus elata*), *CYP71AH18* (*Eucalyptus piperita*), and *CYP71CJ1* (*Eucalyptus elata*) (D. Nelson, P450 nomenclature committee). *CYP71AP17*, *CYP71AH18*, and *CYP71CJ1* were heterologously expressed in *S. cerevisiae*, but the extracted microsomes showed no monooxygenase activity when fed with limonene, the expected precursor for piperitone (data not shown). Alternative monoterpene substrates were tested as well, i.e.  $\gamma$ -terpinene,  $\alpha$ -terpinene, *p*-cymene, sabinene, and 1,8-cineole, but they were not metabolized by the respective P450 enzymes.

In order to classify the isolated cytochrome P450 sequences, a dendrogram analysis with other members of the CYP71 family was conducted (Fig. 41). In general, the enzymes clustered according to their subfamilies. The *Eucalyptus* cytochrome P450s from subfamily CYP71AP showed the closest relationship to CYP71A1, an enzyme isolated from *Persea americana*. Although the function of CYP71A1 *in planta* is still unknown, it has been shown to epoxidate the monoterpene nerol *in vitro* (Hallahan et al., 1994). CYP71AH18 groups with cytochrome P450s involved in aldoxime metabolism and thus may have other functions than terpene modification. CYP71CJ1 showed the closest relationship with cytochrome P450s which play a role in sesquiterpene lactone biosynthesis. However, no function can be deduced from sequence data alone, so further analysis is still to be done.



**Fig. 41** Dendrogram analysis by Maximum Likelihood method of the *Eucalyptus* cytochrome P450s CYP71AP17, CYP71AP18, CYP71AP19, CYP71AH18, CYP71CJ1, and further P450s of the CYP71 family from various species. Each branch is designated with the abbreviation of the species, gene name, and the main substrate of the enzyme. Bootstrap values greater than 50 are given as a percentage of 1000 replicates. The branch lengths are measured in the number of substitutions per site. The tree was rooted on CYP706B1. GenBank IDs in parentheses: Aa *Artemisia annua*: CYP71AV1 (DQ315671); At *Arabidopsis thaliana*: CYP71A13 (NM\_128630), CYP71B15 (NM\_113595); Cc *Cynara cardunculus*: CYP71AV9 (KF752448); CYP71BL5 (KF752451); Ed *Eucalyptus dives* CYP71AP17; Ee *Eucalyptus elata* CYP71AP19, CYP71CJ1; Ep *Eucalyptus piperita* CYP71AP18, CYP71AH18; Ga *Gossypium arboreum*: CYP706B1 (AF332974); Ha *Helianthus annuus*: CYP71BL1 (HQ439590); Ms *Mentha spicata*: CYP71D18 (AF124815); Mxp *Mentha x piperita*: CYP71A32 (AF346833), CYP71D13 (AF124816); Nr *Nepeta racemosa*: CYP71A5 (Y09423); Pa *Persea americana*: CYP71A1 (M32885); Pf *Perilla frutescens*: CYP71D174 (GQ120438); Pt *Populus trichocarpa* CYP71B41v2 (KM458849), CYP71B40v3 (KM458850), CYP71B63v2 (KM458851); Sb *Sorghum bicolor*: CYP71E1 (AF029858); Tv *Thymus vulgaris*: CYP71D179v1, CYP71D181; Zm *Zea mays*: CYP71C2 (X81829); Zz *Zingiber zerumbet*: CYP71BA1 (AB331234).

## 4 Discussion

### 4.1 Different pathways for the formation of the phenolic monoterpenes thymol and carvacrol are conceivable

#### 4.1.1 CYP71D179 and CYP71D181 are active cytochrome P450 monooxygenases, but their explicit role in thymol and carvacrol formation remains unclear

In previous studies,  $\gamma$ -terpinene was suggested to be the precursor of thymol and carvacrol, and *p*-cymene was supposed to be an intermediate of this pathway (Introduction II, Fig. 27) (Poulose and Croteau, 1978). Considering that thymol and carvacrol are aromatic and hydroxylated compounds, it is conceivable that cytochrome P450 enzymes are responsible for their formation. However, a definite proof of certain P450 enzymes that are able to catalyze the complete reaction is still awaited.

C. Crocoll isolated the sequences of five cytochrome P450 enzymes from *Origanum vulgare* and *Thymus vulgaris*: CYP71D178-182. It was shown that the expression of *CYP71D178-182* correlated with the occurrence of thymol and carvacrol in the essential oils of thyme and oregano. Therefore a role in monoterpene biosynthesis can be assumed. After heterologous expression in yeast, all of the five enzymes were active and accepted a range of different monoterpenes as substrate, namely  $\gamma$ -terpinene,  $\alpha$ -terpinene,  $\alpha$ -phellandrene, and (*R*)- and (*S*)-limonene (Crocoll, 2011 and C. Crocoll, personal communication). Two of these P450s, CYP71D179 and CYP71D181, were investigated in more detail in the present study. With  $\gamma$ -terpinene as substrate, only *p*-cymene was detected as product despite the fact that C. Crocoll was able to detect traces of thymol and carvacrol as well (Fig. 29). Thus, the P450 enzymes are able to oxidize  $\gamma$ -terpinene by forming the aromatic ring in *p*-cymene, but no hydroxylation of this compound was observed. Based on the results of his investigations, C. Crocoll proposed an adapted pathway from  $\gamma$ -terpinene to thymol and carvacrol, in which not *p*-cymene but an allylic alcohol would be an intermediate (Introduction II, Fig. 28). The fact that *p*-cymene was not accepted as substrate by any of the P450 enzymes supports the assumption that this compound is not an intermediate of the pathway to thymol and carvacrol, and the observed *p*-cymene product could be an artifact due to unstable reaction conditions (Crocoll, 2011). Searching for optimized enzyme assay conditions, varying incubation temperatures were tested. In a study from Novak *et al.* (2010), the influence of the environmental temperature on the essential oil composition in different *Oregano* species was investigated. While the ratio of thymol in the essential oil decreased with increasing temperatures, the concentration of

carvacrol was elevated. This finding led to the conclusion that two separate enzymes might be responsible for thymol and carvacrol formation, respectively. However, the product specificity of CYP71D179 and CYP71D181 was not temperature dependent, as no altered product spectra were observed (Fig. 30). Novak *et al.* (2010) discussed that also the kinetic properties of the enzymes responsible for thymol and carvacrol formation or other processes, e.g. gene expression, could be temperature dependent. Furthermore it is possible, that other enzymes than the P450s involved in thymol and carvacrol formation are influenced by different temperature levels and thus influence the amount of the product. Further analysis is necessary in order to find out whether the conversion of  $\gamma$ -terpinene to *p*-cymene by CYP71D179 and CYP71D181 is an artifact of the *in vitro* assay conditions. Alternative expression systems like *E. coli* or transient expression in *Nicotiana benthamiana* could lead to an optimized enzyme folding or improved biochemical conditions in a physiological plant system. *E. coli* has a membrane system different from eukaryotes. For the expression of membrane-bound P450 enzymes their N-terminal lipophilic membrane anchor needs to be modified to increase the solubility (Li and Chiang, 1991; Scott *et al.*, 2001). This modification not seems to hamper the enzyme activity. For example, CYP71B40 and CYP71B41 from *Populus trichocarpa* showed no activity when heterologously expressed in *S. cerevisiae*, but were fully active after expression in *E. coli* (Irmisch *et al.*, 2014). However, after expression of CYP71D179 and CYP71D181 in *E. coli* strain C41, no enzyme activity was observed. In an SDS-PAGE, no protein band could be detected in the microsomes (data not shown).

### **4.1.2 The transformation from $\gamma$ -terpinene into thymol and carvacrol may require the activity of multiple enzymes**

The pathway from  $\gamma$ -terpinene to thymol and carvacrol consists of two formal oxidation steps. There are plant P450 enzymes reported which catalyze multiple subsequent oxidation steps, e.g. enzymes involved in the biosynthesis of the cyanogenic glycoside dhurrin or involved in steroid biosynthesis (Sibbesen *et al.*, 1995; Halkier, 1996; Bak *et al.*, 1998; De Montellano, 2005). On the other hand, complexes formed by multiple enzymes, organized in so called metabolons, are known (Srere, 2000). In this case, the enzymes are closely associated to each other, e.g. anchored next to each other in the lipid bilayer. This spatial proximity allows a substrate and the resulting intermediates to quickly shift from one active center to another one. This facilitates both the fast metabolization of unstable intermediates and catalyzing of multiple successive reactions in an efficient and stable way. Reaction

pathways performed by multiple P450 enzymes were for example discovered in dhurrin biosynthesis, although the detailed mechanism of enzyme interaction remained unclear (Sibbesen et al., 1995; Bak et al., 1998; Jensen et al., 2011). The newly discovered P450 gene sequences found in the transcriptome of *Thymus vulgaris* (Results II, section 3.1.2) are possible candidates to encode for P450 enzymes being part of a multienzyme complex formed with the already known P450 enzymes CYP71D178-182. They could perform the second oxidation step, which is needed to form thymol or carvacrol from  $\gamma$ -terpinene. Co-expression of these uncharacterized P450 enzymes with CYP71D178-82 in a heterologous system and subsequent feeding of  $\gamma$ -terpinene might reveal whether these cytochrome P450s are able to form thymol and carvacrol in association with CYP71D178-82.

Other enzyme classes such as dioxygenases (Cochrane and Vederas, 2014; Farrow and Facchini, 2014) or dehydrogenases from the short-chain dehydrogenase/reductase family (SDR) (Jörnvall et al., 1999; Ringer et al., 2005) are also able to perform oxidative transformations of plant metabolites, including terpenes. Dehydrogenases, for example, are reported in *Mint* species to oxidize either (-)-*trans*-isopiperitenol or (-)-*trans*-carveol, which yields the corresponding ketones (-)-isopiperitenone or (-)-carvone (Ringer et al., 2005). These reactions are very similar to the second step in the postulated pathway to thymol and carvacrol (Fig. 28) (Crocoll, 2011). In both cases an allylic alcohol is oxidized. Therefore, an alternative pathway to thymol and carvacrol was proposed, in which first a P450 enzyme hydroxylates  $\gamma$ -terpinene at C-3 or C-6, and then a dehydrogenase similar to those from mint would transform this dienol-intermediate to the corresponding ketones (Fig. 31). A gene homologous to the mint dehydrogenase genes was found in the transcriptome of the T- and C-type of *Thymus vulgaris*, which may encode for an enzyme being involved in this alternative pathway. One indication is the high amino acid sequence identity of 80 %. Since the large family of SDRs normally share only low amino acid sequence identities of about 15-30 % (Jörnvall et al., 1995; Jörnvall et al., 1999), it is likely that the putative SDR found in *Thymus vulgaris* performs the same or a similar reaction as the mint dehydrogenases. A dendrogram analysis also confirmed the close relationship of (-)-*trans*-isopiperitenol dehydrogenases (ISPD) from two *Mentha* species and the putative SDR (ISPD-like) from *Thymus vulgaris* (Fig. 33). The alternative pathway including a dehydrogenase could also explain the *p*-cymene formation from  $\gamma$ -terpinene in the enzyme assays with recombinant CYP71D179 and CYP71D181 expressed in *S. cerevisiae* (Fig. 29). If the dienol intermediate is formed from  $\gamma$ -terpinene by the P450 enzymes and then released in the aqueous conditions of the assay buffer, this intermediate could

spontaneously rearrange to *p*-cymene because it is unstable. Another possibility is the rearrangement of the dienol intermediate in the GC injector at high temperatures. To enable the fast transport of the unstable dienol intermediate *in planta*, the dehydrogenase could be associated with the cytochrome P450s. The P450 enzymes are most likely located at the cytosolic surface of the ER (Bak et al., 2011). According to the prediction server SignalP, the putative ISPD-like dehydrogenase lacks a transit signal peptide and is therefore cytosolic, as it is assumed for the (+)-*trans*-carveol dehydrogenase from *Carum carvi* (Bouwmeester et al., 1998). Thus, both enzymes could be connected via non-covalent protein interactions like hydrogen bonds, hydrophobic bonds, or van der Waals forces in form of an enzyme complex (Chothia and Janin, 1975; Massova and Kollman, 1999). This would facilitate an efficient transport of the dienol intermediate and prevent the release into the cytosol. The involvement of a dehydrogenase might also explain another observation: C. Crocoll could show that after overexpression of CYP71D178 and CYP71D180 in *Arabidopsis thaliana* and feeding with  $\gamma$ -terpinene, carvacrol was produced in the transgenic plants and stored as a glycoside. With CYP71D178, also traces of thymol were found (Crocoll, 2011). In these cases, an unspecific endogenous alcohol dehydrogenase could have converted the dienol intermediates into their corresponding ketones to accomplish the thymol and carvacrol formation. Similar unspecific dehydrogenase activity was observed with monoterpene alcohols in *Nicotiana tabacum* (Suga and Hirata, 1990; Lückner et al., 2004). Another study confirmed that monoterpeneol dehydrogenases in general show a low substrate specificity (Bouwmeester et al., 1998).

The new genes found in the transcriptomes of *Thymus vulgaris* gave rise to alternative ideas on how the phenolic monoterpenes thymol and carvacrol could be formed from  $\gamma$ -terpinene. Of course, gene expression data and characterization studies are required to be conducted in the future, in order to define the actual function of these genes *in planta*.

#### **4.1.3 Limonene as a potential origin of thymol and carvacrol**

(*R*)- and (*S*)-limonene are hydroxylated by CYP71D179 and CYP71D181 mainly at either C-3 or C-6, which resembles the same hydroxylation positions in the cyclohexene ring as in thymol and carvacrol, respectively (Fig. 34). Thus, limonene was postulated as a potential origin for thymol and carvacrol in *Thymus vulgaris* (Fig. 35). The essential oil of the T- and C-type of *Thymus vulgaris* contains only low amounts of limonene and no detectable amounts of the proposed intermediates (Schimmel, 2014). It is conceivable that



the metabolization of limonene is very efficient and thus only small amounts can accumulate in the plant. This was shown for various *Mentha* species which contain large amounts of (-)-menthol while the precursor limonene was found, if ever, only in low amounts (Mint Genomics Resource, Mark Lange, <http://langelabtools.wsu.edu/mgr/>). The formation of (*S*)-limonene is presumed to be the limiting step for monoterpene formation in *Mentha x piperita* (Diemer et al., 2001). Also, only trace amounts of the first intermediates of the pathway to (-)-menthol, isopiperitenol and isopiperitenone, were found (Croteau et al., 2005). It is stated that the first enzymatic steps of the pathway are faster and more efficient than the later steps, since (+)-pulegone is the first intermediate which can be found in considerable amounts in the *Mentha x piperita* essential oil (Croteau and Venkatachalam, 1986). Thus it is plausible that in *Thymus vulgaris* the intermediates of the putative pathway to thymol and carvacrol from limonene (Fig. 35) only accumulate with an oversupply of the precursor and are normally not detectable because of fast metabolization.

A limonene feeding experiment with thyme cuttings was conducted in order to get confirmation for this theory. In all investigated chemotypes of *Thymus vulgaris*, intermediates of the putative pathway could be found after feeding of limonene (Fig. 38). The G-type of *Thymus vulgaris* lacks oxygenated monoterpenes and thus the formation of *trans*-isopiperitenol from limonene was not expected. However, the observed *trans*-isopiperitenol could be the result of an unspecific reaction catalyzed by other unidentified P450 enzymes. Alternatively, the specific P450 enzymes could be expressed despite the lack of an available substrate in this chemotype. For the limonene feeding experiment, the cuttings of all chemotypes were incubated together in one glass vessel and thus the detected *trans*-isopiperitenol in the G-type also might be a contamination produced by the other chemotypes. To ensure that no cross-contamination distorted the monoterpene blends of the chemotypes, the test needs to be repeated with every chemotype in single vessels. In the case of the G-type it is also possible that some substances were not detected because of other abundant components in the oil which elute at the same retention time. In order to find definite proof for limonene as an origin for thymol and carvacrol, the feeding of labelled limonene is necessary. Furthermore, the genes for the respective enzymes of the putative pathway from limonene need to be isolated. Expression analysis and enzyme characterization would bring insights about their role in monoterpene biosynthesis. A gene encoding for a dehydrogenase was already identified in the *Thymus vulgaris* transcriptome and its role in a putative pathway from  $\gamma$ -terpinene is discussed in section 4.1.2. Of course,

this enzyme could also be responsible for the oxidation of isopiperitenol/carveol into isopiperitenone/carvone. The only missing enzymes would be a putative isomerase to catalyze the transformation of isopiperitenone or carvone to the dienol intermediates and a limonene synthase able to synthesize the required amounts of the putative precursor. To date, the only monoterpene synthase isolated from the phenolic chemotypes T and C is a  $\gamma$ -terpinene synthase, which does not produce noteworthy amounts of limonene (Schimmel, 2014).

### **4.2 The oxygenation site of cyclic monoterpenes is determined by sterical constraints in the P450 active center**

#### **4.2.1 Alteration of amino acids in SRS 4-6 did not affect the (*R*)-limonene hydroxylation patterns of CYP71D179 and CYP71D181**

Opposite hydroxylation patterns of CYP71D179 and CYP71D181 were observed with (*R*)-limonene as substrate (either at C-3 or C-6) (Fig. 34). Site-directed mutagenesis of both P450 enzymes did not lead to the identification of amino acids determining the hydroxylation site (section 3.2). The mutated amino acids were chosen according to their locations in SRS 4, SRS 5, and SRS 6, which have been reported in several studies to contain amino acid residues crucial for the enzyme activity and/or regiospecificity of plant P450 enzymes. For example, in CYP71D18 the limonene hydroxylation site was altered from C-6 to C-3 after a F363I substitution in SRS 5 (Schalk and Croteau, 2000). In the cinnamate-4-hydroxylase CYP73A1, the amino acid Asn-302 located in SRS 4 was assumed to have a critical function in binding and positioning the substrate in the active center by forming a hydrogen bond with the substrate (Schoch et al., 2003). In CYP93C2, the amino acid residues Ser-310 in SRS 4 and Lys-375 in SRS 5 were shown to be essential for the conversion of (*2S*)-flavanone to 2-hydroxyisoflavanone (Sawada et al., 2002). Critical amino acids located in SRS 6 have already been reported for subfamilies CYP75A and CYP75B (Seitz et al., 2007) or the fatty acid hydroxylase CYP94A2 (Kahn et al., 2001). However, in the present study no change of regiospecificity was observed after exchanging amino acids of SRS 4-6 in CYP71D179 or CYP71D181 (Fig. 40). Thus, amino acids in other regions may determine the regiospecificity of these two enzymes. In a recent study, an amino acid corresponding to Phe-87 in CYP102A1 was suggested to be crucial for substrate binding and stereospecificity. It is located in SRS 1, oriented straight towards the heme center and, according to the authors, can be found in almost every CYP structure

(Sirim et al., 2010). Therefore, this position is a candidate for further mutagenesis studies of the P450 enzymes from thyme and oregano.

#### 4.2.2 The biological function of CYP71D178-182 might be the C-6 hydroxylation of either (*S*)-limonene or $\gamma$ -terpinene

In most of the thyme and oregano chemotypes from which the *CYP71D178-182* genes were isolated, either thymol or carvacrol were predominant (Crocoll, 2011). Therefore, the respective P450 enzymes are assumed to have a clear preference for the hydroxylation at either C-3 or C-6, as it was shown for the C-3- and C-6-hydroxylases from mint: (*S*)-limonene is specifically hydroxylated at either C-3 in *Mentha x piperita* in the pathway to (-)-menthol or at C-6 in *Mentha spicata* in the pathway to (-)-carvone (Lupien et al., 1999). Feeding of the limonene-6-hydroxylase CYP71D18 with the unnatural substrate (*R*)-limonene resulted in a more complex hydroxylation pattern, probably due to a more flexible positioning of this enantiomer in the active site (Wüst et al., 2001). The same phenomenon, even though to a lesser extent, was observed with the thyme and oregano P450 enzymes. The hydroxylation of (*S*)-limonene by CYP71D179 and CYP71D181 was almost completely restricted to C-6, whereas the hydroxylation pattern with (*R*)-limonene was more complex (Fig. 34). CYP71D179 produced *trans*-isopiperitenol and *cis*-carveol almost in a 1:1 ratio, the main product of CYP71D181 was *cis*-carveol. If thymol and carvacrol are of limonene origin, it can be deduced that (*S*)-limonene is the native substrate *in planta* and the hydroxylation pattern of (*R*)-limonene is an artifact due to the feeding of an unnatural substrate. In this case, additional P450 enzymes would be required to catalyze the opposite hydroxylation of limonene explicitly at C-3 in order to form thymol. Potential candidates for these enzymes are the newly identified P450 genes from the transcriptome of *Thymus vulgaris* (chapter 3.1.2). The deduced amino acid sequences contain either Val or Ile at the position corresponding to Phe-362 of CYP71D179 and CYP71D181 (Fig. 42). This position corresponds to the regiospecificity-determining position in SRS 5 of CYP71D13 and CYP71D18 from peppermint and spearmint, respectively (Schalk and Croteau, 2000). CYP71D18 is a limonene-6-hydroxylase, bearing a Phe at this position (Phe-363), just as CYP71D178-182, which *in planta* might act as C-6 hydroxylases as well. On the other hand, in the limonene-3-hydroxylase CYP71D13, an Ile was found at the corresponding position and in the new P450 enzymes CYP71D\_u1, CYP71D\_u2, and CYP71D\_u3 there was also an Ile or Val found (Fig. S5). Therefore it is possible that these enzymes perform the C-3 hydroxylation of limonene. A detailed characterization of the



The substrate  $\gamma$ -terpinene, the originally suggested precursor for thymol and carvacrol, was oxidized to *p*-cymene by CYP71D179 and CYP71D181. A hydroxylating activity could not be observed *in vitro* (Fig. 29). However, the overexpression of two other CYP71D enzymes from thyme and oregano (CYP71D178 and CYP71D180) in *Arabidopsis thaliana* and subsequent feeding of the transgenic plants with  $\gamma$ -terpinene resulted almost exclusively in the formation of carvacrol, the C-2 hydroxylation product of  $\gamma$ -terpinene (corresponding to C-6 in limonene) (Crocoll, 2011). As (*S*)-limonene is predominantly hydroxylated by these enzymes at C-6 too, the steric conformation of the cyclohexene ring of  $\gamma$ -terpinene bound in the active center of the P450 enzymes might be similar to the one from (*S*)-limonene, which would direct the reaction to the same molecule site. Regardless of whether  $\gamma$ -terpinene or limonene is the precursor for thymol and carvacrol, both molecules are monocyclic, structural related monoterpenes. The main influence of the regio- and stereochemistry of the hydroxylation is the positioning and the extent of flexibility of the cyclohexene ring in the active center of the P450 enzymes (Halkier, 1996; Wüst et al., 2001). Considering the oxygenation patterns of (*S*)-limonene and  $\gamma$ -terpinene (C-6 (C-2) hydroxylation), the binding geometries of these substrates seem to be similar in CYP71D178-182 and to differ with (*R*)-limonene, which in contrast is mainly hydroxylated at C-3.

### **4.3 The *Eucalyptus* cytochrome P450s identified in this study belong to the large and diverse CYP71 family**

The essential oils of Australian *Myrtaceae* are dominated by monoterpenes such as 1,8-cineole,  $\alpha$ -pinene, and terpinen-4-ol, which are all products of terpene synthase enzymes (Padovan et al., 2013). Piperitone is a commercially valuable compound as well and found in considerable amounts in *Eucalyptus* species. As it is an oxygenated monoterpene, the biosynthesis of piperitone is thought to require P450 enzyme activity, like in mint, but to date no information about the P450 enzymes responsible is available (Keszey et al., 2008; Padovan et al., 2013). The isolation of P450 genes from *Eucalyptus* species in this study was a first attempt to learn more about P450-mediated modification of plant secondary metabolites in eucalyptus. Five P450 ORFs were isolated, all of them were assigned to the P450 enzyme family CYP71 and designated as *CYP71AP17* (*Eucalyptus dives*), *CYP71AP18* (*Eucalyptus piperita*), *CYP71AP19* (*Eucalyptus elata*), *CYP71AH18* (*Eucalyptus piperita*), and *CYP71CJI* (*Eucalyptus elata*) (D. Nelson, P450 nomenclature committee). CYP71 is the largest and most diverse plant P450 family. In completely

sequenced genomes, up to 85 family members were found (rice), whereas in papaya, for example, only 12 genes were present (Nelson and Werck-Reichhart, 2011). The CYP71 family is part of the CYP71 clan, from which with one exception all of the monoterpene metabolizing P450 enzymes known to date were isolated (Hamberger and Bak, 2013). The P450s within the CYP71 clan are involved in the biosynthesis of many specialized compounds, but especially enzymes from the subfamilies CYP71A, CYP71D, CYP71AR, and CYP76B were reported to be involved in monoterpene metabolism (Bak et al., 2011; Hamberger and Bak, 2013). The monoterpene metabolizing P450s from thyme and oregano which are discussed in this study are also members of CYP71D. However, the function of the eucalyptus P450 enzymes could not be revealed in this study. The recombinant enzymes CYP71AP17, CYP71AH18, and CYP71CJ1 did not metabolize limonene and other tested monoterpene substrates *in vitro*. Structure-function relationships in P450 enzymes are diverse and poorly explored, thus no function can be deduced from sequence data alone. Since no information of the enzymatic function of members of the subfamilies CYP71AH, CYP71AP, and CYP71CJ is available, a dendrogram analysis was used to narrow possible functions of the eucalyptus P450 down (Fig. 41). According to this analysis, CYP71CJ1 seems to be the most promising candidate to be involved in monoterpene metabolism. It showed 50 % amino acid sequence identity to CYP71BE10 from *Vitis vinifera* and 50 % identity to CYP71D226 from *Solanum tuberosum* and was therefore placed in a new subfamily (D. Nelson, personal communication). CYP71AP17-19 are also good candidates for monoterpene-oxidizing enzymes since they clustered with CYP71A1, a P450 enzyme able to hydroxylize nerol. CYP71AH was most distantly related to all the monoterpene metabolizing P450s and may therefore accomplish other functions. Alternative substrates need to be tested and other expression systems, such as *E. coli* or *Nicotiana benthamiana* could improve enzyme activity, as already discussed for the P450 enzymes from thyme and oregano (chapter 4.1.1)

## Summary

Thyme (*Thymus vulgaris*) is a valued medicinal plant from the *Lamiaceae*. Its essential oil is a natural resource for a wide range of monoterpenes. Due to the antibacterial and spasmolytic effects of these monoterpenes, the essential oil is used in many pharmaceuticals. The oil composition varies significantly between thyme plants. Each chemotype is characterized by one high abundant monoterpene. In this study, the chemotypes containing high amounts of (*E*)-sabinene hydrate,  $\alpha$ -terpineol, thymol, and carvacrol were investigated.

*Melaleuca* and *Eucalyptus* species belong to the Australian *Myrtaceae*. They are especially appreciated for the high content of 1,8-cineole and terpinen-4-ol in their essential oil. These monoterpenes are applied for their mucolytical, antimicrobial, and antifungal activities in many medications.

The emphasis of the present study was the elucidation of stereospecific and regiospecific mechanisms of the enzymes involved in the formation of the oxygenated, cyclic monoterpenes found in *Thyme*, *Melaleuca*, and *Eucalyptus* essential oils. Therefore, genes encoding for the respective enzymes were isolated. Subsequently, the enzymes were characterized after heterologous expression in *Escherichia coli* or *Saccharomyces cerevisiae*.

In plants, monoterpenes are formed from the precursor GPP by monoterpene synthases. From the (*E*)-sabinene hydrate chemotype of *Thymus vulgaris*, the monoterpene synthases TvTPS6 and TvTPS7 were isolated. They were multiproduct enzymes, with (*Z*)-sabinene hydrate and (*E*)-sabinene hydrate as main products, respectively. Although the product spectra of both enzymes were similar, they formed the opposite enantiomers of all chiral monoterpenes. Mutagenesis studies revealed one amino acid residue in the active site of both enzymes which had great impact on the stereospecificity: Asn-350 in TvTPS6 and Ile-346 in TvTPS7. In TvTPS5, an  $\alpha$ -terpineol synthase isolated from the  $\alpha$ -terpineol chemotype of *Thymus vulgaris*, the same amino acid position (Ile-346) had a similar influence on the stereospecificity. Feeding studies of the enzymes with (*R*)-LPP and (*S*)-LPP demonstrated that the configuration of this intermediate determined the configuration of the subsequent chiral terpene products. For this reason, an influence of the critical amino acids Asn-350 (TvTPS6) and Ile-346 (TvTPS7) on the binding conformation of GPP via steric interactions was assumed.

The  $\alpha$ -terpineol synthase TvTPS5 and the (*E*)-sabinene hydrate synthase TvTPS7 showed a high sequence identity, only two differing amino acids were found in the active center. Further investigation is needed to elucidate structural elements of the enzymes which influence the product outcome and water capture during the carbocationic reaction.

From *Melaleuca* species, four ORFs of putative monoterpene synthases were isolated: *Mttps1*, *Mttps2*, *Mttps3*, and *Mltps1*. They showed a high similarity to other 1,8-cineole synthases and sabinene hydrate synthases from *Myrtaceae*. The activity and function of the corresponding enzymes is yet to be determined.

Terpenes can be modified through oxygenation reactions, for example, by cytochrome P450 monooxygenases. In this study, two P450 enzymes, CYP71D179 and CYP71D181, were investigated in order to clarify their role in the formation of thymol and carvacrol from  $\gamma$ -terpinene. In a recent study, a pathway via an allylic alcohol was postulated. When fed with  $\gamma$ -terpinene, the main product of the P450 enzymes was *p*-cymene. Thymol or carvacrol formation could not be observed. Therefore, alternative pathways were suggested for the formation of thymol and carvacrol, including a pathway based on limonene as origin. Furthermore, additional enzymes could be involved in the reaction process, such as another P450 enzyme or dehydrogenases. Genes encoding for the respective candidate enzymes were found in the transcriptome of the thymol and carvacrol chemotypes of *Thymus vulgaris*.

CYP71D179 and CYP71D181 showed an opposite hydroxylation activity when fed with (*R*)-limonene. CYP71D179 had a slight preference for the C-3 hydroxylation to form *trans*-isopiperitenol. CYP71D181 was more regiospecific in producing almost only *cis*-carveol, the C-6 hydroxylation product. Site-directed mutagenesis was performed in order to identify amino acids critical for the differing hydroxylation regiospecificity. All resulting mutants showed an unaltered regiospecificity or a loss of activity.

From the piperitone-rich *Eucalyptus* species *Eucalyptus elata*, *Eucalyptus dives*, and *Eucalyptus piperita*, five ORFs encoding for P450 enzymes from family 71 were isolated. As it is an oxygenated monoterpene, the piperitone formation is thought to involve P450 enzyme activity. However, the heterologously expressed P450 enzymes did not show any activity when fed with limonene, the likely precursor of piperitone. Further investigation is needed to determine the activity of these enzymes *in planta*.



## Zusammenfassung

Thymian (*Thymus vulgaris*) ist eine geschätzte Arzneipflanze der *Lamiaceae*. Sein ätherisches Öl ist eine natürliche Quelle einer Vielfalt an Monoterpenen. Aufgrund der antibakteriellen und spasmolytischen Wirkungen dieser Monoterpene wird das ätherische Öl in vielen Pharmazeutika verwendet. Die Ölkombination variiert deutlich zwischen einzelnen Thymianpflanzen. Jeder Chemotyp ist dabei durch ein dominantes Monoterpen charakterisiert. In dieser Arbeit wurden Chemotypen mit hohem Gehalt an (*E*)-Sabinenhydrat,  $\alpha$ -Terpineol, Thymol und Carvacrol untersucht.

*Melaleuca*- und *Eucalyptus*-Arten gehören zu den australischen *Myrtaceae*. Sie werden besonders für ihren hohen Anteil an 1,8-Cineol und Terpinen-4-ol in ihrem ätherischen Öl geschätzt. Diese Monoterpene werden aufgrund ihrer mukolytischen, antimikrobiellen und antifungalen Eigenschaften in vielen Arzneimitteln eingesetzt.

Der Schwerpunkt der vorliegenden Arbeit war die Aufklärung von stereospezifischen und regiospezifischen Mechanismen der Enzyme, welche an der Bildung der oxygenierten zyklischen Monoterpene aus den ätherischen Ölen von *Thymus*-, *Melaleuca*- und *Eucalyptus*-Arten beteiligt sind. Dafür wurden die Gene, die für die entsprechenden Enzyme kodieren, isoliert. Anschließend wurden die Enzyme nach heterologer Expression in *Escherichia coli* oder *Saccharomyces cerevisiae* charakterisiert.

In Pflanzen werden Monoterpene von Monoterpensynthasen aus der Vorstufe GPP gebildet. Aus dem (*E*)-Sabinenhydrat-Chemotyp von *Thymus vulgaris* konnten die Monoterpensynthasen TvTPS6 und TvTPS7 isoliert werden. Diese Multiproduktenzyme bildeten (*Z*)-Sabinenhydrat bzw. (*E*)-Sabinenhydrat als Hauptprodukt. Außerdem bildeten sie gegensätzliche Enantiomere aller chiralen Monoterpenprodukte. Durch Mutagenese-Studien konnten in den aktiven Zentren der Enzyme die Aminosäurereste identifiziert werden, welche einen großen Einfluss auf die Stereospezifität haben: Asn-350 in TvTPS6 und Ile-346 in TvTPS7. In TvTPS5, einer aus dem  $\alpha$ -Terpineol-Chemotyp isolierten  $\alpha$ -Terpineolsynthase, hatte dieselbe Aminosäureposition (Ile-346) einen ähnlichen Einfluss auf die Stereospezifität des Enzyms. Fütterungsstudien der Enzyme mit (*R*)-LPP und (*S*)-LPP verdeutlichten, dass die Konfiguration dieses Intermediates die Konfiguration der nachfolgend gebildeten chiralen Monoterpene bestimmt. Aus diesem Grund wurde ein Einfluss der kritischen Aminosäuren auf die Bindungskonformation von GPP durch sterische Interaktionen angenommen.

Die  $\alpha$ -Terpineolsynthase TvTPS5 und die (*E*)-Sabinenhydratsynthase TvTPS7 zeigten hohe Sequenzidentitäten. Nur zwei unterschiedliche Aminosäuren wurden im aktiven Zentrum gefunden. Weitere Untersuchungen sind notwendig, um die strukturellen Elemente der Enzyme, welche die Produktbildung und Wasserbindung während der Reaktion beeinflussen, zu identifizieren.

Aus *Melaleuca*-Arten wurden vier ORFs von putativen Monoterpensynthasen isoliert. Sie zeigten eine große Ähnlichkeit zu anderen 1,8-Cineolsynthasen oder Sabinenhydratsynthasen aus *Myrtaceae*. Die Aktivität der jeweiligen Enzyme konnte noch nicht bestimmt werden.

Terpene können weiterhin durch Oxygenierungen modifiziert werden, z.B. durch Cytochrom-P450-Monooxygenasen. In dieser Arbeit wurden zwei P450-Enzyme, CYP71D179 und CYP71D181, untersucht, um ihre Rolle bei der Bildung von Thymol und Carvacrol aus  $\gamma$ -Terpinen zu klären. In einer vorhergehenden Arbeit wurde ein Reaktionsweg über einen allylischen Alkohol postuliert. Das Hauptprodukt dieser P450-Enzyme aus dem Substrat  $\gamma$ -Terpinen war *p*-Cymen, es konnte keine Thymol- oder Carvacrol-Bildung beobachtet werden. Darum wurden alternative Reaktionswege für die Bildung von Thymol und Carvacrol vorgeschlagen, einschließlich einer Reaktion ausgehend von Limonen. Außerdem könnten mehrere Enzyme an der Reaktion beteiligt sein, wie z.B. ein weiteres P450-Enzym oder eine Dehydrogenase. Gene, welche für die jeweiligen Enzyme kodieren, wurden im Transkriptom der Thymol- und Carvacrol-Chemotypen von *Thymus vulgaris* gefunden.

CYP71D179 hatte eine geringe Präferenz für die C-3-Hydroxylierung von (*R*)-Limonen, resultierend in *trans*-Isopiperitenol. CYP71D181 produzierte fast nur *cis*-Carveol, das C-6-Hydroxylierungsprodukt. Mutagenesestudien sollten die kritischen Aminosäuren für die unterschiedliche Regiospezifität aufdecken. Alle Mutanten zeigten jedoch eine unveränderte Regiospezifität oder einen Verlust der Aktivität.

Aus den Piperiton-reichen *Eucalytus*-Arten *Eucalyptus elata*, *Eucalyptus dives* und *Eucalyptus piperita* wurden fünf ORFs isoliert, welche für P450-Enzyme aus Familie 71 kodieren. Da Piperiton ein oxygeniertes Monoterpen ist, geht man davon aus, dass bei der Biosynthese ein P450-Enzym involviert ist. Die heterolog exprimierten P450-Enzyme zeigten jedoch keine Hydroxylierungsaktivität mit dem Substrat Limonen, welches als Vorläufer für Piperiton angenommen wird. Weitere Untersuchungen sind notwendig um die physiologische Rolle der isolierten P450-Enzyme in *Eucalyptus* zu klären.

---

## References

- Adams, R.P.** (2007). Identification of essential oil components by gas chromatography/mass spectrometry. (Allured publishing corporation).
- Alonso, W.R., Rajaonarivony, J., Gershenzon, J., and Croteau, R.** (1992). Purification of 4S-limonene synthase, a monoterpene cyclase from the glandular trichomes of peppermint (*Mentha x piperita*) and spearmint (*Mentha spicata*). *Journal of Biological Chemistry* **267**, 7582-7587.
- Amiot, J., Salmon, Y., Collin, C., and Thompson, J.D.** (2005). Differential resistance to freezing and spatial distribution in a chemically polymorphic plant *Thymus vulgaris*. *Ecology Letters* **8**, 370-377.
- Arnold, K., Bordoli, L., Kopp, J., and Schwede, T.** (2006). The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* **22**, 195-201.
- Atkins, W.M., and Sligar, S.** (1988). The roles of active site hydrogen bonding in cytochrome P-450cam as revealed by site-directed mutagenesis. *Journal of Biological Chemistry* **263**, 18842-18849.
- Ayasse, M., Schiestl, F.P., Paulus, H.F., Löfstedt, C., Hansson, B., Ibarra, F., and Francke, W.** (2000). Evolution Of Reproductive Strategies In The Sexually Deceptive Orchid *Ophrys Sphegodes*: How Does Flower-Specific Variation of Odor Signals Influence Reproductive Success? *Evolution* **54**, 1995-2006.
- Bak, S., Kahn, R., Nielsen, H., Möller, B., and Halkier, B.** (1998). Cloning of three A-type cytochromes P450, CYP71E1, CYP98, and CYP99 from *Sorghum bicolor* (L.) Moench by a PCR approach and identification by expression in *Escherichia coli* of CYP71E1 as a multifunctional cytochrome P450 in the biosynthesis of the cyanogenic glucoside dhurrin. *Plant Molecular Biology* **36**, 393-405.
- Bak, S., Beisson, F., Bishop, G., Hamberger, B., Höfer, R., Paquette, S., and Werck-Reichhart, D.** (2011). Cytochromes P450. *The Arabidopsis Book/American Society of Plant Biologists* **9**.
- Baser, K.H.C., and Demirci, F.** (2007). *Chemistry of Essential Oils. Flavours and Fragrances: Chemistry, Bioprocessing and Sustainability*, 43-86.
- Baser, K.H.C., and Buchbauer, G.** (2009). *Handbook of essential oils: science, technology, and applications*. (CRC Press).
- Bohlmann, J., Meyer-Gauen, G., and Croteau, R.** (1998). Plant terpenoid synthases: molecular biology and phylogenetic analysis. *Proceedings of the National Academy of Sciences* **95**, 4126-4133.
- Bohlmann, J., Steele, C.L., and Croteau, R.** (1997). Monoterpene Synthases from Grand Fir (*Abies grandis*) cDNA Isolation, Characterization, and Functional Expression Of Myrcene Synthase, (-)-(4S)-Limonene Synthase, and (-)-(1S,5S)-Pinene Synthase. *Journal of Biological Chemistry* **272**, 21784-21792.
- Bolwell, G.P., Bozak, K., and Zimmerlin, A.** (1994). Plant cytochrome P450. *Phytochemistry* **37**, 1491-1506.
- Bouwmeester, H.J., Gershenzon, J., Konings, M.C.J.M., and Croteau, R.** (1998). Biosynthesis of the Monoterpenes Limonene and Carvone in the Fruit of Caraway: I. Demonstration of Enzyme Activities and Their Changes with Development. *Plant Physiology* **117**, 901-912.
- Brooker, I.** (2002). Botany of the eucalypts. In *Eucalyptus. The genus Eucalyptus.*, J.J.W. Coppen, ed (Taylor & Francis), pp. 3-35.
- Brophy, J.J., Craven, L.A., and Doran, J.C.** (2013a). Preface. In *Melaleucas. Their botany, essential oils and uses*, J.J. Brophy, L.A. Craven, and J.C. Doran, eds (Australian Centre for International Agricultural Research (ACIAR)), pp. 13-14.
- Brophy, J.J., Craven, L.A., and Doran, J.C.** (2013b). Uses. In *Melaleucas. Their botany, essential oils and uses*, J.J. Brophy, L.A. Craven, and J.C. Doran, eds (Australian Centre for International Agricultural Research (ACIAR)), pp. 34-46.

- Brud, W.S.** (2010). Industrial Uses of Essential Oils. In Handbook of Essential Oils. Science, Technology, Applications., K.H.C. Baser and G. Buchbauer, eds (CRC Press), pp. 843-854.
- Bryant, J.P., Provenza, F.D., Pastor, J., Reichardt, P.B., Clausen, T.P., and du Toit, J.T.** (1991). Interactions between woody plants and browsing mammals mediated by secondary metabolites. Annual Review of Ecology and Systematics, 431-446.
- Burt, S.** (2004). Essential oils: their antibacterial properties and potential applications in foods - a review. International journal of food microbiology **94**, 223-253.
- Cane, D.E., Xue, Q., and Fitzsimons, B.C.** (1996). Trichodiene synthase. Probing the role of the highly conserved aspartate-rich region by site-directed mutagenesis. Biochemistry **35**, 12369-12376.
- Chen, F., Tholl, D., Bohlmann, J., and Pichersky, E.** (2011). The family of terpene synthases in plants: a mid-size family of genes for specialized metabolism that is highly diversified throughout the kingdom. The Plant Journal **66**, 212-229.
- Chothia, C., and Janin, J.** (1975). Principles of protein-protein recognition. Nature **256**, 705-708.
- Christianson, D.W.** (2006). Structural biology and chemistry of the terpenoid cyclases. Chemical reviews **106**, 3412-3442.
- Cochrane, R.V.K., and Vederas, J.C.** (2014). Highly Selective but Multifunctional Oxygenases in Secondary Metabolism. Accounts of Chemical Research **47**, 3148-3161.
- Connolly, J.D., and Hill, R.A.** (1991). Dictionary of terpenoids (Chapman and Hall).
- Coppen, J.J.W.** (2002a). Preface. In Eucalyptus. The genus *Eucalyptus*., J.J. Coppen, ed (Taylor and Francis), pp. xi-xiii.
- Coppen, J.J.W.** (2002b). Production, trade and markets for eucalyptus oils. In Eucalyptus. The genus *Eucalyptus*, J.J.W. Coppen, ed (Taylor & Francis), pp. 365-383.
- Coppen, J.J.W.** (2003). Eucalyptus: The genus *Eucalyptus*. (Taylor & Francis).
- Cornwell, C.P., Leach, D.N., and Wyllie, S.G.** (1995). Incorporation of oxygen-18 into terpinen-4-ol from the H<sub>2</sub><sup>18</sup>O steam distillates of *Melaleuca alternifolia* (Tea Tree). Journal of Essential Oil Research **7**, 613-620.
- Cornwell, C.P., Leach, D.N., and Wyllie, S.G.** (1999). The Origin of Terpinen-4-ol in the Steam Distillates of *Melaleuca argentea*, *M. dissitiflora* and *M. linariifolia*. Journal of Essential Oil Research **11**, 49-53.
- Crocoll, C.** (2011). Biosynthesis of the phenolic monoterpenes, thymol and carvacrol, by terpene synthases and cytochrome P450s in oregano and thyme, Ph.D. Thesis. Biological-Pharmaceutical Department (Friedrich-Schiller-University Jena, Germany), pp. 143.
- Crocoll, C., Asbach, J., Novak, J., Gershenzon, J., and Degenhardt, J.** (2010). Terpene synthases of oregano (*Origanum vulgare* L.) and their roles in the pathway and regulation of terpene biosynthesis. Plant Molecular Biology **73**, 587-603.
- Croteau, R., and Venkatachalam, K.V.** (1986). Metabolism of monoterpenes: Demonstration that (+)-*cis*-isopulegone, not piperitenone, is the key intermediate in the conversion of (-)-isopiperitenone to (+)-pulegone in peppermint (*Mentha piperita*). Archives of Biochemistry and Biophysics **249**, 306-315.
- Croteau, R., Kutchan, T.M., and Lewis, N.G.** (2000). Natural products (secondary metabolites). Biochemistry and molecular biology of plants **24**, 1250-1319.
- Croteau, R., Satterwhite, D., Cane, D., and Chang, C.** (1986). Biosynthesis of monoterpenes. Enantioselectivity in the enzymatic cyclization of (+)- and (-)-linalyl pyrophosphate to (+)- and (-)-bornyl pyrophosphate. Journal of Biological Chemistry **261**, 13438-13445.
- Croteau, R., Davis, E., Ringer, K., and Wildung, M.** (2005). (-)-Menthol biosynthesis and molecular genetics. Naturwissenschaften **92**, 562-577.
- Czygan, F.C., and Hiller, K.** (2002). Thymi herba. Ph. Eur. In Teedrogen und Phytopharmaka, 4. Auflage, M. Wichtl, ed (Wiss. Verlagsgesellschaft mbH Stuttgart).
- Davis, E.M., and Croteau, R.** (2000). Cyclization enzymes in the biosynthesis of monoterpenes, sesquiterpenes, and diterpenes. In Biosynthesis (Springer), pp. 53-95.
- Davis, G.R.** (2002). Cultivation and production of eucalypts in Australia: with special reference to the leaf oils. In Eucalyptus. The genus *Eucalyptus*, J.J.W. Coppen, ed (Taylor & Francis), pp. 183-201.

- De Montellano, P.R.O.** (2005). Cytochrome P450: structure, mechanism, and biochemistry. (Springer Science & Business Media).
- De Montellano, P.R.O.** (2010). Hydrocarbon hydroxylation by cytochrome P450 enzymes. *Chemical reviews* **110**, 932-948.
- Degenhardt, J., Köllner, T.G., and Gershenzon, J.** (2009). Monoterpene and sesquiterpene synthases and the origin of terpene skeletal diversity in plants. *Phytochemistry* **70**, 1621-1637.
- DeLano, W.** (2002). The PyMOL Molecular Graphics System, Version 1.3, Schrödinger LLC, 2002.
- Diemer, F., Caissard, J.-C., Moja, S., Chalchat, J.-C., and Jullien, F.** (2001). Altered monoterpene composition in transgenic mint following the introduction of 4S-limonene synthase. *Plant Physiology and Biochemistry* **39**, 603-614.
- Dudareva, N., Cseke, L., Blanc, V.M., and Pichersky, E.** (1996). Evolution of floral scent in *Clarkia*: novel patterns of *S*-linalool synthase gene expression in the *C. breweri* flower. *The Plant Cell Online* **8**, 1137-1148.
- Eisenreich, W., Schwarz, M., Cartayrade, A., Arigoni, D., Zenk, M.H., and Bacher, A.** (1998). The deoxyxylulose phosphate pathway of terpenoid biosynthesis in plants and microorganisms. *Chemistry & Biology* **5**, R221-R233.
- Fährnich, A., Krause, K., and Piechulla, B.** (2011). Product Variability of the 'Cineole Cassette' Monoterpene Synthases of Related *Nicotiana* Species. *Molecular Plant* **4**, 965-984.
- Farrow, S.C., and Facchini, P.J.** (2014). Functional diversity of 2-oxoglutarate/Fe(II)-dependent dioxygenases in plant metabolism. *Frontiers in Plant Science* **5**.
- Gershenzon, J.** (1998). Plant defenses: surface protectants and secondary metabolites. In *Plant Physiology* (Sinauer), pp. 347-376.
- Gershenzon, J., and Croteau, R.** (1991). Terpenoids. In *Herbivores: Their Interactions With Secondary Plant Metabolites* (second ed.), The Chemical Participants vol. I, G. Rosenthal and M. Berenbaum, eds (San Diego: Academic Press), pp. 165-219.
- Gietz, R.D., and Woods, R.A.** (2002). Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods in enzymology* **350**, 87-96.
- Gill, S.C., and Von Hippel, P.H.** (1989). Calculation of protein extinction coefficients from amino acid sequence data. *Analytical biochemistry* **182**, 319-326.
- Gotoh, O.** (1992). Substrate recognition sites in cytochrome P450 family 2 (CYP2) proteins inferred from comparative analyses of amino acid and coding nucleotide sequences. *Journal of Biological Chemistry* **267**, 83-90.
- Graham, S.E., and Peterson, J.A.** (1999). How similar are P450s and what can their differences teach us? *Archives of Biochemistry and Biophysics* **369**, 24-29.
- Greenhagen, B.T., O'Maille, P.E., Noel, J.P., and Chappell, J.** (2006). Identifying and manipulating structural determinates linking catalytic specificities in terpene synthases. *Proceedings of the National Academy of Sciences* **103**, 9826-9831.
- Guynot, M., Marín, S., Setó, L., Sanchis, V., and Ramos, A.** (2005). Screening for antifungal activity of some essential oils against common spoilage fungi of bakery products. *Food Science and Technology International* **11**, 25-32.
- Hall, T.** (2004). BioEdit version 7.0.0. Distributed by the author, website: [www.mbio.ncsu.edu/BioEdit/bioedit.html](http://www.mbio.ncsu.edu/BioEdit/bioedit.html).
- Halkier, B.A.** (1996). Catalytic reactivities and structure/function relationships of cytochrome P450 enzymes. *Phytochemistry* **43**, 1-21.
- Hallahan, D.L., Lau, S.-M.C., Harder, P.A., Smiley, D.W., Dawson, G.W., Pickett, J.A., Christoffersen, R.E., and O'Keefe, D.P.** (1994). Cytochrome P-450-catalysed monoterpene oxidation in catmint (*Nepeta racemosa*) and avocado (*Persea americana*); evidence for related enzymes with different activities. *Biochimica et Biophysica Acta (BBA)-General Subjects* **1201**, 94-100.
- Hallahan, T.W., and Croteau, R.** (1988). Monoterpene biosynthesis: demonstration of a geranyl pyrophosphate: sabinene hydrate cyclase in soluble enzyme preparations from sweet marjoram (*Majorana hortensis*). *Archives of Biochemistry and Biophysics* **264**, 618-631.

- Hamberger, B., and Bak, S.** (2013). Plant P450s as versatile drivers for evolution of species-specific chemical diversity. *Philosophical Transactions of the Royal Society B: Biological Sciences* **368**, 20120426.
- Hamdane, D., Zhang, H., and Hollenberg, P.** (2008). Oxygen activation by cytochrome P450 monooxygenase. *Photosynthesis Research* **98**, 657-666.
- Hasemann, C.A., Kurumbail, R.G., Boddupalli, S.S., Peterson, J.A., and Deisenhofer, J.** (1995). Structure and function of cytochromes P450: a comparative analysis of three crystal structures. *Structure* **3**, 41-62.
- Hyatt, D.C., and Croteau, R.** (2005). Mutational analysis of a monoterpene synthase reaction: Altered catalysis through directed mutagenesis of (-)-pinene synthase from *Abies grandis*. *Archives of Biochemistry and Biophysics* **439**, 222-233.
- Irmisch, S., Clavijo McCormick, A., Günther, J., Schmidt, A., Boeckler, G.A., Gershenzon, J., Unsicker, S.B., and Köllner, T.G.** (2014). Herbivore-induced poplar cytochrome P450 enzymes of the CYP71 family convert aldoximes to nitriles which repel a generalist caterpillar. *The Plant Journal* **80**, 1095-1107.
- Jensen, K., and Møller, B.L.** (2010). Plant NADPH-cytochrome P450 oxidoreductases. *Phytochemistry* **71**, 132-141.
- Jensen, K., Osmani, S.A., Hamann, T., Naur, P., and Møller, B.L.** (2011). Homology modeling of the three membrane proteins of the dhurrin metabolon: Catalytic sites, membrane surface association and protein-protein interactions. *Phytochemistry* **72**, 2113-2123.
- Jörnvall, H., Höög, J.-O., and Persson, B.** (1999). SDR and MDR: completed genome sequences show these protein families to be large, of old origin, and of complex nature. *FEBS Letters* **445**, 261-264.
- Jörnvall, H., Persson, B., Krook, M., Atrian, S., Gonzalez-Duarte, R., Jeffery, J., and Ghosh, D.** (1995). Short-chain dehydrogenases/reductases (SDR). *Biochemistry* **34**, 6003-6013.
- Kahn, R.A., Bouquin, R.L., Pinot, F., Benveniste, I., and Durst, F.** (2001). A Conservative Amino Acid Substitution Alters the Regiospecificity of CYP94A2, a Fatty Acid Hydroxylase from the Plant *Vicia sativa*. *Archives of Biochemistry and Biophysics* **391**, 180-187.
- Kampranis, S.C., Ioannidis, D., Purvis, A., Mahrez, W., Ninga, E., Katerelos, N.A., Anssour, S., Dunwell, J.M., Degenhardt, J., and Makris, A.M.** (2007). Rational conversion of substrate and product specificity in a *Salvia* monoterpene synthase: structural insights into the evolution of terpene synthase function. *The Plant Cell Online* **19**, 1994-2005.
- Keller, R.K., and Thompson, R.** (1993). Rapid synthesis of isoprenoid diphosphates and their isolation in one step using either thin layer or flash chromatography. *Journal of Chromatography A* **645**, 161-167.
- Keszei, A., Brubaker, C.L., and Foley, W.J.** (2008). A molecular perspective on terpene variation in Australian *Myrtaceae*. *Australian Journal of Botany* **56**, 197-213.
- Kim, J.-K., Kang, C.-S., Lee, J.-K., Kim, Y.-R., Han, H.-Y., and Yun, H.K.** (2005). Evaluation of Repellency Effect of Two Natural Aroma Mosquito Repellent Compounds, Citronella and Citronellal\*. *Entomological Research* **35**, 117-120.
- Köllner, T.G., Schnee, C., Gershenzon, J., and Degenhardt, J.** (2004). The variability of sesquiterpenes emitted from two *Zea mays* cultivars is controlled by allelic variation of two terpene synthase genes encoding stereoselective multiple product enzymes. *The Plant Cell Online* **16**, 1115-1131.
- König, W.A., Fricke, C., Saritas, Y., Momeni, B., and Hohenfeld, G.** (1997). Adulteration or natural variability? Enantioselective gas chromatography in purity control of essential oils. *Journal of High Resolution Chromatography* **20**, 55-61.
- Kreis, P., and Mosandl, A.** (1994). Chiral compounds of essential oils. Part XVI. enantioselective multidimensional gas chromatography in authenticity control of balm oil (*Melissa officinalis* L.). *Flavour and Fragrance Journal* **9**, 249-256.
- Lang, G., and Buchbauer, G.** (2012). A review on recent research results (2008-2010) on essential oils as antimicrobials and antifungals. A review. *Flavour and Fragrance Journal* **27**, 13-39.

- Larkov, O., Dunkelblum, E., Zada, A., Lewinsohn, E., Freiman, L., Dudai, N., and Ravid, R. (2005). Enantiomeric composition of (*E*)- and (*Z*)-sabinene hydrate and their acetates in five *Origanum* spp. *Flavour and Fragrance Journal* **20**, 109-114.
- Lawrence, B.M., and Tucker, A.O. (2002). The genus *Thymus* as a source of commercial products. In *Thyme. The genus Thymus*, E. Stahl-Biskup and F. Sáez, eds (New York: Taylor & Francis), pp. 252-262.
- Li, Y.C., and Chiang, J.Y. (1991). The expression of a catalytically active cholesterol 7 alpha-hydroxylase cytochrome P450 in *Escherichia coli*. *Journal of Biological Chemistry* **266**, 19186-19191.
- Lichtenthaler, H.K. (1999). The 1-Deoxy-D-Xylulose-5-Phosphate Pathway of Isoprenoid Biosynthesis in Plants. *Annual Review of Plant Physiology and Plant Molecular Biology* **50**, 47-65.
- List, S.E. (1995). Functional Anatomy and Clonal Multiplication of *Melaleuca alternifolia* Cheel. in The Central Queensland Region. Thesis for Master of Applied Science. Department of Biology (University of Central Queensland), pp. 95.
- Lücker, J., Schwab, W., Franssen, M.C.R., Van Der Plas, L.H.W., Bouwmeester, H.J., and Verhoeven, H.A. (2004). Metabolic engineering of monoterpene biosynthesis: two-step production of (+)-*trans*-isopiperitenol by tobacco. *The Plant Journal* **39**, 135-145.
- Lupien, S., Karp, F., Wildung, M., and Croteau, R. (1999). Regiospecific Cytochrome P450 Limonene Hydroxylases from Mint (*Mentha*) Species: cDNA Isolation, Characterization, and Functional Expression of (-)-4*S*-Limonene-3-hydroxylase and (-)-4*S*-Limonene-6-hydroxylase. *Archives of Biochemistry and Biophysics* **368**, 181-192.
- Maruyama, T., Ito, M., Kiuchi, F., and Honda, G. (2001). Molecular cloning, functional expression and characterization of *d*-limonene synthase from *Schizonepeta tenuifolia*. *Biological and Pharmaceutical Bulletin* **24**, 373-377.
- Maruyama, T., Saeki, D., Ito, M., and Honda, G. (2002). Molecular cloning, functional expression and characterization of *d*-limonene synthase from *Agastache rugosa*. *Biological & pharmaceutical bulletin* **25**, 661-665.
- Massova, I., and Kollman, P.A. (1999). Computational Alanine Scanning To Probe Protein-Protein Interactions: A Novel Approach To Evaluate Binding Free Energies. *Journal of the American Chemical Society* **121**, 8133-8143.
- McGarvey, D.J., and Croteau, R. (1995). Terpenoid metabolism. *The Plant Cell* **7**, 1015-1026.
- Mondello, F., De Bernardis, F., Girolamo, A., Cassone, A., and Salvatore, G. (2006). *In vivo* activity of terpinen-4-ol, the main bioactive component of *Melaleuca alternifolia* Cheel (tea tree) oil against azole-susceptible and -resistant human pathogenic *Candida* species. *BMC Infectious Diseases* **6**, 158.
- Morales, R. (2002). The history, botany and taxonomy of the genus *Thymus*. In *Thyme. The genus Thymus*, E. Stahl-Biskup and F. Saez, eds (Taylor & Francis), pp. 1-43.
- Nakatsu, T., Lupo, A.T., Chinn, J.W., and Kang, R.K. (2000). Biological activity of essential oils and their constituents. In *Studies in Natural Products Chemistry* (Elsevier), pp. 571-631.
- Nei, M., and Kumar, S. (2000). *Molecular evolution and phylogenetics*. (Oxford University Press).
- Nelson, D., and Werck-Reichhart, D. (2011). A P450-centric view of plant evolution. *The Plant Journal* **66**, 194-211.
- Nelson, D. (2006). Cytochrome P450 Nomenclature, 2004. In *Cytochrome P450 Protocols* (Springer), pp. 1-10.
- Nelson, D., Schuler, M.A., Paquette, S.M., Werck-Reichhart, D., and Bak, S. (2004). Comparative genomics of rice and *Arabidopsis*. Analysis of 727 cytochrome P450 genes and pseudogenes from a monocot and a dicot. *Plant Physiology* **135**, 756-772.
- Novak, J., Lukas, B., and Franz, C. (2010). Temperature Influences Thymol and Carvacrol Differentially in *Origanum* spp. (*Lamiaceae*). *Journal of Essential Oil Research* **22**, 412-415.
- Özek, T., Tabanca, N., Demirci, F., Wedge, D. E., and Başer, K.H.C. (2010). Enantiomeric Distribution of Some Linalool Containing Essential Oils and Their Biological Activities. *Records of Natural Products*, 4(4), 180-192.

- Padovan, A., Keszei, A., Kühlheim, C., and Foley, W.J.** (2013). The evolution of foliar terpene diversity in *Myrtaceae*. *Phytochemistry Reviews* **13**, 695-716.
- Peters, R.J., and Croteau, R.B.** (2003). Alternative termination chemistries utilized by monoterpene cyclases: chimeric analysis of bornyl diphosphate, 1, 8-cineole, and sabinene synthases. *Archives of Biochemistry and Biophysics* **417**, 203-211.
- Phillips, M.A., Wildung, M.R., Williams, D.C., Hyatt, D.C., and Croteau, R.** (2003). cDNA isolation, functional expression, and characterization of (+)-alpha-pinene synthase and (-)-alpha-pinene synthase from loblolly pine (*Pinus taeda*): Stereocontrol in pinene biosynthesis. *Archives of Biochemistry and Biophysics* **411**, 267-276.
- Podust, L.M., and Sherman, D.H.** (2012). Diversity of P450 enzymes in the biosynthesis of natural products. *Natural Product Reports* **29**, 1251-1266.
- Pompon, D., Louerat, B., Bronine, A., Urban, P., Eric, F.J., and Michael, R.W.** (1996). Yeast expression of animal and plant P450s in optimized redox environments. In *Methods in Enzymology* (Academic Press), pp. 51-64.
- Poulos, T.L., and Johnson, E.F.** (2005). Structures of cytochrome P450 enzymes. In *Cytochrome P450: Structure, mechanism, and biochemistry*, 3e, P.R.O.d. Montellano, ed (Kluwer Academic/Plenum Publishers), pp. 87-114.
- Poulose, A.J., and Croteau, R.** (1978). Biosynthesis of aromatic monoterpenes: Conversion of g-terpinene to *p*-cymene and thymol in *Thymus vulgaris* L. *Archives of Biochemistry and Biophysics* **187**, 307-314.
- Raut, J.S., and Karuppayil, S.M.** (2014). A status review on the medicinal properties of essential oils. *Industrial Crops and Products* **62**, 250-264.
- Ringer, K.L., Davis, E.M., and Croteau, R.** (2005). Monoterpene Metabolism. Cloning, Expression, and Characterization of (-)-Isopiperitenol/(-)-Carveol Dehydrogenase of Peppermint and Spearmint. *Plant Physiology* **137**, 863-872.
- Ro, D.-K., Ouellet, M., Paradise, E.M., Burd, H., Eng, D., Paddon, C.J., Newman, J.D., and Keasling, J.D.** (2008). Induction of multiple pleiotropic drug resistance genes in yeast engineered to produce an increased level of anti-malarial drug precursor, artemisinic acid. *BMC biotechnology* **8**, 83.
- Rupasinghe, S., and Schuler, M.A.** (2006). Homology modeling of plant cytochrome P450s. *Phytochemistry Reviews* **5**, 473-505.
- Sanner, M.F.** (1999). Python: a programming language for software integration and development. *J Mol Graph Model* **17**, 57-61.
- Satterwhite, D.M., Wheeler, C., and Croteau, R.** (1985). Biosynthesis of monoterpenes. Enantioselectivity in the enzymatic cyclization of linalyl pyrophosphate to (-)-endo-fenchol. *Journal of Biological Chemistry* **260**, 13901-13908.
- Sawada, Y., Kinoshita, K., Akashi, T., Aoki, T., and Ayabe, S.-i.** (2002). Key amino acid residues required for aryl migration catalysed by the cytochrome P450 2-hydroxyisoflavanone synthase. *The Plant Journal* **31**, 555-564.
- Schalk, M., and Croteau, R.** (2000). A single amino acid substitution (F363I) converts the regiochemistry of the spearmint (-)-limonene hydroxylase from a C6- to a C3-hydroxylase. *Proceedings of the National Academy of Sciences* **97**, 11948-11953.
- Schimmel, J.** (2014). Molekulare Mechanismen der Ausprägung von Chemotypen in *Thymus vulgaris*. Ph. D. thesis. Institute for Pharmacy (Martin-Luther-University Halle-Wittenberg, Germany).
- Schmidt, C.O., Bouwmeester, H.J., Franke, S., and König, W.A.** (1999). Mechanisms of the biosynthesis of sesquiterpene enantiomers (+)- and (-)-germacrene D in *Solidago canadensis*. *Chirality* **11**, 353-362.
- Schnitzler, P., Koch, C., and Reichling, J.** (2007). Susceptibility of drug-resistant clinical herpes simplex virus type 1 strains to essential oils of ginger, thyme, hyssop, and sandalwood. *Antimicrobial agents and chemotherapy* **51**, 1859-1862.
- Schoch, G.A., Attias, R., Le Ret, M., and Werck-Reichhart, D.** (2003). Key substrate recognition residues in the active site of a plant cytochrome P450, CYP73A1. *European Journal of Biochemistry* **270**, 3684-3695.



- Schwab, W., Williams, D.C., Davis, E.M., and Croteau, R.** (2001). Mechanism of monoterpene cyclization: stereochemical aspects of the transformation of noncyclizable substrate analogs by recombinant (-)-limonene synthase, (+)-bornyl diphosphate synthase, and (-)-pinene synthase. *Archives of Biochemistry and Biophysics* **392**, 123-136.
- Schwede, T., Kopp, J., Guex, N., and Peitsch, M.C.** (2003). SWISS-MODEL: an automated protein homology-modeling server. *Nucleic acids research* **31**, 3381-3385.
- Scott, E.E., Spatzenegger, M., and Halpert, J.R.** (2001). A Truncation of 2B Subfamily Cytochromes P450 Yields Increased Expression Levels, Increased Solubility, and Decreased Aggregation While Retaining Function. *Archives of Biochemistry and Biophysics* **395**, 57-68.
- Scott, E.E., He, Y.A., Wester, M.R., White, M.A., Chin, C.C., Halpert, J.R., Johnson, E.F., and Stout, C.D.** (2003). An open conformation of mammalian cytochrome P450 2B4 at 1.6-Å resolution. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 13196-13201.
- Sefidkon, F., Dabiri, M., and Rahimi-Bidgoly, A.** (1999). The effect of distillation methods and stage of plant growth on the essential oil content and composition of *Thymus kotschyanus* Boiss. & Hohen. *Flavour and Fragrance Journal* **14**, 405-408.
- Seifert, A., and Pleiss, J.** (2009). Identification of selectivity-determining residues in cytochrome P450 monooxygenases: A systematic analysis of the substrate recognition site 5. *Proteins: Structure, Function, and Bioinformatics* **74**, 1028-1035.
- Seitz, C., Ameres, S., and Forkmann, G.** (2007). Identification of the molecular basis for the functional difference between flavonoid 3'-hydroxylase and flavonoid 3',5'-hydroxylase. *FEBS Letters* **581**, 3429-3434.
- Sell, C.** (2009). Chemistry of Essential Oils. In *Handbook of essential oils: Science, technology, and applications*, K.H.C. Baser and G. Buchbauer, eds (CRC Press), pp. 121-150.
- Sibbesen, O., Koch, B., Halkier, B.A., and Møller, B.L.** (1995). Cytochrome P-450TYR is a multifunctional heme-thiolate enzyme catalyzing the conversion of L-tyrosine to *p*-hydroxyphenylacetaldehyde oxime in the biosynthesis of the cyanogenic glucoside dhurrin in *Sorghum bicolor* (L.) Moench. *Journal of Biological Chemistry* **270**, 3506-3511.
- Sirim, D., Widmann, M., Wagner, F., and Pleiss, J.** (2010). Prediction and analysis of the modular structure of cytochrome P450 monooxygenases. *BMC structural biology* **10**, 34.
- Southwell, I.** (1999). Introduction. In *Tea Tree. The genus Melaleuca*, I. Southwell and R. Lowe, eds (Amsterdam: Harwood Academic Publishers), pp. 1-7.
- Southwell, I.A., and Stiff, I.A.** (1990). Differentiation between *Melaleuca alternifolia* and *M. linariifolia* by monoterpene comparison. *Phytochemistry* **29**, 3529-3533.
- Srere, P.A.** (2000). Macromolecular interactions: tracing the roots. *Trends in biochemical sciences* **25**, 150-153.
- Stahl-Biskup, E.** (2002). Essential oil chemistry of the genus *Thymus* - a global view. In *Thyme: The genus Thymus*, E. Stahl-Biskup and F. Saez, eds (Taylor & Francis), pp. 75-124.
- Staniek, A., Bouwmeester, H., Fraser, P.D., Kayser, O., Martens, S., Tissier, A., van der Krol, S., Wessjohann, L., and Warzecha, H.** (2013). Natural products – modifying metabolite pathways in plants. *Biotechnology Journal* **8**, 1159-1171.
- Starks, C.M., Back, K., Chappell, J., and Noel, J.P.** (1997). Structural basis for cyclic terpene biosynthesis by tobacco 5-*epi*-aristolochene synthase. *Science* **277**, 1815-1820.
- Steele, C.L., Crock, J., Bohlmann, J., and Croteau, R.** (1998). Sesquiterpene synthases from grand fir (*Abies grandis*) comparison of constitutive and wound-induced activities, and cDNA isolation, characterization, and bacterial expression of  $\delta$ -selinene synthase and  $\gamma$ -humulene synthase. *Journal of Biological Chemistry* **273**, 2078-2089.
- Suga, T., and Hirata, T.** (1990). Biotransformation of exogenous substrates by plant cell cultures. *Phytochemistry* **29**, 2393-2406.
- Tamura, K., Stecher, G., Peterson, D., Filipowski, A., and Kumar, S.** (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular biology and evolution* **30**, 2725-2729.
- Thompson, J.D., Chalchat, J.-C., Michet, A., Linhart, Y.B., and Ehlers, B.** (2003). Qualitative and quantitative variation in monoterpene co-occurrence and composition in the essential oil of *Thymus vulgaris* chemotypes. *Journal of chemical ecology* **29**, 859-880.

- Trapp, S.C., and Croteau, R.B.** (2001). Genomic organization of plant terpene synthases and molecular evolutionary implications. *Genetics* **158**, 811-832.
- Trott, O., and Olson, A.J.** (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of computational chemistry* **31**, 455-461.
- Turnbull, J.W., and Booth, T.H.** (2002). Eucalypts in cultivation: an overview. In *Eucalyptus. The genus Eucalyptus*, J.J.W. Coppen, ed (Taylor & Francis), pp. 52-74.
- Turner, G., Gershenzon, J., Nielson, E.E., Froehlich, J.E., and Croteau, R.** (1999b). Limonene synthase, the enzyme responsible for monoterpene biosynthesis in peppermint, is localized to leucoplasts of oil gland secretory cells. *Plant Physiology* **120**, 879-886.
- Turner, G.W.** (1999a). A brief history of the lysigenous gland hypothesis. *The Botanical Review* **65**, 76-88.
- Vernet, P., Gouyon, R., and Valdeyron, G.** (1986). Genetic control of the oil content in *Thymus vulgaris* L: a case of polymorphism in a biosynthetic chain. *Genetica* **69**, 227-231.
- Weitzel, C., and Simonsen, H.T.** (2013). Cytochrome P450-enzymes involved in the biosynthesis of mono- and sesquiterpenes. *Phytochemistry Reviews*, 1-18.
- Whittington, D.A., Wise, M.L., Urbansky, M., Coates, R.M., Croteau, R.B., and Christianson, D.W.** (2002). Bornyl diphosphate synthase: structure and strategy for carbocation manipulation by a terpenoid cyclase. *Proceedings of the National Academy of Sciences* **99**, 15375-15380.
- Williams, D.C., McGarvey, D.J., Katahira, E.J., and Croteau, R.** (1998). Truncation of limonene synthase preprotein provides a fully active 'pseudomature' form of this monoterpene cyclase and reveals the function of the amino-terminal arginine pair. *Biochemistry* **37**, 12213-12220.
- Wise, M.L., and Croteau, R.** (1999). Monoterpene biosynthesis. In *Comprehensive natural products chemistry*, D.E. Cane, ed (Elsevier), pp. 97-153.
- Worth, H., Schacher, C., and Dethlefsen, U.** (2009). Concomitant therapy with Cineole (Eucalyptole) reduces exacerbations in COPD: a placebo-controlled double-blind trial. *Respir Res* **10**, 69.
- Wüst, M., Little, D.B., Schalk, M., and Croteau, R.** (2001). Hydroxylation of limonene enantiomers and analogs by recombinant (-)-limonene 3- and 6-hydroxylases from mint (*Mentha*) species: evidence for catalysis within sterically constrained active sites. *Archives of Biochemistry and Biophysics* **387**, 125-136.
- Yamaguchi, S.** (2008). Gibberellin Metabolism and its Regulation. *Annual Review of Plant Biology* **59**, 225-251.
- Zabaras, D., and Wyllie, S.G.** (2002). Rearrangement of *p*-menthane terpenes by Carboxen during HS-SPME. *Journal of Separation Science* **25**, 685-690.
- Zaruelo, A., and Crespo, E.** (2002). The medicinal and non-medicinal uses of thyme. In *Thyme. The genus Thymus*, E. Stahl-Biskup and F. Saez, eds (Taylor & Francis), pp. 263-292.
- Zwenger, S., and Basu, C.** (2008). Plant terpenoids: applications and future potentials. *Biotechnology and Molecular Biology Reviews* **3**, 1-7.

## Supplementary Material Chapter I

### DNA sequences of the isolated monoterpene synthases

#### > *Tvtps5* (1704 bp)

ATGCTACCATTAGCATAAATCATGTGGGAATCCTTAGAAATCCTCTCCAATGCAAAAACAAGAGAAGCTTCAATCAATAA  
 ACCATGGAGTCTCAGTCTCCCTCGTTCCGCCACCCGCTCTCGCCTCGTGAAGCCTTGCCGTGTCTCCTCGAAAGTCGATA  
 CCATGCCCGCTGAAATAACCCGACGTTCTGGTAACTATGAGCCTTCGCTTTGGGATTTGATTTTCAATCTCTCGAC  
 GATCATCATCCCTATGTGAAGGAGAAGCAGTTGAAAAGGGAGGAAGAGCTGATTGTGCAGGTGAAGATGCTGCTGGGGAC  
 AAAAATGGAGGCTGTGAAGCAGTTGGAGTTGATCGATGACTTGAAAAATCTCGGATTGTCTATTTTTTTTCGGGAGGAGA  
 TTAAGACGATCTTAACTTCTATATACAATAATTTCTTTGAAAAATAATAATAAAGTAGGGGATTTGTATTTTACCGCTCTT  
 GGATTCAGACTCCTGAGGCAGCACGGTTCAATGTTTCAACAATAATTTGACTGTTTTAAGGGTAATGATTTGACGA  
 AACCATAAATCGGCAAGATACAAAAGGAATTTCCAACCTACGAAGCATCTTTCCACTTGAGGGAAGGGCAAAACACAC  
 TAGAAGTAGCTAGGCAAAATTTCCACCAAGTATCTGCAGAAAAAAGTCGACGAGGGGAAGTATAAATGATGAAAACCTATCG  
 TCGTGGATTTCGACATTCATTGGATCTCCCTCTTCACTGGAGGATTCAAAGGCTCGAGGGCAGATGGTTCTTAGATGCTTA  
 CACAGTGAGGGAAGACAAGAACCCGCTTATTTTCGAGCTCGCCAACTCGACTTCAATATTTATTCAAGCGACTCAACAAC  
 AAGAATCAAAGAAGTCTTAGGTGGTGAATGATTCATGTTTGGCCGAAAAACTTCCCTTCGTGAGGGATAGAGTAGTG  
 GAAAGCTCTTTTGGCCGGTAGGGCTATTTGAGGGTCATGAATTTGGATATCAGAGGAAAAATGACTACTGCTATTATCAT  
 TCTAATTACTGCTATAGATGATCTTTATGATGCTACGGTACATTAGACGAACCTTATCTTTTCACTGACGTCATACGAA  
 GATGGGATACACAATCAATAGACCAACTTCCATACTACATGCAACTCTGTTATTTGGCGCTTACAACCTTTGTCTCGGAT  
 CTGGCTTACGATATTCTCAAAGACCGCGGTCCCAATACTATCCCATATCTACAGGAATCGTGGGTGAAATTTGGTTGAAGC  
 ATATCTGAAGGAGGCAGAGTGGTTTGGAGTGGATACACACCAAGCCTTGAAGAGTATCTACCAACGCCAAGACTTCCGA  
 TAGGCTCTCTTACAATATTAATCCAAGTCGAATTATCATTACCGAATCCACCATTGATCGCGCTGCATTGACTGTTGC  
 CAAAAATACTTTACCTCTCTGCGACGGTTTCAAGGCTTGGCGATGATCTTGAACAACATCGTCTGAGCTGGAGAGAGG  
 GGACGTGCCGAAAGCAATCCAGTGTACATGAAGGATGCAAAATGCATCGGAGGAGGAGGCGGGGAGCAGCTGAGGTTTA  
 TGATCGGAGAGGTGTGGAAAGGAGTTGAACACGGCCATGACGGAGCCGACGATTGCCCGTTTACGGAACAACCTGGTGGAG  
 GCTGCACCTAATCTTGGAAAGAGCAGCACAGTATATTTATAGAGAGGGAGATGGCCATGGACATTTCAAAATTCATCAACA  
 CATCGGAAACTTGTCTTCCACCCATGCGTATGA

#### > *Tvtps6* (1812 bp)

ATGCTACCATTAGCATAAATCATGTGGGACTCCTTAGAAATCCTCTCCACGGCAAAAGCAAGAGAGCTTCGATCAATAA  
 ATCATGGAGTCTCTGTCTCCCTCGTTCCGTCATCCGCTCTCGCCTCGTGAAGCCTTGCCGTGTCTCCTCGAAACCATA  
 CCAAGCCCGCTGAAATGACCCGACGTTCCGGAAACTACGAGCCTTCGCTTTGGGATTTGATTTTCAATCTCTCGAC  
 AATCATCATCCCCACGTGAAGGAGAAGCAGTTGAAAAGGGAGGAAGAGCTGATTGTGGAGGTGAAGATGCTGCTTGGGAC  
 AAAAATAGAGGCTGTAAAGCAGTTGGAGTTAATCGATGACTTGAAAAATCTCGGATTGTCTATTTCTTTTCGGGACGAGA  
 TAAAGATGGTCTTAACTTCTATATATAATAATTTTTTTGAAAAATAAAAAATAAATCAAGTAGGGGATTTGTATTTTACAGCT  
 CTTGGATTTCAGACTCCTTAGGCAACATGGTTTCAATGTTTCAACAAGAAATATTTGATTGTTTCAAGAACGAGAAGGGAAAG  
 TGATTTTCGACGAAACCCTAATCGGCGAAGATACAAAAGCAACCTTACGAAGCATCTTTCCATCTGAGGAAAG  
 GCGAAAACACACTGGAGCTAGCTAGGCAAAATTTCCACCAAGTATCTGCAGAAAAAAGTCGACGAGGGAAAGTATAAATGAT  
 GAAAACCTATCGTCATGGATTTCGACATTCATTGGATCTCCCTCTTCACTGGAGGATCCAAAGGCTCGAGGGCAGATGGTT  
 CTTAGACGCTTACGCGCAAGGGAAGATAAGAACCCGCTTATCTTCGAGCTCACCAACTCGACTTCAATATTTATTCAAG  
 CAACTCAACAAGAAGAACTCAAAGAGGTCTCTAGGTGGTGAATAATTCACGTTTGGCTGAAAAACTCCCTTTCGTGAGG  
 GATAGAGTAGTAGAATGCTACTTTTGGGCGAGTTGGGCTATTTGATGGTTCATGACTATGGATTTTCAGAGGAAAGTTAATGC  
 TGCTGTTAATATTTAATCACTGCCATAGACGAGTTTACGACGCTATGGTACATTAGACGAACCTTCGACTTTTTCAGG  
 ACGTCAATCCGGAGATGGGATACTCAATCGATAGACCAACTTCCATATTACATGCAACTCTGCTATCTGACGCTCTACAAC  
 TATGTATCCGATCTGGCTTACAATATTTCTCAAAGACCGCGCATCAACACTATCCCGCATTTACACCAATCGTGGGTAAA  
 TACAGTTGAAGCATATTTGAAGGAGGCAGAGTGGTACGAGAGTGGATATGCACCAAGCCTTGAAGAGTATCTCAGCATCG  
 CTAGTATTTCAATAGGTGTTATTTCTATAGTAATACCCTCGAAGTATCTATACCAACTCCACTTTTTCATGCTCGCTCC  
 CCATTCGAATATCATCGCTACGATATACTTCACTCTCAGCGATGGTTTTAAGGCTTGCTGATGACCTCGGAACAGCGCA  
 GTATGAGGTGGAGACGGGCGAGCTGCCGAAAGCAGTCCAGTGTACATCAAGGATACAAATGCATCGGAGGAGGAGGCGC  
 GGGAGCAGTGGGTTTATGATCGGAGAGGTGTGGAAGGAGCTGAACACGGCCATGGCGGAGTCCGATGATTGTCCGTTT  
 ACGGAGCAAGGGGCGTGGCTGCAGTTAATATTGGAAGAGCAGCACAGTTTATTTATCTAGAGGGAGATGGGATGGACG  
 TTTCCAAATCCATCAACACATGAAAACCTGTTCTTCCACCCATGCGTATGA

> *Tvtps7* (1794 bp)

ATGTCTACCATTAGCATAAAATCATGTGGGAATCCTTAGAAATCCTCTCCAATGCAAAAACAAGAGAAGCTTCAATAAATAA  
ACCATGGAGTCTCAGTCTCCCTCGTTCGTCACCCGCCCTCGCCTCGTGAAGCCTTGCCGTGTCTCCTCGAAAGTCGATA  
CCATGCCGATGAAATAACCCGACGTTCTGGTAACATGAGCCTTCGCTTTGGGATTTGGATTTCAATCAATCTCTCGAC  
AATCATCATCCATATGTGAAGGAGATGCAGTTGAAAAGGGAGGAAGAGTTGATTGTGCAGGTGAAGATGCTGCTGGGGAC  
AAAAATGGAGGCTGTGAAGCAGTTGGAGTTGATCGATGACTTGAAAAATCTCGGATTGCTTATTTTTTTTCGGGACGAGA  
TTAAGACGATCTTAACCTCTATATACAATAATCTTTTCGAAAATAATAATCAAGTAGGGGATTTGTATTTACGGCTCTT  
GGATTGAGACTCCTGAGGCGACCGGTTTCAATGTTTTCAACAATAATTTGACTGTTTTAAGGATAATGATTTTCGACGA  
AACCTAATCGGCGAAGATACAAAAGGAATTTCTCCAACCTACGAAGCATCTTCCATTTGAGGGAAGGCGAAAACACAC  
TGGAGCTAGCTAGGCAAAATTTCCACCAAGTATCTGCAGAAAAAGATCGACGAGGGAAGTATAAATGATGAAAACCTTACG  
TCGTGGATCCGACATTCATTGGATCTCCCTCTTCACTGGAGGATCCAAAAGGCTCGAGGCGAGATGTTTTCTTAGACGCTTA  
CGCGGCGAGGGAAGACAAGAACCCGCTTATTTTTCAAGCTCGCCGAACCTCGACTTCAATATATTCAAGCAACACAACAAC  
AAGAATCAAAGAGATCTCTAGGTGGTGAATGATTCAAGTTTGGCTGAAAAACTCCCTTCGTGAGAGATAGGGTGGTG  
GAATGCTACTTTTTGGGAGTTGGGCTATTTGAGGGTCAAGATTTGGATTTTCAGAGGAAAATTAAGTCTGCTATATTAT  
CCTAATTACTGCTATAGACGATGTTTACGACGCTACGGTACATTAGATGAACTTCAACTTTTCACTGACGTCATTTCGAA  
GATGGGATACACAATCGATTGACCAACTTCCGATTTACATGCAACTCTGTATTTGGCGCTTACAACCTTTGTCTCCGAT  
CTGGCTTACGATATTTCAAGACCGCGGCTCAATACTATCCCATATACACAGATCGTGGGTAGAATTTGGTTGAAGC  
ATATTTGAAGGAGGCTGGGTGGTACGAGAATGGATACACCAAGCCTTGAAGAGTATCTCACCAACGCCACTATTTCAA  
TAGGCGTTCCCTCTATAGTTTTACCGGTCGAAGTATCCTTACCAACTCTACGATTCATCGCACCCAGTTTCGATCGTCCC  
CACAAAATACTCGACCTTCTGCAAGGTTTTGAGGCTTGTGACGATCTTGAACAGCATCGTCTGAGCTGGAGAGAGG  
GGACGTCGCCGAAAGCAATCCAGTGTACATGAAGGACAACAATGCCCTCGGAGGAGGAGGACGGGAGCACGTGAGGTTTTA  
TGATCAGGGAGGCATGGAAGGAGCTGAACACGGCCATGGCGGAGCCGACAATTTGTCGTTTACGGAACAAACGGTGGAG  
GCTGCAGCTAATCTCGGAAGAGCAGCACAGTTTTATTATCTAGAGGGAGATGGGCATGCATTTCAAATCCATCAACA  
CTTGGAAAACCTGTTCTTCCACCATGCGTATGA

> *Mtts1* (1659 bp)

CGCGTGACATGCTCTTCGACGATTGGAGATCAAGAGATTGTGAGACGTTTCGGCAAATTTGGCAACCTAGCGTTTGGGACTA  
TGGCTTTGTGCAGTCACTCAGTGTGATTACACGGAGAATAAATAACGGAGCAAGTCCAAAGGTTGAAGGAAGAAGTCA  
AGGGTCTATTTGACAAAAAGATGAATCAGGTGGCCAAGCTTGAGTTTCAATGACGCGGTTCAAAGACTAGGACTTGGCTAC  
CAGTTTGAGATGGAGATCAAGAATAGTCTCAGTTCATCTATAACAGCACCGAAGATGCTGGGACTTCGGACGATCTCTA  
TGCCACTTCCCTTCGATTCGACTACTTAGGCAACATGGATACAACATACAGCAAGATGTGTTTTGAAAGTTTTTGAACA  
AGACGGGCACATTCACGAATCGCTCGATAAGGATGTGAAGGGGCTTCTTGGTCTGTATGAAGCTTCATTTTATGATTG  
GAAGGTGAACCTATAGTGGACGAAGCTTGGAACTTTGCTTCCAAGCATCTGATGGATCTAAATCTCAACGAAGTTCACCAC  
TAATTTAGCGAGCAACGTGAGTCACGCGTTAGAAATGCCGATCCACTGGAGGCCAAACAGGTTGGAGGCTCGGTGGTTCA  
TGGACAGTACGGAACAACAAGACATGATCCCTCTTGTGCGATTGGCTAAATTAGATTTCAATTTAGTGCAGTCA  
ATCCACAGGGACGAACCTAGCAATCTGGCAAGGTGGTGGTTGAACTTTGGGCGAACAAGATGACCTTCTTTAGGGACAG  
GCTAGTGGAAAACCTATTTTTGGAGCTGTGTAATGGTTTTTCAACCTCAATATACAGCTTTTCAAGAGTTGGTGCCGAAGC  
TCGGTAATTTGGTGACGCTGATCGATGACGTTTATGATATATATGGGACGCCGGAAGAGCTTGAAGCTCTTAACAGATTTT  
CTTGTACAGTGGGACATCACAGAGGTCGATAAATCTCATCAACAATAAGGGATAGCTACATGGTCTTGTTCATATAAC  
CAACGAAGTTCGGATATTTCTACGATGAGAGAGCGAGGAATCAACCCCATCCCTTATATGCGGAAAAGTGTGGGCGGATGAAT  
GCACGGCGTACATGAAGGAGGTCGGTTGGTACAGCAAGGGCATTAAAGCCGACATTGAAGGAGTACATGGACGTTGCAGTG  
GATTCGATTTGGAGGCTGATTTTGTGTTGGGCAGCTACTTTTTAACCACAGACAAATGACGAAAGAGGACTTGGATGAGTA  
CGTGTCCAATATCCCGAGTGTGATGCTTCTTCCGCCAAGATCCTTCGACTCAACGATGACCTCAGTACATCATCGCATG  
AATTGGCACGAGGAGACAACATAAGGCACTAGAATGTACATGAATGAAAGCGGCGCTTCGGAAAAGGACCGCGGGAA  
CATATCAAGTCTCACGTGCGGAGACTTGGAAAGAAGATGAACAAAGAAGTGTCAAGGACTACCCGTTCCGACGTTTCGA  
CCCTTTCTCCGTGCCTGTTTGAACCTGGCACGAGCTTCTCATTGCTTTTACGACTTTGGGACGGACATGCCCTTCCCG  
GTCACCAAAACCAAGAACCATCTTGAGTGGACTATATTTGAACCCGTGCCCTTCGATTGA

> *Mtts2* (1659 bp)

CGCGTGACGCTGCTCTTCGACGATTGGAGATCAAGAGATTGTGAGACGTTTCGGCAAATTTGGCAACCTAGCGTTTGGGACTA  
TGGCTTTGTGCAGTCACTCAGTGTGATTACACGGAGGATAAATAACGGAGCAAGTCCAAAGGTTGAAGGAAGAAGTGA  
AGGGTCTATTTGACAAAAAGATGAATCAGGTGGCCAAGCTTGAGTTTCAATGACGCGGTTCAAAGACTAGGACTTGGCTAC  
AATTTGAGATGGAGATCAAGAATAGTCTCAGTCCATCTATAACGGCACCGAAGATGCAGGGATTTCCGGACGATCTCTA  
TGCCACTTCCCTTCGATTCGACTACTTAGGCAACATGGATACAACATACAGCAAGATGTGTTTTGAAAGTTTTTGGAGCA  
AGACGGGCACATTCACGAATCGCTCGATAAGGATGTGAAGGGCTTCTTGGTCTGTATGAAGCTTCAATTCATGGATTG  
GAAGGTGAACCTATAGTGGACGAAGCTTGGAACTTTGCTTCCAAGCATCTGATGGATCTAAATCTTAACGAAGTTCACCAC  
TAATTTAGCGAGCAACGTGAGTCACGCATTAGAAATGCCGATCCACTGGAGGCCAAACAGGTTGGAGGCTCGGTGGTTCA  
TGGACAGTACGAGAAAACAACAAGACATGATCCCTCTTGTGCGATTGGCTAAATTAGATTTCAATTTAGTGCAGTCA  
ATCCACAGGGACGAACCTAGCAATCTGGCAAGGTGGTGGTTGAACTTTGGGCGAACAAGATGACCTTCTTTAGGGACAG  
GCTAGTGGAAAACCTATTTTTGGAGCTGTGTAATGGTTTTTCAACCTCAATATACAGCTTTTCAAGAGTTGGTGCCGAAGC  
TCGGTAATTTGGTGACGCTGATCGATGACGTTTATGATATATATGGGACGCCGGAAGAGCTTGAAGCTCTTAACAGATTTT  
CTTGTACAGTGGGACATCACAGAGGTCGATAAATCTCATCAACAATAAGGGATAGCTACATGGTCTTGTTCATATAAC  
CAACGAAGTTCGGATATTTCTACGATGAGAGAGCGAGGAATCAACCCCATCCCTTATATGCGGAAAAGTGTGGGCGGATGAAT  
GCACGGCGTACATGAAGGAGGTCAGTTGGTACAGCAAGGGCATTAAAGCCGACATTGAAGGAGTACATGGACGTTGCGGTG  
GATTCGATTGGAGGGCTGATTTTGTGTTGGGCAGCTACTTTTTAACCACAGACAAATGACGAAAGAGGACTTGGATGAGTA

CGTGTCCAATATCCCGAGTGTGCATGCATCTTCCGCCAAGATCCTTCGACTCAACGATGATCTCAGTACATCATCGCATG  
AATTGGCACGAGGAGACAACACTACAAGGCACTAGAATGCTACATGAATGAAAGCGGCGCTTCGGAAAAGGACGCGCGGGAA  
CATATCAAGTATCACGTGCGCGAGACTTGGAAAGAAGTGAACAAAGAAGTGTTCAGGACTACCCGTTCCGCAGCTTCGA  
CCCTTTTCTCCGTGCCTGTTTGAACCTGGCACGAGCTTCTCATTGCTTTTACGACTTTGGGGACGGACATGGCCTTCCCG  
GTCACCAAACCAAGAACCATCTTGAGTGGACTATATTTGAACCCGTGCCCTCGATTGA

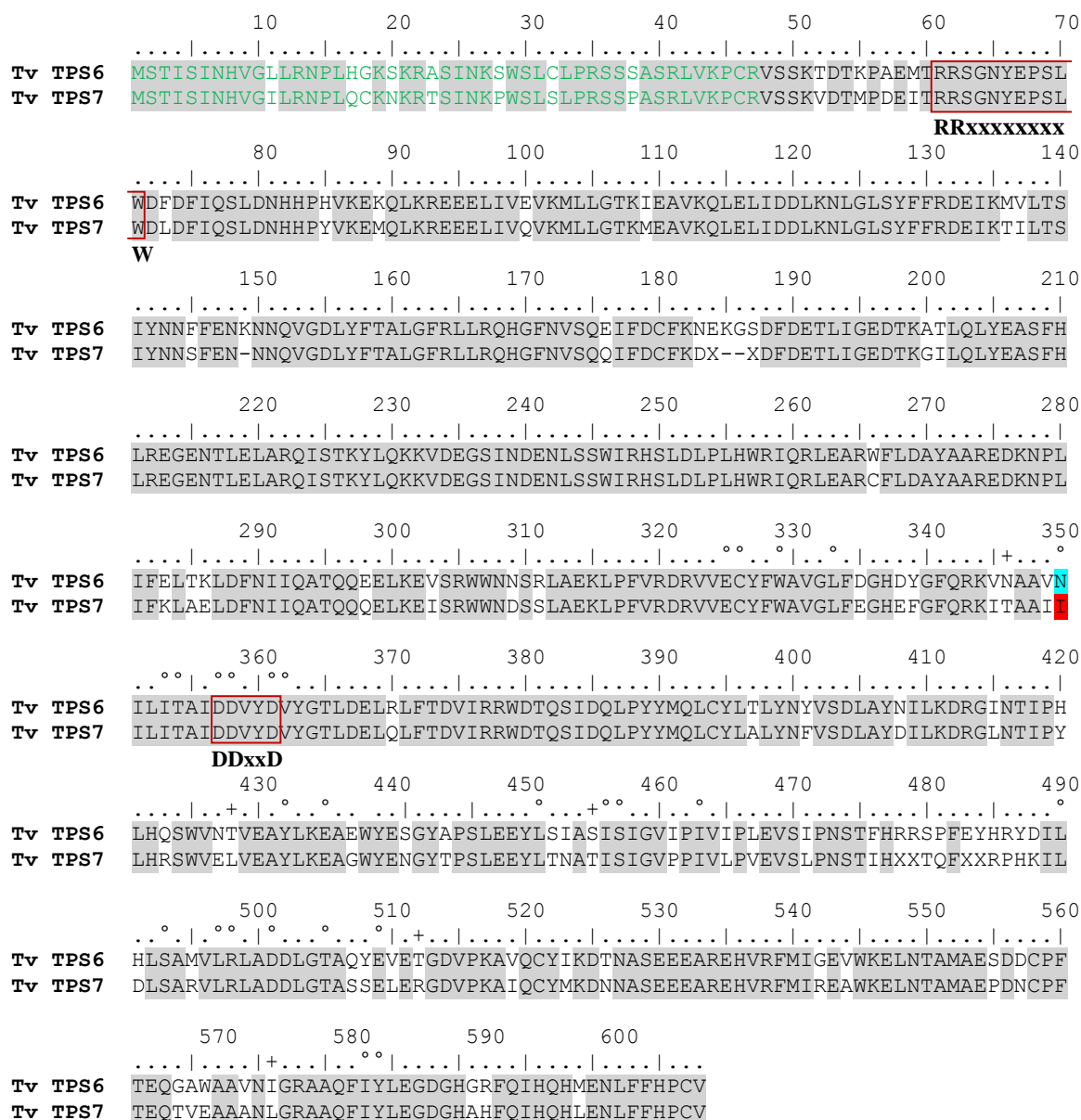
> *Mtps3* (1659 bp)

CGCGTGACATGCTCTTCAACGATTGGAGATCAAGAGATTGTGAGACGTTCCGGCAAATTTGGCAACCTAGCGTTTGGGACTA  
TGGCTTTGTGCAGTCACTCAGTGTGATTACACGGAGGATAAATATACGGAGCAAGTCCAAAGGTTGAAGGAAGAAGTGA  
AGGCTCTATTTGACAAAAAGATGAATCAGGTGGCCAAGCTTGAATTCATTGACGCGGTTCAAAGACTAGGACTTGGCTAC  
AATTTTGGAGATGGAGATCAAGAATAGTCTCAGCTCCATCTATAACGGCACCGAAGATGCAGGGATTTCCGGACGATCTCTA  
TGCCACTTCCCTTCGATTCCGACTACTTAGGCAACATGGATAACAACATACAGCAAGATGTGTTTCGAAAGTTTTTGGAGCA  
AGACGGGCACATTCAACGAATCGCTCGATAAAGGATGTGAAGGGCTTCTTGGTCTGTATGAAGCTTCATTTTCATGGATTG  
GAAGGTGAACCTATAGTGGACGAAGCTTGGAACTTTGCTTCCAAGCATCTGATGGATCTAAATCTTAACGAAGTTCCCCAC  
TAATTTAGCGAGCAACGTGAGTCACGCATTAGAAATGCCGATCCACTGGAGGCCAAACAGGTTGGAGGCTCGGTGGTTCA  
TGGACACGTACGAGAAACAACAAGACATGATCCCCCTTTGCTGCGATTGGCTAAATTAGATTTCAATTTAGTGCAGTCA  
ATCCACAGGGACGAACCTAGCAATCTGGCAAGGTGGTGGGTTGAACCTGGGGCGAACAAGATGACCTTCTTTAGGGACAG  
GCTAGTGGAAAACCTATTTTGGAGCTGTGTAATGGTTTTCGAACCTCAATATACAGCTTTCAGAAGGTTGGTCCCGAAGC  
TCGGTAATTTGGTGACGCTGATCGATGACGTTTATGATATATATGGGACGCCGGAAGAGCTTGAGCTCTTAACAGATTTT  
CTTGTGAGGTGGGACATCACAGAGGTCGATAAACTTCATCCAACAATAAGGGATAGCTACATGGTCTTGTTCATACAAC  
CAACGAAGTCCGATATTCTACGATGAGAGAGCGAGGAATCAACCCATCCCTTATATGCCGAAAAGTGTGGCCGGATGAAT  
GCACGGCGTACATGAAGGAGGTGAGTTGGTACAGCAAGGGCATTAAAGCCGACATTGAAGGAGTACATGGACGTTGCGGTG  
GATTCGATTGGAGGGCTGATTTTGTGTTGGGCAGCTACTTTTTAACCACAGACAAATGACGAAAAGAGGGACTTGAGTA  
CGTGTCCAATATCCCGAGTGTGCATGCATCTTCCGCCAAGATCCTTCGACTCAACGATGATCTCAGTACATCATCGCATG  
AATTGGCACGAGGAGACAACACTACAAGGCACTAGAATGCTACATGAATGAAAGCGGCGCTTCGGAAAAGGACGCGCGGGAA  
CATATCAAGTATCACGTGCGCGAGACTTGGAAAGAAGTGAACAAAGAAGTGTTCAGGACTACCCGTTCCGCAGCTTCGA  
CCCTTTTCTCCGTGCCTGTTTGAACCTGGCACGAGCTTCTCATTGCTTTTACGACTTTGGGGACGGACATGGCCTTCCCG  
GTCACCAAACCAAGAACCATCTTGAGTGGACTATATTTGAACCCGTGCCCTCGATTGA

> *Mtps1* (1659 bp)

CGCGTGACATGCTCTTCGACGATTGGAGATCAAGAGATTGTGAGACGTTCCGGCAAATTTGGCAACCTAGCGTTTGGGACTA  
TGGCTTTGTGCAGTCACTCAGTGTGATTACACGGAGGATAAATATACAAAAGGAAGTCCAAAGGTTGAAGGAAGAAGCCA  
AGGCTCTATTTGGGCAAGGAGATGAATCAGGTGGCCAAGCTCAGATTAATTGACGTGGTTCAAAGACTAGGACTTGGCTAC  
CATTTTGGAGTGGAGATCAAGAGTGGTCTCAGCTCCATCTATAACAGCACTGAAGATGCCCGGATTTCAAACGATCTCTA  
TGCCACTTCCCTTCGATTCCGACTACTTAGGCAACATGGATACGACATACCGCAAGATGTGTTTCGAATGTTTTTGAACA  
AGACGGGCACATTCAACGAATCGCTCAATAAGGATGTGAAGGGCTTCTTGGTCTGTATGAAGCTTCATTTTCATGGATTG  
GAAGGTGAACCTATACCGATGAAGCTTGGAACTTTGCTTCCAAGCATCTGAAGGATCTAAACCTCAACGAAGTTCCCCAC  
CAATATAGCGTGCAGCGTGAGTCACGCATTAGAAATGCCGATCCACTGGAGGCCAAACAGGTTGGAGGCACGGTGGTTCA  
TGGACATGTACGAGAAACAAGAAGACATGATCCCCCTTTTGTGCTGCGATTGGCTAAAATAGATTTTCATTTTAGTGCAGTCA  
ATCCACAGGGACGAACCTAGCAATCTGGCAAGGTGGTGGGTTGAACCTGGGGCGAACAAGATGACCTTCTTTAGGGACAG  
GCTCGTGGAAAACCTATTTTGGGCTGTGATTTCTTTTCGAACCTCAATATACAGATTTCAGAAGATTGGTCCCGAAGC  
TCGGTAATATGGTGACGCTGATCGATGACGTTTATGATATATATGGGACGCCGGAAGAGCTTGAGCTCTTGACAGATTTT  
CTTGTGAGGTGGGACATCACGACGTCGATAAACTTCATCCAACAATAAGGGATAGCTTCAATATCTTGTTCATACGAC  
CAACGAAGTTGGGATTTGGACAATGAGAGAGCGAGGAGTCAACTCCATCCCTTACCTGCCGAAAAGTGTGGCCGGATGAAG  
TCAAGGCGTACATGAAGGAGGTGAGTTGGTACGACAAGGGCATTAAACCACACTGAAGGAGTACATGGACGTTGCGGTG  
GATTCGATCGGAGGGCTGGTATTGCTGTTGGGTTGCTACAGCCTAACCACAGATAAATGACGAAAAGAGGGACTTACTA  
CGTGTCCAAAATCCCGAGTGTGCATGCATCTTCCGCCAAGATCCTTCGACTCAACGATGATCTCAGTACATCATCGCATG  
AATTGGCACGAGGAGACAACCTCAAGGCGCTAGAATGCTACATGAATGAAACTGGCGCGTTCGGAAAAGTTCGCGCGGGAA  
CATATCAAGCATCTGGTGCAGGAGACTTGGAAAGAAGTGAACAAAGAAGTGTTCAGGACTACCCATTCGACGGCTTCGA  
GCCTTTTCTCCGTGCCTGATGAACCTGGCACGAGCTTCTCATTGCTTTTACGATTTTGGAGACGGACACGGCCTTCCCG  
GACACCAAACCAAGAACCATCTCGAGTGGACTATATTTGAACCCGTGCCCTCGATTGA

## Supplementary Material I



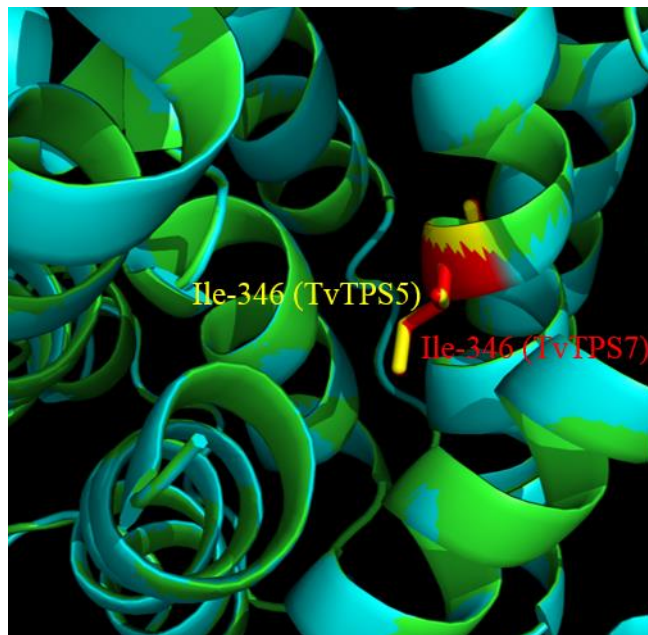
**Fig. S1 Amino acid alignment of TvTPS6 and TvTPS7.** Identical amino acids are shaded in grey. The predicted signal peptide sequence is marked in green, the RR<sub>x</sub>W and DD<sub>xx</sub>D sequence motifs are designated. Amino acids contributing to the active site are marked with a circle, differing amino acids among them are marked with a plus (according to the 3D model, Results I, section 3.1.5). The amino acids important for the stereospecificity of the enzymes are marked in blue (Asn-350 of TvTPS6) and red (Ile-346 of TvTPS7).

```

      10      20      30      40      50      60      70
Tv TPS5  MSTISINHVGI LRNPLQCKNKRTSINKPWSLSLPRSPASRLVKPCRVSSKVDTMPAEITRRSGNYEPSL
      RRxxxxxxxx
      80      90      100     110     120     130     140
Tv TPS5  WDFDFIQSLDDHHPYVKEKQLKREEELIVQVKMLLGTKMEAVKQLELIDDLKKNLGLSYFFREEIKTILTS
      W
      150     160     170     180     190     200     210
Tv TPS5  IYNNSFENNNKVGDLYFTALGFRLLRQHGFNVSQQIFDCFKGNDFDETIIGEDTKGILQLYEASFHLREG
      220     230     240     250     260     270     280
Tv TPS5  ENTLELARQISTKYLQKKVDEGSINDENLSSWIRHSLDPLHWRIQRLEARWFLDAYTVREDKNPLIFEL
      290     300     310     320     330     340     350
Tv TPS5  AKLDFNI IQATQQQELKEVSRWWNDSCLAEKLPFVRDRVVESSFFWAVGLFEGHEFGYQRKMTAI IILIT
      360     370     380     390     400     410     420
Tv TPS5  AIDDLVDVYGTLDLHFLFTDVIRRWDTSIDQLPYMQLCYLALYNFVSDLAYDILKDRGPNPTIPYLQES
      DDxxD
      430     440     450     460     470     480     490
Tv TPS5  WVKLVEAYLKEAEWFESGYTPSLEEYLTNAKTSIGSLTILLQVELSLPNSTIDRAAFDCCHKILYLSATV
      500     510     520     530     540     550     560
Tv TPS5  SRLADDLGTTSSELERGDVPKAIQCYMKDANASEEEAREHVRFMIGEVWKELNAMTEPDDCPFTEQLVE
      570     580     590
Tv TPS5  AAANLGRAAQYIYREGDGHGHFKIHQHIGHNLFHPCV

```

**Fig. S2 Amino acid sequence of TvTPS5.** The signal peptide sequence is marked in green, the RR<sub>x</sub>W and DDxxD sequence motifs are designated. The amino acid corresponding to the amino acids important for the stereospecificity in TvTPS6 and TvTPS7 is marked in yellow (Ile-346).



**Fig. S3 Model of the TvTPS5 and TvTPS7 active sites.** The steric position of Ile-346 (yellow) in TvTPS5 resembles that of Ile-346 (red) in TvTPS7.





**Table S1 Primers used in this study - Chapter I**

| <b>Primer name</b>  | <b>Primer sequence (5'-3')</b>                               |
|---|--|
| Degenerate primers used for isolation of <i>Tvtps6</i> and <i>Tvtps7</i> fragments: |  |
| Lab1-fwd ( <i>Tvtps7</i> ) –outer nested primer                                     | CARTTGGAGTTGATTGATGACTTG                                     |
| Lab2-fwd ( <i>Tvtps6</i> )  | CTCTTGGATTCAGACTYCTCAGACAACATGGTTT                           |
| Lab3-fwd ( <i>Tvtps7</i> ) – inner nested primer                                    | GTGAGGGATAGGCTNGTGGAAWGCTACTTTTGG                            |
| Primers used for isolation of the 5' ends of <i>Tvtps6</i> and <i>Tvtps7</i> :      |  |
| Tv TPS6 rev 9   | GCCCGTCTCCACCTCATACTG  |
| Tv TPS7 rev 7   | CCTCCTCCGAGGCATTGTTG   |
| Primers used for qRT-PCR:   |  |
| Tv TPS6 fwd 3   | TCCTCGAAAACCGATACCAAG  |
| Tv TPS6 rev 8   | CACGTGGGGATGATGATTGTC  |
| Tv TPS7 fwd 4   | ATCGTCCCCACAAAATACTC   |
| Tv TPS7 rev 7   | CCTCCTCCGAGGCATTGTTG   |
| Tv HG 18S fwd (Housekeeping gene)   | ATGA TAACTCGACGGATCGC  |
| Tv HG 18S rev (Housekeeping gene)   | CTTGGATGTGGTAGCCGTTT   |
| Primers used for subcloning into pASK-IBA37+:                                       |  |
| MTPS6-IBA1a fwd ( <i>Tvtps6</i> )   | ATGGTAGGTCTCAGCGCAAGCCTTGCCGTGTCTCCTTGA                      |
| MTPS6-IBA3a fwd ( <i>Tvtps7</i> )   | ATGGTAGGTCTCAGCGCATGCCCGCTGAAATAACCCGAC                      |
| MTPS6-IBA2 rev ( <i>Tvtps6</i> & <i>Tvtps7</i> )                                    | ATGGTAGGTCTCATATCATACGCATGGGTGGAAGAACAAGT                    |
| SfCinS1 Iba37 fwd   | ATGGTAGAAGACAAGCGCCGACGAACTGGAGGCTACCAG                      |
| SfCinS1 Iba37 rev   | ATGGTAGAAGACAATATCACTCATAGCGGTGGAACAGCAAG                    |
| Primers used for site-directed mutagenesis:   |  |
| TPS6-U MUT-2 fwd  | AGTTAATGCTGCTGTTA <sup>t</sup> TATTCTAATCACTGCC              |
| TPS6-U MUT-2 rev  | GGCAGTGATTAGAATA <sup>a</sup> TAACAGCAGCATTA <sup>a</sup> CT |
| TPS7-U MUT-1 fwd  | TTACTGCTGCTATTA <sup>a</sup> TATCCTAATTACTG                  |
| TPS7-U MUT-1 rev  | CAGTAATTAGGATA <sup>t</sup> TAATAGCAGCAGTAA                  |
| TPS10-A MUT-4-fwd   | TGACTACTGCTATTA <sup>a</sup> CATTCTAATTACTGC                 |
| TPS10-A MUT-4-rev   | GCAGTAATTAGAATG <sup>t</sup> TAATAGCAGTAGTCA                 |
| TPS6 MUT-4 fwd  | AATACAGTTGAAGCAT <sup>t</sup> cTTGAAGGAGGCAGAG               |
| TPS6 MUT-4 rev  | CTCTGCCCTCCTTCA <sup>a</sup> gaATGCTTCAACTGTATT              |
| TPS6 MUT-5 fwd  | GAAGCATATTTGAAG <sup>c</sup> tgcGAGAGTGGTACGAG               |
| TPS6 MUT-5 rev  | CTCGTACCACTCTGC <sup>c</sup> agCTTCAAATATGCTTC               |
| TPS6 MUT-6 fwd  | CAGCATCGCTAGTATT <sup>g</sup> tAATAGGTGTTATTC                |
| TPS6 MUT-6 rev  | GAATAACACCTATT <sup>a</sup> cAATACTAGCGATGCTG                |
| TPS6 MUT-7 fwd  | TCAGCGATGGTTTT <sup>a</sup> ctgCTTGCTGATGACCTCG              |
| TPS6 MUT-7 rev  | CGAGGTCATCAGCAA <sup>g</sup> cagTAAAACCATCGCTGA              |
| TPS6 MUT-8 fwd  | GGTTTTAAGGCTTGCT <sup>c</sup> tgcGACCTCGGAACAG               |
| TPS6 MUT-8 rev  | CTGTTCCGAGGTC <sup>c</sup> agAGCAAGCCTTAAAACC                |
| TPS6 MUT-9 fwd  | CGGAACAGCGCAGTAT <sup>c</sup> tgcGTGGAGACGGGC                |
| TPS6 MUT-9 rev  | GCCCGTCTCCAC <sup>c</sup> agATACTGCGCTGTTCCG                 |

## Supplementary Material I

---

---

|                |                                    |
|----------------|------------------------------------|
| TPS7 MUT-4 fwd | GAATTGGTTGAAGCATtcTTGAAGGAGGCTGGG  |
| TPS7 MUT-4 rev | CCCAGCCTCCTTCAAgAATGCTTCAACCAATTC  |
| TPS7 MUT-5 fwd | GAAGCATATTTGAAGctgGCTGGGTGGTACGAG  |
| TPS7 MUT-5 rev | CTCGTACCACCCAGCagCTTCAAATATGCTTC   |
| TPS7 MUT-6 fwd | CACCAACGCCACTATTgtAATAGGCGTTCCTC   |
| TPS7 MUT-6 rev | GAGGAACGCCTATTacAATAGTGGCGTTGGTG   |
| TPS7 MUT-7 fwd | TCTGCAAGGGTTTTGctgCTTGCTGACGATCTTG |
| TPS7 MUT-7 rev | CAAGATCGTCAGCAAGcagCAAAACCCTTGCAGA |
| TPS7 MUT-8 fwd | GGTTTTGAGGCTTGCTctgGATCTTGGAAACAG  |
| TPS7 MUT-8 rev | CTGTTCCAAGATCagAGCAAGCCTCAAAACC    |
| TPS7 MUT-9 fwd | TGGAACAGCATCGTCTctgCTGGAGAGAGGG    |
| TPS7 MUT-9 rev | CCCTCTCTCCAGcagAGACGATGCTGTTCCA    |
| SfCinS1 N338I  | atgctcaccaaaataaTtgctctgttacaac    |
| SfCinS1 N338I  | ggtgaacaagagcaAttatttggtagcat      |

---

Primers used for sequence isolation and subcloning *Tvtps5* into pASK-IBA37+:

---

|                 |  |
|-----------------|--|
| MTPS6-IBA1b fwd | ATGGTAGGTCTCAGCGCAAGCCTTGCCGTGTCTCCTCGA    |
| MTPS6-IBA2 rev  | ATGGTAGGTCTCATATCATAACGCATGGGTGGAAGAACAAGT |

---

Primers used for isolation of *Mttps1*, *Mttps2*, *Mttps3*, *Mltps1*:

---

|  |   |
|--|---|
| CS15C1-2B rev (all)                                | ATGGTAACCTGCATTATATCAATCGAGGGGCACGGGTCAAAT  |
| CS15C3-6A fwd ( <i>Mttps2</i> )                    | ATGGTAACCTGCATTAGCGCGTGACGTGCTCTTCGACGATTG  |
| CS15C4-1A fwd ( <i>Mttps3</i> )                    | ATGGTAACCTGCATTAGCGCGTGACATGCTCTTCAACGATTGG |
| CS15C1-2A fwd ( <i>Mttps1</i> ,<br><i>Mltps1</i> ) | ATGGTAACCTGCATTAGCGCGTGACATGCTCTTCGACGATTGG |

---

## Supplementary Material Chapter II

### DNA sequences of the isolated cytochrome P450 monooxygenases

#### > *CYP71AP17* (1506 bp)

ATGGCTCTCCTTCAGCATCTGTTTCTCTAACCTTGGTAATAGCGGCTGTAATTGCAGTTCTCTTGAAGAACAAGCCTAG  
 GAAAAGAAAGAACAACCTCCCTCCGAGCCCTCCCAAGTTACCGATCCTCGGCAATCTCCATCAGCTAGGCAAATTGCCAC  
 ACATATCCCTCCACCACCTTGCAGAGAACATACGGCCCAATCATCTTCTTACAGCTTGGTGAATCCCAACCGTAGTTGTT  
 TCCTCAGCTAGAATGGCGAAGGAGGTTATGAAAACCCACGACCTCGCACTTTCGAGCCGTCACAGATCTTCTCCGCCAA  
 ACACCTGTCTATGATTGCACTGACGTGGTCTTCTCCCCCTATGGTGCTTATTGGAGGCATATAAGAAAAATCTGCATAC  
 TTGAATTGCTTAGTGTAAAACGGGTTCAATCATACAGTCATGTCAGGGTAGAAGAAGTTGCTCGGTTAGTTGACCGGATT  
 ACAGAGTCTTTTCCAGGCACCACAAATCTAACTAAACTGTTTGGTTTTATATGCAAATGATGTTCTCTGTCTCGGATAGCTTT  
 TGGGAGGGATTTTTTCAGCAGGAGGGGACTACGACAGGGCATGGGTTTTCAAAAGATGCTGGAGGAGTATCAGGAGTTGCTCG  
 GAGGATTCAGCATAGGAGATTTCCCTTCCCTTCCATGGAATTCATCCACAGCCTGACCGGGATGAAGTCCAGACTTCAGCAC  
 ACATTTCCGGCGCTTTGATAACCTTTTCGACGAGATACGAGATCATCAAAAAGAAAACAAAGTACGGAGTTGCCCAA  
 GTATCTTGTGGATGTTCTACTTGTATGTGAAGAACGGTTATGGTGATATGCCTCTGACCAGGGACAATGTTAAAGCCA  
 TCATCTGGATATGTTTGCAGCAGGAACGATACAACCTTCATAACCCCTTGATTGGGGAATGACAGAGCTTGTCAATCAAT  
 CAAAAGCCATGGAAGAGCACAAGCCGAAGTACGAAGCATCGTTGGAGAGAGAAAATTCGTGCAAGAACTGACCTACC  
 TCAACTGCGATACTTGAAGGCTGTCATTAAGGAGATATCCGCTTACACCTCCTGCTCCAGTGTTAGTCCCCAGAGAAT  
 CTATGGAGGACGTAACCTATTGATGGGTACTCCATCCCAGCAAAAACACGTTTCTTCGTCAATGCTTGGTCAATTGGGCGG  
 GACCCAGAATCCTGGGTCGATCCAGAATCGTTTCGACCAGAAAGATTTCTCAATGAAAATAGCAACATTGACTTCAAAGG  
 GCAGGATTTTGAAGTGAACCTTCCGTTGCAGGTAGAAGGAGTTGCCCGCTATTGCATTTGGAAATGCCAGTGTGAGC  
 TTGCTTTAGCTCAACTTCTTACAGTTTCGATTGGGAGCTTCTGATGGGATCCAGCCTAGGGACTTGGATATGACCGAA  
 GTTTTTGGCATCACAATGCACAGAATTGCCAACCTCATGGTTGTAGCCAAACCCACTTCTCTAG

#### > *CYP71AP18* (1503 bp)

ATGGGTCTCCTTCAGCATCTGTTTCTCCCAACATGCTTGGTTCGTAGCAGGTTTGGCTGATTGTTCTCTTGGAGAGCAAGTC  
 TAGAAAAGAAAGAGCAACCTCCACCCGAGCCCTCCTAAGTTGCCGATCATCGGCAATCTTACCAGCTTGGCAAATCGC  
 CACACATATCTCTCCATGGCCTCGCGAGAAAGTACGGGCCAATCATGTCTTGCAGCTCGGCCAAGTCCCGACCATAGTC  
 GTTTCCTCAGCCGCAATGGCCAAGGAGGTGGTAAAACCCATGACCTAGTGTCTCGCAACCCGCTCAGATCTTCTCTGC  
 CAAGCACTTGTTTTATGACTGCACAGACATGGCCTTCTTCCCTACGGCATTATTGGAGGCACATAAGGAAAATCTGCA  
 TACTCGAAGTGTCTAGCGCAAAACGGGTTTCAGTCATTTAGTCATGTCAGGGAGGAAGAAGTTGCTCGTTTTAGTTCATCGT  
 ATTTGCAGAGTCTTTCCAGGCACCACAAATCTTTCCAACTACTCGGTTTATATGCAAATGACGTTCTATGCCGTACTGC  
 TTTGGGAGGGACTTTTCAGCAGGAGGGGATTATGATAAGCATGGATTTCAAAGATGCTGGAGGAGTATCAGGAGCTGC  
 TTGGGGGATTTCAGCATAGGAGATTTCTTCCCTTCCATGGAATTCATTCATAGCCTGACCGGGATCAAATCTAGACTTCAA  
 CACACTTTCAAGCGCTTCGATAATCTTTTTGACCAGATACTGAGAGAACATTCAAATGAAAGCAGAGCTACGGAGGTGCA  
 GAAGGATCTAGTGGATGTTTTGATCGATATCGAGAAGAATGGTTACGGTGACATGCCCTTACGAGAGCCAAATGTTAAAG  
 CCCTCATCTGGATTTTTTTCAGCAGGAAACAGATACAACATTCATCACTCTTGACTGGGCAATGACAGAGCTTGTCTATG  
 AACCTAAAGCCATCAAAAAGAGCACAAGCCGAATTACGGAGCATCGTGGGAGAGACAAAATACGTGCAAGAGACTGACCT  
 ACCTCAATTACCATACTGAAGGCTATCATCAAGGAGGATTTCCGACTGCATCCTCTGCTCTGTGTACTCCCAAGAG  
 AATCCATGGAGACGTAATAATTGATGGATACTGCATTTCCGGCAAAAACCGGTTTCTTTGTTAATGCTTGGTCAATAGGG  
 CGTGATCCGCAATCCTGGGTTGATCCTGAAGCATTCTACCACAAAGGTTTTTGGATAGCACCATCGACTTCAAGGGACA  
 GGACTTCGAGCTGATACCGTTTGGCGCAGGCCGAAGAAGCTGCCAGCTATTGCATTTGGAAATGCCAGTGTAGAGATTG  
 CTTTAGCTCAACTTCTTACAGTTTCGATTGGCAACTTCTCCGGGTGTCCAGCCAAAAGATTTGGACATGGAAGAAGCT  
 TTTGGCATCACAATGCATAGGATCGAAAACCTCATTTGTCATTGCCAAACCCACTTCTCTAG

#### > *CYP71AP19* (1506 bp)

ATGGCTCTCCTTCAGCATCTGTTTCTCTAACCTTGGTAATAGCGGCTGTAATTGCAGTTCTCTTGAAGAACAAGCCTAG  
 GAAAAGAAAGAACAACCTCCCTCCGAGCCCTCCCAAGTTACCGATCCTCGGCAATCTCCATCAGCTAGGCAAATTGCCAC  
 ACATATCCCTCCACCACCTTGCAGAGAACATACGGCCCAATCATCTTCTTACAGCTTGGTGAATCCCAACCGTAGTTGTT  
 TCCTCAGCTAGAATGGCGAAGGAGGTTATGAAAACCCACGACCTCGCACTTTCGAGCCGTCACAGATCTTCTCCGTCAA  
 ACACCTGTCTATGATTGCACTGACGTGGTCTTCTCCCCCTATGGTGCTTATTGGAGGCATATAAGAAAAATCTGCATAC  
 TTGAATTGCTTAGTGTAAAACGGGTTCAATCATACAGTCATGTCAGGGTAGAAGAAGTTGCTCGGTTAGTTGACCGGATT  
 ACAGGGTCTCTTCCAGGCACCACAAATCTAACTAAACTGCTTGGTTTTATATGCAAATGATGTTCTCTGTCTCGGATAGCTTT  
 TGGGAGGGATTTTTTCAGCAGGAGGGGACTACGACAAGCGTGGGTTTTCAAAAGATGCTGGAGGAGTATCAGGAGTTGCTCG  
 GAGGATTCAGCATAGGAGATTTCTTTCCCTTCCATGGAATTCATCCACAGCCTGACCGGGATGAAGTCCAGACTTCAGCAC  
 ACATTTCCAGCGCTTTGATAACCTTTTCGACGAGATACTGAGAGATCATCAAAAAGAAAACAAAGCTACGGAGTTGCACAA  
 GTATCTTGTGGATGTTCTACTTGTATATACTGAAGAACGGTTATGGTGATATGCCTCTGACCAGGGACAATGTTAAAGCCA  
 TCATCTGGATATGTTTGCAGCAGGAACGATACAACCTTCATAACCCCTTGATTGGGGAATGACAGAGCTTGTCAATCAAT  
 CAAAAGCCATGGAAGAGCACAAGCCGAAGTACGAAGCATCGTTGGAGAGAGAAAATTCGTGCAAGAACTGACCTACC  
 TCAACTGCGATACTTGAAGGCTGTCATTAAGGAGATATCCGCTTGCACCCTCCTGCTCCAGTGTTAGTCCCCAGAGAAT  
 CTATGGAGGACGTAACCTATTGATGGGTACTCCATCCCAGCAAAAACACGTTTCTTCGTCAATGCTTGGTCAATTGGGCGG

## Supplementary Material II

---

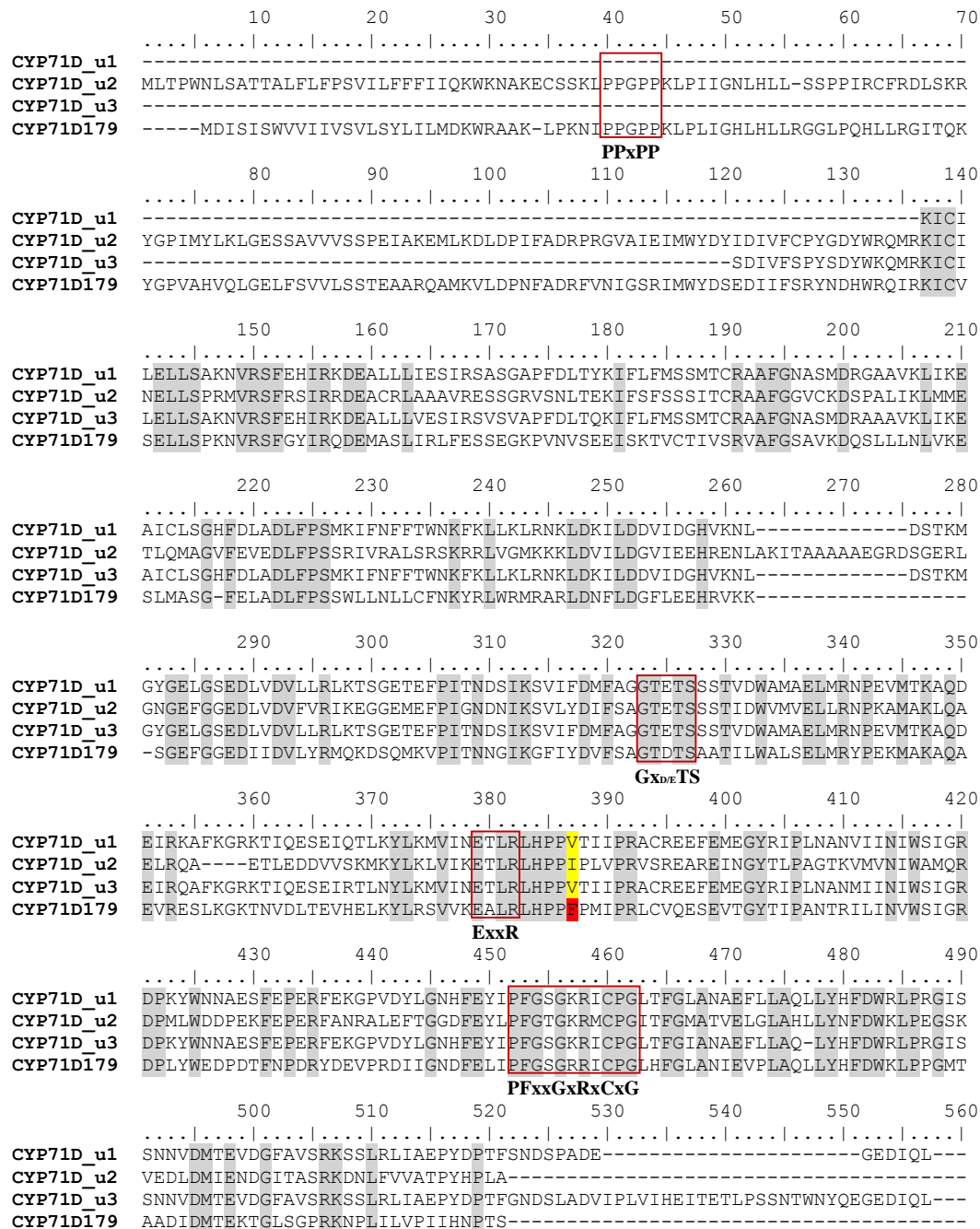
GACCCAGAATCCTGGGTGATCCAGAATCGTTTCGACCAGAAAGATTTCTCAATGAAAATAGCAACATTGACTTCAAAGG  
GCAGGATTTTGAGCTGATACCCTTCGGTGCAGGTAGAAGGAGTTGCCCGCTATTGCGTTTGGAAATGCCAGTGTGAGC  
TTGCTTTAGCTCAACTTCTTACAGTTTCGATTGGGAGCTTCTGATGGGATCCAGCCTAGGGACTTGGATATGACCGAA  
GTTTTTGGCATCACAATGCACAGAATTGCCAACCTCATGGTTGTAGCCAAACCCCACTTCTCCTAG

### > *CYP71AH18* (1503 bp)

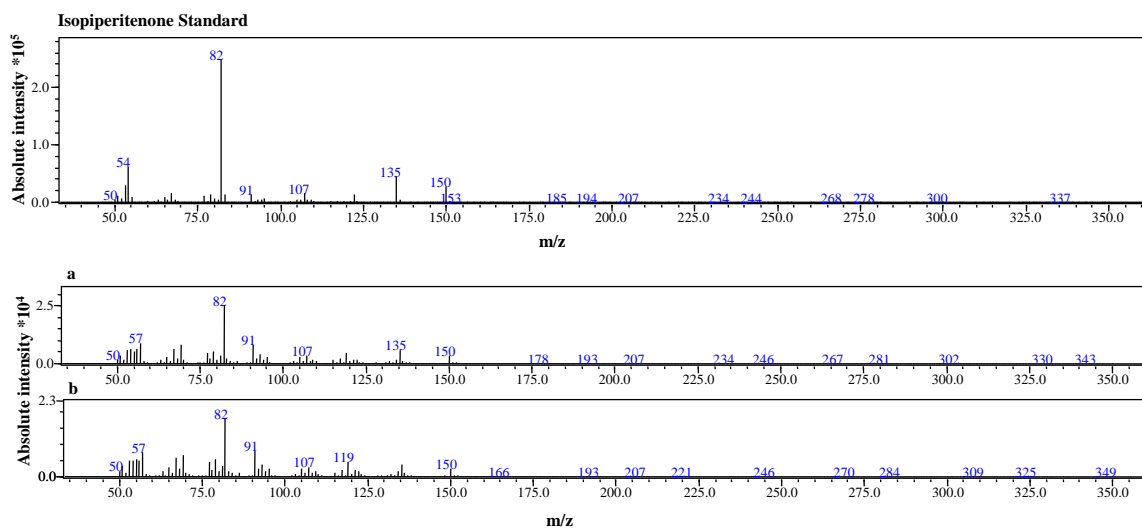
ATGGGTGTTTCAGCTTATGTAGCCCTCATATGCGCTCTTGTGATACCGTGTACTAGTCTTTTTGATTGACACGAAAA  
GAAGACACAAGCATCTAGGAATCTCCCTCCCGGTCTTAGGAAGTTGCCGATCATCGGGAATTTGCACCAGCTAGGCTCCT  
TGCCTCACCGATCGCTCACACGCCTCTCAAAACAGTATGGCCAGATCATGTTACTCCAGCTCGGGTCTGTTCCACATTA  
GTCATCTCGTGAAGATGTCGCCAGAGAAGTATTCAAGGCCACGACCGCTTTCTCCGGCAGACCTGTCTTTTATGC  
AGTGAAGAAGTTGACATACAATTGTTCTGATGTAAGTTTTGGTGCCTATGGCGATTCTTGAGAGAGCTCAGGAAAAATAG  
TGATCCTAGAACTCCTGAGCAGTAAGAGAGTCCAGTTGTTCAATCTGTGAGGAAGGAAGAGGTTAAGCTCATGCTAGAC  
GCAATCACTTCTTCTCCAGGTCCAGTCAATATTGGTGAATGGCACTTTTATTGGCTAACAATCTCGTGTGCCGAGTGGC  
TTTTGGTAAGAAATGGCAGGCTGAAGGTGGTGGTCGCAAGAGCAAATTTACGAGACAGCCACGAAATACAGAGCATT  
TAGGGGGCTTTTGTGTTGCGGATCTGTTTCCACAGATAGCTTGGTTTAAACAGGCTCAATGGCTTCAAGGCGAAGGTTGAA  
AAGAACTTCGGGGAGTTAGACAAGTTGTACGATGAAGTGATCGAGGAGCACCAAGACCTGAAAGGCCATAACCCGATCA  
CGAAGATCTTGTGACGTGTGCTTCGGTTACAAAGGGATCCAAATCAAATGATTGCCCTCACCGGGGAGCAAATTAAG  
GAGTATTAAGTACATGTTCAATGCAGGAACAGACACCTCTGCAACCACTATACTCTGGACAATGACAGAGCTCGTCCGC  
AATCCAGCTATGATGAGAAAAGCACAAGAAGAGGTTGAGAAAGCAGCAAAGGAAAGTACAGGTTGAAGAGACTGACCT  
TCTAGGACTCATCTACCTTAGATCAGTCAAAAGAGGATTAAGACTCCATCCACCGCTCCCTCTTCTAGTACCAAGAG  
CGACGATTGAGGATTGCAAGATAAGGGGATACGAAGTTCTAGAGGAACACTGTATTCAATGTGACAGCAATCTCC  
ACAGACCCAAAATCTTGGGAAAACCCGGAAGAGTTGAGCCCGAGAGGTTCTTGAACAGCTCCATCGATTTAAGGGACA  
GAACTATGAGTTTTTGCATTTGGTTCTGGCAGGCGGGGATGCCCCGGCATTAACTTCAGTGTGATAGTTGAGCTTG  
CATTGGCAAATCTTCTCATCGGTTTGTGATTGAAACTGCCTGAAGGGATGAGTATCGAAGATGTCGACATGGAAGAAGCT  
TATGGCCTTACGACTCGCAAGAGGACTCCTCTTTCCTGATAGCAACCCCGTGAGTGGTTGA

### > *CYP71CJ1* (1463 bp)

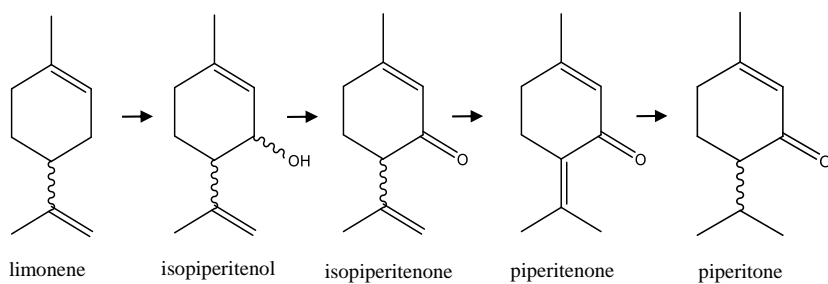
ATGCTCATCAAAATTTTCAGAGCAAAAACGTTCCAATGCTTTGAAGTTACCCCCGGGGCCAAAGAAACTGCCCTTATCGG  
AAATTTGCCCCAGTTGATAAGTCCCGAGCTGCCACACCATCATCTGAATAACCTAGCAAAAACCTACGGTCCGATTTATC  
ACCTCAGGGTCGGCGAGCTCATGCTGGTCGTTCTTACGACACCTGAAATTGCCGAAGAGGTCCTAAAGACGCACGAGATC  
AATTTGCTCAGAAGCCACTATTCCCTCTTTAGAGATGATGTTGACACATGCTCAAGCTTCATGTATGCACCTTATGG  
AGAACATTTGAGGCAGTTTCGCAAGATTTGTGCTTGGAACTTCTTAGCATCAAACGCGTTCCGGTCTTTCCAATCAATAA  
GAGAAGATGAGATGAGTAATCTGGTCCAGTTCGTGCGCTCGTCGAGGTTGCCCTCTCAATCTTACGAAAAGATCTCT  
TCATGCTTAAATAACATAATATCCAAAGCAGCATTGTTGGTGAAGCTATAAGTGCCAGGAGTACCTTTTCAGCTCTTAAAGA  
CGCAATCGAGCAGGCGGGGTTAAGTGTGCTGGATGTTTTCCCATCCCTCAAAATCCTTCGTTACATTAGTGGGAGGA  
AATGGAGGCTAGACAGTGTTCGGCGAAGATGTGATCAAATCTCAAATCATATCCTCAATGATCATAAGAGGCAGAGAGAA  
AACTTAAGAACCAGGAAGATAGCTTTGTCAAGGAAGATTTGTGATGACTTTTACGGATTCAAGAGTCAAACGAAC  
TGGATTTTCATCTTACCAGCACCACATTAATACTATCATCTGGAATTCATGGTGGTGGAAACCGATGCTTCATCAGTGA  
CAGCAGAAATGGCAATGTCCGAACTGCTCAAGAACCACAAGCGATGGCAAAGCGCAGGCTGAGGTACGCATGGCATTG  
AAAGGAAAGGATCGAGTCAAGAATCGGATCTCGAGGATCTTAAAGTACTTAAATCAGTGATCAAAGAGACTCTTCGGTT  
ACACCCTCCAGTACCATTAGTCCCCAGAGAAGCGAGGAAGGCATGCAAGATCAGGGGCTACGATATACCGGTGAAATCGA  
GGTGCTTATTCATGCCGGGCATTTGGAAGGATCCAAATCATTGGGAAGACCCAGAAAAATTTGTGCCGGAGAGGTTT  
CTTGAATCATCAGTAGACTTTATCGGGACACATTACCATTTTCGTTCCATTTGGGTTTGGCAGGAGGTTGTCCGGGAT  
CGCATTTGCTGTGGCAATATTGAGCTTACTAGCACTTTTGTGTACCATTTTCGACTGGGCACTTCAAATGGCCAAA  
CTCCTGAAGAACTCGACATGACCGAGGCTTTCGCGGCCACGGTCACGAGAAAGAGTGATCTTTATGTGGTTGCCACCCCT  
CATCGTCATGGCCATGCCTG



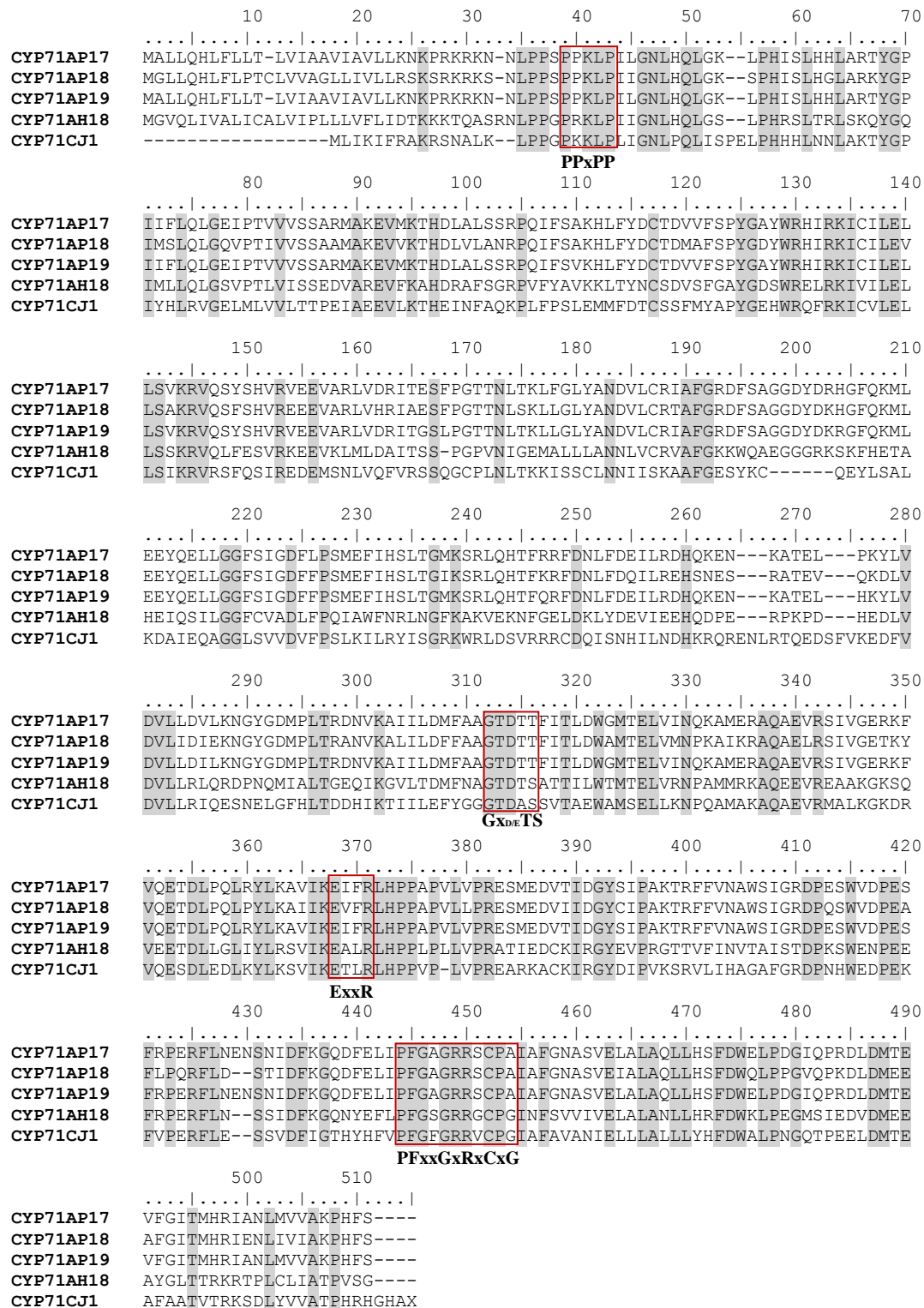
**Fig. S5 Amino acid alignment of the identified (partial) P450 enzyme sequences CYP71D\_u1-u3 and CYP71D179.** Identical amino acids are shaded in grey. Important sequence motifs are designated: The proline-rich region (PPxPP), the Gx(D/E)TS motif, the ExxR motif, and the heme binding sequence (PFxxGxRxCxG). The amino acids putatively important for the regiospecificity of the enzymes are marked in yellow (Val/Ile in CYP71D\_u1-u3) and red (Phe-362 in CYP71D179).



**Fig. S6** Comparison of the recorded mass spectra of the isopiperitenone GC standard and substances "a" and "b" (recorded in the limonene feeding experiment with *Thymus vulgaris*, Results II, section 3.1.4).



**Fig. S7** Proposed pathway from limonene to piperitone. The first step to isopiperitenol is most likely performed by a cytochrome P450 enzyme (Diemer et al., 2001).



**Fig. S8** Amino acid alignment of the *Eucalyptus* cytochrome P450 enzymes CYP71AP17, CYP71AP18, CYP71AP19, CYP71AH18, and CYP71CJ1. Identical amino acids are shaded in grey. Important sequence motifs are designated: The proline-rich region (PPxPP), the Gx(D/E)TS motif, the ExxR motif, and the heme binding sequence (PFxxGxRxCxG).

Table S2 Primers used in this study - Chapter II

| Primer name   | Primer sequence (5'-3')                           |
|---|---|
| Primers used for site-directed mutagenesis:   |   |
| 181 L364M fwd-2   | CATCCACCATTCCCAATGATTCCAAGATTG                    |
| 181 L364M rev-2   | CAATCTTGGGAATCATTGGGAATGGTGGATG                   |
| 179 M364L fwd-2   | CATCCACCATTCCCAATGATTCCAAGATTG                    |
| 179 M364L rev-2   | CAATCTTGGGAATCAATGGGAATGGTGGATG                   |
| 179 V370K fwd   | TTCCAAGATTGTGT <sup>aaa</sup> CAAGAATCCGAAGTTACC  |
| 179 V370K rev   | GGTAACTTCGGATTCTT <sup>g</sup> ttACACAATCTTGGAA   |
| 179 Q371E fwd   | TTCCAAGATTGTGT <sup>g</sup> TTAAGAATCCGAAGTTACC   |
| 179 Q371E rev   | GGTAACTTCGGATTCTT <sup>c</sup> AAACACACAATCTTGGAA |
| 179 V370K Q371E fwd   | TTCCAAGATTGTGT <sup>aaag</sup> AAGAATCCGAAGTTACC  |
| 179 V370K Q371E rev   | GGTAACTTCGGATTCTT <sup>cttt</sup> ACACAATCTTGGAA  |
| 181 K370V fwd   | TTCCAAGATTGTGT <sup>gtt</sup> GAAGAATGCGAAGTTACC  |
| 181 K370V rev   | GGTAACTTCGCATTCTT <sup>Caac</sup> ACACAATCTTGGAA  |
| 181 E371Q fwd   | TTCCAAGATTGTGT <sup>AAA</sup> cAAGAATGCGAAGTTACC  |
| 181 E371Q rev   | GGTAACTTCGCATTCTT <sup>g</sup> TTTACACAATCTTGGAA  |
| 181 K370V E371Q fwd   | TTCCAAGATTGTGT <sup>gttc</sup> AAGAATGCGAAGTTACC  |
| 181 K370V E371Q rev   | GGTAACTTCGCATTCTT <sup>gaac</sup> ACACAATCTTGGAA  |
| 179 Y292F fwd   | CAAGGGTTTCATCT <sup>t</sup> CGATGTTTTTCCGC        |
| 179 Y292F rev   | GCGGAAAAACATCG <sup>a</sup> AGATGAAACCCTTG        |
| 179 A304T fwd   | CTGATACTTCTGCT <sup>a</sup> CTACTATTTTGTGGG       |
| 179 A304T rev   | CCCACAAAATAGTAG <sup>t</sup> AGCAGAAGTATCAG       |
| 179 K474T fwd   | GATATGACTGAAA <sup>c</sup> GACTGGTTTGTCCGGTCC     |
| 179 K474T rev   | GGACCGGACAAACCAGTC <sup>g</sup> TTTCAGTCATATC     |
| 179 T475P fwd   | GATATGACTGAAA <sup>Ag</sup> cCTGGTTTGTCCGG        |
| 179 T475P rev   | CCGGACAAACCAG <sup>g</sup> CTTTTCAGTCATATC        |
| 181 F292Y fwd   | GTATTAAGGGTTTCATCT <sup>a</sup> CGATGTTTTCTCC     |
| 181 F292Y rev   | GGAGAAAACATCG <sup>t</sup> AGATGAAACCCTTAATAC     |
| 181 T304A fwd   | CTGAAACTTCTGCT <sup>g</sup> CTACTATTCAATGG        |
| 181 T304A rev   | CCATTGAATAGTAG <sup>c</sup> AGCAGAAGTTTCAG        |
| 181 T474K fwd   | GGATATGTCTGAAA <sup>ag</sup> CCAGGTTTGTCTGG       |
| 181 T474K rev   | CCAGACAAACCTGG <sup>ct</sup> TTTCAGACATATCC       |
| 181 P475T fwd   | GATATGTCTGAAACT <sup>a</sup> CAGGTTTGTCTGGTC      |
| 181 P475T rev   | GACCAGACAAACCTG <sup>t</sup> AGTTTCAGACATATC      |
| Primers used for isolation of cytochrome P450 enzymes from <i>Eucalyptus</i> species: |   |
| Eg D01189 fwd ( <i>CYP71CJ1</i> )   | ATGCTCATCAAAATTTTCAGAGC                           |
| Eg D01189 rev   | TCAGGCATGGCCATGACGATG                             |
| Eg B00198 fwd ( <i>CYP71AP17-19</i> )   | ATGGCTCTCCTTCAGCATCTG                             |
| Eg B00198 rev   | CTAGGAGAAGTGGGGTTTGGC                             |
| Eg B00184 fwd ( <i>CYP71AH18</i> )  | ATGGGTGTTTCAGCTTATTGTAGC                          |
| Eg B00184 rev   | TCAACCACTCACGGGGTTTGC                             |
| Primers used for subcloning in pESC-LEU2d yeast expression vector:                    |   |
| BO0198 pESC fwd ( <i>CYP71AP17</i> )  | GTCGAC <sup>atg</sup> GCTCTCCTTCAGCATCTG          |
| BO0198 pESC rev   | AAGCTTCTAGGAGAAGTGGGGTTTGG                        |
| BO0184 pESC fwd ( <i>CYP71AH18</i> )  | GGGCC <sup>gatg</sup> GGTGTTCAGCTTATTG            |



---

|                              |                            |
|------------------------------|----------------------------|
| BO0184 pESC rev              | CTCGAGTCAACCACTCACGGGGGTTG |
| BO0198 pESC fwd2 (CYP71AP18) | GTCGACatgGGTCTCCTTCAGCATC  |
| BO0198 pESC rev2             | CTCGAGGGAGAAGTGGGGTTTGG    |
| DO1189 pESC fwd (CYP71CJ1)   | GTCGACATGCTCATCAAAATTTTCAG |
| DO1189 pESC rev              | GCTAGCCAGGCATGGCCATGACGATG |

---

**Table S3 Amino acid sequence identity of CYP71D\_u1-u3 and other P450 enzymes from the CYP71D subfamily**

|                  | <b>CYP71D412</b><br>( <i>Salvia miltiorrhiza</i> ) | <b>CYP71D413</b><br>( <i>Salvia miltiorrhiza</i> ) | <b>CYP71D374</b><br>( <i>Salvia miltiorrhiza</i> ) | <b>CYP71D51v3</b><br>( <i>Nicotiana tabacum</i> ) |
|------------------|--|--|--|---|
| <b>CYP71D_u1</b> |  | <b>81 %</b>  | <b>56 %</b>  | <b>56 %</b>                                       |
| <b>CYP71D_u2</b> | <b>71 %</b>  |  |  |   |
| <b>CYP71D_u3</b> |  | <b>81 %</b>  | <b>55 %</b>  | <b>56 %</b>                                       |



## Danksagung

An erster Stelle möchte ich mich ganz besonders bei Herrn Prof. Dr. Jörg Degenhardt bedanken, für die Bereitstellung des interessanten Themas und die Betreuung und stetige Unterstützung während meiner gesamten Doktorandenzeit.

I would like to thank Prof. William Foley for the great cooperation and the opportunity to work on the tea tree and eucalyptus projects. A special thank you also for organizing the field trip along the East Coast of Australia in November 2012.

Ich danke ganz herzlich den Mitgliedern meiner Arbeitsgruppe für die großartige Arbeitsatmosphäre, die Unterstützung im Labor und die hilfreichen wissenschaftlichen Diskussionen. Ein besonderer Dank geht dabei an Annett Richter, Claudia Schaff, Marko Walkowiak, Jette Schimmel und Natalie Arndt.

Jette Schimmel danke ich außerdem für das Gegenlesen meiner Arbeit und Natalie Arndt danke ich für die Vorbereitung der *Thymus vulgaris* Transkriptom-Sequenzierung.

Weiterhin danke ich Kathrin Thomasch für die Hilfe bei allen technischen Angelegenheiten im Labor und Bestellungen in letzter Minute.

Kathrin Kunth möchte ich für die Bewältigung aller bürokratischer Angelegenheiten danken.

Ich danke Peter Lindemann für die technische Hilfe, besonders bei GC-Angelegenheiten.

Ein großer Dank geht an Tobias Köllner für die stetige Unterstützung, die unzähligen hilfreichen wissenschaftlichen Diskussionen und die mehrmalige Durchsicht dieser Arbeit.

Ein besonderer Dank geht an Christoph Crocoll für die Einführung in das Thema der Thymol- und Carvacrolsynthese in Thymian, die Überlassung der Sequenzen von CYP71D179 und CYP71D181 zur weiteren Untersuchung und die vielen hilfreichen Diskussionen.

Julia Asbach möchte ich für die Einführung in die Welt der Terpensynthesen danken und außerdem für die Bereitstellung erster Teilsequenzen von Monoterpensynthesen (TvTPS6 und TvTPS7) aus *Thymus vulgaris*.

I would like to thank Prof. Dr. Jonathan Gershenzon for helpful discussions, supporting advice and the opportunity to conduct P450 expression experiments at the MPI for Chemical Ecology, Jena.

Ich danke ganz herzlich Sandra Irmisch, besonders für die Hilfe bei den P450-Expressionen und die vielen nützlichen Ratschläge.

I would like to thank Amanda Padovan for the great cooperation in the field of terpene synthases in *Eucalyptus* and *Melaleuca* species, and the help with the GC analyses of

eucalyptus and tea tree leaf samples. Furthermore, a big thank you for the pleasant time during my stay in Canberra and on the field trip in November 2012.

I would like to thank Andras Keszei and Carsten Kühlheim for the great cooperation in the field of terpene syntheses in *Eucalyptus* and *Melaleuca* species, the pleasant time in Canberra and helpful discussions.

Ich danke Bettina Rahfeld für die Bereitstellung von Mikroskopieaufnahmen und hilfreiche Diskussionen.

Vielen Dank an Nicole Reinhardt für die Durchsicht des Dehydrogenase-Kapitels und die nützlichen Anmerkungen.

Ein herzlicher Dank geht an Prof. Dr. Peter Imming für die Hilfe bei chemischen Fragestellungen zu den P450-Aktivitätstests.

I thank John Thompson for providing the *Thymus vulgaris* chemotypes.

Ich danke Fabian Bull für die Erstellung des BLAST-Programmes für das *Thymus vulgaris*-Transkriptom.

Ich danke Mark Lange für die Überlassung von GC-Standards.

Ich möchte Alexandra Irrgang und Stefanie Müller ganz herzlich danken für das Gegenlesen meiner Arbeit und die freundschaftliche und moralische Unterstützung während Höhen und Tiefen meiner Doktorandenzeit.

Nicht zuletzt geht der größte Dank an meine Eltern und meine Schwestern. Eure Unterstützung in jeglicher Hinsicht war stets das Wichtigste für mich.

## Scientific Articles:

**Krause, S.T., Köllner T.G., Asbach, J., and Degenhardt, J.** (2012). Stereochemical mechanism of two sabinene hydrate synthases forming antipodal monoterpenes in thyme (*Thymus vulgaris*). Archives of Biochemistry and Biophysics **529**, 112-121.

**Vezzaro, A., Krause, S.T., Nonis, A., Ramina, A., Degenhardt, J., and Ruperti, B.** (2012). Isolation and characterization of terpene synthases potentially involved in flavor development of ripening olive (*Olea europaea*) fruits. Journal of Plant Physiology 169(9), 908-914.

**Irmisch, S., Krause, S.T., Kunert, G., Gershenzon, J., Degenhardt, J., and Köllner, T.G.** (2012). The organ-specific expression of terpene synthase genes contributes to the terpene hydrocarbon composition of chamomile essential oils. BMC Plant Biology 12(1), 84.

**Padovan, A., Patel, H.R., Chuah, A., Huttley, G.A., Krause, S.T., Degenhardt, J., Foley, W.J., and Kühlheim, C.** (2015). Transcriptome sequencing of two phenotypic mosaic *Eucalyptus* trees reveals large scale transcriptome re-modelling. PLoS ONE, in press.

**Kühlheim, C., Padovan, A., Hefer, C.A., Krause, S.T., Degenhardt, J., and Foley, W.J.** (manuscript formally accepted). The *Eucalyptus* terpene synthase gene family. BMC Genomics.

**Schimmel, J., Krause, S.T., Asbach, J., and Degenhardt, J.** (manuscript in preparation). Chemotype regulation in thyme is based on terpene synthase expression patterns.

**Keszei, A., Padovan, A., Hassan, Y., Krause, S.T., Köllner, T.G., Degenhardt, J., Gershenzon, J., and Foley, J.W.** (manuscript in preparation). Expression of terpene synthases explains chemotypic variation in *Melaleuca alternifolia*.

## Oral presentations:

**Krause, S.** (2011). "Stereochemical mechanism of two sabinene hydrate synthases forming antipodal monoterpenes in thyme (*Thymus vulgaris*)". Terpnet 2011 - The 10th International Meeting: Biosynthesis and Function of Isoprenoids in Plants, Microorganisms and Parasites, Kalmar, Sweden.

## Poster presentations:

**Krause, S.T., Crocoll, C., Gershenzon, J., and Degenhardt, J.** (2012). "The stereospecificity of a P450 family involved in the biosynthesis of thymol and carvacrol." 11<sup>th</sup> International Symposium on Cytochrome P450 Biodiversity&Biotechnology, Torino, Italy.



## Curriculum vitae

### Personal:

Name: Sandra Krause  
 Date of birth: 21.09.1984  
 Place of birth: Karl-Marx-Stadt, jetzt Chemnitz  
 Address: Bernhardystraße 31, 06110 Halle (Saale), Germany  
 Citizenship: German

### Education School:

08/1994 – 06/2003 Dr.-Wilhelm-André-Gymnasium, Chemnitz  
 Abitur 2003

### Education University:

10/2004 – 10/2009 University Studies of Pharmacy,  
 Martin-Luther-University Halle-Wittenberg  
 03/2010 License to practice pharmacy (Approbation)  
 11/2008 – 04/2009 Diploma thesis, Institute for Pharmacy, Dep. of  
 Pharmaceutical Biotechnology, Martin-Luther-University  
 Halle-Wittenberg  
 Prof. Dr. Jörg Degenhardt  
 07/2010 Diploma in pharmacy

### Professional Carrer:

12/2009 – 11/2013 PhD Student, Dep. of Pharmaceutical Biotechnology,  
 Martin-Luther-University Halle-Wittenberg  
 Prof. Dr. Jörg Degenhardt  
 05/2011 – 10/2013 Pharmacist, "Apotheke am Händelhaus", Halle (Saale)  
 12/2013 – 10/2014 Pharmacist, "Ginkgo-Apotheke", Lutherstadt Eisleben  
 03/2015 until now Research Associate, Dep. of Pharmaceutical Biotechnology,  
 Martin-Luther-Universität Halle-Wittenberg  
 Prof. Dr. Jörg Degenhardt  
 Prof. William Foley

Halle (Saale), 01.06.2015

Sandra Krause





## **Selbstständigkeitserklärung**

Hiermit versichere ich, dass diese Arbeit in gleicher oder ähnlicher Form bisher keiner anderen Prüfbehörde zur Erlangung des Doktorgrades vorgelegt wurde. Ferner erkläre ich, gemäß der Promotionsordnung der Naturwissenschaftlichen Fakultät I – Biowissenschaften der Martin-Luther-Universität, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Alle Stellen, die wörtlich oder sinngemäß aus Veröffentlichungen entnommen sind, habe ich als solche kenntlich gemacht.

Halle (Saale), den 01.06.2015

Sandra Krause