Martin-Luther-Universität Halle-Wittenberg
Fachbereich Biochemie/Biotechnologie

Expression and characterization of terpene synthases from *Cannabis sativa* L. and *Salvia sclarea* L.

Doctoral Thesis

prepared at the
Gottfried Wilhelm Leibniz Institut für Pflanzenbiochemie
in Halle (Germany)

and at the
Donald Danforth Plant Science Center
St. Louis, MO (United States of America)

submitted to the
Naturwissenschaftliche Fakultät I Biowissenschaften
of the Martin-Luther Universität
Halle-Wittenberg

from
Nils Günnewich
born at the 05.23.1977 in Minden/Westphalia

Referee/Gutachter

1.

2.

3.

Halle an der Saale (Germany), November 2008

Ich bin das Licht der Welt;
wer mir nachfolgt, der wird nicht wandeln in der Finsternis,
sondern wird das Licht des Lebens haben.
Johannes 8 Vers 12

Widmen möchte ich diese Dissertation meinen Eltern, die mir die Möglichkeit gegeben haben meinen Traum ein Wissenschaftler zu werden ermöglicht haben. Meiner Mutter, die mir mein ganzes Leben Ihre Liebe, Wärme und Geborgenheit gegeben hat. Meinem Vater, der mir immer ein Vorbild in Aufrichtigkeit, Ehrlichkeit und im Umgang mit Mitmenschen ist.

Danken möchte ich meiner Gefährtin Andrea Barbara Ringel, für ihre Liebe, Geborgenheit und Stärke. Du hast mir Sinn gegeben und meine Leere gefüllt.

Publications follows are part of this thesis:

Paper:

Günnewich N, Page JE, Köllner TG, Degenhardt J, Kutchan TM. (2007) Functional expression and characterization of trichome-specific (-)-limonene synthase and (+)-α-pinene synthase from *Cannabis sativa*. *Natural Product Communications.* **2,** 223-232.

Brandt W, Weber R, Günnewich N, Bräuer L, Schulze D, Rausch F, Page JE, Schmidt J, Kutchan TM, Wessjohann LA. (2008) Modelling of a limonene/pinene synthase and product specifity of the two monoterpene synthases of *Cannabis sativa* L.. (in progress)

Poster:

Botanikertagung (botany meeting) Braunschweig 2004

Unravelling the biosynthesis of the fragrance compound sclareol in *Salvia sclarea* L.
Nils Günnewich, Kumboo Choi and Toni M. Kutchan

The studies of geranylgeranyl diphosphate phosphatase in *Croton stellatopilosus* ohba.
N. Nualkaew, W. De-Eknamkul, N. Günnewich, Toni M. Kutchan and M.H. Zenk

55th International Congress and Annual Meeting of the Society for Medicinal Plant Research (Graz, Austria; 2007), published in:

N Nualkaew, N Guennewich, K Springob, W De-Eknamkul, MH Zenk, TM Kutchan. (2007) cDNA cloning of prenyl diphosphate phosphatase from *Croton stellatopilosus* Ohba. *Planta Medica,* **73 (9),** 1019-1020.

Sequences**:**

DQ 839404 - Cannabis sativa (-)-limonene synthase mRNA, complete cds
DQ 839405 - Cannabis sativa (+)-alpha-pinene synthase mRNA, complete cds

Table of contents

Table of contents I-IV

List of abbreviations V-VI

1. Introduction 1

1.1. The plants 1

1.1.1. *Cannabis sativa* L. 2

1.1.2. *Salvia sclarea* L. 2

1.1.3. Glandular Trichomes 4

1.2. Terpenoids 5

1.2.1. Terpenoid biosynthesis 6

1.2.2. Isopentenyl pyrophosphate biosynthesis 6

1.2.2.1. The mevalonate pathway (MVA-Weg) 8

1.2.2.2 The methylerithritolphosphate-pathway (MEP-Pathway) 8

1.3. Prenyltransferases 9

1.3.1. Short-chain prenyltransferases 10

1.3.2. Long-chain prenyltransferases 11

1.3.3. Aromatic prenyltransferases 12

1.3.4. Protein prenyltransferases 13

1.4. Terpenoid synthases or cyclases 14

1.4.1. Monoterpene synthases 14

1.4.2. Sesquiterpene synthases 15

1.4.3. Diterpene synthases 16

1.5. Compartmentation of the terpenoid-pathways 17

1.6. Preliminary results and aim of the thesis 18

2. Materials and methods 20

2.1. Materials 20

2.1.1. Organisms 20

2.1.1.1. Plants 20

2.1.1.2. Bacterial hosts 20

2.1.2. Plasmids and oligonucleotides 21
2.1.2.1. Plasmids 21
2.1.2.2. Oligonucleotides 22
2.1.2.3. Nucleotides 24
2.1.3. Biochemicals (enzymes, proteins, antibiotics and antibodies) 24
2.1.4. Chemicals 24
2.1.5. Miscellaneous (columns, reagents etc.) 27
2.1.6. Equipment 28
2.2. Microbiological methods 29
2.2.1. Media and agar plates 29
2.2.2. Transformation 29
2.2.3. Preparation of competent cells 29
2.3. Gelelectrophoresis 30
2.3.1. DNA/RNA-agarose gels 30
2.3.2. Protein polyacrylamide gel electrophoresis (SDS-PAGE) 31
2.3.2.1. SDS-PAGE for sequencing with a mass-spectrometer 31
2.3.2.2. SDS-PAGE drying 33
2.4. Isolation of nucleic acids 34
2.4.1. DNA mini and midi preparation 34
2.4.2. Purification of PCR products 34
2.4.3. Nucleic acid purification 35
2.4.4. TRIzol RNA purification 36
2.5. Concentration determinations 37
2.5.1. Nucleic acid concentration determination with a UV-spectrometer 37
2.5.2. Protein concentration determination 37
2.6. Polymerase chain reaction (PCR) 37
2.6.1. RACE-PCR 37
2.7. Sequencing of DNA 38
2.8. Cloning techniques 38
2.8.1. TOPO ligation 38
2.8.2. *Taq*-tailing and pGEM-T Easy cloning of PCR-products 39

2.8.3. Restriction digest 39
2.8.4. Ligation and Blunting of PCR-products 40
2.9. Trichome isolation 40
2.10. Blotting 40
2.10.1. Transfer of RNA on nylon membranes 40
2.10.1.1. Amplifying the cDNA hybridization probe 41
2.10.1.2. Hybridization probe preparation of the RNA-Blot 42
2.10.2. Transfer of proteins on nylon membranes 43
2.11. Protein expression and purification 44
2.11.1. Bacterial culture 44
2.11.2. Bacterial lysis 45
2.11.3. Protein purification with Talon resin 45
2.11.4. FPLC-protein purification 47
2.11.4.1. His-tag column regeneration 47
2.11.5. Size exclusion chromatography (SEC) 48
2.12. Enzyme assays 48
2.12.1. Monoterpene synthase assay 48
2.12.2. Diterpene synthase assay 49
2.13. Gas chromatography-mass spectrometry (GC-MS) 50

3. Results 51
3.1. *Cannabis sativa* L. 51
3.1.1. Purification of monoterpene synthases 51
3.1.2. Characterization of CsTPS1 and CsTPS2 54
3.1.2.1. pH-optimum of the CsTPSs 54
3.1.2.2. Temperature optimum of the CsTPSs 56
3.1.2.3. k_m and V_{max} of the CsTPSs 56
3.1.2.4. Size exclusion chromatography (SEC) of the CsTPSs 59
3.1.3. Phylogenetic (identity tree) analysis of the CsTPSs 65
3.1.4. Graphical overview of the CsTPS's 66

3.2. *Salvia sclarea* L. 67
3.2.1. EST-Sequencing of the putative diterpene synthase 67
3.2.2. RACE-PCR of the putative diterpene synthase 67
3.2.3. RNA-Blot of the putative diterpene synthase 70
3.2.4. Purification of the putative diterpene synthase 71
3.2.5. Protein sequencing of the putative diterpene synthase 72
3.2.6. Diterpene synthase assay for the characterization 74
3.2.7. GC-MS analysis of the putative diterpene synthase 75
3.2.8. Phylogenetic (identity tree) analysis and sequence comparison of SscTPS1 85
3.2.9. Size exclusion chromatography of SscTPS1 88
3.2.10. Graphical overview of SscTPS1 89

4. Discussion **90**
4.1. Disscussion of *C. sativa* L. monoterpene synthases 90
4.2. Disscussion of *S. sclarea* L. diterpene synthase 95

5. Summary **100**
5.1. Summary of *C. sativa* L. monoterpene synthases 100
5.2. Summary of *S. sclarea* L. diterpene synthase 100

6. Literature (cited by numbers) **101**

7. Appendix **115**

8. Acknowledgement **140**

9. *Curriculum Vitae*/Lebenslauf **141**

10. Statement/Erklärung **142**

List of abbreviations:

All abbreviations of chemicals, proteins, enzymes, *etc.* are explained in the Materials and Methods section.

aa	amino acid
Bam	*Bacillus amyloliquefaciens*
bp	base pair
bc	before Christ
C	carbon
cDNA	complementary DNA
°C	degree centrigrade or Celsius
CTS	chloroplast transit sequence
Ci	Curie
cm	centimeter
Da	dalton
dATP	deoxy-adenosine-triphosphate
dTTP	deoxy-thymidine-triphosphate
dUTP	deoxy-uridine-triphosphate
dCTP	deoxy-cytidine-triphosphate
dGTP	deoxy-guanidine-triphosphate
dNTP	deoxy-nucleotide-triphosphate
^{32}P-α-dATP	alpha-32-phosphorus- adenosine triphosphate
d.	day
DNA	deoxyribonucleic acid
dH_2O	distilled water
DSB	diterpene synthase buffer
EI	electron ionization
Eco	*Escherichia coli*
EST	expressed sequence tag
°F	degree Fahrenheit
ft	feet
g	gram or acceleration of gravity
GC	gass chromatography
ha	hectar (German unit of an area)
His-tag	histidin-tag
h.	hour
-OH	hydroxyl
in	inch
IPB	Leibniz Institute for Plant Biochemistry Halle/Saale
kb	kilo base
kDa	kilo dalton
l	liter
m	meter
mg	milligram
min	minute
ml	milliliter
mm	millimeter

mM	millimolar (millimoles/liter)
M	molar (moles/liter)
M_r	relative molecular mass or molecular weight
MCS	multiple cloning site
mRNA	messenger RNA
MS	mass spectrometry
µg	microgram
µl	microliter
µm	micrometer
µmol	micromole
ng	nanogram
nm	nanometer
nmol	nanomole
Nde	*Neisseria denitrificans*
OD	optic density
PAGE	poly acrylamide gel electrophoresis
ppm	parts per million
%	percentage
pI	isoelectric point
pmol	picomole
PCR	polymerase chain reaction
pH	pondus hydrogenii
Primer	oligonucleotide
Pfu	*Pyrococcus furiosus*
PPi	pyrophosphate
R_t	retention time
RT	room temperature
sec.	second
Taq	*Thermophilus aquaticus*
TPS	terpene synthase
rpm	revolutions per minute
Sac	*Streptomyces achromogenes*
SDS	sodium dodecyl sulfate
RACE	rapid amplification of cDNA ends
RNA	ribonucleic acid
TIC	total ion chromatogram
U	units
UTR	untranslated region
UV	ultra violet
V	volt
[v/v]	volume percentage
W	watt
[w/v]	mass per volume percentage
Xba	*Xhantomonas badrii*

1. Introduction

1.1. The plants

Many volatile oil (or essential) oil-producing plants are particular interesting because of their high economic value. The term ‘volatile oil’ is preferred because it refers to the fact that most components of the oils, which are stored in extra cellular spaces in the epidermis or mesophyll, have low boiling points and can be recovered from the plant tissues by steam distillation. Volatile oils are quite distinct from more common triglyceride oils and fats, which are known as ‘fixed’, because of their high boiling point [1]. Vast amounts of volatile oils are terpenoids that are made up of five-carbon isoprene or isopentenoid units. Other biological molecules belong to this family, are for example; porphyrins, chlorophylls, carotenoids, steroids, gibberellins and natural rubber [2]. The production and accumulation of volatile, low molecular weight terpenes [mono-(C_{10}), most sesqui-(C_{15}), and some diterpenes (C_{20})] are not restricted to one specialized taxonomic group but occur throughout the plant kingdom. Volatile oils can be obtained from various tissues of Gymnosperms (Pinaceae: conifers; Taxaceae: the yew family; Cupressaceae: cypresses and junipers; Cycadaceae: the tree fern family). The ability to accumulate terpenes is widely distributed in Angiosperms (subphylum: Magnoliophytina), too. By temperate and tropical dicotyledons (Chenopodiaceae: beets and goosefoots; Compositae: the daisy family; Geraniaceae: the geranium family; Guttiferae/Hypericaceae: including St. Johns worts; Lamiaceae: the mint family; Cannabaceae*:* hemp and hop; Lauraceae: including bays; Myricacea; Myristicaceae: nutmeg and mace; Myrtaceae: myrtles, eucalypts and clove; Scrophulariaceae: olives and lilacs; Piperaceae: the pepper family; Rosaceae: the rose family; Rutaceae: the citrus family; Santalaceae: sandalwood; Apiaceae: the carrot family; Verbenaceae: including verbenas; Violaceae: violets and pansies) and monocotyledons (Araceae: the aroid family; Cyperaceae: the sedge family; Poaceae: the grass family; Zingiberaceae: the ginger family). This indicates that the accumulation of terpenes was a feature of the earliest seed plants [1]. In this doctoral thesis two volatile oil producing plants were investigated, due to their ability to produce economically valuable terpenoids, *Cannabis sativa* L. and *Salvia sclarea* L..

1.1.1. *Cannabis sativa* L.

Cannabis sativa L.; family: Cannabaceae; order: Rosales, eurosids I, rosids; subphylum Magnoliopsida; phylum: Spermatophyta. *Cannabis* (hemp) and *Humulus* (hop) were previously subdivided into mulberry plants (Moraceae) and stinging-nettle plants (Urticaceae). Today both plants belong to the family of the hemp plants (Cannabaceae). *Cannabis* is comprised of three subspecies (*C. sativa* L., *C. indica, C. ruderalis*); this is controversial and some scientists suggest that all belong to one species, like hop (*Humulus lupulus, H. japonicus, H. yunnanensis*). Mankind has used the plant since the Stone Age (6500 bc) as evidenced by fossil plant remains such as pollen, fibers, seeds and half-burned trichomes. The origin of the plant is until now unknown but three regions are possible due to the long cultivation there; China (4200-1150 bc) in the region of the Huanghe-river (Yellow-river) and the Jangtse, Central Asia from the Caucasus- to the Altai-mountains or south Asia in the Himalaya and Hindukush regions. As a hemerophilous species, the plant is now found all over Asia, Africa, Europe and the Americans. *Cannabis* is a 1-5 m tall, cultivated annual plant, and is used for fiber, crop, oil, and medicinal purposes, but also as a stress relaxation and religious drug. [3, 4] Beside the fiber industry, the pharmaceutical industry is interested in the use of the plant. *Cannabis* produces a variety of medicinally active cannabinoids like THC (Δ^9-tetrahydrocannabinol) [5]. THC is used as a prescription treatment for nausea that can accompany chemotherapy as well as for spasms and tremors associated with multiple sclerosis [6, 7, 8].

1.1.2. *Salvia sclarea* L.

Salvia sclarea L., family: Lamiaceae; order: Lamiales, asterids, Magnoliopsida, Spermatophyta. Pliny the Elder, a roman scientist and historian, was the first to use the Latin name *Salvia*, derives from salvare, to heal or save, and salvus, meaning uninjured or whole, and refers to the several species of *Salvia* with medicinal properties. His interest in plants was utilitarian, and his encyclopaedic compilation, Natural History, included salvias in the vegetable kingdom. He dealt almost exclusively with the agricultural and medicinal attributes of plants. Information on the virtues of sage is also being found in the old herbals of medieval and renaissance Europe, usually illustrated with woodcuts or engravings. Not only are medicinal recipes given, but charms and spells are also de-

scribed. The common name sage originated in England and is probably a corruption of the French sauge. *Salvias* are members of the mint family, *Lamiaceae*, and comprise the largest genus in that family. The fragrant foliage of many types of *Salvia* has been used for more than 20 centuries to heal the minds and bodies of people of many different climates and cultures. *Salvias* are described by growth habit, perennial, biennial, or annual herbs, or as evergreen or deciduous shrubs. Some species are scandent and appear to climb, but they lack organs such as tendrils for support. The genus is distributed throughout the temperate and sub-tropical regions of the world, occurring from sea level to elevations of 3,400 m (11,000 ft) or more. The temperature ranges where they are found are equally great. Some habitats experience -18°C (0°F) or lower, while others may have readings of more than 38°C (100°F). *Salvias* are easy to determine. First look for opposite leaves and square stems that with age sometimes become round. Next, observe an individual flower closely. The corolla, the colorful tube, can have a variety of shapes, but it must have two lips of unequal length; the upper one is variable in shape and the lower is usually spreading. The calyx must also be two lipped. The upper one may have two or three teeth or it may be undivided. The lower one is often two-toothed. There are always two fertile stamens and sometimes two infertile ones called staminodes. Usually four seeds are produced; they are frequently mucilaginous when dampened. More than 900 species of *Salvia* exist worldwide, with well over half occurring in the Americas, but none occur in Australasia. Add to that number cultivated hybrids and natural hybrids from the wild as well as gardens, and the total figure increases by several hundred more. *Salvia sclarea* Linnaeus (clary sage) is a striking herbaceous plant that is classified as both biennial and perennial; *Salvia sclarea* L. has been recognized for its essential oils and was used extensively since well before the birth of Christ. Theophrastus, Discorides and Pliny, all wrote at length about its useful properties. In the wild, it is found throughout the northern Mediterranean region and in limited parts of North Africa and central Asia. It is commonly known as toute-bonne or sauge sclarée in France, sclarea in Italy, hierba de los ojos (herb of the eyes) in Spain, Muskateller or Römischer Salbei in Germany, and clary in Britain. Requiring little water or attention and being adaptable to temperatures below -18°C (0°F), this *Salvia* has become naturalized in central Europe. An early summer-blooming plant, *Salvia sclarea* L. can develop in a year's time from a

seedling to a plant 1-1.3 m (3-4 ft) tall when in flower. Stems are square and covered with hairs and oil globules (trichomes). Leaves vary in size from 30 cm (1 ft) long at the bottom of the plant to less than half that size at the top. They may be sessile or have a short petiole when growing at the top of the plant. The length of the petiole varies too. The leaf surface is rugose and covered with short hairs and few oil globules. Leaves are grassy green on the surface, and on the underside there are pronounced creamy veins. The entire edge is saw-toothed. Flowers number between two and six in each verticil and are held in large, colorful floral bracts that are pale mauve to lilac or white to pink with a pink marking at the edge. The lilac or pale blue corolla is about 2.5 cm (1 in) in length, and the lips are held wide open in the shape of a scythe blade. *Salvia sclarea* L. is not suitable as a cut flower because of the powerful odor of penetrating oils. These essential oils are widely used in making perfume and in imparting a muscatel flavor to wines, vermouths, and liqueurs. The oils are one source of the plant's healing qualities; the seeds are another. The plant seeds have traditionally been placed in the eye as a means of producing thick mucilage; this practice is said to clean the eye of impurities. The name "sclarea" carries the connotation of clear and bright, and it is said the "clary" is an English corruption of clear-eye [9].

1.1.3 Glandular Trichomes

Volatile oils are synthesized, stored and released by a variety of epidermal or mesophyll structures. These structures include: oil cells (ginger, turmeric, vanilla and pepper) secretory glands (bay, clove, *Citrus* and *Eucalyptus*); secretory ducts (vittae) or canals (anis, *Angelica*, *Pinus* and *Artemisia)*; and glandular trichomes (Labiatae, Cannabaceae, Compositae, Geraniceae and Solanaceae). Here the report of glandular trichomes (hairs, scales) is focused on the Lamiaceae family, but the Cannabaceae family feature the same elements. Labiates carry different epidermal hairs, glandular or non-glandular, sometimes on the same leaf. Most of them are capitate or peltate type. Each type of glandular trichome is made up of a basal cell, located in the epidermis, supporting a stalk of one to three cells, and either a single head cell (capitate trichome) or a more complex head, typically of four to ten cells (peltate trichome). These two main types of trichomes can be further subdivided according to the number of stalk cells, and whether the basal cell is

continuous with the epidermal surface or sinks into an epidermal pit. A glandular trichome develops from a single epidermal cell that is similar in size to epidermal cells, but has a larger nucleus, a smaller central vacuole and a more electron-dense cytoplasm. Successive divisions in the direction of the long axis of the future trichome give a structure made up of a basal cell, up to three stalk cells in series, culminating in a head cell. This completes cell division for capitate trichomes, but in the development of peltate trichomes, the head cell then undergoes a series of further divisions to give the appropriate number of head cells. Once the synthesis and accumulation of the terpenes begin, the Golgi apparatus and the endoplasmatic reticulum within trichome cells become more prominent, and the vacuoles become more osmiophilic the mature trichome's sub cellular integrity deteriorates. The oils produced by capitate trichomes appear to be lost to the surrounding environment, possibly through pores in the cuticle of the head cell. In peltate trichomes, the oil accumulates in spaces formed by separation of the head cell walls from their covering cuticle, from an expansion in its surface area. Oil storage capacity varies amongst trichomes and species [1].

1.2 Terpenoids

Terpenoids form the largest class of plant-derived specialized secondary metabolites, exceeding the alkaloids, cyanogenic glucosides and other secondary metabolites [10]. More than 36,000 different terpenoids are currently known [11]. You can find them in primary and secondary metabolism. These compounds have a wide range of functions in the plant kingdom. The volatile terpenoids are known to be involved in direct defence, feeding, oviposition deterrents and toxins [12], or indirect defence to attract predators of herbivores [13, 14]. They are available in relatively large amounts as resins, waxes and essential oils, are important renewable resources and provide a range of commercially useful products including solvents, flavorings and fragrances, coatings, adhesives and synthetic intermediates. Some of them are industrially useful polymers (rubber, chicle or bubblegum), pharmaceuticals (artemisinin, taxol) and agrochemicals (pyrethrins, azadirachtin) [15].

1.2.1. Terpenoid biosynthesis

Biosynthesis of terpenoids is based in general on the condensation of two activated isoprene C_5-units; isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP). To understand terpenoid biosynthesis it is necessary to understand the biosynthesis of these two C_5-units. An exemption is isoprene (hemiterpenoid), which is the simplest terpenoid, that is derived from DMAPP by diphosphate elimination by a divalent metal dependent (Mg^{2+}, Mn^{2+}) isoprene synthase, known from aspen (*Populus subsp.*) and velvet beans (*Mucuna pruriens*). Isoprene is produced and emitted by leaves of many plants and accounts for a significant proportion of atmospheric hydrocarbons [16].

1.2.2. Isopentenyl pyrophosphate biosynthesis

Since the enlightenment of the mevalonate-pathway to IPP and its isomerization to DMAPP (by Bloch, Conforth and Lynen in the 50s and 60s) which works on the sterol biosynthesis in animal and yeast cells, it has been widely believed that all organisms use only this pathway to build up huge amounts of isoprenoids [15, 17, 18]. But early feeding studies in this field using labeled mevalonate show that only small amounts of integration rates were found in the biosynthesis of monoterpenes and carotenoids [19, 20]. Feeding experiments with $NaH^{14}CO_3$[U-^{14}C]3-phosphoglycerate and [2-^{14}C]-pyruvate and isotope dilution studies showed labeling of carotenoids in purified spinach chloroplasts. And beside this, mevinolin that normally blocks the mevalonate-pathway (hydroxymethylglutaryl-CoA-reductase) had no effect on carotenoid and plastochinone biosynthesis [21]. This led to the speculation that the cytoplasm was not the only place for isopentenylpyrophosphate biosynthesis and a second location had to exist in plants. The alternative terpenoid biosynthesis pathway was unravelled, using ^{13}C-labeling experiments in bacteria and plants, by the independent groups of D. Arigoni and M. Rohmer. Application experiments showed labeling deviation from the acetate/mevalonate-pathway [22, 23]. Rohmer and Sahm hypothesized that the basic modules of the terpenoid biosynthesis were based on the condensation of pyruvate and a triose. Instead Arigoni's working group thought that the basic modules were based on a condensation of an activated acetaldehyde, derived from thiaminpyrophosphate mediated decarboxylation of pyruvate, with glyceraldehyde-3-phosphate. Later Arigoni could show that the condensation product 1-desoxy-

D-xyulose after application to bacteria and plants was found in the corresponding terpenoids [24, 25]. Since this time the alternative terpenoid biosynthesis pathway has been named after the first intermediate the DOXP- or desoxyxylulosephosphate-pathway. It was shown that desoxyxylulosephosphate was the precursor of thiamin (vitamin B_1) [26] and pyridoxol (vitamin B_6) biosynthesis [27]. Until now we know that animals and fungi use only the acetate-mevalonate pathway to synthesize terpenoids [28]. In bacteria the pathway is different; we have here for example the acetate-mevalonate pathway in *Chloroflexus aurantiacus* [29] and the Archeabacteria, but in *Pseudomonas aeruginosa*, *Salmonella typhimurium*, *Heliobacter pylori*, Mycobacteria and the Cyanobacteria the DOXP-pathway is followed [30]. Instead in plants both pathways may exist together in the same organism; the DOXP-pathway in plastids, where carotenoids, plastoquinone, phytol and mono- and diterpenes are synthesized, and the acetate-mevalonate pathway in the cytosol, where sesqui- and triterpenoids, and the ubiquinone side chains in the mitochondria are synthesized [31]. This is not always the case, for example in glandular cells of *Mentha x piperita* mono- and sesquiterpenoids are synthesized by leucoplasts [32]. Rhodophyta (*Cyanidium caldarium*) synthesize phytol over the DOXP-pathway and sterols over the mevalonate-acetate pathway, whereas chlorophyta (*Scenedesmus obliquus*, *Chlorella fusca*, *Chlamydomonas reinhardtii*) synthesize sterols, phytol and β-carotin over the DOXP-pathway and *Euglena gracilis* use exclusively the mevalonate-acetate pathway [33]. The appearance of the DOXP-pathway in plastids of plants and bacteria is another sign for the endosymbiotic theory, that the plastids are cyanobacteria-derived and that the eukaryotes are methanobacteria-derived [34]. There is evidence of a cross talk of both pathways in the biosynthesis of certain metabolites. As an example the *Matricaria recutila* (chamomile) sesquiterpenes are composed of two C5 isoprenoid units formed via the DOXP-pathway, with the third unit being derived from the MVA-pathway [35]. An interaction of both pathways is reported too, in lima bean for the biosynthesis of mono- and sesquiterpenoids [36]. In Arabidopsis the inhibition of the MVA-pathway by lovastatin (mevinolin) shows an upregulation of the DOXP-pathway metabolites, it is the same vice versa with inhibition of the DOXP-pathway by fosmidomycin. With the difference that sterol levels in lovastatin treated plants recovered to normal, but fosmidomycin treated plants show chlorophyll deficiency. This regulation is based on protein level but

not on significant alterations in the expression levels of the pathway genes [37]. Lovastatin treated Arabidopsis shows a change of the phenotype to dwarf growth, short hypocotyls and roots, and contracted cotyledons. DOXP-pathway deficient mutant show similar phenotypes in growth plus an albino form, but feeding of them with mevalonic acid showed only a weak recovery [38]

1.2.2.1. The mevalonate pathway (MVA-pathway)

This pathway is localized in the cytosol in which acetyl-CoA is converted into isoprene C_5-units (IPP and DMAPP) by a series of reactions that begins with the formation of hydroxymethylglutaryl-CoA (HMG-CoA). The synthesis of HMG-CoA in animals and in yeast requires two enzymes, a thiolase (acetyl-CoA acetyltransferase) and a HMG-CoA Synthase (HMGS). The thiolase condenses two molecules of acetyl-CoA to acetoacetyl-CoA, the HMG-CoA synthase condenses acetoacetyl-CoA with a third acetyl-CoA to HMG-CoA. There are situations in plants where these steps require only one enzyme, which uses Fe^{2+} and quinone cofactors, to yield HMG-CoA [39, 40]. The CoA thioester group of the HMG-CoA is reduced to an alcohol in an NADPH-dependent four-electron reduction catalyzed by HMG-CoA reductase (HMGR), yielding R-mevalonate. The 5-OH group is phosphorylated, with ATP, by mevalonate-5-phosphotransferase (or kinase, MVK) to phosphomevalonate. The phosphate group is then phosphorylated, with ATP, by phosphomevalonate kinase (PMK) to 5-pyrophosphomevalonate (MVPP). This molecule is then decarboxylated by pyrophosphomevalonate decarboxylase (PMD), in an ATP-dependent and divalent metal ion-dependent step, to isopentenyl pyrophosphate. The isopentenyl pyrophosphate isomerase (IDI; divalent metal ion-dependent) interconverts IPP to DMAPP by a concerted protonation/deprotonation reaction thus keeping equilibrium between both products [15, 18, 41, 42].

1.2.2.2. The methylerythritolphosphate-pathway (MEP-pathway)

The reaction begins with the formation of 1-deoxy-D-xylulose-5-phosphate (DXP). The deoxy-xylulose-phosphate-synthase (DXS) leads to an acyloin condensation reaction of pyruvate and glyceral-aldehyde-3-P (G3P), whereas pyruvate reacts with thiamin pyrophosphate (TPP) to yield a two-carbon fragment, hydroxyethyl-TPP, which condenses

with G3P. TPP is released to form the five carbon intermediate DXP, which is also the precursor of pyridoxal and thiamin. Further studies revealed that the DXS exists in different isoforms within the plant kingdom, one is responsible for the biosynthesis of housekeeping genes, and the other is responsible for the biosynthesis of secondary metabolites like monoterpenes, diterpenes and carotenoids. The DXP-reductase (DXR) rearranges and reduces DXP to form 2-C-methyl-D-erythritol-4-phosphate (MEP). It is common to name this pathway according to the intermediate MEP, because DOXP/DXP is the precursor of another pathway, but not the initial compound of this pathway. The cytidine diphosphate methylerithritol-synthase (CMS) phosphorylates with CTP MEP, which is then phosphorylated an additional time by the cytidyl-methyl-kinase (CMK) to 4-diphosphocytidyl-2C-methyl-D-erythritol-2-P (CDP-MEP). The following reaction is a cyclization, by which the dissociation of cytidine monophosphate (CMP) leads to a phosphodiester between C2 and C4. The cyclization to 2-C-methyl-D-erythritol-2,4-cyclodiphosphate (ME-cPP) is done by the methyl-erithritol-cyclo-diphosphate-synthase (MCS). The hydroxyl-methyl-butenyl-diphosphate-sythase (HDS) then catalyze the turnover to (E)-4-hydroxy-3-methylbut-2-enyldiphosphate (HMBPP), and the following enzyme the IPP/DMAPP-synthase (IDS) produces IPP and DMAPP in a 5:1 ratio. The hint that all of the enzymes above are part of this pathway is the fact that all contain a chloroplast -transit -peptide [43, 44]

1.3. Prenyltransferases

Prenyltransferases or polyprenyldiphosphate synthases are responsible for a prenyl transfer reaction. This biological process, in which a terpene fragment is attached to another moiety, is a bi-substrate reaction in which one of the partners is an allylic, terpene pyrophosphate and the other can be one of a variety of terpene or nonterpene compounds. This alkylation reaction provides the polyprenyl pyrophosphate precursors for the various terpenoid families [45, 46]. The prenyltransferases are subdivided into short-chain, long-chain, aromatic, and protein prenyltransferases.

1.3.1. Short-chain prenyltranferases

Under short-chain prenyltransferases or prenyltransferases type **I**, we include the geranyl pyrophosphate synthase (GPPS), farnesyl pyrophosphate synthase (FPPS) and the geranylgeranyl pyrophosphate synthase (GGPPS) which require either Mg^{2+} or Mn^{2+} to support enzyme activity as cofactors. They bear the characteristic $DD(X)_{2-4}D$-motif and are on the average 70-80 kDa big in size. They consist mainly of homodimer, heterodimer and heterotetramer which are tightly bound together, and are found in gymnosperms and angiosperms. Small subunits of GPPS lack the characteristic $DD(X)_{2-4}D$-motif. The enzymes bind the pyrophosphates in the absence of the divalent cations and only bind the cations when a pyrophosphate moiety is present; it is reasonable to assume that the metal-ion is assisting the catalytic step. GPPS catalyzes the condensation of IPP and DMAPP to give GPP (C_{10}), FPPS catalyzes the successive condensations of IPP with both DMAPP and GPP to give FPP (C_{15}), whereas GGPPS catalyzes the condensation step between FPP and IPP, GPP and two IPP or DMAPP and three IPP to give GGPP (C_{20}) [15, 45, 46, 47, 48, 49, 50]. The prenyltransferase reaction is initiated by a head-to-tail (1´-4) condensation of DMAPP and IPP to form GPP and so on. For this reaction two different mechanisms are in discussion. The first is an $\mathbf{S_N1}$ mechanism in which an allylic carbocation forms by the elimination of PPi (pyrophosphate). IPP then condenses with this carbocation, forming a new carbocation, and then a proton is eliminated to form a product. The second mechanism is an $\mathbf{S_N2}$ reaction in which the allylic PPi is displaced in a concerted manner. In this case, an enzyme nucleophile, X, assists in the reaction. This group is eliminated in the second step with the loss of a proton to form a product. Stereochemical studies were not sufficient to permit a distinction between ionization-condensation-elimination ($\mathbf{S_N1}$), and condensation-elimination ($\mathbf{S_N2}$) mechanisms [42, 46]. The chain elongation proceeds in two ways, cis- or (*Z*)-chain and trans- or (*E*)-chain elongation. Addition of an IPP occurs with stereo-specific removal of a proton at the 2-position, *pro*-R for (*E*)-type condensation or *pro*-S for (*Z*)-type condensation, respectively, to form a new C-C and a new double bound in the product, in which one C_5-isoprene unit elongation occurs [47].

1.3.2. Long-chain prenyltransferases

Long-chain prenyltransferases are subdivided into three different groups prenyltransferases **II-IV**. Prenyltranferases **II** synthesize medium chain prenyl diphosphates like (all-*E*)-hexaprenyl- (HexPPi, C_{30}) or (all-*E*)-heptaprenyl diphosphates (HepPPi, C_{35}). They comprise two non-identical protein components that do covalently bind with each other. Neither of them has any catalytic activity by itself but the activity returns when the two components are mixed together. Prenyltransferases **III** are octaprenyl diphosphates synthases, solanesyl (nonaprenyl diphosphate synthase) and decaprenyl diphosphate synthase, which catalyze the formation of (all-*E*)-long chain prenyl diphosphates. The prenyl chain length C_{40}-C_{50}, require additional protein activation factors that may be involved in removing the hydrophobic polyprenyl products from their active sites to facilitate and maintain the turnover of catalysis. The (*Z*)-prenyl chain elongation (Prenyltransferases **IV**) requires phospholipids or detergents for their catalytic function. All the enzymes in this group as well as in Prenyltransferases **III** have homodimer structures like Prenyltransferases **I** [47]. The majority of the polyprenyl diphosphates are the side chains of ubiquinones, phylloquinones, plastoquinones and menaquinones that are components of the electron-transfer system of the respiratory chain and vitamins (Vitamin K group). This group will be discussed in more detail under the aromatic-prenyltransferases [51, 52, 53]. A very important group of the polyprenyl diphosphate synthases is responsible for the biosynthesis of the precursors from steroids and carotenoids. The squalene synthases synthesizes by a head to head condensation from two FPPs, squalene, which is the initial component of the steroid biosynthesis [42]. Phytoene is the carotenoid precursor, in this reaction the squalene synthase complex (IPP-isomerase, GPPS, geranylgeranyl pyrophosphate-geranylgeranyl transferase and phytoene synthase) synthesize in a head- to-head condensation from two GGPPs, phytoene [54]. The steroid and carotenoid biosynthesis is not further discussed because of the complexity of this issue. The biggest known terpenoid structure is natural rubber, which consists of 400-100,000 isoprene units and is synthesized by a *cis*-prenyltransferase, the rubber transferase [52]

1.3.3. Aromatic-prenyltransferases

The products of the short- and long-chain prenyltransferases are mostly intermediates for other specific prenyltransferases, the aromatic-prenyltransferases. One important aromatic moiety is homogentisate (HGA), which is derived from chorismate over tyrosine to *p*-hydroxyphenyl-pyruvate in plants (shikimate pathway). Tocopherols and tocotrienols comprise the vitamin E family of antioxidants and are synthesized by plants and other photosynthetic organisms. The molecule consists of a chromanol head group; formed by a tocopherol/tocotrienol cyclase from a benzoquinol intermediate, linked to an isoprenoid-derived hydrophobic tail. The aliphatic tail of tocopherols is fully saturated (phytyl-PP), while the side chain of tocotrienols contains three trans double bonds (GGPP). Phytyl-PP is synthesized by geranylgeranyl reductase, and is part of chlorophyll biosynthesis, too. The prenyl transfer reaction (condensation) to homogentisate is done by the homogentisate geranylgeranyl- and phytyl-transferase, further modifications by synthases and methyltransferases lead then to various Vitamins E [55]. Plastoquinone as part of the photosystem II respiratory-chain is derived from HGA as well, which is prenylated with solanyl-PP (C45) by a homogentisate-solanyl-transferase, a methylation of the formed intermediate leads then to plastoquinone [56]. Whereas phylloquinone (Vitamin K_1), menaquinone (Vitamin K_2) and ubiquinone, as part of the photosystem I and bacterial/eukaryotic-respiratory-chain, are synthesize in a different manner. All of them are derived from chorismate as well. In animals ubiquinone biosynthesis starts with the uptake of the essential amino acid phenylalanine to form tyrosine, which is transaminated into 4-hydroxy-phenylpyruvate and then prenylated. The *E. coli* and yeast ubiquinone is derived from a prenylation of *p*-hydroxybenzoate (PHB) with PHB-hexaprenyl or octaprenyl diphosphate synthase with later modifications (decarboxylations, methylations for example) [51, 57]. Vitamin K_1 is an essential component (not endogenous) of the animal/human diet, because it serves as a cofactor for γ-carboxylation of glutamyl residues in different proteins, such as a blood coagulation factor. The phylloquinone component is composed of a naphthoquinone ring, derived from chorismate, which is converted into 1,4-dihydroxy-2-naphthonate (DHNA) and a prenyl side chain (phytyl-PP). The prenyl transfer reaction is mediated by the DHNA phytyl transferase, later methylation of the naphthoquinone ring leads to Vitamin K_1. The bacterial menaquinone (Vitamin K_2), car-

ries an unsaturated prenyl-side-chain with up to 15 isoprene units [53, 57]. In the secondary metabolism of plants we can find for example the geranylpyrophosphate olivetolate-geranyltransferase, which is a part of the biosynthesis of cannabinoids [58]. And an aromatic prenyltransferase from *H. lupulus* is responsible for the prenylation of the polyketide intermediates in humulone and cohumulone biosynthesis [59].

1.3.4. Protein-prenyltransferases

Protein prenyltransferases have only been found up to now in eukaryotic cell systems. Protein prenylation is a posttranslational process in which either a geranylgeranyl or a farnesyl group is attached via a thioether linkage (-C-S-C-) to a cysteine at or near the carboxyl terminus of the protein. In humans we have three different protein prenyltransferases: a FarnesylTransferase (FT) and GeranylGeranylTransferase 1 (GGT1). Both of them share the same motif (the CaaX box) around the cysteine and their substrates, which gives them the name CaaX prenyltransferases, whereas the third one the GeranylGeranylTransferase 2 (GGT2, also called Rab geranylgeranyltransferase) recognizes a different motif and is thus called the non-CaaX prenyltransferase. The proteins contain α and β subunits, these subunits are found in many other prenyltransferases, like squalene-hopene synthases and lanosterol and cycloartenol synthases. Lipid anchors are common posttranslational modifications and are mainly important for attachment of the proteins to membranes, but lipid modifications by protein prenyltransferases seem to have a more complex role: the farnesyl and geranylgeranyl moieties attached to the substrates are directly involved in protein-protein interactions as well as in protein-membrane interactions. The importance of protein prenyltransferases is illustrated by the involvement of their substrates in critical cellular pathways and diseases. The FT farnesylates include many members of the Ras superfamily of small GTPases (H-Ras, K-Ras, N-Ras, Ras2, Rap2, RhoB (which is also geranylgeranylated), RhoE, Rheb, TC10 and TC21), as well as the nuclear lamina proteins lamin A and B, the kinetochore proteins CENP-E and CNEP-F, fungal mating factors, cGMP phosphosphodiesterase α, γ subunit variants of G proteins, DnaJ heat-shock protein homologs, rhodospsin kinase, the peroxisomal membrane proteins Pex19 and PxF and paralemmin (neural protein). GGT1 geranylgeranylates small GTPases (Rac1, Rac2, RalA, Rap1A, Rap1B, RhoA, RhoB, RhoC, Cdc42, Rab8 (which is

also geranylgeranylated by GGT2), Rab11, and Rab13, and some γ-subunits variants of G proteins, cGMP phosphodiesterase β and the plant calmodulin CaM53). GGT2 geranylgeranylates proteins of the Rab family in general, the largest group of small GTPases in the Ras superfamily. There are at least 60 different Rabs in humans and they interact with the Rab escort protein REP, which is required for the prenylation of Rabs by GGT2, and are involved in the docking of transport vesicles to their specific target membranes. There are many diseases and cancer types that are affected by the protein prenylation factory, less effective drugs (FTinhibitors) have been found to treat these diseases [60].

1.4. Terpene synthases or cyclases

Various amounts of cyclic and acyclic plant terpenoids are formed by terpene synthases (TPS) through a divalent metal ion-assisted (Mg^{2+}, Mn^{2+} etc.) generation of enzyme-bound allylic carbocation intermediates from the prenyl pyrophosphate precursors (GPP, FPP and GGPP). This ionization step is common among most of the terpene synthases. The amino acids (aa) that are likely involved in the ionization are usually conserved in terpene synthases, for example the DDXXD-motif, and serve as an indicator of this enzyme class. The allylic carbocation formed in the ionization leads to cyclization reactions by internal double bonds, methyl migrations, hydride shifts, or Wagner-Meerwein rearrangements before deprotonation or nucleophile capture terminates the reaction [61, 62]. Terpenoids can be further modified by cytochrome P450 enzymes that hydroxlate, for example, monoterpenes in *Mentha* [63], and phophatases solvolyze diphosphate esters for example in *Croton stellatopilosus* Ohba [64]. Less is known about the transport of terpenoids out of the cell compartment, an ATP binding cassette-type transporter is suitable to transport diterpenes (sclareolide) out of the cell [65].

1.4.1. Monoterpene synthases

GPP the major precursor of the monoterpenoids undergoes an ionization and isomerization to enzyme-bound linalyl diphosphate (LPP) to permit subsequent cyclization. For gymnosperms, it has been proposed that the enantiomeric *p*-menthane type monoterpenes undergo a divergent, mirror image pathway in which enzymes that produce the (+)-series of products convert GPP to the (3*R*)-linalyl intermediate. Enzymes that produce the (-)-

series of products involve the (3*S*)-linalyl intermediate [66]. After LPP intermediate is formed, a second ionization and C1-C6 cyclization generates the corresponding α-terpinyl cation intermediate common to all cyclic monoterpenes. Alternatively, a deprotonation can occur that leads to the acyclic monoterpenes such as myrcene. The biosynthetic pathways diverge after these initial steps to produce the various monoterpene products. (-)-Limonene is formed by a double bond formation from (3*S*)-linalyl to the (4*S*)-α-terpinyl cation intermediate with a subsequent deprotonation, whereas (+)-α-pinene is formed by a double bond formation from (3*R*)-linalyl intermediate to the (4*R*)-α-terpinyl cation intermediate with a subsequent Markovnikov addition (C2-C8 ring closure) to a pinyl cation followed by a deprotonation [61, 66, 67, 68, 69]. The molecular weight of plant derived monoterpene synthases are in general between 50 kDa and 100 kDa (either monomers or homodimers), have a pI of ~5 and the pH optimum in the neutrality area. They bear chloroplast transit sequences (CTS), which are known to be 50-60 aa long and have a high serine and threonine and low acidic aa content. Monoterpene synthases contain a tandem ariginine motif (RR(X_8)W-motif), which was predicted to play a role in the isomerization of the monoterpene cyclization and to approximate the CTS, and a conserved tandem tryptophan-motif (WW-motif) [61, 70, 71]. The use of GPP as the only substrate for the monoterpene biosynthesis isn't a strict rule. A Chrysanthemyl diphosphate synthase catalyzes the condensation of two molecules DMAPP to produce chrysanthemyl diphosphate, a monoterpene with a non-head-to-tail or irregular c1'-2-3 linkage between isoprenoid units. The reaction is similar to the mechanism of the squalene or phytoene synthase complex discussed above (**1.3.2.**) Irregular monoterpenes are common in *Chrysanthemum cinerariaefolium* and related members of the Asteraceae family; chrysanthemyl diphosphate is an intermediate in the biosynthesis of the pyrethrin ester insecticides. [72]

1.4.2. Sesquiterpene synthases

Sesquiterpene synthases produce acyclic and cyclic terpenoids out of FPP. An acyclic sesquiterpene synthase is the (E)-β-farnesene synthase from *Mentha piperita*, which is an aphid alarm pheromone. It is produced in the presence of Mg^{2+} from FPP via cation formation and proton loss, with (Z)-β-farnesene and the bicyclic δ-cadinene via nerolidyl-

PP, as co-products [73]. A cyclic sesquiterpene synthase like 5-*epi*-aristocholene synthase (TEAS) from *Nicotiana tabacum*, the recombinant crystal structure was unravelled, cyclizes FPP to the germacryl cation and by a subsequent rearrangement via the triene germacrene A and the eudesmyl cation to (+)-aristocholene. Aristocholene synthases are known from *Aspergillus tereus* and *Capsicum annuum*, whereas the latter one is part of a sequence to phytoalexins e.g. capsidiol, when plants are challenged by fungi. A related enzyme is the vetispiradiene synthase from *Hyoscamus muticus*, using eudesmyl cation as an intermediate [74, 75, 76, 77].

1.4.3. Diterpene synthases

One of the largest subgroups of diterpenoids is the labdane-related group, which contains over 5000 compounds, defined here as minimally containing bicyclic hydrocarbon structure. This core structure can be further cyclized and/or rearranged, as in the related/derived structural classes like gibberellins, abietanes, (iso)-pimaranes, and kauranes. Their biosyntheses requires the initial (bi)-cyclization of (*E,E,E*)-GGPP into the characteristic bicyclic backbone in forming labdadienyl/copalyl pyrophosphate (CPP). Two different kinds of CPP-synthases exists, a (-)- or *ent*-CPP-synthase, also known as kaurene synthase A, which uses the chair-chair conformation of GGPP to produce (-)-CPP, and a (+)- or syn-CPP-synthase which uses the chair-boat conformation to produce (+)-CPP. Both synthases could be identified in *Oryza sativa*; (-)-CPP and (+)-CPP are precursors of phytoalexins and allelophatic natural products, with *ent*-kaurene as the key precursor of gibberellins [78]. The biosynthesis of abietadiene involves two distinct intermediates and two active sides. The first step of the sequence is a proton-initiated cyclization of (E,E,E)-GGPP which leads to the closure of the A and B ring to yield (+)-CPP at the first active site, in a reaction analogous to that mediated by kaurene synthase, of the gibberellin biosynthesis pathway but with opposite enantiospecifity. After internal channeling of (+)-CPP to the second active site an ionization leads to the formation of the C ring of a pimarane-type intermediate, in a reaction that corresponds to that mediated by kaurene synthase B. The final step, a 1,2-methyl shift in the pimarane intermediate generates the C13 isopropyl group characteristic for the abietan skeleton, and a deprotonation completes the reaction to abietadiene [67, 79].

1.5. Compartmentation of the terpenoid-pathways

The pathways of terpenoids were summarized and described from McGarvey and Croteau in 1995 [15]. Since that time, changes in the pathways have occurred due to the obtainment of new data e.g. the unravelling of IPP-pathways. Based on this review, changes and new data are presented in this chapter. Two individual pathway exists leading to IPP and DMAPP, the mevalonate- and the MEP-pathway, which are found in the cytosol or in the plastids [44]. The synthesis of the prenyltransferase products seem to be clear. Synthesis of FPP and its derived sesquiterpenes takes place in the cytosol and the cytosol/ER boundary, as does the synthesis of triterpenes, including the phytosterols. Plastids are the exclusive site of synthesis for GPP, GGPP, monoterpenes, diterpenes, tetraterpenoids (carotenoids), chlorophylls, prenylquinones and tocopherols. The mitochondria appear to have an independent system for the biosynthesis of ubiquinones from IPP. Ubiquinone and plastoquinone formation appears to take place in microsomes as well, possibly in the Golgi apparatus. They summarize that at least three distinct semiautonomous subcellular compartments that segregate terpenoid biosynthesis exist: the cytosol/ER for sesquiterpenoid and triterpenoid biosynthesis; plastids for monoterpenoid, diterpenoid, and tetraterpenoid biosynthesis (as well as for the prenyl moieties of chlorophyll, plastoquinones, and tocopherols); and mitochondria (and/or Golgi apparatus) for ubiquinone biosynthesis. They mention that a transfer of metabolites from plastids to sites of secondary transformations, such ER-bound cytochrome P450 oxygenases and various cytosolic redox enzymes exists. New studies revealed that FPP synthases from *A. thaliana* are localized to the mitochondria [80] and in *O. sativa,* wheat and tobacco to the plastid [81], *A. thaliana* GGPP synthases are located in addition to plastids, to mitochondria and to the endoplasmic reticulum [82]. A maize FPP synthase is capable of producing FPP and GGPP [83]. In *Fragaria vesca* a cytosolic pinene synthase exists. And terpene synthases, with a high sequence identity to monoterpene synthases, of *F. vesca* and *F. ananassa* synthesize linalool, a monoterpene, beside the sesquiterpene nerolidol and are located to the mitochondria and/or plastids, depends on the isoform. There are hints that monoterpenes may be produced in the cytosol in non photosynthetic tissues such as fruit, glands, or other secretory tissues. Because an *A. thaliana* GPP synthase encoded two isoforms, one targeted to the plastid and the other to the cytosol. Similarly, the formation of shikonin, a

monoterpene derivative, was reported to occur in the cytosol of *Lithospermum erythrorhizon* cells [84]

1.6. Preliminary results and aim of the thesis

During my Diploma thesis two putative *Cannabis sativa* L. cv. Skunk monoterpene synthases, *CsTPS1* and *CsTPS2* were isolated from cDNA prepared from trichome-specific RNA. The pET101/D-TOPO vector (Invitrogen) was used for expression in *E. coli*. Oligonucleotides with a partial vector sequence and a gene specific sequence were used to truncate the possible chloroplast transit sequence (CTS) preceding the RR(X_8)W-motif, and some clones were constructed to contain a C-terminal hexa histidine extension. Since the transit sequence prediction was only an approximation, two different N-terminal starting points were used, immediately preceding the RR-motif; as already mentioned the RR motif was not required for the catalysis, but lack of it leads to a decrease in monoterpene synthase activity (**1.4.1.**) [70, 85]. The analysis of the full-length terpene synthases showed for *CsTPS1* 2046 bp with an ORF of 1869 nt encoding 622 aa, and for *CsTPS2* 2016 bp with an ORF of 1848 nt encoding 615 aa. To determine the tissue-specific expression of each *CsTPS*, RNA was isolated from different parts of the female and male *Cannabis* plant, including roots, stems, old leaves, young leaves, female blossoms, female blossoms without trichomes, trichomes, mature male blossoms, and young male blossoms. The RNA gel blot analysis indicated a trichome-specific expression of each CsTPS. GC-MS analysis of *C. sativa* var. “skunk” trichomes was performed. The major monoterpenes were limonene, α-pinene, β-pinene, trans-β-ocimene, myrcene and α-terpinolene, with the latter two being dominant. The sesquiterpenes trans-β-caryophyllene, γ-elemene, α-guaiene, trans-β-farnesene, α-humulene, E-E α-farnesene, Δ-guaiene, selina-3,7(11)-diene, germacrene, guaiol, γ-eudesmol, α-eudesmol and 5-azulenemethanol were found The terpenophenolic compounds detected were Δ-9-tetrahydrocannabinol and cannabigerol. All recombinant CsTPSs were assayed in a crude lysate and the reaction mixtures were analyzed by HSPME-(headspace solid-phase micro extraction)-GC-MS and were shown to be enzymatically active. CsTPS1 produced (-)-limonene, in addition to minor amounts of α-pinene, β-pinene, β-myrcene and α-terpinolene. CsTPS2 produced, (+)-α-pinene, as well as minor amounts of β-pinene, β-myrcene and limonene.

Although limonene and pinene synthases are common to many plant species, until now few have been characterized with respect to kinetic parameters [71]. The aim of the thesis was at this time point the determination the kinetic parameters of both terpene synthases, with respect to temperature-optima, pH-optima and Michaelis constant (K_m and V_{max}). A size exclusion chromatography was performed and the evolutionary relationship of each with respect to other plant families was also demonstrated.

During a research internship (2002-2003) Dr. Kum-Boo Choi (Korea) established a trichome specific EST-project of *Salvia sclarea* L. cv. Trakystra at the Leibniz Institute of Plant Biochemistry, to unravel the biosynthesis of the flavor and fragrance compound sclareol. Several terpenoid metabolism-related EST's were found by the end of the internship. The project was continued as part of my Ph.D. thesis. This work here represents the cloning and characterization of the novel mainly trichome expressed, sclareol diterpene synthase. Size exclusion chromatography and the evolutionary relationship with respect to other plant families were demonstrated too.

2. Materials and Methods

2.1. Materials

2.1.1 Organisms

2.1.1.1. Plants

Cannabis sativa L. cv. Skunk plants were grown in a greenhouse or in a growth chamber. Plants were grown in the greenhouse under a high pressure sodium lightning system (Philips, SON-T AGRO 400 W) in a 18 h. light (20°C -24°C)/6 h. dark (18°C-20°C) cycle with a relative humidity of ≈60%, or for the flower induction in a growth chamber in a 12 h. light (25°C)/12 h. dark (23°C) cycle at ≈60% humidity.

Salvia sclarea L. cv. Trakystra from Pharmasaat with high yields and high essential oil content was used for the research. The recommendations for cultivation of the plant were for the drilling; 6-10 kg/ha in April, and for the planting; 6 plants/m^2 also in April, the distance for the rows was 50 cm and the harvesting time was during the flowering period. The plants were grown in an open field at the Leibniz Institute for Plant Biochemistry in Halle/Saale (Germany; IPB) or in a greenhouse at the Donald Danforth Plant Science Center in St. Louis, MO (United States of America). Growing conditions of *Salvia sclarea* in St. Louis: Plants were grown at 21°C with 16 hours of supplemental light. In order to get the plants to flower they were placed outside when it was around 0°C until the plant froze back. The plants were watered as needed and were fertilized using Excel 15-5-15 Cal-Mag with a target of around 200 ppm N 2-3 times per week.

2.1.1.2. Bacterial hosts

DH5α (Clontech): F^-, *deo*R, *end*A1, *gyr*A96, *hsd*R17, ($r_k^-m_k^+$), *rec*A1, *rel*A1, *sup*E44, Φ80*lac*ZΔM15, *thi*-1, Δ(*lac*ZYA-*arg*FV)U169

BM25.8 (Novagen): [F' *traD36 lacI*^{q}*lacZΔM15 proA*$^+$*B*$^+$] *supE thi* Δ(*lac-proAB*) P1 Cm^R*hsdR* (rK_{12}^- mK_{12}^+) (λ*imm*^{434}Kan^R)

TOP10 (Invitrogen): F^-, *mc*rA, Δ(*mrr-hsd*RMS-*mcr*BC), Δ*lac*X74, Φ80*lac*ZΔM15, *deo*R, *rec*A1, *end*A1, *gal*K, *nup*G, *ara*D139, Δ(*ara-leu*)7697, *rps*L, (Str^R), *gal*U

TOP10F´ (Invitrogen): *mcr*A, Δ(*mcr*BC-*hsd*RMS-*mrr*), *end*A1, *rec*A1, *rel*A1, *gyr*A96, Φ80*lac*ZΔM15, *deo*R, *nup*G, *ara*D139,F{*lac*I^{q}, *Tn*10(Tet^{r})}, *gal*U, Δ*lac*X74, *gal*K, Δ(*ara-leu*)7697

BL21(DE3) (Stratagene): *F – dcm ompT hsdS (r B– m B–)gal .(DE3)*

BL21(DE3)pLysS (Novagen): *F – dcm ompT hsdS (r B– mB–)gal .((DE3)[pLysS Cam r]*

BL21(DE3)RIL (Stratagene): *F – ompT hsdS (r B– mB–)dcm+Tet r gal . (DE3)endA Hte [argU ileY leuW Cam r]*

XL1-Blue (Stratagene): *recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F´ *proAB lacI*qZΔ*M15* Tn*10* (Tetr)]

2.1.2 Plasmids and oligonucleotides

2.1.2.1. Plasmids

pTripleEX2 (Clontech): This 3.6 kb vector has the *E. coli* lac promoter and operator to provide regulated expression of inserts in *E. coli* hosts expressing the lac repressor (lacIq). The 5' untranslated region (UTR) from the *E. coli* ompA gene stabilizes the mRNA, thereby increasing expression. pTriplEx2 incorporates a triple-reading-frame translation cassette consisting of translation initiation signals from the *E. coli* ompA and lacZ genes, in two different reading frames, followed by a transcription/translation slip site. Downstream of this cassette is the pTriplEx2 MCS, which is embedded within the lacZ α-peptide allowing, clones with inserts to be identified by blue/white screening in an appropriate host strain. The T7 RNA polymerase promoter downstream of the MCS allows production of single-stranded RNA in vitro for use as a probe. In the presence of helper phage, the f1 origin in pTriplEx packages the non coding strand of the lacZ gene into phage particles, and this single-stranded DNA can be used for sequencing or mutagenesis procedures. The ampicillin resistance gene and pUC origin of replication allow selection and propagation, respectively, of pTriplEx2 in *E. coli*.

pGEM/T-Easy (Promega): This 3.0 kb linear cloning-vector has a 3'Thymidin overhang, which allows to clone *Taq* DNA-Polymerase amplified PCR-products. The general specifications are: a T7 and SP6 transcriptions- and promotor site, a MCS, lacO, a lacZ start codon, a β-lactamase sequence plus pUC/M13 forward and reverse sequencing starting points.

pET101/D-TOPO (Invitrogen): This 5.7 kb vector contains a TOPO cloning site with a 5´ "sticky-end" and GTGG antisense overhang and a 3' "blunt-end". The phosphorus rest is covalently linked with a topoisomerase I (*Vaccinia* virus), which activates the vector. The amplified PCR-products attackes with the 5´-OH group to the 3´-phosphorus of the vector, the ligation works with the help of the binding enthalpy. This composes a directed ligation. The vector also contains a T7-promotor and a transcription-termination site, a T7r-Primer site, a ribosome binding site (RBS), a lac-operator (lacO), a V5-epitope, a C-terminal 6x His-tag, a pBR322 origin of replication, and an ampicillin resistance.

pET100/D-TOPO (Invitrogen): This 5.7 kb vector bearing the same features like the pET101/D-TOPO vector, except that we have here an N-terminal 6 x His-tag.

2.1.2.2. Oligonucleotides

All primers below are written in 5'- to -3' direction. The temperatures mentioned are the melting temperatures of the primers.

C. sativa L. TPS-cloning oligonucleotides:

7G10 5'60 TOPO	CAC CAT GTG TAC TGT GGT CGA TAA CCC	66,5°C
7G10 5'RR TOPO	CAC CAT GCG AAG ATC AGC CAA CTA TGG	66,5°C
7G10 3'stop	TTA CAT GGG GAT AGG AGT AAT A	54,7°C
7G10 3'nostop	CAT GGG GAT AGG AGT AAT AAT C	56,5°C
14A9 5'55 TOPO	CAC CAT GTG TAG TTT GGC CAA AAG CCC	66,5°C
14A9 5'RR TOPO	CAC CAT GAG AAG ATC AGC CAA CTA TGA T	63,7°C
14A9 3'stop	TTA TAA AGG AAT AGG ATT AAT AAT TA	50,6°C
14A9 3'nostop	TAA AGG AAT AGG ATT AAT AAT TAA AT	50,6°C

C. sativa L. TPS-sequencing oligonucleotides:

7G10 5'sp3	TCC CAT TTG GTC TTT TGA TT	51.2°C
7G10 5'sp5b	CAA GAA ACT GAT GGT AGA GTT GTG	59.3°C
7G10 5'sp4	AGG CTT CTA CGC CAA TAT GG	57.3°C
7G10 3'sp1	ATT GTG GCT CAA ATC TTA CTC C	56.5°C
7G10 3'sp2	CAG ACA ACC ATC CAT TCT CA	55.3°C
14A9 sp1	TTT ATA TGC CAC TGC TCT CG	55.3°C
14A9 sp2	GGT TGG TTT GTA TCC GCT AT	55.3°C
14A9 5'sp5	TGC TCT AAT TAC AAT CAT AGA TGA CA	56.9°C
14A9 3'sp6	TGA ACC ACT TAG CTT CTG CTC	57.9°C

S. sclarea L. TPS- RACE oligonucleotides:

5'RACECop1antisense	GGT GAC ACG CTT ATT CGC CCG TCG CCC G	73.9°C
5'RACECop2antisense	TCG TCG CCC CAT GAG CCA TCA GCC AGT TGG TGG T	>75.0°C
3'RACECop1sense	TGC GCC GGC CTC AAC GAA GAC GTG TTG	71.0°C

S. sclarea L. TPS-cloning:

5'CopfullTOPO-2	CAC CAT GAC TTC TGT AAA TTT GAG CAG AGC	65.4°C
3'CopfullTOPO-stop	TCA TAC AAC CGG TCG AAA GAG TAC	61.0°C
3'CopfullTOPO-ns	TAC AAC CGG TCG AAA GAG TAC	59.0°C
3'CopfullTOPO-ns2	TAC AAC CGG TCG AAA GAG TAC TTT G	61.3°C
5'CopfullTOPO+RR	CAC CAT GCG CAG GCT GCA GCT ACA G	69.5°C
5'CopfullTOPO-RR	CAC CAT GCA GCT ACA GCC GGA ATT TC	66.4°C
5'CopfullTOPO-W	CAC CAT GCT GAA AAG CAG CAG CAA ACA C	66.6°C
5'CopfullTOPO-SP	CAC CAT GCA CGC GCC CTT GAC CTT G	69.5°C
5'CopTOPOipsort	CAC CAT GCC AGC AGC GAT TAC CCG G	65.9°C

S. sclarea L. TPS-sequencing oligonucleotides:

5'RACE_K1_SP6	AGG TTG CCA GCA GAG TC	55.2°C
5'RACE_K2_SP6	ATC TCA TCA AAT AAG TGG TCG G	56.5°C
5'RACE_K3_SP6	TAT CGA TTG CCC AAA GTC TTG	55.9°C
5'RACE_K4_SP6	ATG AGC CAT CAG CCA GTT G	56.7°C
CopfullGap1	GTG TCA CCC TAT GAC ACG	56.0°C
CopfullGap2	AGT GGA ACA ACA CTA TCT TTG G	56.5°C

Additional oligonucleotides:

T7	GAA TTG TAA TAC GAC TCA CTA TAG	55.9°C
T7 reverse	TAG TTA TTG CTC AGC GGT GG	57.3°C
pHIS8 T7	CGA AAT TAA TAC GAC TCA CTA TAG	55.9°C
TriplEx_Insert	CTC GGG AAG CGC GCC ATT GTG TTG GT	69.5°C
pTriplex 5'long	CAA GCT CCG AGA TCT GGA CGA GC	66.0°C
pTriplex 3'long	ATA CGA CTC ACT ATA GGG CGA ATT GGC C	66.6°C

2.1.2.3. Nucleotides

^{32}P-α-dATP (ICN, PerkinElmer); dATP, dTTP, dCTP, dGTP (Promega, Sigma); 1 kb Base ladder [New England Biolabs (NEB)].

2.1.3. Biochemicals (enzymes, proteins, antibiotics and antibody)

Enzymes: *Pfu*-Polymerase (Promega, Stratagene); *Taq*-Polymerase (Jörg Ziegler IPB, Promega, NEB, Sigma); Herculase Hotstart DNA Polymerase (Stratagene); Lysozyme (Sigma, PIERCE); T4-DNA Ligase (Roche, NEB); Restriction Endonucleases: *Eco*RI, *Sac*I, *Xba*I, *Bam*HI, *Nde*I (NEB); Wheat germ acid phosphatase (Sigma), Trypsin sequencing grade modified (Promega); Benzoase (DNase).

Proteins: Proteinmarker Fermentas (SM0671; PageRuler Prestained Protein Ladder, SM0431; Unstained Protein Molecular Weight Marker); Prestained SDS-PAGE Standards (BIO RAD 161-0305; Broad Range), bovine serum albumin (BSA; Fisher Scientific, Serva), ovalbumin, aldolase, catalase, ferritin (all Serva).

Antibiotics: Ampicillin disodium salt 50 mg/ml (Sigma, Fisher Scientific, Roth); chloramphenicol 50 mg/ml or 34 mg/ml (Sigma, Fisher Scientific, Serva); kanamycinsulfate 50 mg/ml (Roth).

Antibody: monoclonal Anti-polyhistidine-Alkaline Phosphatase antibody produced in mouse (Sigma)

2.1.4. Chemicals

Acetic acid, glacial	Roth, ACROS Organics
Acetone	ACROS Organics
Acetonitrile (ACN, CH_3CN)	ACROS Organics, Fisher Scientific
Acrylamide/Bisacrylamide 37.5 : 1 (40%)	ACROS Organics
Agar	Serva, Fisher Scientific

APS (ammoniumpersulfate)	Serva, ACROS Organics
Ammoniumhydrogen carbonate	Sigma
p-Anisaldehyde	Fluka
ATA (aurintricarboxylicacid or aluminon reagent)	Sigma
β-Mercaptoethanol (β-EtSH)	Roth, ACROS Organics
Bromophenolblue	Sigma, Fisher Scientific
CaCl (Calcium chloride)	ACROS Organics
BCIP (5-Bromo-4-chloro-3-Indolyl phosphate)	ACROS Organics
Chloroform	Roth, ACROS Organics
Coomassie Brilliant Blue G-250	Serva, ACROS Organics
Dichlormethane	Roth, ACROS Organics
Dimethylformamide	ACROS Organics
DTT (Dithiothreitole)	Roth, ACROS Organics
EDTA (Ethylen-diamin-tetra-acetate)	Roth, ACROS Organics
Ethanol	Merck, AAPER Alcohol and Chemical Co.
Ethidium bromide	Sigma, PerBio, Pierce
FPP (Farnesyl pyrophosphate)	Marco Dessoy (IPB)
Formaldehyde	Merck, ACROS Organics
Formic acid	Sigma
Formamide	Fluka, Fisher Scientific
Ficoll 400	Sigma
GPP (Geranyl pyrophosphate)	BioTrend, Roman Weber (IPB)
1-^{3}H-GPP	ARC, American Radiolabeld Chemicals
Geranylgeraniol	MPI
GGPP (Geranylgeranyl pyrophosphate)	Echelon, Sigma, Günther Scheid (IPB)
1-^{3}H-GGPP	PerkinElmer
Glucose	Merck, ACROS Organics
Glycerol	Roth, ACROS Organics
Glycine	Roth, Fisher Scientific

HCl (Hydrochloric acid)	Roth, ACROS Organics
HEPES (2-[2-Hydroxyethyl) -1-piperazinyl]-ethansulfonacid)	Merck, ACROS Organics
Hexane	Roth, ACROS Organics
Imidazole	Roth, ACROS Organics
IPTG (Isopropyl-1-thio-β-galactopyranoside)	Peqlab, Gold Bio, Roth
Isopropanole (2-Propanol)	Merck, ACROS Organics
LiCl (Lithiumchloride)	Merck, ACROS Organics
Manool	Sigma-Aldrich
MES (2-(N-Morpholino)ethansulfoneacid)	Serva, ACROS Organics
Methanol	Merck, ACROS Organics
Methylcellulose	Sigma
$MnCl_2$ x 6 H_2O (Manganese chloride)	ACROS Organics
$MgCl_2$ x 4 H_2O (Magnesium chloride)	Roth, Sigma
$MgCl_2$ x 6 H_2O (Magnesium chloride)	Roth, ACROS Organics
$MgSO_4$ (Magnesium sulfate)	Sigma, Fisher Scientific
MOPS (3-(N-Morpholino)-2 -propansulfonacid	Roth, ACROS Organics
MOPSO(3-(N-Morpholino)-2 -hydroxypropansulfonacid)	Roth, ACROS Organics
NBT (Nitrobluetetrazolium)	RPI, Research Product International Corp.
NPP (neryl pyrophosphate)	Roman Weber (IPB)
$NiSO_4$ (Nickel sulfate)	Merck
H_2SO_4 (Sulfuric acid)	ACROS Organics
Orange G	Sigma, ACROS Organics
Peptone	Difco, Fisher Scientific
Phenol	ACROS Organics
Phenol-Chloroform	Roth, ACROS Organics
PVPP (Polyvinylpolypyrrolidon)	Sigma, Fisher Scientific
PVP (Polyvinylpyrrolidon)	Sigma, ACROS Organics

Potassium acetate	ACROS Organics
KCl x 2 H_2O (Potassium chloride)	ACROS Organics
Sclareol	Collection Prof. M.H. Zenk, Fluka
Sodium acetate	Merck, Fisher Scientific
NaCl (Sodium chloride)	Roth, ACROS Organics
Sodium citrate(citric acid trisodium hydrate)	Roth, Fisher Scientific
NaOH (Sodium hydroxide)	Roth, ACROS Organics
Saccharose	Roth, Schnucks
SDS (Sodium lauryl sulfate)	Roth, Fisher Scientific
Sephadex G-50 Superfine	Amersham Pharmacia
Sorbitole	Roth, Sigma
TEMED (N',N',N',N'-Tetramethylethylen-diamine)	Roth, ACROS Organics
Thiourea	Sigma
Tris (Tris-(hydroxymethyl)-aminomethane)	Roth, ACROS Organics
TRIzol reagent	Invitrogen
Tryptone	Fisher Scientific, BD Biosciences
Tween 20	Roth, Fisher Scientific
XAD-4	Sigma
X-Gal (5-Brom-4-chlor-indolyl-β–D-galacto pyranosid)	Roth, Gold Bio
Yeast extract	Difco, Fisher Scientific

2.1.5. Miscellaneous (columns, reagents, etc.)

Phosphor imager screen	Molecular Dynamics, FUJI Film
Filter paper	"common"
GB 004 Gel Blotting paper	Schleicher & Schuell (Whatman)
Sterile Filter 25 µm and 45 µm GD/X	Schleicher & Schuell (Whatman)
Synthetic filter floss	Common from a pet shop
Biodyne® B 0.45 µm (nylon membrane)	Pall Corporation

BigDye Terminators Version 1.1	Applied Biosystems
Talon Resin	Clontech
Talon 2 ml Disposable Gravity Column	Clontech
PD10 Column (Sephadex™ G-25M)	Amersham Biosciences (GE)
HisTrap™ HP 1 ml or 5 ml	Amersham Biosciences (GE)
HisTrap™ FF crude 1 ml or 5 ml	Amersham Biosciences (GE)
Superdex 200 16/60 SEC-column	GE Healthcare
Bio-Rad Protein Assay solution	BIO-RAD
Microcentrifuge tubes siliconized	Sigma-Aldrich
NEBlot® Kit	NEB

2.1.6. Equipment

PCR-Machines; GenAmp PCR System 9700 PE and 2720ThermalCycler (Applied Biosystems), PTC 200 Peltier Thermal Cycler (MJ Research), Mastercycler gradient (Eppendorf). Power supplies; Standard Power Pack P25 (Biometra), E455 (Consort), PHEROstab 500 (Biotec-Fischer), FB300 (Fisher Scientific). Centrifuges, shakers and stirrer/mixer; Centrifuge 5810R and 5415D or C (Eppendorf), 4K10 and 3K18 (Sigma), Sorvall RC 26 Plus (DuPont), Avanti J-25 and J-20 XP (Beckman/Coulter). Shaker I26 and C25KC (New Bruinswick), Incubator Shaker Orbit (LabLine), Platform Shaker Isotemp series (FisherScientific). Stirrer and Mixer Isotemp series (Fisher Scientific). Autoclaves and heat systems; Varioklav Dampfsterilisator Typ 250T (H+T), Primus (Sterilizer CO. INC.), Heatblock Isotemp series (Fisher Scientific). Sequencer; ABI 310 and ABI 3100-*Avant* Genetic Analyzer (Applied Biosystems). Sterile technology; Hera Safe Laminar Box (Heraeus), Incubator Isotemp series (Fisher Scientific) MilliQ Biocel A10 (Millipore) Analyzer; Storm Phosphor imager (Molecular Dynamics), Scintillation counter (Beckman LS 6000 TA), Fluorimeter (Molecular Devices), UV Analyzer Ultrospec 3000 (Pharmacia Biotech). pH-Meter; inoLab pH Level 1 (WTW) and SevenEasy (Mettler-Toledo). Bead beater (Biospec Products). Cool techniques; -80˚C (Sanyo, NewBrunswickScientific), -20˚C (Fridgidare Commercial), +4˚C (Kenmore Elite). Balances; PB3002-S/FACT Delta Range and XS105 Dual Range (MettlerToledo), TE3102S (Sartorius), Omnilab. Sonic Dismembrator; Model 100 or 550 (Fisher Scientific).

2.2. Microbiological methods

2.2.1. Media and agar plates

For *E. coli* cultures mostly LB-media (Luria-Bertani-media) was used, with 1% (w/v) Peptone/Tryptone, 0,5% (w/v) yeast extract, and 1% (w/v) NaCl dissolved in 1 liter H_2O at a pH of 7.0 adjusted with NaOH (Germany) or HCl (USA). For agars 1.5% (w/v) Agar was added to the LB-media. For the preparation of competent cells SOB-Medium was used. This contains 2% (w/v) peptone/tryptone, 0.5% (w/v) yeast extract, 10 mM NaCl, 2,5 mM KCl, pH 7.0, after autoclaving 10 mM $MgCl_2$ x $6H_2O$, 10 mM $MgSO_4$ x 7 H_2O was added, previously autoclaved or sterile filtered. For transformations of *E. coli* cells SOC-Medium was used. This is SOB-Medium plus 20 mM glucose; previously sterile filtered [86].

2.2.2. Transformation

An aliquot of competent cells was taken out of a -80°C freezer, and thawed on ice. 100 µl of the cells were transferred into a falcon tube, and 2 µl plasmid or ligation reaction was added. The preparation was swirled and incubated for 30 min on ice. After the incubation the cells were heat shocked for 25 sec. in a 42°C water bath, and chilled for 2 min on ice. 900 µl pre warmed SOC-media were added and the cells were incubated at 37°C by 200 rpm on a shaker for 1 h. Different aliquots were then plated on agar plates with the appropriate antibiotics and incubated at 37°C overnight.

2.2.3. Preparation of competent cells

To transform bacteria with recombinant plasmids, we needed competent cells with an efficiency of transformation of 10^6-10^7 colonies/µg plasmid. The modified preparation of the cells was done according to the $CaCl_2$ method of Hanahan [87]. 3 ml of SOB-media (**2.2.1.**), supplemented with the appropriate antibiotic was inoculated with the glycerol stock of the used bacterial host/cells, and incubated at 37°C overnight. The following day 100 ml SOC-media (**2.2.1.**) was inoculated with 10 µl of the overnight culture (1:10,000 dilution factor), and incubated under vigorous shaking at 37°C, until the culture reached an OD_{550nm} of 0.280. The cells were then centrifuged for 5 min at 1000 x g at 4°C, the cells were then resuspended in 15 ml of ice cooled TfB I-Buffer (30 mM potassium ace-

tate; 100 mM KCl x 2 H_2O; 50 mM $MnCl_2$ x 6 H_2O; 10 mM $CaCl_2$; 15% glycerol; adjusted with 10% acetic acid to pH 5.8; sterile filtrated or autoclaved). After a 10 min incubation on ice the cells were centrifuged again, as above, and the pellet was carefully resuspended during a time period of 45 min in an ice water bath with 4 ml of TfB II-Buffer (10 mM MOPS; 10 mM KCl x 2 H_2O; 75 mM $CaCl_2$; 15% glycerol; adjusted with 0.1 M NaOH to pH 7.0; sterile filtrated or autoclaved). Afterwards the cells were aliquoted in dry-ice cooled reaction tubes and were frozen at -80°C. The transformation efficiency was tested with a test plasmid (pUC18); the efficiency had to be 10^6-10^7 colonies/µg plasmid.

2.3. Gel electrophoresis

2.3.1. DNA/RNA-agarose gels

Agarose gel electrophoresis was used to separate and identify DNA fragments, with a length of 0.5-25 kb. Agarose (0.5%-2%, dependent on the size of the DNA fragments) and 1 x TAE-buffer (50 x TAE: 242 g Tris-Base, 57.1 ml glacial acetic acid, 100 ml 0.5 M EDTA pH 8.0 add 1000 ml H_2O, adjust to pH 8.5; diluted) was boiled, and ethidium bromide (0.4 µg/ml) was added, and poured into a gel chamber. To the samples 5 x loading buffer (sucrose (20% w/v), Orange G (0.035% w/v)) was added, and the gel was separated between 60-100 V. To identify the length of the fragments a length-marker was also added. For the analysis of RNA, a 1.2% formaldehyde-agarose gel was used. 1.2% agarose (w/v) was added into 1 x FA-buffer (10 x FA buffer diluted: 200 mM MOPS; 50 mM sodium acetate; 10 mM EDTA; adjust with NaOH pellets the pH to 7.0 and autoclaved) and was then boiled. When the gel was cooled down to ~50°C, formaldehyde (0.22 M) plus ethidium bromide (0.4 µg/ml) was added, and poured into a gel chamber. The samples were diluted with 5 x loading buffer (80 µl 500 mM EDTA, pH 8.0; 16 µl saturated aqueous bromphenol blue solution; 720 µl 37% (12.3 M) formaldehyde; 2 ml glycerol; 3,084 ml formamide; 4 ml 10 x FA-buffer ad 10 ml) and denatured for 5 min at 65°C, chilled on ice, centrifuged and then loaded. The electrophoresis was done at 60-80 V in a 1 x running buffer (100 ml 10x FA-buffer; 880 ml RNase free water; 20 ml 37% (12.3 M) formaldehyde).

2.3.2. Protein polyacrylamide gel electrophoresis (SDS-PAGE)

To determine protein expression a modified denaturing SDS-polyacrylamide-gel-electrophoresis (SDS-PAGE**)** [88] was used. The SDS-PAGE was composed of a 7%-15% separating acrylamide/bis-acrylamide gel (0.375 M Tris-HCl pH8.8, 0.1% SDS (w/v), 0.05% APS, 0.05% [v/v] TEMED), overlaid with a 4% stacking or acrylamide/bis-acrylamide gel (0.125M Tris pH6.8, 0.1% SDS (w/v), 0.05% APS, 0.1% [v/v] TEMED). The expression pellets were charged with 1 x SDS-PAGE loading-buffer (50 µl for the pre- and 100 µl for the post-induction; according to the QIAexpressionist; 0.045 M Tris HCl pH 6.5, 10% [v/v] glycerol, 1% SDS (w/v), 0.01% bromphenol blue (w/v) and 0.05 M DTT); protein samples were charged and diluted according to their volume with 5 x loading-buffer. The sample was mixed, at 95°C 5 min denatured and later chilled for 2 min on ice, then centrifuged at 16,100 x g. Some samples contained DNA, which disturbed the loading procedure; to reduce this problem an ultrasonication was performed. 10-20 µl of the supernatant were loaded on the gel, and separated first by 80-90 V during the passage of the stacking gel and later at 120-135 V in the separating gel. A protein marker was included in each run (Fermentas #SM0431 116.0 kDa, 66.2 kDa, 45.0 kDa, 35.0 kDa, 25.0 kDa, 18.4 kDa, 14.4 kDa; #SM0671 170 kDa, 130 kDa, 100 kDa, 70 kDa, 55 kDa, 40 kDa, 35 kDa, 25 kDa, 15 kDa, 10 kDa; BIO RAD #161-0305 103 kDa, 77 kDa, 50 kDa, 34.4 kDa, 28.8 kDa, 20.7 kDa). The production of the gel was dependent on the apparatus used, here equipment from Biometra was used, the gels were run with 1 x running buffer (5 x running buffer 1:5 diluted; 15 g Tris-Base ph 8.8, 72 g glycine and 5 g SDS ad 1 L H_2O). The gel was stained for 40 min on a platform shaker with stain solution (100 ml glacial acetic acid, 300 ml methanol, 600 ml H_2O and 1 g Coomasie Brilliant Blue G-250), and destained with the destain solution (stain solution without Coomasie Brilliant Blue G-250), or destained with 7% glacial acetic acid.

2.3.2.1. SDS-PAGE for protein-sequencing with a mass-spectrometer

During the whole procedure nitrile gloves had to be worn, to protect samples from contamination with skin proteins (i.e. keratin). A normal 10% SDS-Gel with the protein samples was run (**2.3.3.**). After the run, the gel was rinsed 3 times in MilliQ-H_2O and then 3 times 30 min in 100 ml Sypro Fix/Destain buffer (10% methanol and 7% acetic

acid); the buffer was disposed in non-chlorinated hazardous waste. Next the gel was stained with 50 ml Sypro Ruby Gel Stain™ (Invitrogen), at least 3 h, but best overnight. After this the gel was placed in the dark. Next the gel was rinsed again 3 times in the Fix/Destain buffer and the MilliQ-H_2O, as above. For bigger SDS-Gels (~ 17 cm) 200ml Fix/Destain buffer and 200 ml Sypro Ruby Gel Stain™ was used. For imaging the with UV detectable dye stained gel the GEL PIX spot picker of GENEPIX was used, and the picture was taken with a cooled CCD-Camera using blue light. Afterwards the appropriate band was cut out of the gel with a scalpel under UV-light, cut in ~1 mm pieces and placed into a reaction tube. Only USA Scientific 1.5 ml Seal-Rite microcentrifuge tubes (polypropylene) catalog #1615-5500 were used for all procedures, other tubes might have leached plasticizers into the sample, which would adversely affect mass spec analysis. 300 µl of Milli-Q H_2O was added to the tube which was then shaken for 15 min The H_2O was removed by pipetting: A 10 µl pipet tip attached to a 1000 µl pipette tip was used (the small opening of the 10 µl tip made it easier to avoid aspirating the gel pieces). The Sypro stained bands were destained by adding 300 µl ACN (CH_3CN; acetonitrile) to the tube and shaking for 15 min The ACN was discarded, 300 µl of 50 mM NH_4HCO_3/50% ACN was added, and then shaken again for 15 min. Supernatant was removed and discarded, and the washing step above was repeated at least three times. After destaining was complete, the supernatant was removed and discarded. 200 µl ACN was added and the sample was shaken again 5 min, removed and the supernatant was discarded, then the gel pieces were air-dried for a few minutes. For the digestion a trypsin digestion buffer (5 µl of 100 ng/ µl trypsin, 40 µl H_2O and 40 µl 100 mM ACN) was prepared. For gel bands, 30-40 µl of the buffer was used, after 5 min, if the gel pieces were not completely covered by the buffer, wait 5 min; wait time was added as needed (to allow more trypsin to be absorbed), then enough 100 mM ACN was added to cover it. The tube was placed in a 37°C incubator for 8-12 hrs or was stored at 4°C until it was convenient to prepare the sample. At this time point the sample could be stored at -20°C for later analysis. It was important that the tube was sealed tightly and the gel pieces were settled in the bottom of the tube. To extract the proteins 30 µl 1% formic acid/2% ACN were added to the digest, and the reaction was incubated for 30 min on a platform shaker. The supernatant was removed to a new tube with the tip to tip method described above to avoid transfer-

ring gel pieces. 24 µl 60% ACN was added to the gel pieces, which were then shaken for 30 min, the supernatant was removed and combined with that of the previous extraction, the gel pieces were discarded. The extracted digests were placed, into a SpeedVac. There they were lyophilized until all of the liquid has evaporated, leaving a tiny pellet (~1h). Dried samples were stored until needed, or, if ready to proceed, went directly to ZipTipping, rehydrated with 10 µl 1% formic acid/2% ACN. For the following step the ZipTip sample clean up was done by the DDPSC MS-facility (Jeanne Sheffield). A µ-C18 ZipTip (Omix Tips are Varian's brand of ZipTip; Millipore) was used for cleaning up peptide samples. The sample had to be in an aqueous buffer to bind to the ZipTip. If the sample was dry it was rehydrated first in 1% formic acid/2% ACN, freshly prepared. A 4 µl aliquot of elution buffer (60% ACN/0.1% formic acid) was prepared for the sample before beginning (to avoid contamination). It was important to avoid drawing air through the tip during the procedure (from equilibration to elution). If you found that there are too many bubbles in the tip, try pulling the buffers in more slowly. The pipette was set to 10 µl and the ZipTip was attached. The ZipTip was equilibrated using 3 x 100% ACN and 3 x 0.1% formic acid, the buffer was pipetted and then discarded it in the waste, these steps act as a gradient for the mini column, which wets the resin and conditions it to be ready to bind peptides. The next step was then, loading the peptides by pipetting the protein digest 10 x up and down (it was discarded back into its tube). The ZipTip was then washed by pipetting 6 x 0.1% formic acid buffer. The peptides were eluted by pipetting the ZipTip 10 x up and down in the elution buffer (4 µl aliquot of elution buffer; 60% ACN/0.1% formic acid). The organic phase elutes the peptides off the resin and into the buffer. Now the sample is desalted and is concentrated in 4 µl of solution appropriate for analysis by MALDI-TOF-MS (Matrix-assisted laser desorption/ionisation-time of flight mass spectrometry) and ESI-Q-TOF-MS (electrospray ionization quadrupole-time of flight mass spectrometry). If mass spec analysis will not be performed on the same day, freeze the samples at –80˚C until needed.

2.3.2.2. SDS-PAGE drying

The drying method used was based on TUT'S TOMB Gel Dryer (RPI; Research Products International Corp.), but it is further described here because it has been modified from the

original user's manual. A dual sheet method was used. This is the most commonly used method for long-term storage. The gel was stained and destained as normal. Two pieces of cellophane were cut to the size of the solid backing plate, the cellophane and the gel were soaked in gel drying buffer (20% [v/v] methanol and 10% [v/v] glycerol) for around 1-2 min. The curly nature of the cellophane was allowed to re-roll the cellophane. The cellophane was smoothed over the backing plate, and then the gel plate was centered. A second sheet of soaked cellophane was smoothed over the gel, then the frame clamped over this sheet with several binder clips, and was dried horizontally overnight. The dried gel was removed from the frame, excess cellophane was cut off and the gel stored flat. Gels could be stored indefinitely if not exposed to humidity fluctuations.

2.4. Isolation of nucleic acids

2.4.1. DNA mini and midi preparation

The DNA mini preparation was used to isolate plasmids out of bacterial cultures. For this 6 ml of LB-media containing the appropriate antibiotics were inoculated with a glycerol stock or a bacterial colony, and incubated at 37°C overnight at around 180 rpm. The overnight culture was then centrifuged at 4°C and 3220 x g for 5 min. The plasmid purification was then done with the QIAprep Spin Miniprep Kit (Qiagen), according to the manufacturer's protocol for micro centrifuges. The DNA midi preparation procedure was used to increase the yield of plasmids. For this a 50 ml overnight culture was used. The solution was centrifuged in a 50 ml falcon-tube at 4°C by 1811 x g, for 10 min. The plasmid purification was done with the QIAfilter Plasmid Midi Kit (Qiagen), according to the manufacturer's protocol.

2.4.2. Purification of PCR products

The purification of PCR products was done to get rid of the dNTP's, salts, and primer. Here the QIAquick PCR Purification Kit (Qiagen) was used according to the manufacturer's protocol. For some applications like restriction digests, ligations, and PCR, an agarose gel extraction was done. For this a preparative gel was poured, which allowed the loading of a reaction volume of 50 µl and more. After the separation, appropriate bands were cut out under UV-light, and transferred into a tube. The purification out of the gel

was then done with the MinElute Gel Extraction Kit (Qiagen), according to the manufacture's protocol.

2.4.3. Nucleic acid purification

This procedure was used to isolate DNA and RNA out of plant-specific tissues. 1g of plant tissue was placed in liquid N_2, ground with mortar and pestle to a homogenous powder, and transferred to a 15 ml falcon-tube, 3.5 ml cooled lysis-buffer (10 mM Tris-HCl pH 7.5; 50 mM NaCl; 1% SDS; 4% PVPP; 1 mM EDTA pH 8.0; 14 mM β–EtSH (freshly added)) was added + 3,5 ml phenol/chloroform and extracted 30 min on a shaker at RT, centrifuged 10 min 3220 x g at RT, the aqueous phase was taken + 3,5 ml phenol/chloroform, 10 min 3220 x g at RT, the aqueous phase was taken + 3 ml chloroform, 10 min 3220 x g at RT, the aqueous phase was taken + 0,1 [v/v] 3 M Na-acetate pH 5,2 + 1 [v/v] isopropyl alcohol placed at -20°C for 1h - 2h, 20 min 3220 x g, the supernatant was discarded and the pellet was washed with 500 µl 70% EtOH (do not resolve the pellet), 10 min 3220 x g at RT, the supernatant was discarded and the pellet was dried, the pellet was resolved with 300µl TE-buffer (10 mM Tris-HCl pH 8.0 ; 1 mM EDTA pH 8.0), a pipet-tip was used (do not vortex), 300 µl 6 M LiCl was added, RNA precipitated overnight at 4°C; on the next day the RNA was centrifuged for 15 min by 16.100 x g at 4°C, **IMPORTANT** supernatant contains the genomic DNA. To precipitate the genomic DNA 0.1 volumes 3 M Na-acetate pH 5,2 + 1 volume isopropyl alcohol was added and later stored at -20°C for 20 min, **IMPORTANT** pellet contains the total RNA, the pellet was washed with 500 µl 70% EtOH (do not resolve the pellet), 10 min 3220 x g at 4°C, the pellet was dried and resolved in 50 µl TE-buffer by mixing and pipetting = **total RNA,** to quantify the RNA with a photometer (OD_{260nm}) it was diluted 1:20 – 1:50 (**2.5.1.**). The genomic DNA supernatant was centrifuged for 15 min (16.100 x g) at 4°C, the pellet was washed with 500 µl EtOH (do not resolve the pellet) and centrifuged for 10 min (16.100 x g) at 4°C, the pellet was dried and resuspended in 50 µl TE-buffer (do not vortex) = **genomic DNA,** to quantify the DNA (OD_{260nm}) dilute it 1:50 - 1:100 (**2.5.1.**). RNA and DNA were stored at -80°C.

2.4.4. TRIzol RNA purification

Plant tissue frozen in liquid N_2 was homogenized, using a mortar and pestle, ground to a homogenous powder. Each sample was dissolved in 1 ml of TRIzol Reagent (0.1 M sodium acetate pH 5.0, 0.8 M guanidinium thiocyanate, 0.4 M ammonium thiocyanate, 5% [v/v] glycerol, 38% [v/v] aqueous phenol per 50-100 mg of plant tissue [89]. The reagent freezes at this step, wait until the mixture is thawed and ground it in between, until you can pour it into a reagent tube. The homogenized phase was incubated for 5 min at 15˚C to 30˚C to permit the complete dissociation of nucleoprotein complexes, 0.2 ml of chloroform per 1 ml TRIzol Reagent was added. Sample tubes were capped securely. The tube was vigorously shaken for ca. 15 sec. and then incubated for 2-3 min at 15˚C to 30˚C. The samples were centrifuged for 15 min at 12,000 x g by 2-8˚C. In the following centrifugation, the mixture separated into a lower red, phenol chloroform phase, an interphase, and a colorless upper aqueous phase. RNA remained exclusively in the aqueous phase. The volume of the aqueous phase was about 60% of the volume of TRIzol Reagent used for homogenization. The aqueous phase was transferred to a fresh tube; the organic phase was saved if isolation of DNA or protein was desired (optional). The RNA from the aqueous phase was precipitated by mixing with 0.5 ml isopropyl alcohol per 1 ml TRIzol Reagent. The samples were incubated for 10 min at 15-30˚C, and then the samples were centrifuged for 15 min (12,000 x g) at 2-8˚C. The RNA precipitation was often invisible before centrifugation, and formed a gel-like pellet on the side and bottom of the tube. The supernatant was removed and the pellet was washed once with 1 ml 75% ethanol per 1 ml TRIzol Reagent. The sample was mixed using a Vortex or similar equipment and centrifuged for 5 min (7,500 x g) by 2-8˚C. At the end of the procedure, the RNA pellet was briefly dried, for not longer than 10 min, and the RNA samples were dissolved into 100 µl of RNase free H_2O, 0.5% SDS solution or TE-Buffer (**2.4.5.**) and the concentration was determined (2.5.1.). The RNA was incubated for 10 min at 55-60˚C, prior to running it onto a gel or storing it for later use at -80˚C.

2.5. Concentration determinations

2.5.1. Nucleic acid concentration determination with a UV-spectrophotometer

The concentration of nucleic acids was performed photometrically at 260 nm. Absorption of 1.0 was equivalent to the following concentrations:

DNA (double strand)	50 µg/ml
RNA	40 µg/ml

2.5.2. Protein concentration determination

For this a calibration curve with BSA (0-30 µg/ml) was done at 595 nm, using 800 µl of the BSA standards and 200 µl Bio-Rad protein assay solution [90].

2.6. Polymerase chain reaction (PCR)

This procedure was used to duplicate DNA fragments in an exponential way, with primer and thermo stable DNA-polymerases [91]. A PCR program was based on 3 steps, a de-naturing-, an annealing, and an elongation part.

General PCR-cycle:

denaturing	94°C	10 sec.-5 min
25-45 cycles:		
denaturing	94°C	30-45 sec.
annealing	50-60°C	30-45 sec.
elongation	72°C	1-5 min
final elongation	72°C	5-7 min
reaction stop	4°C	∞

Preparation:

10x polymerase buffer	3 - 5 µl
dNTP's	0.2 mM each
5' and 3' Primer	0.2 µM
DNA	100-200 ng/1 µl *E. coli* miniculture
DNA-Polymerase	2.5 U
dH_2O	ad 30-50 µl

2.6.1. RACE-PCR

Is a method for performing both 5'- and 3'-rapid amplification of cDNA ends (RACE). For this the "BD SMART^{TM} RACE cDNA amplification kit" (BD-biosciences) was used.

2.7. Sequencing of DNA

The DNA was sequenced according to Sanger *et al*. 1977 [92]. Laser detectable fluorescence marked dNTP's were used. The plasmid sequencing was performed according to the BigDye V 1.1. protocol of the manufacturer (Applied Biosystems).

Preparation:	Sequencing cycle (25cycles):	
1-5 µl plasmid DNA (500 ng)	denaturing 96°C	10 sec.
1 µl primer (10 pmol)	annealing 50-55°C	5 sec.
4 µl BigDyeMix V 1.1.	extension 60°C	4 min
ad 10 µl dH_2O	reaction stop 4°C	∞

The sequencing reaction was purified over Sephadex and loaded on the ABI-sequencer. A 96-well multiscreen plate (Millipore) was loaded with Sephadex G-50 Superfine. To the Sephadex 300 µl H_2O was added and the gel was placed at 4°C, to allow swelling, for 3-4 hrs. The plate was then centrifuged for 5 min at 910 x g in another plate, afterwards 10 µl H_2O was added to the sequencing reaction, and centrifuged over the Sephadex for 5 min at 910 x g in a sequencing plate, or the SigmaSpin™ Post-Reaction Purification Columns were used according to the manufacturer's protocol. The sequencing reaction was then placed in an ABI 310 Genetic Analyzer or in an ABI 3100-*Avant* Genetic Analyzer (both Applied Biosystems). Otherwise, the sequencing was done by Washington University in St. Louis.

2.8. Cloning techniques

2.8.1. TOPO-ligation

The ligation of PCR products into a TOPO vector was based on the system described under **2.1.2.**

Preparation:	
PCR-product	10 ng
1.2 M NaCl, 60 mM $MgCl_2$	1 µl
TOPO vector	1 µl
dH_2O	ad 6 µl; incubate for 30 min at RT and chill on ice

2.8.2. *Taq*-tailing and pGEM-T Easy cloning of PCR products

The *Taq*-polymerase has a matrices-independent activity; making it capable to bind adenine on a double stranded blunt-end DNA. For this a plasmid was used for PCR and the product was purified.

General reaction preparation:

Purified PCR product	30 µl
10 x *Taq*-buffer	5 µl
dATP	1 µl
Taq-polymerase	0.5 µl
dH_2O	ad 50 µl; incubated 30 min at 72°C.

The reaction was purified (2.4.2.) and cloned into a thymine overhang containing pGEM-T Easy vector.

General reaction preparation:

2 x rapid ligation buffer	5 µl
pGEM-T Easy vector	1 µl
PCR-product	3 µl
T4-ligase	1 µl
dH_2O	ad 10 µl; incubated overnight at 4°C.

Afterwards 2 µl of the product were transformed into DH5α and plated on LB-amp (50 µg/ml) with X-Gal (80 µg/ml) and IPTG (1mM).

2.8.3. Restriction digest

Double stranded DNA could cut on palindrome-like sequences with the help of specific restriction enzymes. With this technique DNA fragments were yielded to clone in other vector systems. As an example *Eco*RI and *Bam*HI were used as restriction endonucleases.

Preparation:

DNA	1-2 µg
*Eco*RI	10 U
*Bam*HI	10 U
*Eco*RI 10 x buffer	5 µl
dH_2O	ad 50 µl; incubated for 2 h. at 37°C, than to 4°C.

The digest was purified (2.4.2.), and the concentration was determined (2.5.1.).

2.8.4. Ligation and Blunting of PCR products

With the help of T4-ligase a DNA fragment was cloned into a vector. The reaction was done under a usage of dATP (contained the reaction buffer); whereas the 3'-hydroxyend led a nucleophilic attack on the 5'phosphorus activated AMP rest, which leads to the binding.

Preparation:

Vector-DNA	50 ng
Insert-DNA	x ng
10 x Ligase-buffer	1 µl
T4 DNA-Ligase	1 µl
H_2O	ad 10 µl, incubated overnight at 4°C.

For determination of the insert, the molar measure vector : insert ratio is 1 : 3. For the calculation the following formula was used: (ng vector x kb insert / kb vector) x molar measure vector:insert = ng insert. DNA polymerases such as *Taq* or DNA polymerase mixtures like Herculase Hotstart or BD Advantage 2, containing *Taq*- and proof reading polymerases, produced adenosine overhangs which had to be removed for cloning into blunt-end vectors, for this blunting is recommended before ligation.

Preparation:

PCR product	x
10 x *Pfu* Buffer	y
dNTP's	0.2 mM each
Pfu-polymerase	2.5 U
dH_2O	ad 30-50 µl; incubated at 72°C for 30 min and placed on ice.

2.9. Trichome isolation

The trichome isolation was done according to Gershenzon *et al.*, 1992 [93]

2.10. Blotting

2.10.1. Transfer of RNA on nylon membranes

The RNA blot (northern blot) was used to determine the tissue-specific expression of researched genes, by radioactive labelling of DNA and hybridization on the RNA. The following buffer components were used: 20 x SSC buffer: 0.3 M Na-Citrate, pH 7.3, 3 M NaCl; 2 x SSC buffer: 20 x SSC 1:10 diluted with H_2O; 2 x SSC buffer with 0.1% [v/v] SDS: 20 x SSC 1:10 diluted with H_2O 0.1% [v/v] SDS; 0.1 x SSC buffer with 0.1% [v/v] SDS: 20 x SSC 1:200 diluted with H_2O 0.1% [v/v] SDS. A 1.2% RNA agarose gel was

loaded with 10 µg of the isolated RNA (**2.4.5.**), which was previously denatured at 65°C for 5 min and chilled on ice. The electrophoresis was done at 60-80 V, until the front reached 4 cm. The gel was then placed on a filter paper, which was previously rinsed in 20 x SSC buffer. The filter paper built a bridge between two 20 x SSC containing reservoirs. The gel was then sealed on every side with parafilm. Afterwards the following components were placed on the gel: a nylon membrane, first rinsed in H_2O and then rinsed in 20 x SSC; three slides of GB 004 Gel Blotting Paper; approx. 5 cm high filter paper; approx. 10-15 cm high paper towels; finished with a glass plate on which a weight of 250-500 g was laid. The nylon membrane had to have the same size as the gel. The blot was done overnight at RT. All components used had to be washed previously with ethanol, to reduce contamination with RNases.

2.10.1.1. Amplifying the cDNA hybridization probe

To produce the cDNA a plasmid, that contained the appropriate gene in the MCS, was used for a PCR. Two gene-specific 5'- and 3'-primers were used.

Common cycle:			Preparation:	
denaturing	93°C	3 min	10 x *Taq*-buffer	5 µl
30 cycles			dNTP's 10 mM	1 µl
denaturing	94°C	30 sec.	5'-primer 10 pmol	1 µl
annealing	50°C	30 sec.	3'-primer 10 pmol	1 µl
elongation	72°C	1-1.5 min	*Taq*-polymerase	0.5 µl
final elongation	72°C	7 min	dH_2O	40.5 µl
reaction stop	4°C	∞		
plasmid	1 µl			

The PCR product was then purified (**2.4.3.**), and the concentration was determined (**2.5.1.**)

2.10.1.2. Hybridisation probe preparation of the RNA-Blot

To hybridize an RNA-Blot a radioactive labelled hybridization probe has to be prepared.

General reaction preparation:

x µl	25 ng DNA-sample
5 µl	random primer (nonameres)
ad 22 µl	dH_2O; denatured for 5 min at 95°C and placed on ice
4 µl	of each nucleotide (dCTP, dTTP, dGTP)
5µl	10 x reaction buffer
2 µl	klenow fragment
y µl	[^{32}P-α-dATP] 30-50 µCi (3000 Ci/mmol), 3-5 µl
ad 50 µl	dH_2O; incubated for 10 min at 37°C

For the purification of the reaction SampleQuant G-50 microtubes were used; the tubes were mixed to resuspend the column material. After decapping of the column-end the tube was transferred to a reaction tube and centrifuged (735 x g) for 1 min and the flow through was discarded. The hybridization probe was then transferred to the column and centrifuged (735 x g) for 2 min; afterwards the hybridization probe was denatured for 1 min at 95°C and was placed on ice. Here it was possible to determine the incubation rate of radioactivity with liquid scintillation counting, 1 µl diluted in 49 µl H_2O prior and post column purification. The nylon membrane was short rinsed in H_2O and the lines were labelled, then the blot was checked with UV-light. To fix the RNA on the nylon-membrane a UV-crosslinking was carried out for 1 min and the membrane was placed on a humid chromatography paper. Afterward the membrane was washed in H_2O or 2 x SSC-buffer. Then the nylon-membrane was transferred into a hybridisation glass tube; for this step there were two choices of procedures. Adding ~15 ml hybridisation solution (BD Biosciences) prewarmed and incubated for 30 min at 68°C. Or self-made hybridisation solution (6.3 ml H_2O, 2.5 ml 20 x SSC, 100 µl 10% SDS, 1 ml 50 x Dehnhard's reagent (2 g Ficoll 400, 2 g Polyvinylpyrrolidone, 2 g BSA ad 200 ml H_2O), 125µl 100 µg/ml herring sperm DNA) prior denatured for 5 min at 95°C and then 3-4 hrs incubated at 68°C. To label the membrane 25 µl of the hybridization probe were added, and incubated overnight in a hybridisation oven at 68°C. The remaining DNA-sample was stored at -20°C, for a possible rehybridisation. On the next day the membrane was washed 3 x at RT with 2 x SSC; 0.1% SDS and then 2 x at 50°C with 0.1 SSC; 0.1% SDS. The mem-

brane was then sealed in plastic wrap and placed overnight on a phosphor-imaging screen. The visualisation of the screen was done with a phosphor imager.

2.10.2. Transfer of proteins on nylon membranes (western blot)

A 10% SDS-PAGE was run (**2.3.3.**), but instead of the normal protein-ladder a pre stained ladder was used (Prestained SDS-PAGE Standards, Broad Range: BIO RAD). The SDS-sample-buffer front was allowed to run out of the gel. The Trans-Blot® SD Semi-Dry Electrophoretic Transfer Cell (BIO RAD) was used for the transfer of the proteins to the nylon membrane (Biodyne® B 0.45 µm; Pall Corporation). The blotting buffer was prepared according to Bjerrum and Schafer-Nielsen 1986 [94] (for 1 liter: 48 mM Tris (5.82 g), 39 mM glycine (2.93 g), 20% methanol HPLC-grade and 3.75 ml 10% SDS, the pH was between pH 9.0 and pH 9.5, with no adjustment necessary). The SDS-PAGE, extra thick GB-blotting paper and the nylon membrane was incubated for 1 h, at RT, in the blotting buffer; afterwards blotting paper, nylon membrane, SDS-PAGE and additional time blotting paper were placed on the moist (blotting-buffer) surface of the Trans-Blot® SD Semi-Dry Electrophoretic Transfer Cell. The apparatus was assembled and the transfer of proteins to the nylon membrane was performed for 1 h. at 25 Volt. The nylon membrane or blot was blocked (to avoid unspecific labelling of proteins) with TBS-T buffer [20 mM Tris HCl (2.4 g) pH 7.5, 150 mM NaCl (8.8 g), 0.05% Tween 20 (500 µl)/liter] containing 3-5% milk powder, and incubated for 1 h. at RT or overnight at 4°C, on a platform shaker. Prior to the labelling of the searched His-tagged protein with the antibody, the blot was washed 3 x 5 min with 20 ml TBS-T buffer. Monoclonal Anti-polyhistidine-Alkaline Phosphatase antibody produced in mouse (Sigma; primary and secondary antibody are here together) was used for the labelling, then the nylon membrane was incubated for 2 hrs at RT with 20 ml TBS-buffer [20 mM Tris-HCl pH 9.5 (2.4 g), 150 mM NaCl (8.8 g)/liter] containing 1:2000 diluted antibody (10 µl) and 1% BSA (0.2 g); afterwards the blot was washed an additional 3 x 5 min with 20 ml TBS-T buffer. To start the color reaction 20 ml TBS-buffer [20 mM Tris-HCl pH 9.5 (2.4 g), 150 mM NaCl (8.8 g)] containing 90 µl NBT (Nitrobluetetrazolium 75mg/ml in 70% DMF), 70 µl BCIP (5-Bromo-4-chloro-3-Indolyl phosphate 50mg/ml in dimethylformamide or as a salt in H_2O) and 200 µl 0.5 M MgCl were added. The color reaction was

stopped with 20 ml TBS buffer containing 800 µl 0.5 M EDTA, afterwards the western blot was dried and scanned.

2.11. Protein expression and purification

Volumes and amounts given for 1 l expression cultures, scaled accordingly. The lysis and the purification steps were done at 4°C, if not else stated.

2.11.1. Bacterial culture

A 50 ml LB media culture (starter culture) was inoculated out of the glycerol stock with the appropriate antibiotics (1% glucose, was included to the starter culture if the expressable protein seem to be toxic for the host, to inhibit basal transcription by catabolite repression (diauxie), over decreasing cAMP level). The culture was grown to a density between OD_{600nm} = 0.6-0.7 at 28-37ºC by 200 rpm shaking. Starter culture could be stored at 4ºC until use. 1 liter of LB and antibiotics was used in a 2 l flask to initiate expression culture by transferring 50 ml starter culture to the media. If more than 1 liter of expression culture was used, the starter cultures were combined and spread prior to the inoculation. The culture was incubated at 37ºC with 200 rpm shaking until OD_{600nm} = 0.6, which usually took about 4 hours. A shaker/incubator was cooled to 28ºC or to the appropriate expression temperature. During the expression, many expressed proteins could have toxic effects on the host, by decreasing the expression temperature this effect was minimized. After the culture reached OD_{600nm} = 0.6 the flask containing the culture was placed on ice for 5 min or in a fridge. Two 1 ml aliquots of the culture were transferred to separate 1.5 ml micro centrifuge tubes as 'preinduction samples'. The tubes were centrifuged for 5 min at 16,100 x g to pellet the cells, the supernatant was discarded and the cell pellet was frozen at -20ºC for later SDS-PAGE analysis. Add IPTG to a concentration of 1 mM/ml (i. e. 1 ml of 1 M solution) to the culture to start the expression (the vector contains a partial lacOperon with the LacI repressor gene. The 1,6-allolactose normally inhibits the repressor and IPTG mimics this molecule). The culture was incubated for 4 hours to overnight at 28ºC (16-22 hrs) with 200 rpm shaking. After the expression was complete 2 × 1 ml aliquots of the culture were transferred to separate 1.5 ml micro centrifuge tubes as 'postinduction samples', stored as above for 'preinduction sam-

ples'. The culture was centrifuged for 10 minutes at 10,000 × g at 4ºC to pellet the cells. The supernatant decanted and remaining media was removed by pipette. At this point you have three options: a) freezing the cells at -80ºC; b) resuspending the cells in HLB (His-Tag-lysis-buffer; 20 ml/L of expression culture, 50 mM Tris-HCl pH 7, 500 mM NaCl, 10 mM imidazole, 10% [v/v] glycerol, 10 mM β-EtSH, 1% Tween 20, 750 µg/ml lysozyme) and then freezing the cell suspension at -80ºC; or c) resuspending the cell pellets in HLB (20 ml/l of expression culture) and continuing with the purification protocol. Glycerol neutralizes non-specific hydrophobic interactions and stabilized the enzyme during freezing process; β-EtSH is a protease inhibitor as or like PMSF and imidazole is identical to the histidine side chain (active site competitor).

2.11.2. Bacterial lysis

The frozen cells or the cell suspension was placed on ice, and mixed occasionally to accelerate the thawing process, then transferred to a centrifuge tube that has been precooled on ice. Subsequently the cell suspension was incubated on ice or in a cold room for 1 h., to allow lysozyme to lyse cells. The cell suspension was sonicated with an ultrasonic sample for 2-3 × 30 sec at ~75% power, and cooled on ice for 1 min between sonications. Then the bacterial lysate was centrifuged for 25 min at 12,000 × g at 4°C. The supernatant was decanted into a new 50 ml falcon tube that has been precooled on ice. A 500 µl aliquot was removed as a 'crude lysate sample' and stored at -20°C for later SDS-PAGE analysis.

2.11.3. Protein purification with Talon resin

The Talon resin suspension was removed from 4°C and inverted to resuspend. Two ml of the suspension (= 1 ml Talon resin) were transferred by pipette into a 15 ml falcon tube, centrifuged at 700 × g for 2 min and the supernatant was discarded. The pelleted resin was resuspended with 2 x ~10 ml of HWB (His-tag-wash-buffer; 50 mM Tris-HCl pH 7, 500 mM NaCl, 10 mM imidazole, 10% glycerol, 10 mM β-EtSH), centrifuged like above and the supernatant was discarded. Afterwards the pelleted resin was resuspended in 5 ml of crude bacterial lysate and then transferred into a 50 ml tube, containing the remaining crude lysate. The resin and lysate were gently agitated for 60 min on a platform shaker to

allow the His-tag to bind the resin. Then the resin was centrifuged at 700 × g for 5 min and the supernatant decanted, a 'Talon supernatant fraction' was kept for later SDS-PAGE analysis. The resin was then transferred into a new 15 ml tube and resuspended in ~12-14 ml of HWB and agitated it on platform shaker for 10 min, and then centrifuged at 700 × g for 5 min The supernatant was discarded and this process was repeated a second time. 1-2 ml of HWB were added to the resin and resuspended by mixing. All of the resin was transferred to a 2 ml gravity flow column (end-capped) and the resin was allowed to settle, or the column was short spinned down at 700 x g to avoid this time consuming step, but this could lead to a decrease in the purification amount. As the resin settles, the column was tapped to force air bubbles to the top. The end-cap was removed, and buffer was drained to the top of the resin. The resin was washed with 5 ml HWB. The protein was eluted by adding 5 ml of HEB (His-tag-elution-buffer; 50 mM Tris-HCl pH 7, 500 mM NaCl, 500 mM imidazole, 10% glycerol, 10 mM β-EtSH) to the column. 5 × 1ml fractions were collected; fractions where kept (20 µl) as 'Talon fraction' for later SDS-PAGE analysis. Usually His-tagged protein elutes in fraction 1.-3. The protein concentrations of the fractions were determined, only to know which fractions contain proteins, this is known as short-Bradford assay, without a real measurement of the concentration. 20 µl of each fraction, 780 µl of HEB and 200 µl of Bio-Rad protein assay solution were mixed. A blank is made by mixing 800 µl of HEB with 200 µl of Bio-Rad protein assay solution. The Abs_{595nm} of each fraction is then determined. Protein containing fractions were combined to a volume of ~2.5 ml, with a storage buffer appropriate for your enzyme. Protein was then applied to a PD10 column that has been equilibrated in 5 column volumes of enzyme storage buffer. The protein was eluted in 1 ml fractions and the protein concentration of each fraction was determined like above, PD10 column aliquots (20 µl) were kept for later SDS-PAGE analysis, and the protein concentration was determined (**2.5.2.).** Store the protein as appropriate at 4ºC, -20ºC or -80ºC. I stored the protein at -80ºC, frozen prior in liquid N_2. Optional: the Talon resin was washed with 10-column volumes (10 ml) of 20 mM MES pH 5.0 to remove bound imidazole and β-EtSH. Fix capped and end-capped; stored at 4ºC or throw it away. Talon resin can be re-used 3-4 times.

2.11.4. FPLC-protein purification

The separation was done using an ÄKTA-FPLC-system (Explorer), at 4°C with a detection wavelength of 280nm. The affinity chromatography columns used, were HisTrap™ HP (1 ml, 5 ml) or HisTrap™ FF crude (1 ml, 5 ml) from Amersham Biosciences. The whole Talon purification protocol was followed, bacterial lysis included (**2.11.2.**, **2.11.3.**). In general the crude lysate was too viscous for the His-tag column because of DNA remains and mostly it clogged during the loading process, on this step there were some choices. Incubate the lysate supernatant additional with a DNase (benzoase). Add to the supernatant 5 mM $MgCl_2$ and 0.5 U/ml benzoase, incubate the reaction slightly shaking 30 min at 4°C, and centrifuge again 20 min at 12 000 × g . To avoid this time and money consuming step pass the lysate supernatant through a syringe half filled with synthetic filter floss, purchased at a pet store, this will be sufficient to remove all of the viscous material [95], the supernatant should be clear without disturbing components before loading on the ÄKTA-system. The ÄKTA system and the column stand normally under 20% ethanol (degassed by a sonicator-bath) conditions to minimize bacterial growing. The first step is to wash the ethanol away, for this we use Milliq H_2O, which is degassed like the ethanol. Later the system was equilibrated with sterile filtrated HWB (50 mM Tris/HCl pH 7.0, 500 mM NaCl, 10 mM imidazole, 10% [v/v] glycerol, 10 mM β-EtSH) before loading the crude lysate on the system. The terpene synthase activity was eluted by a linear gradient between HWB and HEB (50 mM Tris/HCl pH 7.0, 500 mM NaCl, 500 mM imidazole, 10% [v/v] glycerol, 10 mM β-EtSH), collected in 1 ml fractions. In the appropriate storage or assay buffer the fractions where pooled and desalted over a PD10column. Protein concentration was determined (**2.5.2.**). The system was washed afterwards with Millq H_2O and placed under storage conditions with 20% ethanol.

2.11.4.1. His-tag column regeneration

The column was recharged after 5-7 purifications. The column was stripped by washing with at least 5-10 column volumes of stripping buffer (20 mM sodium phosphate pH 7.4, 0.5 M NaCl, 50 mM EDTA). Before recharging, the column was equilibrated with at least 5-10 column volumes of binding buffer (HWB; **2.11.4.**) and 5-10 column volumes

of distilled water. The equilibrated column was recharged then by loading 0.5 ml or 2.5 ml of 0.1 M $NiSO_4$ in MilliQ H_2O on HisTrap HP 1 ml and 5 ml column, respectively. Salts of other metals, chlorides, or sulphates may also be used. The column was subsequently equilibrated with 5 column volumes distilled water, and 5 column volumes of HWB before storage in 20% ethanol.

2.11.5. Size exclusion chromatography (SEC)

A SEC was used to determine the right molecular mass of a protein, or better the hydrodynamic or Stokes radius, and to distinguish between monomeric and multimeric enzymes (homodimer, heteromer etc.). For this a ÄKTA-FPLC-system with a HiLoad Superdex 200 16/60 (GE Healthcare) column, with a total volume of 120.63 ml (V_t), was equilibrated with the TPS-Buffer (**2.12.1.**) containing 150 mM NaCl and 5% glycerol instead of 10% glycerol, or with DSB (**2.12.2.**) containing 150 mM NaCl. The void volume of 41.5 ml (V_0) was determined using Blue Dextran (2000 kDa). The buffer was sterile filtered and degassed in a sonicator-bath. To calibrate the ÄKTA-FPLC-system calibration proteins with a defined molecular mass were used, bovine serum albumin (BSA; 67 kDa), catalase (240 kDa), aldolase (158 kDa), ovalbumin (45 kDa) and ferritin (450 kDa). A mixture of 2 mg/ml BSA and 1 mg/ml catalase (in H_2O) was first applied on the system. Then a mixture of 1 mg/ml aldolase and 1 mg/ml ovalbumin (in 150 mM NaCl, 5% glycerol, 20 mM KPi, pH 7.5). Ferritin was applied alone in a concentration of 2 mg/ml (in H_2O), ferritin appears in 3 different forms (69.4% monomer, 19.5% dimer and 11.1% trimer) [96]. The purified and enriched (Amicon 30 kDa - 50 kDa) protein of interest was centrifuged (5 min, 16.100 x g) and then applied in a concentration of 1 - 2 mg/ml.

2.12. Enzyme assays

2.12.1. Monoterpene synthase assay

The monoterpenes synthase buffer (TPS-buffer) was based on the buffer described by Schnee *et al*. 2002 [97] and modified for the characterization of CsTPS1 and CsTPS2. The buffer used here contained 10 mM MOPSO pH 7.0, 1 mM DTT, 20 mM $MgCl_2$ x $6H_2O$ and 10% [v/v] glycerol. The original buffer contained additional 0.2 mM $MnCl_2$,

0.2 mM $NaWO_4$ and 0.1 mM NaF. The assay buffer was used without $MnCl_2$, because it would disturb competition inhibition studies, if planned. $NaWO_4$ and NaF are phosphatase inhibitors, but cannot be used because during the purification of the enzyme both components, shown by the protein concentration measurement, interfere with the Bio-Rad Protein Assay solution. The enzyme assay was performed in siliconized polypropylene tubes to avoid the substrate and product binding to walls of the tube. The assay was done in a volume of 500 µl containing TPS-buffer, the recombinant purified enzyme and unlabeled and/or 1-^{3}H labelled GPP (in a negligible amount). The assay was then overlaid with 500 µl n-hexane, to trap the volatile terpenes, incubated for various time points (1-20 min the previously determined linear production time) at 30°C, and then the enzymatic reaction was stopped on ice. After incubation, the assay was mixed vigorously and centrifuged down (0.5 min, 16,100 x g), 300 µl of the hexane phase were taken for liquid scintillation counting. During each assay time point, a boiling control was performed, too, and subtracted from the average of an enzyme assay triplicate. The substrate conversion was measured by scaling up the GPP conversion (1-^{3}H GPP).

2.12.2. Diterpene synthase assay

The diterpene synthase buffer (DSB) was based on the buffer described by Xu *et al*. 2004 [78]. The buffer used here contained 50 mM HEPES pH 7.2, 5 mM DTT, 100 mM KCl, 7.5 mM $MgCl_2$ x $6H_2O$ and 5% [v/v] glycerol. The diterpene synthase activity assay was performed in a final volume of 500 µl DSB including 400 µM GPP, NPP, FPP and GGPP, overlaid with 500 µl hexane to trap the products. After incubation for ~1-3 h. at 30°C, the assay was hydrolyzed for 30 min at 30°C using 20µl 3 N HCl or 500 µl 100 mM sodium acetate buffer plus 5 units of a wheat germ acid phosphatase (Sigma) were added to the assay for the hydrolysis, and incubated for 2 h. at 30°C. Mixtures were extracted 3x with 500 µl hexane, by mixing at (0.5 min, 16,100 x g). The hexane phases were combined and evaporated under an N_2-stream to a final volume of ~200 µl to analyze the reaction by GC-MS.

2.13. Gas chromatography–mass spectrometry (GC-MS)

The GC-MS measurements were performed on a Voyager/Trace GC 2000 (Thermo Quest CE Instruments) under the following conditions: EI 70 eV, source temperature 200-°C; column DB-5MS (J&W, 30 m x 0.25 mm, 0.25 µm film thickness); injection temperature 250-°C, interface temperature 300-°C; carrier gas He, flow rate 1.0 ml/min, constant flow mode; splitless injection, column temperature program: 60-°C for 1 min, then raised to 300-°C at a rate of 10-°C min^{-1} and then hold on 300-°C for 10 min. All products were identified by comparison of their EI-MS spectra with those of the NIST library (V 1.6d), respectively, or purchased authentic standards.

3. Results

3.1. *Cannabis sativa* L.

At the Leibniz Institute of Plant Biochemistry in Halle/Saale Germany, a group under the guidance of Prof. Toni M. Kutchan, was established to unravel the biosynthesis pathways of terpenophenolics (Cannabinoids) and terpenoids in *Cannabis sativa* L.. Dr. Jonathan Page took over this *C. sativa* trichome specific EST (expressed sequence tag) project. 1202 ESTs were isolated; a sequence comparison (blastx) showed that 23% were putative involved in the metabolism (primary and secondary) of the plant, and 7% of them were most likely in the cannabinoid and terpenoid metabolism. Out of these three putative monoterpene synthases and three putative prenyltransferases EST's were used for further research. During my diploma thesis two of the putative monoterpene synthases and the putative prenyltransferases were used for cloning and expression experiments. It could be shown, that the monoterpene synthases were a limonene and a pinene synthase, and the prenyltransferases were a FPP-synthase, a GGPP-synthase and most likely a subunit of a GPP-synthase. Although limonene and pinene synthases are common to many plant species, until now few have been characterized with respect to kinetic parameters [71]. The evolutionary relationship of the *C. sativa* L. terpene synthases with respect to other plant families has also been reported, and a SEC chromatography was performed to determine the size. For the characterization of the monoterpene synthases from *C. sativa* L. one clone of each synthase generated during the diploma thesis was used. Bacterial strain *E. coli* BL21(DE3)RIL/CsTPS1c and BL21(DE3)RIL/CsTPS2c in pET101/D-TOPO expression vector truncated after aa 60 respectively aa 55 preceding the RR-motif, both containing a hexahistidine extension were chosen.

3.1.1. Purification of the monoterpene synthases:

BL21(DE3)RIL/CsTPS1c or BL21(DE3)RIL/CsTPS2c was grown to A_{600} = 0.6 at 37°C in 1-4 L LB medium supplemented with 100 µg/ml ampicillin and 50 µg/ml chloramphenicol as determined by the vector and the host strain. Cultures were then induced by addition of IPTG to a final concentration of 1 mM and grown for another 10-14 h at 28°C (**2.11.1**). Cells were harvested by centrifugation and lysed (**2.11.2.**). The lysate of CsTPS1c was purified by column chromatography using an ÄKTA-FPLC-system, with a

gradient between HWB and HEB. The FPLC-purification started with HWB (flow rate: 5 ml/min for 3 min), followed by a gradient of 0-100% HEB (flow rate: 0.7 ml/min; 75 ml; ~107 min) (**Figure 1**). The activity was eluted in general at a maximum conductivity of 43 mS cm^{-1} (≈65% HEB). Due to limited protein quantities caused by low expression levels, CsTPS2c was enriched by affinity chromatography using Talon purification (BD Biosciences) according to the manufacturer's instructions, with HWB (wash/binding buffer) and HEB (elution buffer) (**Figure 1**). The terpene synthase fractions were combined and desalted using a PD-10-column (Amersham) equilibrated with 5 column volumes of TPS-assay-buffer (10 mM MOPSO pH 7.0, 1 mM DTT, 20 mM $MgCl_2$, 10% [v/v] glycerol) (**2.11.3.**, **2.11.4.**). Protein concentration was determined (**2.5.2.**), and aliquots of the purification were checked on an SDS-PAGE (**Figure 2** +**3**).

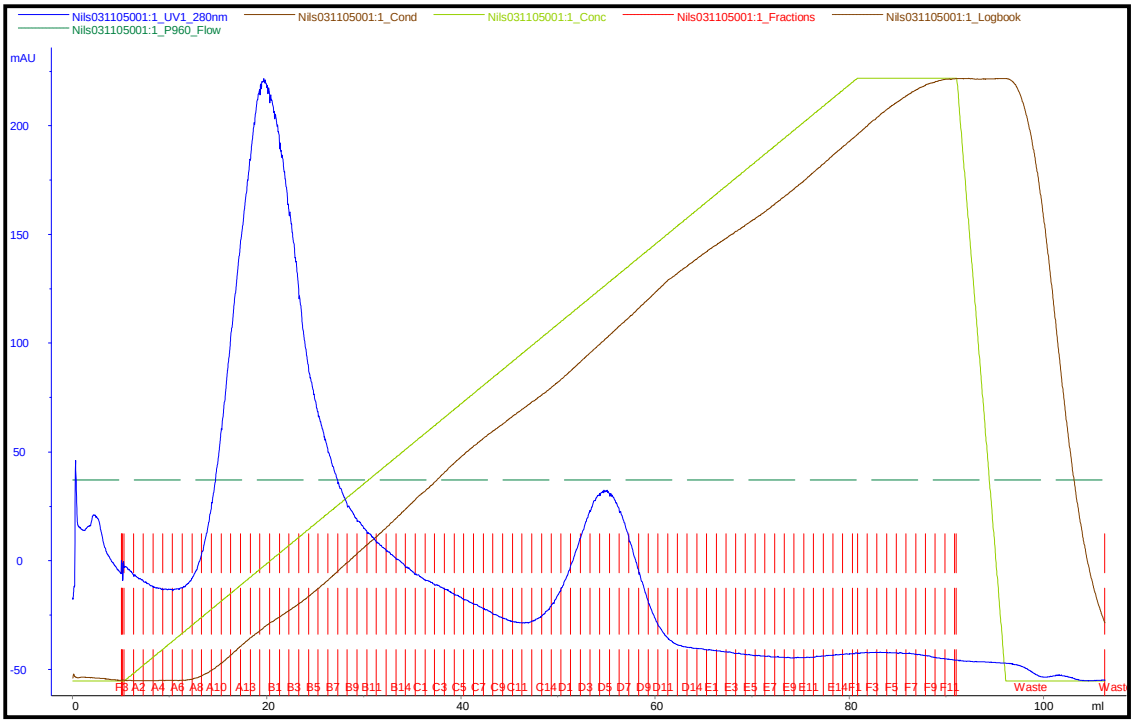


Figure 1: CsTPS1c FPLC-purification-chromatogram, the first peak corresponds to unspecifically bound proteins, the second peak corresponds to the enriched CsTPS1c. CsTPS1c is eluted at a maximum conductivity of 43 mS cm^{-1} (≈65% HEB).

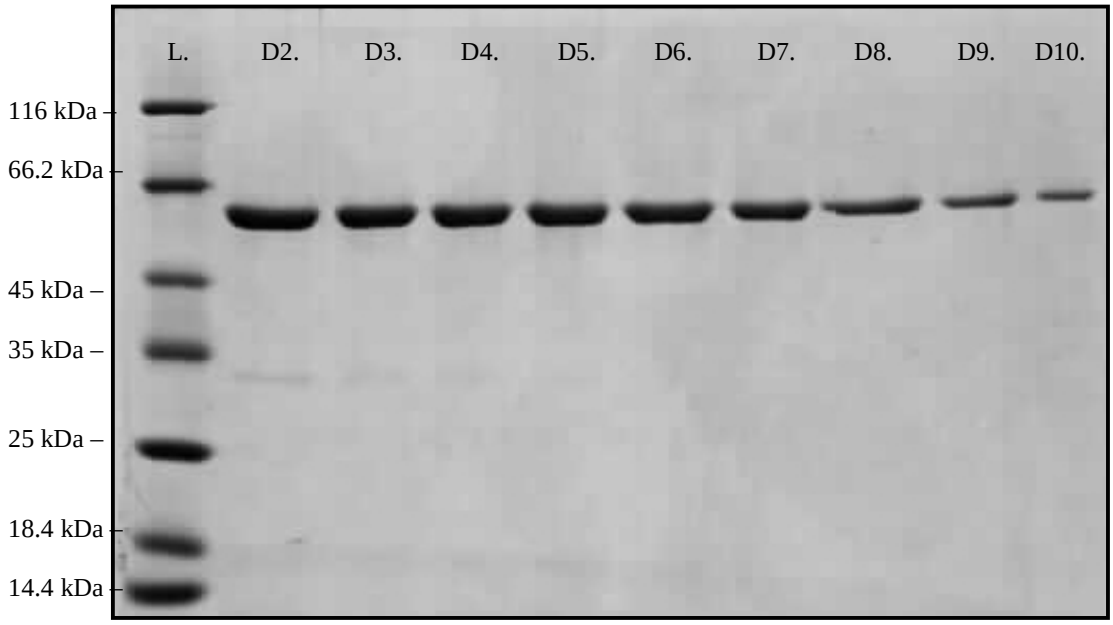


Figure 2: CsTPS1c FPLC-purification-SDS-PAGE, [L]adder and Fractions D2-D10 of the FPLC-purification (**Figure 1**) show an enriched and clean CsTPS1c.

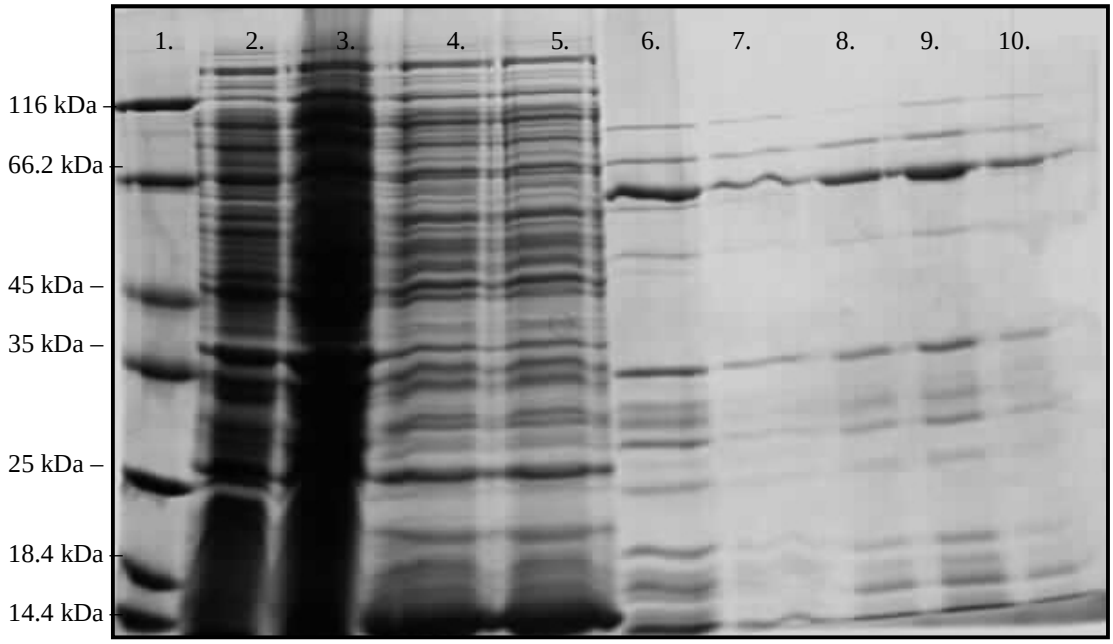


Figure 3: SDS-PAGE of CsTPS2c purification over a Talon^{TM}column: 1. ladder; 2. pre-induction; 3. post-induction (overnight); 4. crude lysate; 5. Talon^{TM}supernatant; 6.+7. Talon^{TM}fraction 2 and 3; 8.,9.+10. PD10column desalting fractions 4-6.

3.1.2. Characterization of CsTPS1 and CsTPS2

The monoterpene synthase activity assay was performed in a final volume of 500 µl TPS-assay-buffer including 10 µM of GPP and 1-^{3}H-GPP (10-20 Ci/mmol, 5 pmol/assay, ARC), overlaid with 500 µl hexane to trap the volatile products, in siliconized tubes (Sigma-Aldrich). After incubation at 30°C at various time points within the previously determined linear range (1 min; 4min; 8min; 12 min; 16 min and 20 min), the mixture (1.25 µg CsTPS1c or CsTPS2c per 500 µl) was mixed vigorously for ~10 sec. and centrifuged 30 sec. at 16,100 x g, the reaction was stopped on ice and 300 µl of the hexane phase were added to 4 ml Ultima gold MV scintillation cocktail and counted in a Beckman LS 6000 TA scintillation counter. For each time point, a boiled enzyme control was included to determine the background value of the assay (**2.12.1**).

3.1.2.1. pH-Optimum of the CsTPSs

The pH has an important effect on the activity of enzymes. The acidity of a solution alters the charge of functional groups from different amino acids, which can lead to a conformational change of the structure and the active site of enzymes, influencing the activity. For CsTPS1c the pH optimum was measured between pH 4.0 and pH 8.5 in 0.5 pH steps with TPS-assay-buffer containing sodium acetate, MES, Bis-Tris, sodium acetate or MOPSO (at 10 mM concentration), the pH optimum was determined at pH 6.5, with a half-maximal velocity at pH 5.7 and 7.4 (**Figure 4**), and a predicted isoelectric point (pI) of 6.7 (DNAstar Inc.). For CsTPS2c the pH optimum was measured between pH 5.5 and pH 8.5 in 0.5 pH steps, here with TPS-assay-buffer containing only MES and MOPSO, because of the now known range from CsTPS1c. The pH optimum was determined at pH 7.0, with a half-maximal velocity at pH 5.9 and 7.5 (**Figure 5**), and a predicted pI of 6.1 (DNAstar Inc.).

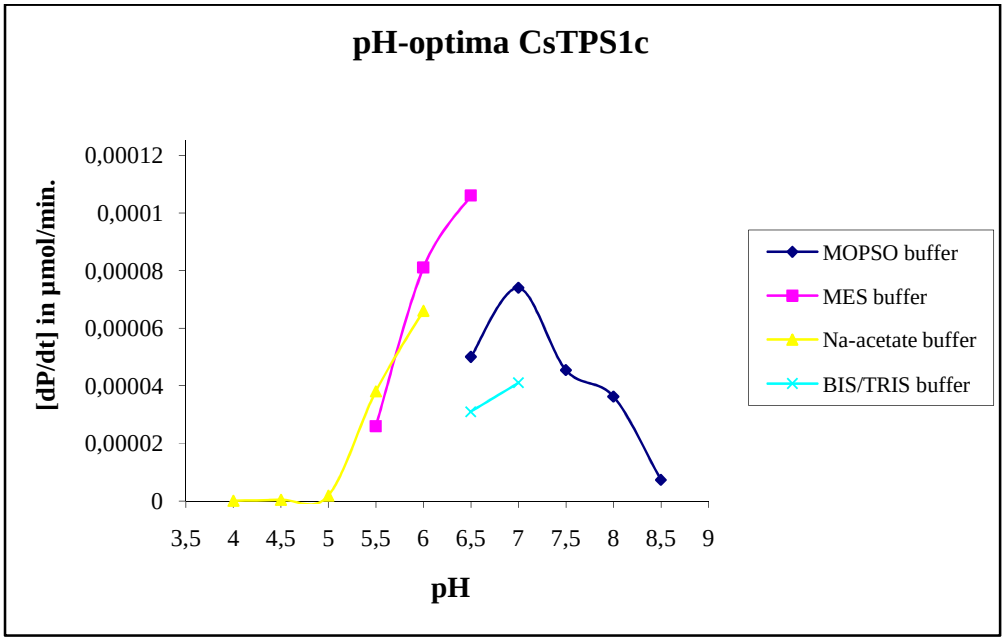


Figure 4: pH-Optimum determination of CsTPS1c at pH 6.5 using MOPSO, MES, Na-acetate and BIS/Tris as buffers.

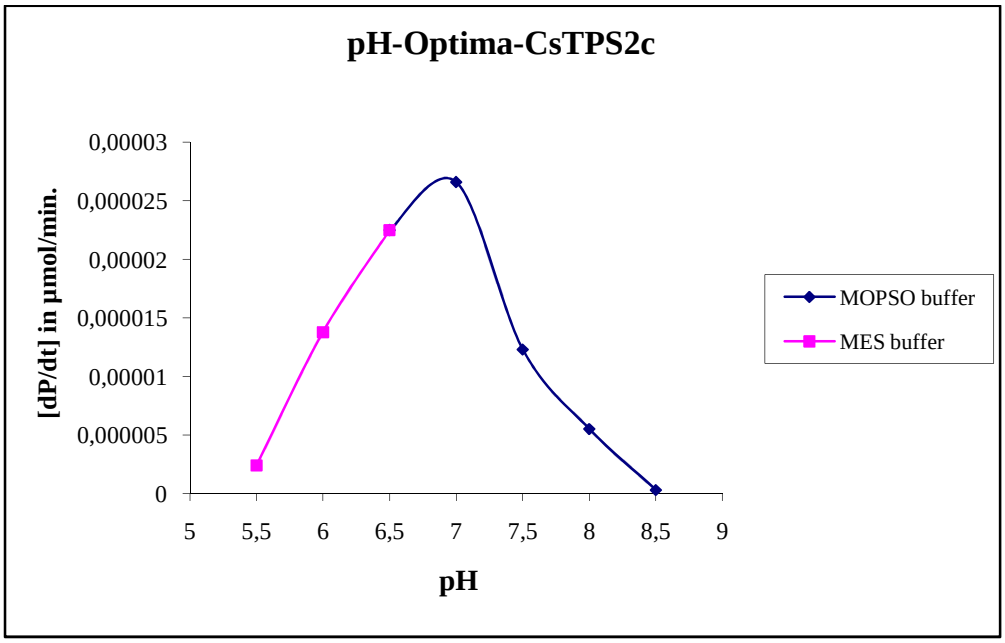


Figure 5: pH-Optimum determination of CsTPS2c at pH 7.0 using MOPSO and MES as buffers.

3.1.2.2. Temperature optimum of the CsTPSs

Each enzyme shows a temperature activity dependency. With increasing temperature the activity of the enzyme increases as well. Until characteristic optima, at which point the activity decreases. The temperature dependence measurement was performed in the TPS-assay-buffer containing MOPSO, between 10-60°C for CsTPS1c, and between 10-50°C for CsTPS2c in 10°C steps. For CsTPS1c the temperature optimum was determined at 40°C, with a half-maximal velocity at ca. 14.5°C and 47.5°C. For CsTPS2c the temperature optimum was determined at 30°C, with a half-maximal velocity at ca 21°C and 42.5°C (**Figure 6**).

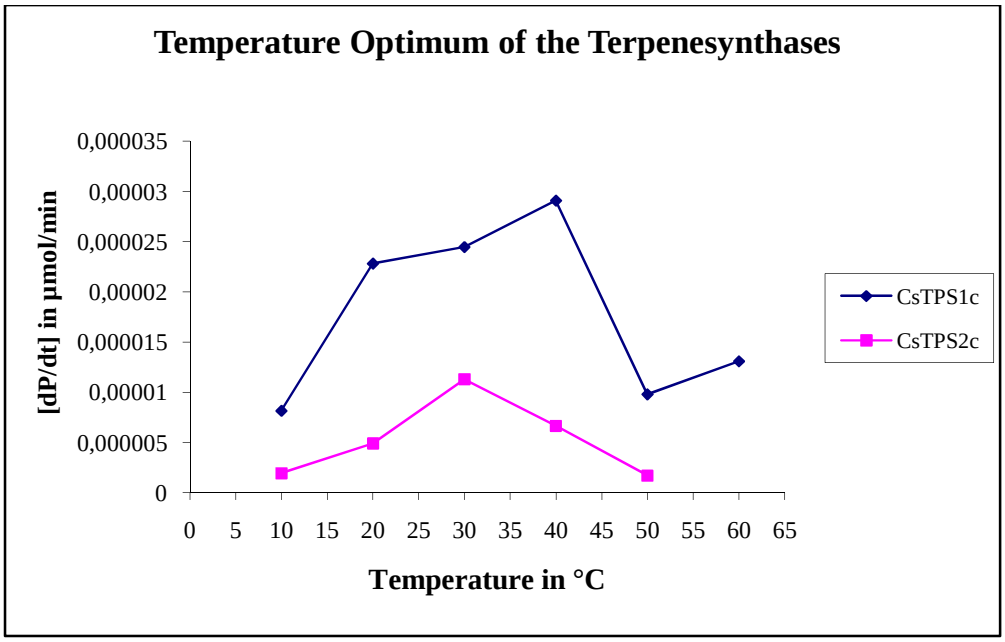


Figure 6: Temperature optimum of CsTPS1 at 40°C and CsTPS2 at 30°C.

3.1.2.3. K_m and V_{max} of the CsTPSs

For the determination of the K_m and V_{max} the standard TPS-assay-buffer with 10 mM MOPSO at pH 7.0 was used. The final substrate concentrations were 0.5; 1.0; 2.5; 5.0; 7.5; 10.0; 12.5; 15.0; 20.0; 25.0 µM GPP respectively. At some concentrations the arithmetic mean of several triplicates or the initial reaction rate of an individual triplicate were taken. For CsTPS1c (**Figure 7**), the initial reaction rates were plotted against the concentrations of GPP. A hyperbolic fit according to the Michaelis-Menten-mechanism (shown as a dashed line) yielded a V_{max} value of 0.12 ± 0.01 µmol min^{-1} mg^{-1} and a K_m of 11.76 ±

2.45 µM, resulting in a k_{cat} value of 0.13 s^{-1} and a catalytic efficiency (k_{cat}/K_m) of 1.1 × 10^4 s^{-1} M^{-1}, respectively. Although the hyperbolic fit yielded reasonable kinetic parameters, the visual inspection of the data suggests a rather sigmoid curve progression. This assumption was supported by various linearizations' procedures. A double reciprocal plot revealed a strong deviation from linearity, even if the values at the lowest and, hence most error prone substrate concentrations are omitted (not shown). Moreover, in a Hanes-Woolfe plot of the data a systematic deviation from linearity was obvious (insert **Figure 7**). A sigmoid behavior, regardless of the exact mechanism, implies the presence of a second substrate-binding site. The modeled structure of CsTPS1c revealed, however, no such allosteric binding site [98]. But CsTPS1c adopts a dimeric state in solution, as was shown by gel filtration experiments (**3.1.3.4.**). Thus, an activation mechanism involving two subunits seems conceivable, in terms of a positive cooperativity. A regression according to the Hill mechanism was applied to the data (**Figure 7**, solid line). This treatment yielded a better fit (in visual inspection and regression statistics by evaluation software; SigmaPlot 8.0, Systat Software Inc. Chicago), resulting in a V_{max} value of 0.08 ± 0.003 µmol min^{-1} mg^{-1}, a half saturation concentration of 6.25 ± 0.41 µM, resulting in an k_{cat} of 0.09 s^{-1}, a apparent value of k_{cat}/K_m 1.5× 10^4 s^{-1} M^{-1} and a Hill coefficient of 1.7. The initial reaction rates of CsTPS2c plotted against the concentrations of GPP are presented in (**Figure 8**). Although the apparent deviations from Michaelis-Menten-behavior seem less pronounced (and at least to a small degree due to data scattering) than in the case of CsTPS2c, the same considerations regarding the sigmoid curve progression and an appropriate substrate activation mechanism hold true. The hyperbolic fit according to Michaelis-Menten yielded a V_{max} value of (0.16 ± 0.02) µmol min^{-1} mg^{-1} and a K_m of 7.91 ± 2.23 µM, resulting in a k_{cat} value of 0.19 s^{-1} and a k_{cat}/K_m of 2.4 × 10^4 s^{-1} M^{-1}, respectively. Alternatively, a sigmoid regression according to Hill resulted in a V_{max} value of 0.13 ± 0.01 µmol min^{-1} mg^{-1}, a half saturation concentration of 4.96 ± 0.38 µM, resulting in an k_{cat} of 0.14 s^{-1}, a apparent value of k_{cat}/K_m 2.9× 10^4 s^{-1} M^{-1} and a Hill coefficient of 1.8. Generally, GPP seems to be a slightly better substrate for CsTPS2c than for CsTPS1c.

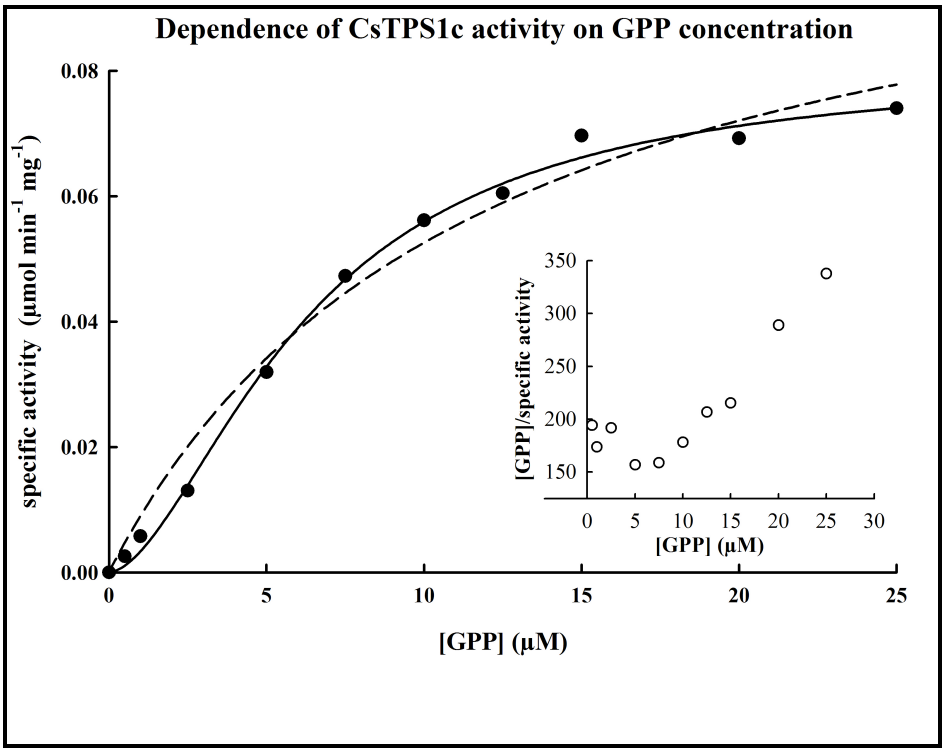


Figure 7: v/[S] characteristic of CsTPS1c including a Hanes-Woolfe plot (small figure insert).

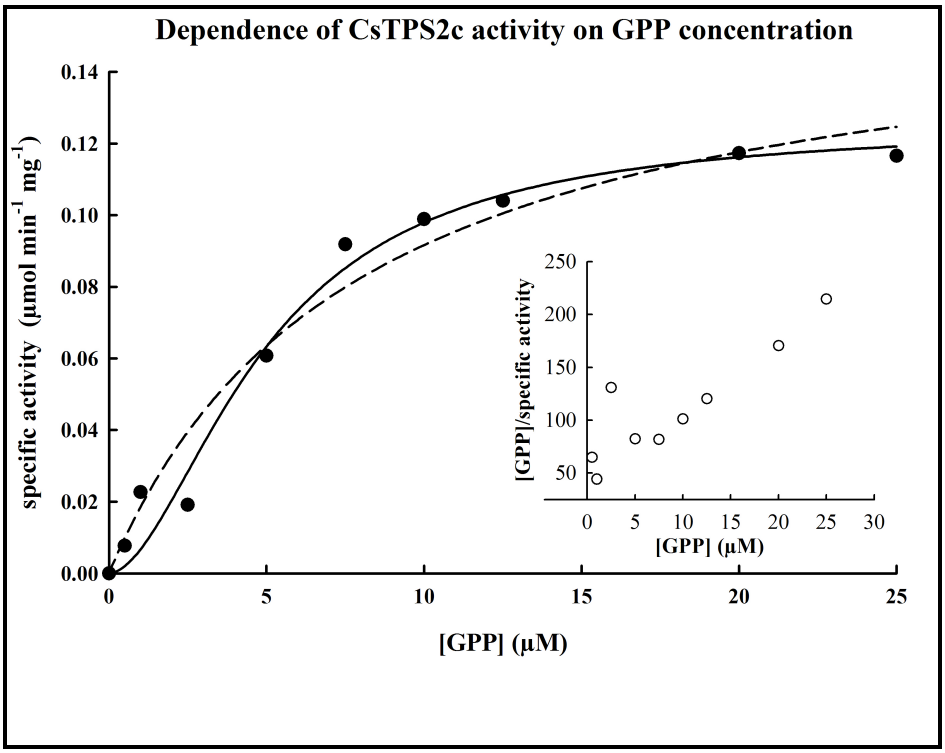


Figure 8: v/[S] characteristic of CsTPS2c including a Hanes-Woolfe plot (small figure insert).

3.1.2.4. Size exclusion chromatography (SEC) of the CsTPSs

A SEC is used to estimate the molecular mass and the hydrodynamic or Stokes radius of a protein and to distinguish between monomeric and multimeric enzymes (dimer, trimer etc.). ÄKTA-FPLC-system (Explorer) with a HiLoad Superdex 200 16/60 (GE Healthcare) column was equilibrated with the TPS-Buffer containing 150 mM NaCl and 5% glycerol instead of 10% glycerol. To calibrate the ÄKTA-FPLC-system calibration proteins with known molecular mass and Stokes radius where used: aldolase (158 kDa; 4.81 nm) and ovalbumin (45 kDa; 3.05 nm) (**Figure 9**), bovine serum albumin (BSA; 67 kDa; 3.55 nm) and catalase (240 kDa; 5.22 nm) (**Figure 10**) as well as ferritin (450 kDa monomer; 6.1 nm) (**Figure 11**) [99, 100]. The elution volumes of the calibration proteins were 45 ml (trimeric ferritin), and 56 ml (monomeric ferritin), 77 ml (BSA), 66 ml (catalase), 69 ml (aldolase) and 83 ml (ovalbumin) (**2.11.5.**). Ferritin appeared in 3 different forms (69.4% monomer, 19.5% dimer and 11.1% trimer), and most likely the elution peaks corresponds to the monomer and trimer (1350 kDa), which is supported by the calibration curve (**Figure 12**) [101]. Dimer did not appear in this run under the buffer conditions used here or is represented under the peak area [96]. The Stokes radius was determined using the distribution coefficient. The distribution coefficient is determined by $K_{av} = (V_e-V_o)/(V_t-V_0)$ (V_0 = void volume; V_e = elution volume and V_t = total volume) (**Figure 13**) [102].

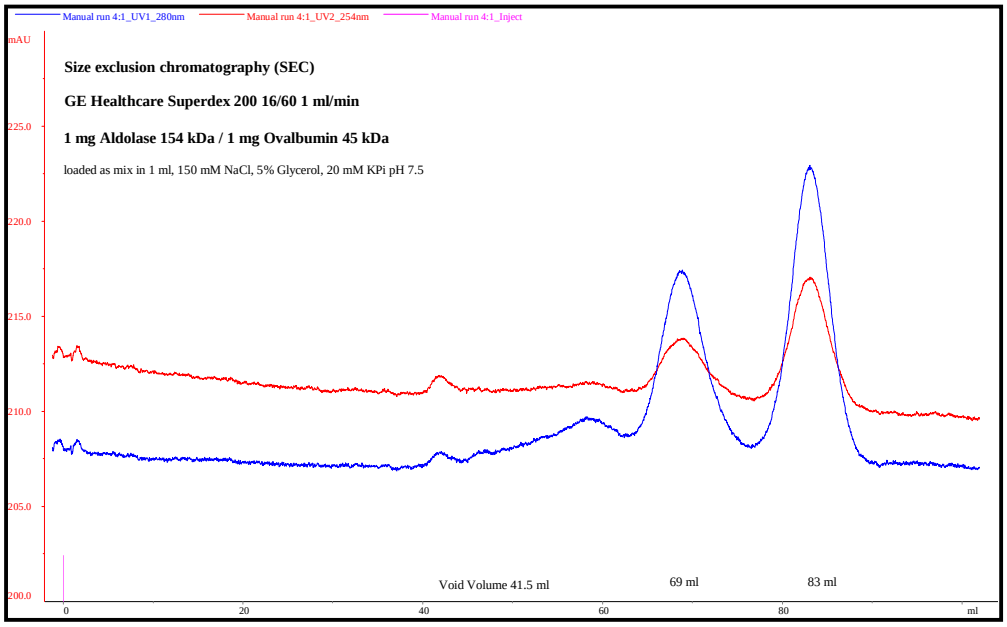


Figure 9: Calibration of the SEC-column with aldolase and ovalbumin as calibration proteins.

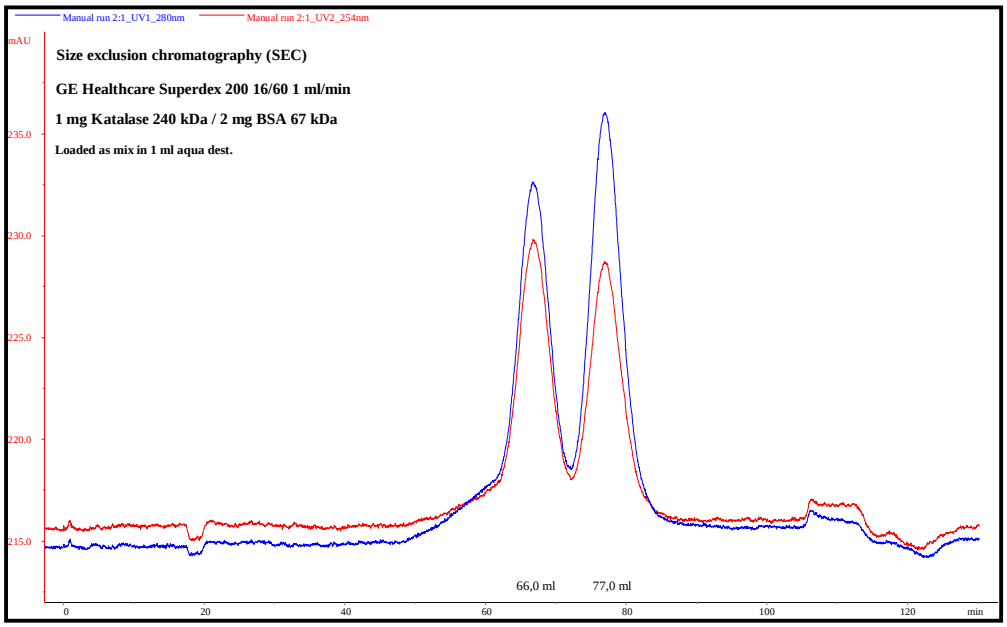


Figure 10: Calibration of the SEC-column with catalase and BSA as calibration proteins.

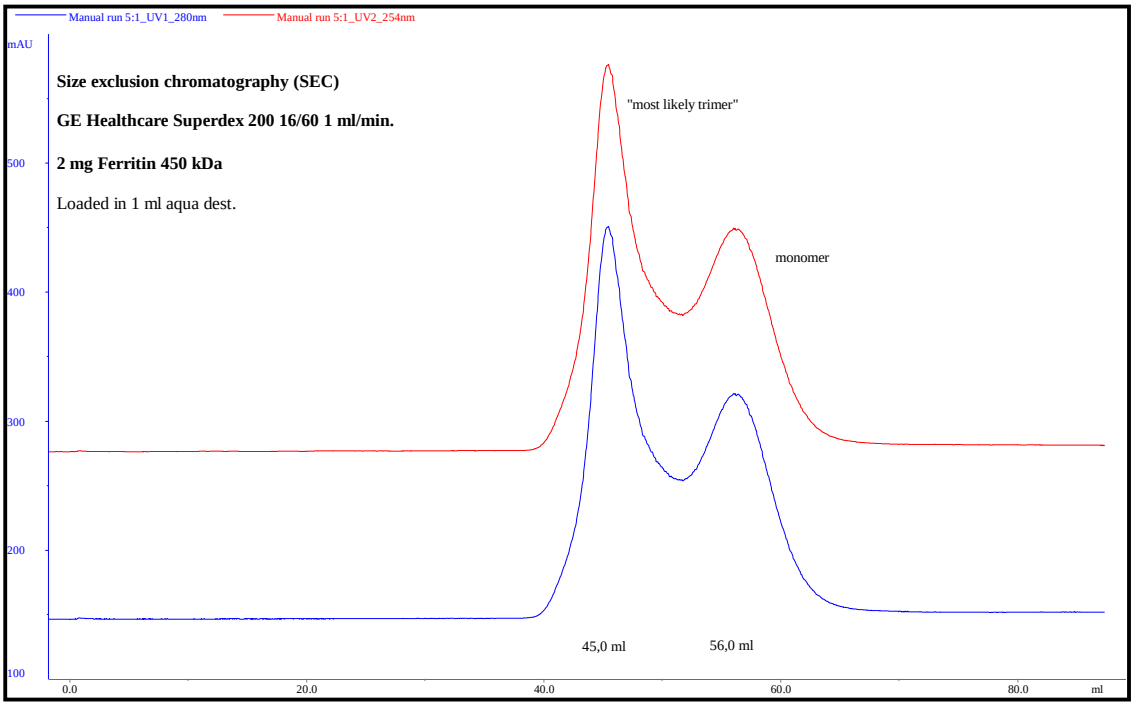


Figure 11: Calibration of the SEC-column with ferritin as calibration protein.

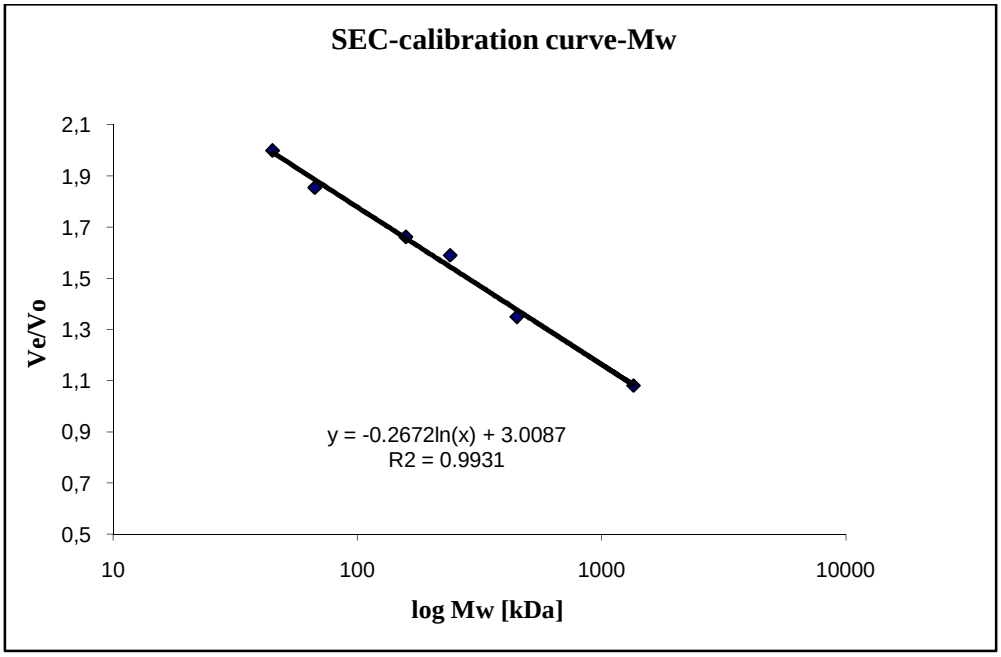


Figure 12: SEC-calibration curve for the determination of the molecular mass (M_r), using ovalbumin (45 kDa); BSA (67 kDa); aldolase (158 kDa); catalase (240 kDa) and ferritin$_{mon.}$ (450 kDa) and Ferritin$_{trim.}$ (1350 kDa).

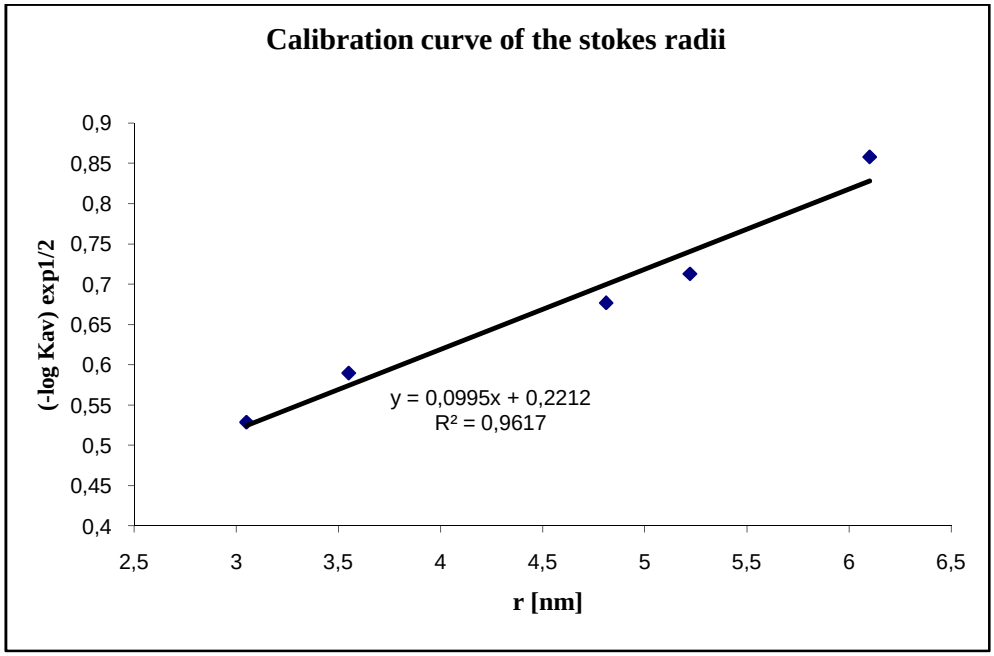


Figure 13: SEC-calibration curve for the determination of stokes or hydrodynamic radius-, using ovalbumin (3.05 nm); BSA (3.55 nm); aldolase (4.81 nm); catalase (5.22 nm) and monomeric ferritin (6.1 nm).

CsTPS1c and CsTPS2c have a calculated molecular mass of 70.03 kDa respectively 69.29 kDa (DNAstar Inc.). They elute both by a maximum at ≈70 ml (69.941 ml; 69.995 ml) (**Figure 14** + **16**), checked by SDS-PAGE (**Figure 15** + **17**), according to the calibration curve, they would have a molecular mass of ≈154 kDa which leads to the speculation that in vitro and under the buffer conditions used here both enzymes seem to be dimers, with the Stokes radius of 4.47 nm (**Table 1**).

Protein	M_r [kDa]	Stokes radius [nm]	V_e [ml]	V_e/V_o	$(-\log K_{av})^{1/2}$
Ovalbumin	45	3.05	83	2.0	0.529
BSA	67	3.55	77	1.588	0.590
Aldolase	158	4.81	69	1.622	0.677
Catalase	240	5.22	66	1.59	0.713
Ferritin$_{mon.}$	440	6.10	56	1.349	0.858
Ferritin$_{trim.}$	1350	9.46	45	1.08	1.163
CsTPS1c	~154	4.47	70	1.698	0.666
CsTPS2c	~154	4.47	70	1.698	0.666

Table 1: Determination of the calculated stokes radius and molecular weight of the *C. sativa* L. monoterpene synthases. In the table are the calibration proteins with their molecular weight (kDa), Stokes radius, elution volume (V_e) and the determined parameters V_e/V_0 and $(-\log K_{av})^{1/2}$. The possible stokes radius of the ferritin trimer is determined by the calibration curve.

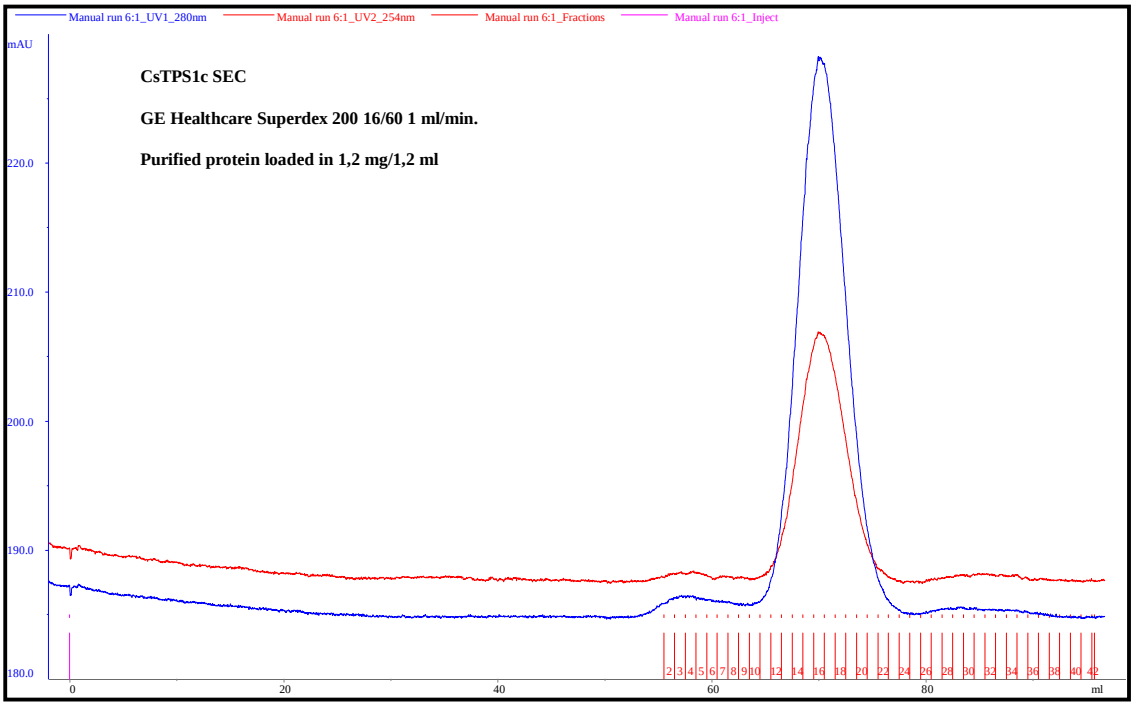


Figure 14: SEC of CsTPS1c.

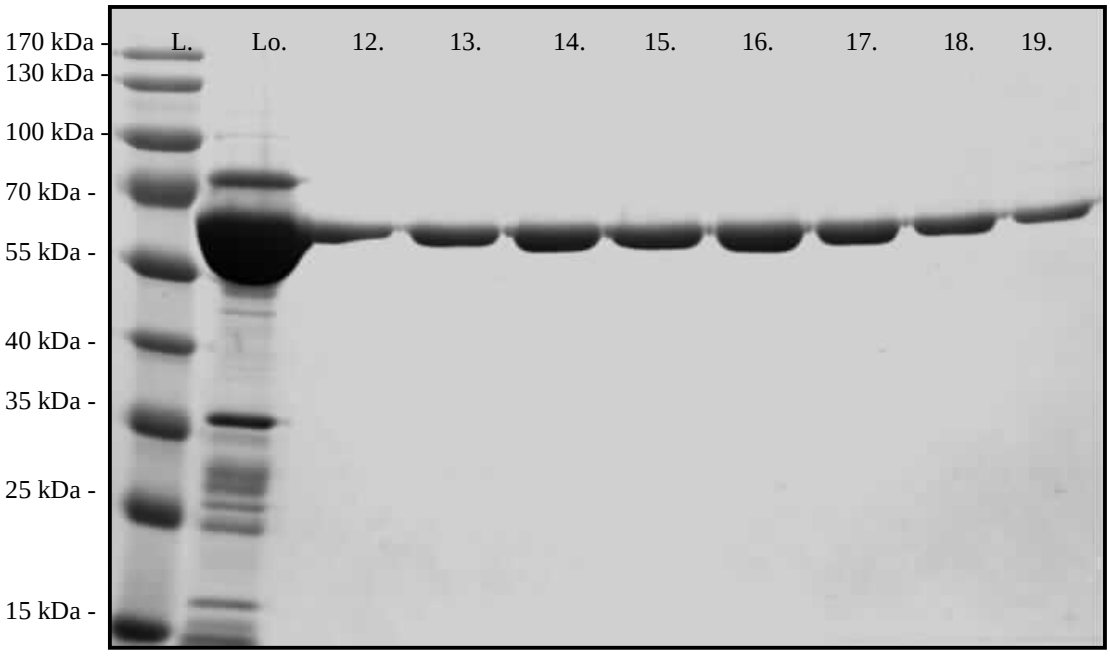


Figure 15: [L]adder [Lo]ad and fractions 12-19 of the CsTPS1c SEC (**Figure 14**)

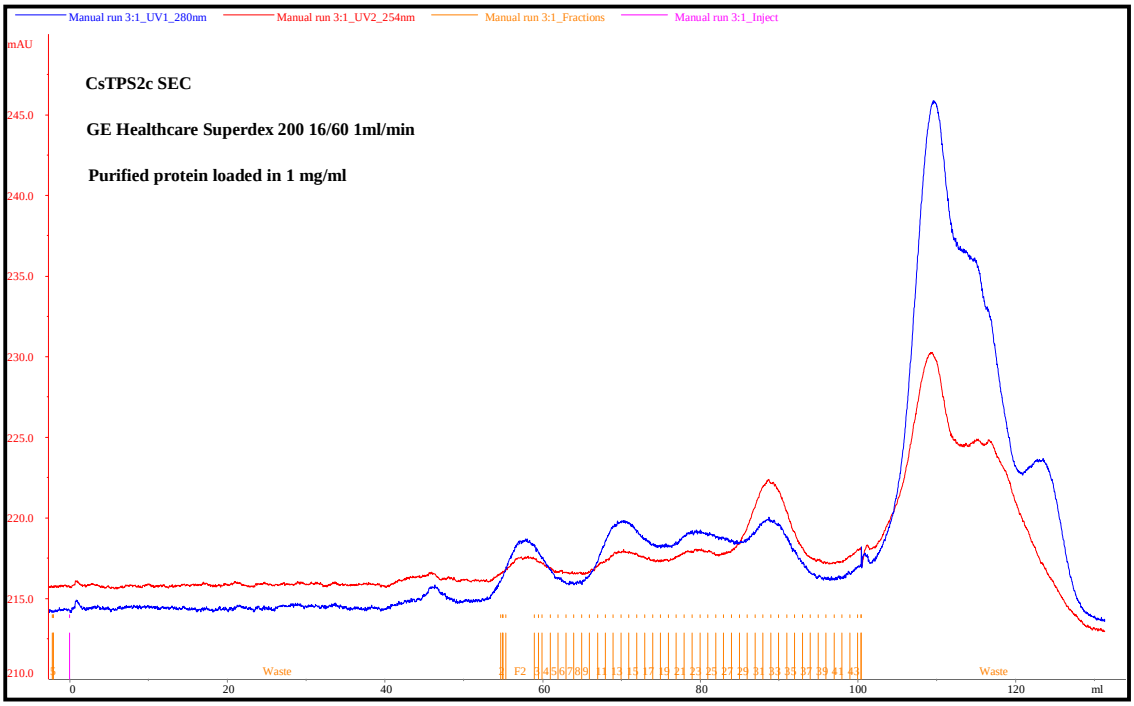


Figure 16: SEC of CsTPS2c.

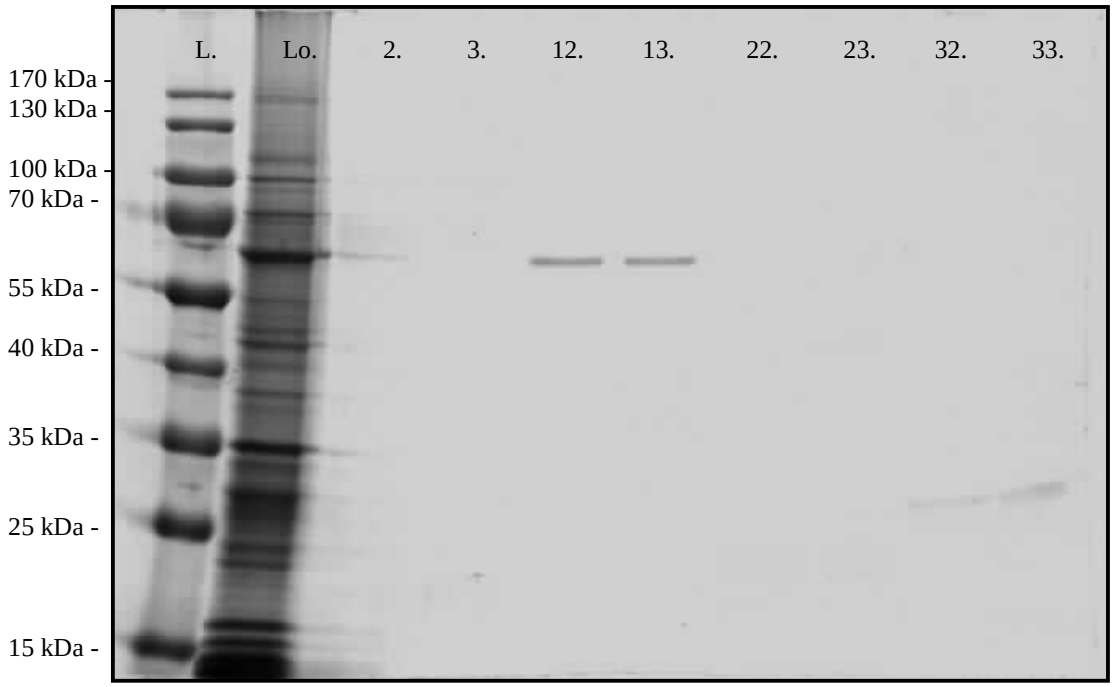


Figure 17: [L]adder, [Lo]ad and fractions 2+3, 12+13, 22+23 and 32+33 of the CsTPS2c SEC (**Figure 16**).

3.1.3. Phylogenetic (identity tree) analysis of the CsTPSs

The TPSs have a high homology (http://www.ncbi.nih.gov/ ; blastx) to both (-)-α-terpineol synthases of *Vitis vinifera* and to γ-terpinene and β-pinene synthases of *Citrus limon* and *Citrus unshiu*. On the basis of this, an identity tree (**Figure 18**) was prepared with 37 aa sequences, including CsTPS1 and CsTPS2, from gymnosperm and angiosperm monoterpene-; sesquiterpene-, and diterpene synthases. ClustalW (http://www.ebi.ac.uk/clustalw/) in fasta format and GeneBee in the Clastw .phb format file (http://www.genebee.msu.su/services/phtree_full.html) were used. The conditions were: scale-random; algorithm-cluster and topological; matrix-blosum 62; with bootstrap values. Amino acid sequence relatedness of plant terpenoid synthases allowed subdivision of the *Tps* gene family into six subfamilies, designated *Tpsa* through *Tpsf*, each distinguished by sharing a minimum of 40% identity among members. Each CsTPS placed within the *Tpsb* subfamily of terpene synthases alongside other diverse monoterpenoid synthases from angiosperms [61]. The new *Tpsg* familiy, that represents monoterpene synthases lacks the $RR(X_8)W$-motif, and could have within the CTS instead a $XXR(X_7)W$-motif, is not represented in this identity tree [103]

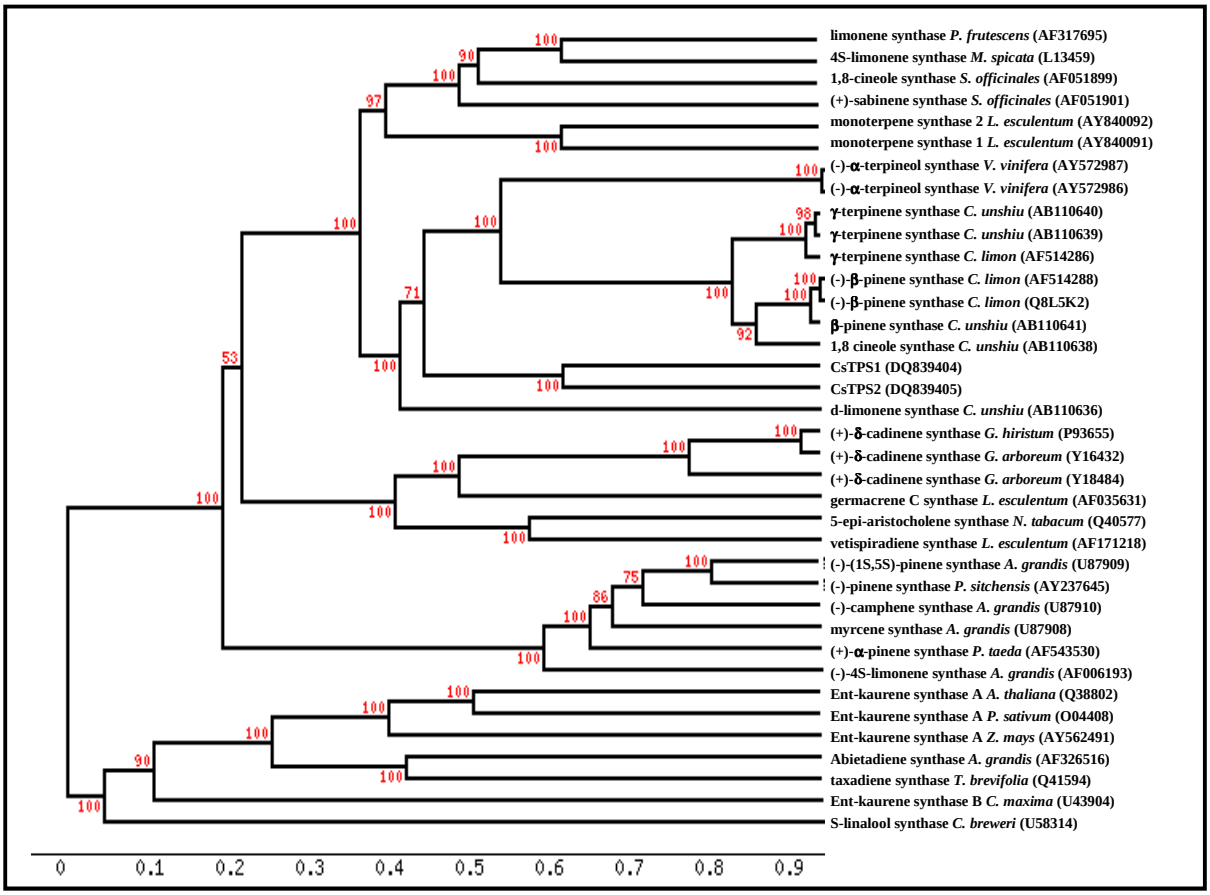


Figure 18: Identity tree analysis of 37 aa sequences of monoterpene-; sesquiterpene-; and diterpene synthases from gymnosperms and angiosperms using ClustalW and GeneBee (algorithm-cluster and topological; matrix-blosum 62) with bootstrap values.

3.1.4. Graphical summary of the CsTPSs

Parameter	CsTPS1	CsTPS2
size (nt)/truncated version	1.896 kb/1.785 kb	1.848 kb/1.78 kb
size (M_w)/truncated version	72.383 kDa/70,03 kDa (CsTPS1c)	71.84 kDa /69,29 kDa (CsTPS2c)
size (aa)/truncated version	622/601	615/593
pH-optimum	6.5	7.0
isoelectric point (pI)	6.7	6.1
temperature-optimum	40°C	30°C
K_m	6.25 ± 0.41 µM	4.96 ± 0.38 µM
V_{max}	0.08 ± 0.003 µmol min^{-1} mg^{-1}	0.13 ± 0.01 µmol min^{-1} mg^{-1}
k_{cat}	0.09 s^{-1}	0.14 s^{-1}
k_{cat}/K_m	1.5 10^4 s^{-1} M^{-1}	2.9 10^4 s^{-1} M^{-1}
Hill coefficient	1.8	1.7
3D structure	most likely dimer	most likely dimer
stokes radius	4.47 nm	4.47 nm
phylogenetic subfamily	*Tpsb*	*Tpsb*

Table 2: Graphical summary of the CsTPS's.

3.2. *Salvia sclarea* L.

3.2.1. EST Sequencing of the putative diterpene synthase

Purified mRNA (**2.4.3.**) isolated from *S. sclarea* trichomes was reverse transcribed and used for construction of a directional cDNA library following the instructions for Clontech SMART cDNA synthesis and introduction into pTriplEx2 with an ABI Prism Big-Dye Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems) with the T7 primer. Sequencing reactions were run after excess dye removal on the ABI Prism 3100-*Avant* Genetic Analyzer Sequencer (**2.7.**). After trimming for vector, the sequences were clustered and assembled using the SeqManII (DNAstar Inc.) application using a minimum match size of 20 and a minimum match percentage of 85. A total amount of 53 putative EST's involved in the terpenoid metabolism were found. One candidate a putative diterpene synthase with a high homology to copalyl pyrophosphate synthases (from here on named in this manner) was taken for further RACE-PCR trials (all this done by Kum-Boo Choi).

3.2.2. RACE-PCR of the putative diterpene synthase

The putative copalyl pyrophosphate RACE clones of Kum-Boo Choi were resequenced with the primers (**2.1.2.2.**) SP6, T7, 5'RACE-K1-SP6, 5'RACE-K2-SP6, 5'RACE-K3-SP6, and 5'RACE-K4-SP6 an insert of ≈1.9 kb was found. The final RACE was then done with 5'RACECop1antisense, 5'RACECop2antisense, and 3RACECop1sense primer, the latter two primers were sufficient to get the remaining part of the full-length clone. 3'RACECop1sense primer RACE-PCR showed bands of 0.5 kb, 1.1 kb and 1.2 kb, and 5'RACECop2antisense primer RACE-PCR showed a band of 0.6 kb (**2.6.1.**). The bands were gel-purified (**2.4.2.**), taq-tailed, and cloned into pGEM T-Easy (**2.8.2.**), and then transformed into DH5α *E. coli* hosts (**2.2.2.**). The 0.5 kb fragment of the 3'RACE and the 0.6 kb-fragment of the 5'RACE show after sequencing the putative 3' and 5' region of the copalylpyrophosphate synthase (**2.7.**). A sequence of 2.756 bp was found with an ORF of 2.358 bp coding 785 aa (Figure). The predicted molecular weight was 89.758 kDa with a pI at 5.708 (DNAstar Inc.). Several PCR attempts (5'CopfullTOPO-2, 3'CopfullTOPO-stop, 3'CopfullTOPOns, 3'CopfullTOPOns-2) to clone the full-length clone out of a cDNA with a *Pfu* polymerase were unsuccessful, until an Advantage 2 polymerase (mix-

ture of a proofreading- and a *Taq*-polymerase) gradient PCR obtained the right fragment size (**2.6.**). The fragment was gel-purified (**2.4.2.**) and the sample was blunted (**2.8.3.**) for a TOPO cloning into the pET101 D-TOPO vector (**2.8.1.**) and transformed into TOP10 cells (**2.2.2.**). In general a miniculture PCR (**2.6.**) was done to verify the correctness of the transformation; positives were then chosen for a plasmid preparation (**2.4.1.**) with a following sequencing reaction (**2.7.**). The full-length clone (5'CopfullTOPO-2 x 3'CopfullTOPOns-2) was used in an expressiontest, but no expression was detectable (**2.11.**). In the literature it is often described that the possible CTS is bound by *E. coli* chaperons resulting in inclusion body formation or the protein folding was altered. Due to this additional clones were designed to truncate the CTS from its original sequence. To determine the truncation site, transit sequence prediction programs from the Internet were used (http://www.cbs.dtu.dk/services/; ChloroP, TargetP, SignalP and TatP and http://hc.ims.u-tokyo.ac.jp/iPSORT/) [61, 70]. All of the constructs (5'CopfullTOPO+RR x 3'CopfullTOPOns-2, 5'CopfullTOPO-RR x 3'CopfullTOPOns-2, 5'CopfullTOPO-W x 3'CopfullTOPOns-2, 5'CopfullTOPO-SP x 3'CopfullTOPOns-2) were cloned into pET101 D-TOPO (**2.8.1.**), containing a C-terminal His-tag, and transformed into BL21(DE3)RIL (**2.2.2.**), in neither case a expression was detectable (**2.11.**). A clone with predicted CTS from the IPSORT program, preceding the putative RR-motif, was designed and cloned into pET101 D-TOPO (5'CopTOPOipsort x 3'CopfullTOPOns-2) and pET100 D-TOPO (5'CopTOPOipsort x 3'CopfullTOPO-stop). Only the clone in the latter vector with an N-terminal His-tag showed an inducable expression and was suitable for purification over a Talon column. The start of the coding region of 5'CopTOPOipsort is shown underlined in **Figure 19**. The calculated molecular weight was 88.929 kDa with a pI at 5.647; the expressable clone with pET100 D-TOPO vector sequence had a molecular weight of 92.94 kDa with a pI at 5.794 (DNAstar Inc.).

```
1     ATGACTTCTGTAAATTTGAGCAGAGCACCAGCAGCGATTACCCGGCGCAGGCTGCAGCTA   60
  1   M  T  S  V  N  L  S  R  A  P  A  A  I  T  R  R  R  L  Q  L   20
61    CAGCCGGAATTTCATGCCGAGTGTTCATGGCTGAAAAGCAGCAGCAAACACGCGCCCTTG  120
 21   Q  P  E  F  H  A  E  C  S  W  L  K  S  S  S  K  H  A  P  L   40
121   ACCTTGAGTTGCCAAATCCGTCCTAAGCAACTCTCCCAAATAGCTGAATTGAGAGTAACA  180
 41   T  L  S  C  Q  I  R  P  K  Q  L  S  Q  I  A  E  L  R  V  T   60
181   AGCCTGGATGCGTCGCAAGCGAGTGAAAAAGACATTTCCCTTGTTCAAACTCCGCATAAG  240
 61   S  L  D  A  S  Q  A  S  E  K  D  I  S  L  V  Q  T  P  H  K   80
241   GTTGAGGTTAATGAAAAGATCGAGGAGTCAATCGAGTACGTCCAAAATCTGTTGATGACG  300
```

```
  81  V  E  V  N  E  K  I  E  E  S  I  E  Y  V  Q  N  L  L  M  T  100
301   TCGGGCGACGGGCGAATAAGCGTGTCACCCTATGACACGGCAGTGATCGCCCTGATCAAG  360
  101 S  G  D  G  R  I  S  V  S  P  Y  D  T  A  V  I  A  L  I  K 120
361   GACTTGAAAGGGCGCGACGCCCCGCAGTTTCCGTCATGTCTCGAGTGGATCGCGCACCAC  420
  121 D  L  K  G  R  D  A  P  Q  F  P  S  C  L  E  W  I  A  H  H  140
421   CAACTGGCTGATGGCTCATGGGGCGACGAATTCTTCTGTATTTATGATCGGATTCTAAAT  480
  141 Q  L  A  D  G  S  W  G  D  E  F  F  C  I  Y  D  R  I  L  N  160
481   ACATTGGCATGTGTCGTAGCCTTGAAATCATGGAACCTTCACTCTGATATTATTGAAAAA  540
  161 T  L  A  C  V  V  A  L  K  S  W  N  L  H  S  D  I  I  E  K  180
541   GGAGTGACGTACATCAAGGAGAATGTGCATAAACTTAAAGGTGCAAATGTTGAGCACAGG  600
  181 G  V  T  Y  I  K  E  N  V  H  K  L  K  G  A  N  V  E  H  R  200
601   ACAGCGGGGTTCGAACTTGTGGTTCCTACTTTTATGCAAATGGCCACAGATTTGGGCATC  660
  201 T  A  G  F  E  L  V  V  P  T  F  M  Q  M  A  T  D  L  G  I  220
661   CAAGATCTGCCCTATGATCATCCCCTCATCAAGGAGATTGCTGACACAAAACAACAAAGA  720
  221 Q  D  L  P  Y  D  H  P  L  I  K  E  I  A  D  T  K  Q  Q  R  240
721   TTGAAAGAGATACCCAAGGATTTGGTTTACCAAATGCCAACGAATTTACTGTACAGTTTA  780
  241 L  K  E  I  P  K  D  L  V  Y  Q  M  P  T  N  L  L  Y  S  L  260
781   GAAGGGTTAGGAGATTTGGAGTGGGAAAGGCTACTGAAACTGCAGTCGGGCAATGGCTCC  840
  261 E  G  L  G  D  L  E  W  E  R  L  L  K  L  Q  S  G  N  G  S  280
841   TTCCTCACTTCGCCGTCGTCCACCGCCGCCGTCTTGATGCATACCAAAGATGAAAAATGT  900
  281 F  L  T  S  P  S  S  T  A  A  V  L  M  H  T  K  D  E  K  C  300
901   TTGAAATACATCGAAAACGCCCTCAAGAATTGCGACGGAGGAGCACCACATACTTATCCA  960
  301 L  K  Y  I  E  N  A  L  K  N  C  D  G  G  A  P  H  T  Y  P  320
961   GTCGATATCTTCTCAAGACTTTGGGCAATCGATAGGCTACAACGCCTAGGAATTTCTCGT 1020
  321 V  D  I  F  S  R  L  W  A  I  D  R  L  Q  R  L  G  I  S  R  340
1021  TTCTTCCAGCACGAGATCAAGTATTTCTTAGATCACATCGAAAGCGTTTGGGAGGAGACC 1080
  341 F  F  Q  H  E  I  K  Y  F  L  D  H  I  E  S  V  W  E  E  T  360
1081  GGAGTTTTCAGTGGAAGATATACGAAATTTAGCGATATTGATGACACGTCCATGGGCGTT 1140
  361 G  V  F  S  G  R  Y  T  K  F  S  D  I  D  D  T  S  M  G  V  380
1141  AGGCTTCTCAAAATGCACGGATACGACGTCGATCCAAATGTACTAAAACATTTCAAGCAA 1200
  381 R  L  L  K  M  H  G  Y  D  V  D  P  N  V  L  K  H  F  K  Q  400
1201  CAAGATGGTAAATTTTCCTGCTACATTGGTCAATCGGTCGAGTCTGCATCTCCAATGTAC 1260
  401 Q  D  G  K  F  S  C  Y  I  G  Q  S  V  E  S  A  S  P  M  Y  420
1261  AATCTTTATAGGGCTGCTCAACTAAGATTTCCAGGAGAAGAAGTTCTTGAAGAAGCCACT 1320
  421 N  L  Y  R  A  A  Q  L  R  F  P  G  E  E  V  L  E  E  A  T  440
1321  AAATTTGCCTTTAACTTCTTGCAAGAAATGCTAGTCAAAGATCGACTTCAAGAAAGATGG 1380
  441 K  F  A  F  N  F  L  Q  E  M  L  V  K  D  R  L  Q  E  R  W  460
1381  GTGATATCCGACCACTTATTTGATGAGATAAAGCTGGGGTTGAAGATGCCATGGTACGCC 1440
  461 V  I  S  D  H  L  F  D  E  I  K  L  G  L  K  M  P  W  Y  A  480
1441  ACTCTACCCCGAGTCGAGGCTGCATATTATCTAGACCATTATGCTGGTTCTGGTGATGTA 1500
  481 T  L  P  R  V  E  A  A  Y  Y  L  D  H  Y  A  G  S  G  D  V  500
1501  TGGATTGGCAAGAGTTTCTACAGGATGCCAGAAATCAGCAATGATACATACAAGGAGCTT 1560
  501 W  I  G  K  S  F  Y  R  M  P  E  I  S  N  D  T  Y  K  E  L  520
1561  GCGATATTGGATTTCAACAGATGCCAAACACAACATCAGTTGGAGTGGATCCATATGCAG 1620
  521 A  I  L  D  F  N  R  C  Q  T  Q  H  Q  L  E  W  I  H  M  Q  540
1621  GAATGGTACGACAGATGCAGCCTTAGCGAATTCGGGATAAGCAAAAGAGAGTTGCTTCGC 1680
  541 E  W  Y  D  R  C  S  L  S  E  F  G  I  S  K  R  E  L  L  R  560
1681  TCTTACTTTCTGGCCGCAGCAACCATATTCGAACCGGAGAGAACTCAAGAGAGGCTTCTG 1740
  561 S  Y  F  L  A  A  A  T  I  F  E  P  E  R  T  Q  E  R  L  L  580
1741  TGGGCCAAAACCAGAATTCTTTCTAAGATGATCACTTCATTTGTCAAAATTAGTGGAACA 1800
  581 W  A  K  T  R  I  L  S  K  M  I  T  S  F  V  K  I  S  G  T  600
1801  ACACTATCTTTGGACTACAATTTCAATGGCCTCGATGAAATAATTAGTAGTGCCAATGAA 1860
  601 T  L  S  L  D  Y  N  F  N  G  L  D  E  I  I  S  S  A  N  E  620
1861  GATCAAGGACTGGCTGGGACTCTGCTGGCAACCTTCCATCAACTTCTAGACGGATTCGAT 1920
  621 D  Q  G  L  A  G  T  L  L  A  T  F  H  Q  L  L  D  G  F  D  640
1921  ATATACACTCTCCATCAACTCAAACATGTTTGGAGCCAATGGTTCATGAAAGTGCAGCAA 1980
  641 I  Y  T  L  H  Q  L  K  H  V  W  S  Q  W  F  M  K  V  Q  Q  660
```

```
1981  GGAGAGGGAAGCGGCGGGGATGACGCGGTGCTCCAAGCGAACGCGCTCAACATCTGCGCC 2040
  661 G  E  G  S  G  G  D  D  A  V  L  Q  A  N  A  L  N  I  C  A  680
2041  GGCCTCAACGAAGACGTGTTGTCCAACAACGAATACACGGCTCTGTCCACCCTCACAAAT 2100
  681 G  L  N  E  D  V  L  S  N  N  E  Y  T  A  L  S  T  L  T  N  700
2101  AAAATCTGCAATCGCCTCGCCCAAATCCAAGACAATAAGATTCTCCAAGTTGTGGATGGG 2160
  701 K  I  C  N  R  L  A  Q  I  Q  D  N  K  I  L  Q  V  V  D  G  720
2161  AGCATAAAGGATAAGGAGCTAGAACAGGATATGCAGGCGTTGGTGAAGTTAGTGCTTCAA 2220
  721 S  I  K  D  K  E  L  E  Q  D  M  Q  A  L  V  K  L  V  L  Q  740
2221  GAAAATGGCGGCGCCGTAGACAGAAACATCAGACACACGTTTTTGTCGGTTTCCAAGACT 2280
  741 E  N  G  G  A  V  D  R  N  I  R  H  T  F  L  S  V  S  K  T  760
2281  TTCTACTACGATGCCTACCACGACGATGAGACGACCGATCTTCATATCTTCAAAGTACTC 2340
  761 F  Y  Y  D  A  Y  H  D  D  E  T  T  D  L  H  I  F  K  V  L  780
2341  TTTCGACCGGTTGTATGA   2358 nucleotides
  781 F  R  P  V  V  stop 785 amino acids
```

Figure 19: Nucleotide and amino acid sequence of the SscTPS1 cDNA clone from *S. sclarea* L.. The underlining show the gene specific region of the 5'CopTOPOipsort primer, bold amino acids are motif which will be discussed in chapter **3.2.8.**.

3.2.3. RNA-Blot of the putative diterpene synthase

To determine the tissue-specific expression of the putative diterpene synthase, RNA was isolated from different parts of the *Salvia* plant (stems, leaves, sepals, buds and trichomes). The isolated trichomes were monitored for purity by light microscopy [93]. The RNA blot (**Figure 20**) analysis indicated a mainly trichome-specific expression of the putative diterpene synthase; we saw a slight expression in stems, sepals and buds too. This data is supported by RT-PCR, where the same results of a slight expression were seen (data not shown).(**2.4.3.**; **2.4.4.**).

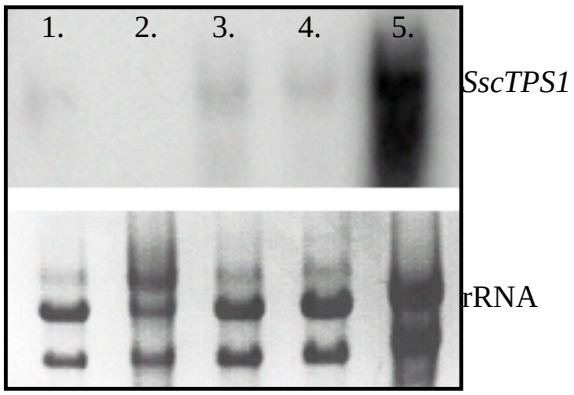


Figure 20: RNA blot of the putative *Salvia sclarea* diterpene synthase. For the blot, 10 µg of a total RNA-extraction of following tissues were used: stems (1), leaves (2), sepals (3), buds (4), trichomes (5).

3.2.4. Purification of the putative diterpene synthase

Bacterial strain *E. coli* BL21(DE3)RIL/Ssc 5'Copipsort x 3'stop in pET100/D-TOPO (SscTPS1) was grown to A_{600} = 0.6 at 37°C in 1 L LB medium supplemented with 100 µg/ml ampicillin and 50 µg/ml chloramphenicol as determined by the vector and the host strain. Cultures were then induced by addition of isopropyl-1-thio-β-D-galactopyranoside to a final concentration of 1 mM and grown for another 10-18 h. at 28°C. Cells were harvested by centrifugation (10,000 x g, 10 min) and resuspended in HLB (50 mM Tris/HCl pH 7.0, 500 mM NaCl, 10 mM imidazole, 10% [v/v] glycerol, 1% [v/v] Tween 20, 10 mM β-EtSH, 75 µg/ml lysozyme) and incubated for 1 h on a platform shaker. After 3 x of mild sonication (~30 sec, cool between on ice) the crude lysate was filtered through synthetic filter floss placed in a syringe [95] and centrifuged (12, 000 x g, 25 min). The lysate was purified by column chromatography using the Talon purification (BD Biosciences) according to the manufacturer's instructions, with HWB (wash/binding buffer) and HEB (Elution buffer). The Copipsort fractions were combined and desalted over a PD-10-colum (Amersham) equilibrated with 5 column volumes of DSB (50 mM HEPES pH 7.2, 5 mM DTT, 100 mM KCl, 7.5 mM $MgCl_2$ x 6 H_2O, 5% [v/v] glycerol [78] (**2.11.1.** +**2.** +**3.**). The protein concentration was determined (**2.5.2.**) and aliquots of the purification were checked on an SDS-PAGE (**Figure 21**). To verify the correctness of the expression and cloning a positive western blot with Monoclonal Anti-polyhistidine-Alkaline Phosphatase antibody was performed (data not shown; **2.10.2.**), and the protein sequence was confirmed using a mass spectrometer (**2.3.2.1.**).

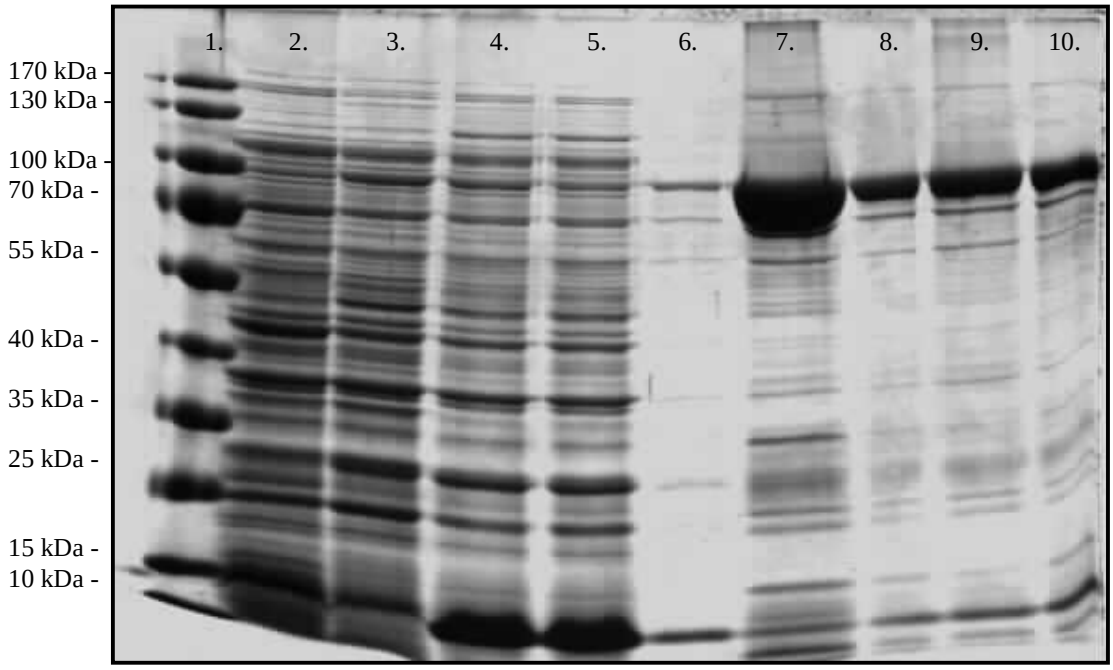


Figure 21: SDS-PAGE of SscTPS1 over a Talon™column: 1. ladder; 2. pre-induction; 3. post-induction (overnight); 4. crude lysate; 5.Talon™supernatant; 6. + 7. Talon™fraction 1 and 2; 8. - 10. PD10column desalting fractions 4. - 6.

3.2.5. Protein sequencing of the putative diterpene synthase

The putative diterpene synthase Ssc 5'Copipsort x 3'stop was digested out of a SDS-PAGE (**Figure 22;** Ssc 5'Copipsort x 3'stop protein bands in fraction 8+9 which were marked with an arrow were cut out of the gel) according to the modified in-gel protein digestion and ZipTip sample clean-up protocol (**2.3.3.1.**), sequenced with an ESI-MS/MS and blasted against the full-length sequence of the putative diterpene synthase. The sequencing and the mass spectrometry was done at the Donald Danforth Plant Science Center Proteomics and Mass Spectrometry Facility (http://www.danforthcenter.org/msb/) by Leslie Hicks, Ph.D. and Jeanne Sheffield, Msc. (**2.3.2.1.**). The protein identification was confirmed through tryptic digestion (cuts C-term side of KR unless next residue is P) and ESI-MS/MS with 37% sequence coverage. Amino acid sequences of 34 internal peptides were found (**Table 3**), overlapped it leads to 18 confirmed internal peptides, shown in bold red (**Figure 23**). Nominal mass of 89.703 kDa and pI of 5.58 were calculated, which is in the range of native full-length protein 89.758 kDa and the cloned construct 92.94 kDa (**3.2.2.**; DNAstar Inc.).

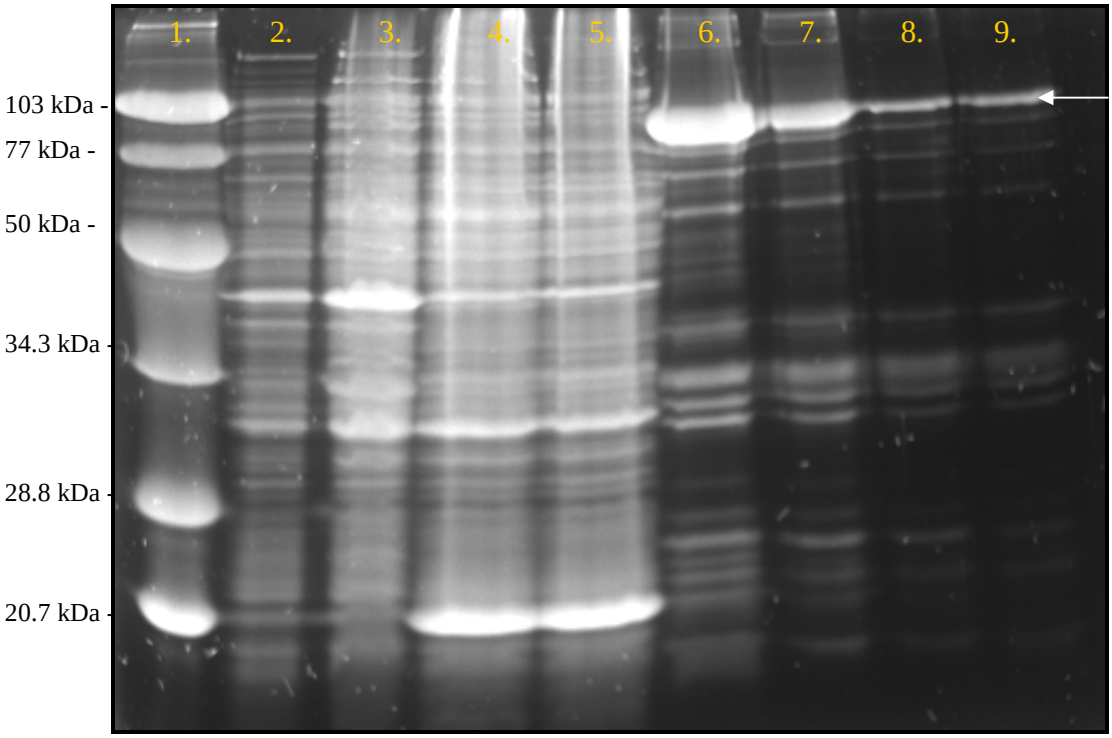


Figure 22: SDS-PAGE of SscTPS1 for sequencing stained with Sypro Ruby Gel Stain™ (Invitrogen), purified over a Talon™column: 1. Ladder; 2. pre-induction; 3. post-induction (overnight); 4. crude lysate; 5. Talon™supernatant; 6. + 7. Talon™fraction 1 and 2; 8. + 9. PD10column desalting fractions 4. and 5. The arrow represents the two bands that are cut out of the gel.

1	LLWAK
2	LWAIDR
3	MITSFVK
4	VLFRPVV
5	MITSFVK + Oxidation (M)
6	YIENALK
7	HTFLSVSK
8	LAQIQDNK
9	FFQHEIK
10	FFQHEIK
11	QLSQIAELR
12	ILQVVDGSIK
13	ELAILDFNR
14	MPWYATLPR + Oxidation (M)
15	MPEISNDTYK + Oxidation (M)
16	VTSLDASQASEK
17	LVLQENGGAVDR
18	LVLQENGGAVDR
19	ILQVVDGSIKDK
20	SWNLHSDIIEK
21	FSDIDDTSMGVR + Oxidation (M)
22	MHGYDVDPNVLK + Oxidation (M)
23	WVISDHLFDEIK

24	FAFNFLQEMLVK + Oxidation (M)
25	DKELEQDMQALVK + Oxidation (M)
26	ISVSPYDTAVIALIK
27	ISVSPYDTAVIALIK
28	SYFLAAATIFEPER
29	FAFNFLQEMLVKDR + Oxidation (M)
30	VEAAYYLDHYAGSGDVWIGK
31	MPEISNDTYKELAILDFNR + Oxidation (M)
32	LQSGNGSFLTSPSSTAAVLMHTK + Oxidation (M)
33	TFYYDAYHDDETTDLHIFK
34	LQSGNGSFLTSPSSTAAVLMHTKDEK + Oxidation (M)

Table 3: Amino acid sequences of 34 internal peptides from the recombinant SscTPS1 of *S. sclarea* L.. With a calculated nominal mass of **89702 kDa**, and pI value of **5.58**. The variable modifications were Oxidation of M. The cleavage was done by Trypsin: (cuts C-term side of KR unless next residue is P Sequence). The sequence coverage was **37%**

001 MTSVNLSRAP AAITRRRLQL QPEFHAECSW LKSSSKHAPL TLSCQIRPK**Q**
051 **LSQIAELRVT SLDASQASEK** DISLVQTPHK VEVNEKIEES IEYVQNLLMT
101 SGDGR**ISVSP YDTAVIALIK** DLKGRDAPQF PSCLEWIAHH QLADGSWGDE
151 FFCIYDRILN TLACVVALK**S** **WNLHSDIIEK** GVTYIKENVH KLKGANVEHR
201 TAGFELVVPT FMQMATDLGI QDLPYDHPLI KEIADTKQQR LKEIPKDLVY
251 QMPTNLLYSL EGLGDLEWER LLK**LQSGNGS FLTSPSSTAA VLMHTKDEK**C
301 LK**YIENALK**N CDGGAPHTYP VDIFSR**LWAI DR**LQRLGISR **FFQHEIK**YFL
351 DHIESVWEET GVFSGRYTK**F SDIDDTSMGV R**LLK**MHGYDV DPNVLK**HFKQ
401 QDGKFSCYIG QSVESASPMY NLYRAAQLRF PGEEVLEEAT K**FAFNFLQEM**
451 **LVKDR**LQER**W VISDHLFDEI K**LGLK**MPWYA TLPRVEAAYY LDHYAGSGDV**
501 **WIGK**SFYR**MP EISNDTYKEL AILDFNR**CQT QHQLEWIHMQ EWYDRCSLSE
551 FGISKRELLR **SYFL**AAA**TIF EPER**TQER**LL WAK**TRILSK**M ITSFVK**ISGT
601 TLSLDYNFNG LDEIISSANE DQGLAGTLLA TFHQLLDGFD IYTLHQLKHV
651 WSQWFMKVQQ GEGSGGDDAV LQANALNICA GLNEDVLSNN EYTALSTLTN
701 KICNR**LAQIQ DNKILQVVDG SIKDKELEQD MQALVKLVLQ ENGGAVDR**NI
751 R**HTFLSVSKT FYYDAYHDDE TTDLHIFKVL FRPVV**

Figure 23: Overlapped amino acid sequences, shown in **bold red**, of the 34 internal peptides.

3.2.6. Diterpene synthase assay for the characterization

The putative diterpene synthase activity assay was performed in DSB including GPP, NPP, FPP and GGPP in siliconized tubes (Sigma-Aldrich), overlaid with hexane, to trap the hydrophobic products. After incubation, the mixture was hydrolized using HCl or a wheat germ acid phosphatase (Sigma), to solvolyze the allylic diphosphate esters present. The assay was mixed vigorously and extracted in total 3 times with hexane. Boiling controls were performed too. The hexane phase was concentrated under N_2-stream and were analyzed using GC-MS. Boiling controls were performed too.

3.2.7. GC-MS analysis of the putative diterpene synthase

The aim of the herewith-presented work was to find a diterpene synthase being responsible for the formation of sclareol. The unambiguous detection of sclareol demonstrated that putative diterpene synthase, now named SscTPS1, forms it from GGPP. Additional it is shown that the enzyme has a promiscuous behavior with respect to different substrates; mono-, sesqui- and diterpenes are formed. The GC-MS procedure is described under chapter **2.13.**.In the following figures and tables are the GC-MS results of the SscTPS1 assay with GPP, NPP, FPP and GGPP, using a wheat germ acid phosphatase (**Figure 24-28**; **Table 4-7**) or HCl (**Figure 29-32**; **Table 8-11**) as hydrolysis agents. All products were identified and determined by comparison of the EI-MS spectra with those of the NIST library (V 1.6b), respectively, or purchased authentic standards. A comparison of the formed sclareol ((**27), Figure 33**) with authentic standard is shown in **Figure 34**, and found in the literature [104]. The structures of the products were shown in the Appendix (**Figure 41**+**42**) [105, 106, 107].

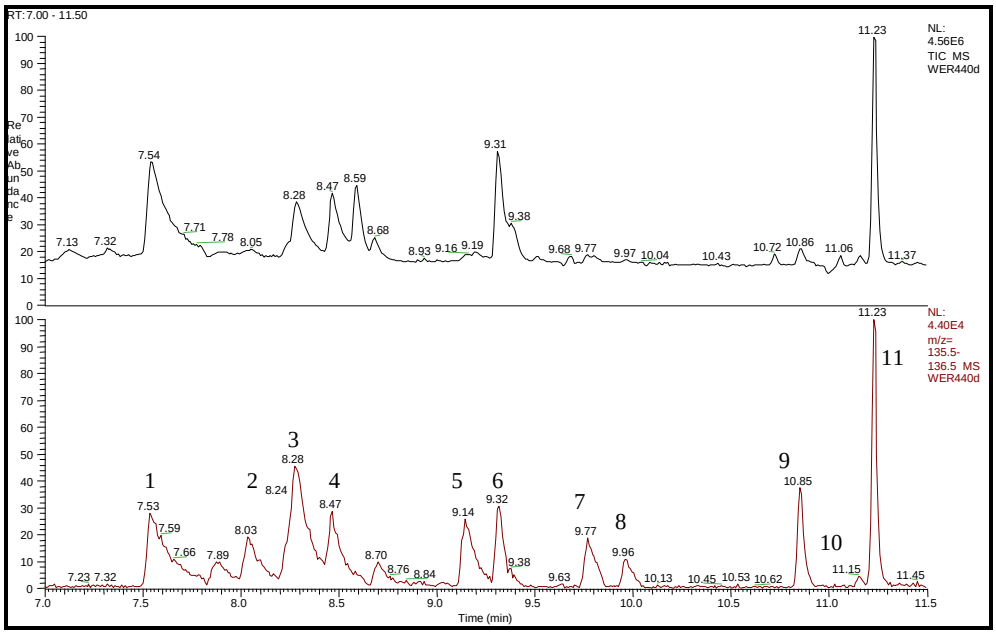


Figure 24: GC-MS analysis (R_t = 7 - 11.5 min) of the wheat germ acid phosphatase hydrolized GPP monoterpene products of SscTPS1: β-pinene (**1**), limonene (**2**), α-pinene (**3**), β-cis-ocimene (**4**), terpinolene (**5**), β-linalool (**6**), not-identified monoterpene (**7**), α-pyronene (**8**), α-terpineol (**9**), not-identified monoterpene (**10**), cis-geraniol (**11**), Appendix (WER440d).

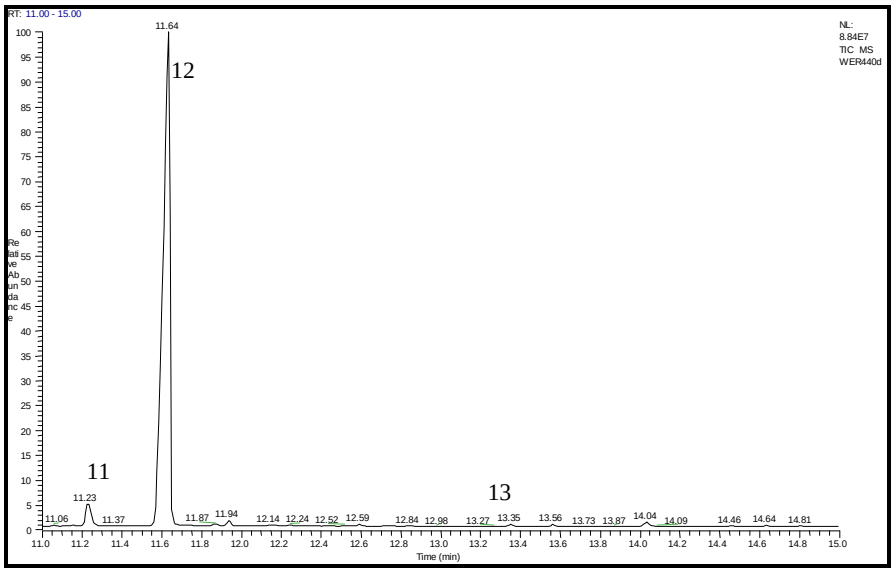


Figure 25: GC-MS analysis (R_t = 11 - 15 min) of the wheat germ acid phosphatase hydrolized GPP monoterpene products of SscTPS1: cis-geraniol (**11**), trans-geraniol (**12**), and nerol-acetate (**13**), Appendix (WER440d).

No.	Compound	R_t (min)	EI-MS (*m/z*, rel int.)
1	β-pinene	7.53	136 (M^+, 4), 121 ($[M\text{-}Me]^+$, 4), 93 ($[M\text{-}C_3H_7]^+$, 100), 79 (21), 69 (81), 67 (2), 53 (18)
2	limonene	8.24	136 (M^+, 22), 121 ($[M\text{-}Me]^+$, 20), 93 ($[M\text{-}C_3H_7]^+$,39), 79 (34), 69 (5), 68 (100) 67 (92), 53(26)
3	α-pinene	8.28	136 (M^+, 10), 121 ($[M\text{-}Me]^+$, 11), 93 ($[M\text{-}C_3H_7]^+$,100), 79 (42), 69 (3), 67 (16), 53 (15)
4	β-cis-ocimene	8.47	136 (M^+, 6), 121 ($[M\text{-}Me]^+$, 16), 93 ($[M\text{-}C_3H_7]^+$, 100), 79 (54), 69 (3), 67 (13), 53 (21)
5	terpinolene	9.14	136 (M^+, 72), 121 ($[M\text{-}Me]^+$, 91), 93 ($[M\text{-}C_3H_7]^+$, 100), 79 (44), 69 (-), 67 (12), 53 (16)
6	β-linalool	9.32	154 (M^+, 54), 139 ($[M\text{-}Me]^+$, 1), 136 ($[M\text{-}H_2O]^+$,6), 93 ($[M\text{-}H_2O\text{-}C_3H_7]^+$, 60), 79 (15), 71 (100), 69 (39), 67 (22), 53 (16)
7	not-identified	9.77	136 (M^+, 40), 121 ($[M\text{-}Me]^+$, 100), 136 ($[M\text{-}H_2O]^+$, -), 93 ($[M\text{-}C_3H_7]^+$, 30), 79 (41), 69 (-), 67 (9), 53 (9)
8	α-pyronene	9.96	136 (M^+, 42), 121 ($[M\text{-}Me]^+$, 100), 93 ($[M\text{-}C_3H_7]^+$, 24), 79 (41), 69 (-), 67 (7), 53 (14)
9	α-terpineol	10.85	154 (M^+, -), 139 ($[M\text{-}Me]^+$, 8), 136 ($[M\text{-}H_2O]^+$, 49), 93 ($[M\text{-}H_2O\text{-}C_3H_7]^+$, 71), 79 (30), 69 (6), 67 (36), 59 (100), 53 (18)
10	not-identified	11.15	136 (M^+, 5), 121 ($[M\text{-}Me]^+$, 18), 93 ($[M\text{-}C_3H_7]^+$, 23), 79 (8), 69 (100), 67 (41), 53 (19)
11	cis-geraniol	11.23	154 (M^+, 1), 139 ($[M\text{-}Me]^+$, 5), 136 ($[M\text{-}H_2O]^+$, 7), 121 (15), 93 ($[M\text{-}H_2O\text{-}C_3H_7]^+$, 45), 84 (15), 79 (16), 69 (100), 67 (30), 53 (19), 41 (89)
12	trans-geraniol	11.64	154 (M^+, 9), 139 ($[M\text{-}Me]^+$, 28), 136 ($[M\text{-}H_2O]^+$, 25), 93 ($[M\text{-}H_2O\text{-}C_3H_7]^+$, 82), 79 (35), 69 (100), 67 (71), 53 (44)
13	nerol acetate	13.35	196 (M^+, 1), 154 ($[M\text{-}CH_2CO]^+$, 1), 136 ($[M\text{-}HOAc]^+$, 13), 93 ($[M\text{-}Ac\text{-}H_2O\text{-}C_3H_7]^+$, 58), 79 (19), 69 (100), 67 (29), 53 (21)

Table 4: EI mass spectra of the monoterpene product profile of SscTPS1, using GPP as the substrate and a wheat germ acid phosphatase as the hydrolysis agent, showing four cyclic, one olefinic, two oxygenated, two not identified, two oxygenated olefinic compounds, and an acetylated products. TIC's shown in the Appendix (WER440d).

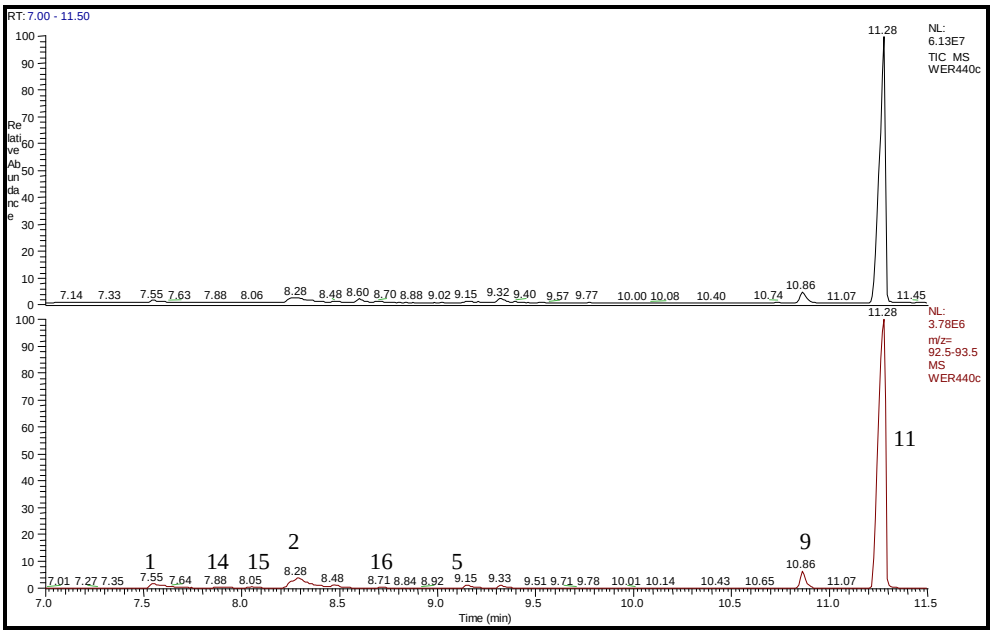


Figure 26: GC-MS analysis (R_t = 7 - 11.5 min) of the wheat germ acid phosphatase hydrolized NPP monoterpene products of SscTPS1: β-pinene (**1**), α-phellandrene(**14**), α-terpinene (**15**), limonene (**2**), γ-terpinene (**16**), terpinolene (**5**), α-terpineol (**9**), cis-geraniol (**11**), Appendix (WER440c).

No.	Compound	R_t (min)	EI-MS (*m/z*, rel int.)
1	β-pinene	7.55	136 (M^+, 4), 121 ($[M-Me]^+$, 5), 93 ($[M-C_3H_7]^+$, 86), 79 (19), 69 (67), 67 (17), 53 (19), 41 (100)
14	α-phellandrene	7.88	136 (M^+, 29), 121 ($[M-Me]^+$, 1), 93 ($[M-C_3H_7]^+$, 100), 79 (9), 69 (-), 67 (-), 53 (4)
15	α-terpinene	8.05	136 (M^+, 52), 121 ($[M-Me]^+$, 96), 93 ($[M-C_3H_7]^+$, 100), 79 (31), 69 (-), 67 (39), 53 (12)
2	limonene	8.26	136 (M^+, 27), 121 ($[M-Me]^+$, 28), 93 ($[M-C_3H_7]^+$, 81), 79 (52), 69 (3), 67 (100), 53 (35)
16	γ-terpinene	8.71	136 (M^+, 38), 121 ($[M-Me]^+$, 29), 93 ($[M-C_3H_7]^+$, 95), 79 (22), 69 (11), 67 (5), 53 (10), 43 (100)
5	terpinolene	9.15	136 (M^+, 76), 121 ($[M-Me]^+$, 90), 93 ($[M-C_3H_7]^+$, 100), 79 (48), 69 (-), 67 (13), 53 (18)
9	α-terpineol	10.86	154 (M^+, 1), 139 ($[M-Me]^+$, 9), 136 ($[M-H_2O]^+$, 50), 93 ($[M-H_2O-C_3H_7]^+$, 74), 79 (32), 69 (6), 67 (38), 59 (100), 53 (19)
11	cis-geraniol	11.28	154 (M^+, 6), 139 ($[M-Me]^+$, 18), 136 ($[M-H_2O]^+$, 23), 93 ($[M-H_2O-C_3H_7]^+$, 84), 79 (31), 69 (100), 67 (59), 53 (27)

Table 5: EI mass spectra of the monoterpene product profile of SscTPS1, using NPP as the substrate and a wheat germ acid phosphatase as the hydrolysis agent, showing six cyclic, one oxygenated and one oxygenated olefinic products. TIC's shown in the Appendix (WER440c).

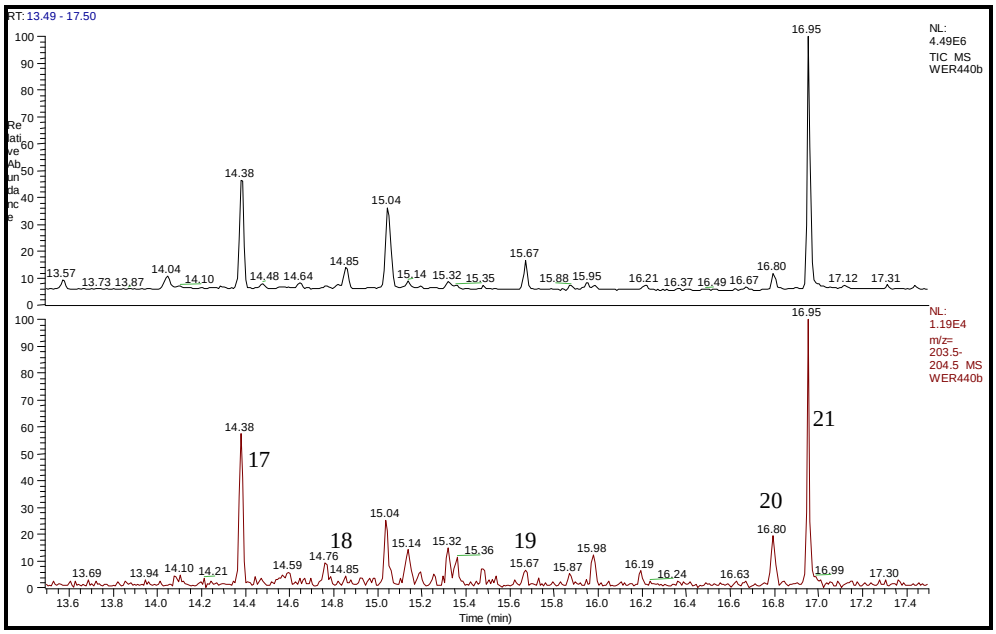


Figure 27: GC-MS analysis (R_t = 13.55 - 17.45 min) of the wheat germ acid phosphatase hydrolized FPP sesquiterpene products of SscTPS1: β-farnesene (**17**), (Z,E)-α-farnesene (**18**), nerolidol (**19**), not-identified (**20**), farnesol (**21**), Appendix (WER440b).

No.	Compound	R_t (min)	EI-MS (*m/z*, rel int.)
17	β-farnesene	14.38	204 (M^+, 2), 189 ($[M\text{-}Me]^+$, 2), 161 ($[M\text{-}C_3H_7]^+$, 12), 148 (2) 136 (-), 135(1), 121 (4), 107 (10), 93 (50), 79 (30), 55 (17), 43 (7), 41 (100)
18	(Z,E)-α-farnesene	14.85	204 (M^+, 1), 189 ($[M\text{-}Me]^+$, 1), 161 ($[M\text{-}C_3H_7]^+$, 1), 148 (1) 136 (1), 135(2), 121 (4), 107 (35), 93 (85), 79 (45), 55 (39), 43 (20), 41 (100)
19	nerolidol	15.67	222 (M^+, -), 204 ($[M\text{-}H_2O]^+$, 1), 189 (3), 161 (10), 148 (2) 136 (8), 135(4), 121 (9), 107 (34), 93 (56), 79 (27), 55 (33), 43 (56), 41 (100)
20	not-identified*	16.80	222 (M^+, -), 204 ($[M\text{-}H_2O]^+$, 6), 189 (3), 161 (7), 148 (2) 136 (2), 135(2), 121 (8), 107 (19), 93 (44), 79 (27), 55 (23), 43 (21), 41 (100)
21	farnesol	16.95	222 (M^+, -), 204 ($[M\text{-}H_2O]^+$, 1), 189 (2), 161 (7), 148 (2) 136 (6), 135(3), 121 (8), 107 (20), 93 (43), 79 (25), 69 (100), 55 (19), 43 (12), 41 (91)

Table 6: EI mass spectra of the sesquiterpene product profile of SscTPS1 using FPP as the substrate and a wheat germ acid phosphatase as the hydrolysis agent showing two olefinic and three oxygenated olefinic products. TIC's shown in the Appendix (WER440b). * Mass spectra suggest a similarity to farnesol.

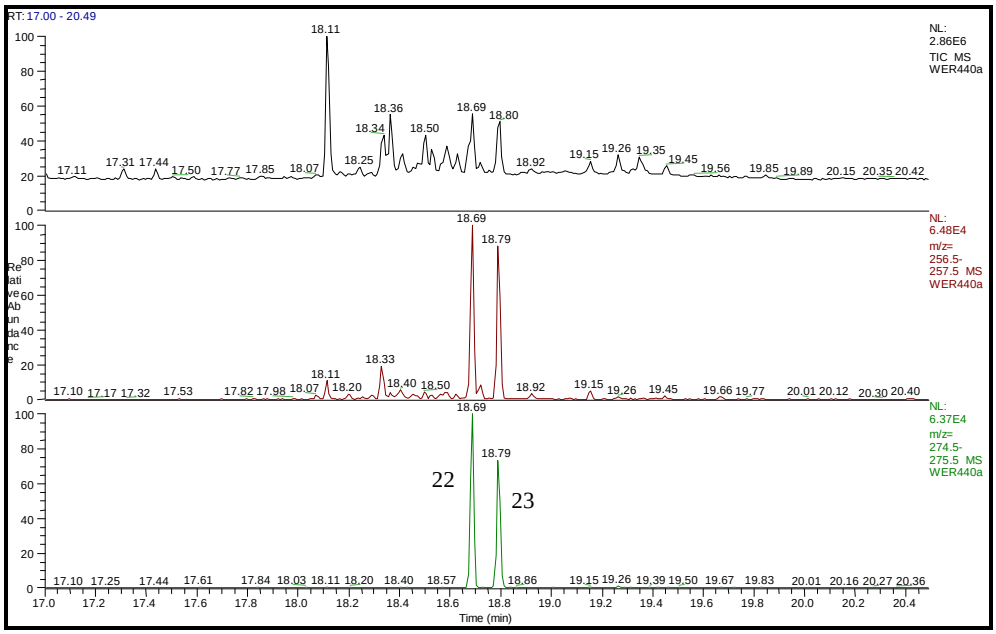


Figure 28: GC-MS analysis (R_t = 17 min - 20.45 min) of the wheat germ acid phosphatase hydrolized GGPP diterpene products of SscTPS1: (13R)-manoyl oxide (**22**), (13S)-manoyl oxide (**23**), Appendix (WER440a).

No.	Compound	R_t (min)	EI-MS (*m/z*, rel int.)
22	(13R)-manoyl oxide	18.69	290 (M^+, -), 275 ($[M\text{-}Me]^+$, 100), 257 ($[M\text{-}H_2O\text{-}Me]^+$, 99), 204 (2), 191(27), 189 (9), 175 (8), 163 (9), 149 (22), 121 (25), 109 (30), 95 (44), 81 (66), 69 (98),
23	(13S)-manoyl oxide	18.79	290 (M^+, -), 275 ($[M\text{-}Me]^+$, 85), 257 ($[M\text{-}H_2O\text{-}Me]^+$, 100), 204 (-), 191(44), 189 (7), 175 (6), 163 (7), 149 (14), 121 (22), 109 (36), 95 (38), 81 (46), 69 (35),

Table 7: Diterpene product profile of SscTPS1 using GGPP as the substrate and a wheat germ acid phosphatase as the hydrolysis agent showing two cyclic products. TIC's shown in the Appendix (WER440a).

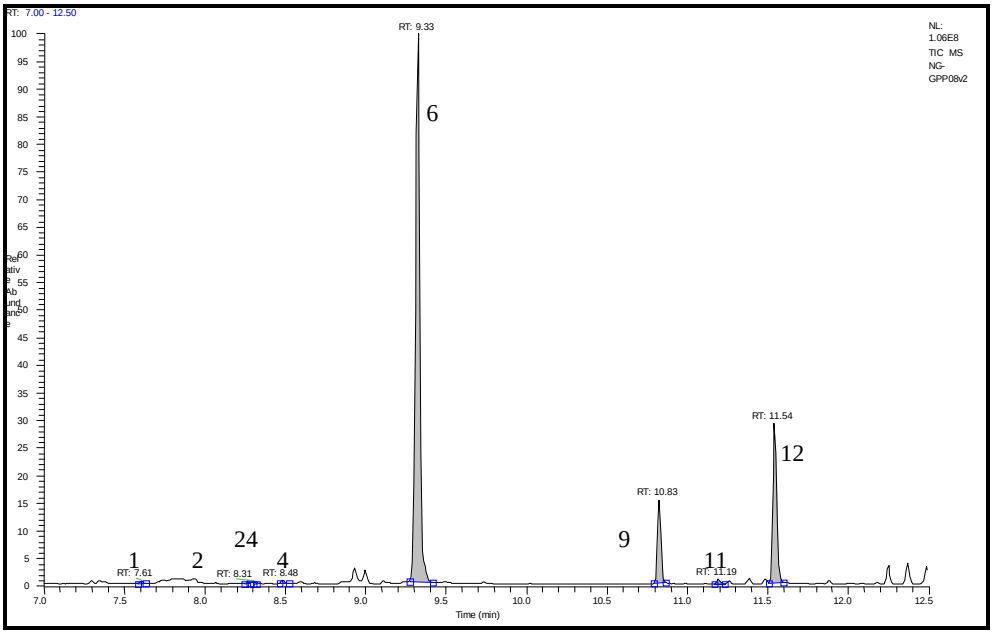


Figure 29: GC-MS analysis (R_t = 7- 12.5 min) of the acid (HCl) hydrolized GPP monoterpene products of SscTPS1: β-pinene (**1**), limonene (**2**), β-trans-ocimene (**24**), β-cis-ocimene (**4**), β-linalool (**6**), α-terpineol (**9**), cis-geraniol (**11**). trans-geraniol (**12**), Appendix (NG_GPP08v2).

No.	Compound	R_t (min)	EI-MS (*m/z*, rel int.)
1	β-pinene	7.61	136 (M^+, 3), 121 ([M-Me]$^+$, 5), 93 ([M-C_3H_7]$^+$, 100), 79 (17), 69 (91), 67 (14), 53 (10)
2	limonene	8.28	136 (M^+, 17), 121 ([M-Me]$^+$, 23), 93 ([M-C_3H_7]$^+$, 78), 79 (37), 69 (6), 68 (100), 67 (88), 53 (22)
24	β-trans-ocimene	8.31	136 (M^+, 2), 121 ([M-Me]$^+$, 10), 93 ([M-C_3H_7]$^+$, 100), 79 (45), 69 (2), 67 (12), 53 (14)
4	β-cis-ocimene	8.48	136 (M^+, 7), 121 ([M-Me]$^+$, 17), 93 ([M-C_3H_7]$^+$, 100), 79 (51), 69 (2), 67 (14), 53 (15)
6	β-linalool	9.31	154 (M^+, -), 139 ([M-Me]$^+$, 3), 136 ([M-H_2O]$^+$, 14), 93 ([M-H_2O-C_3H_7]$^+$, 83), 79 (27), 70 (100), 69 (66), 67 (38), 53 (19)
9	α-terpineol	10.83	154 (M^+, -), 139 ([M-Me]$^+$, 7), 136 ([M-H_2O]$^+$, 40), 93 ([M-H_2O-C_3H_7]$^+$, 58), 79 (19), 69 (4), 67 (14), 59 (100), 53 (8)
11	cis-geraniol	11.19	154 (M^+, 1), 139 ([M-Me]$^+$, 3), 136 ([M-H_2O]$^+$, 3), 93 ([M-H_2O-C_3H_7]$^+$, 27), 79 (8), 69 (100), 67 (20), 53 (10)
12	trans-geraniol	11.54	154 (M^+, 1), 139 ([M-Me]$^+$, 4), 136 ([M-H_2O]$^+$, 4), 93 ([M-H_2O-C_3H_7]$^+$, 26), 79 (9), 69 (100), 67 (44), 53 (16)

Table 8: EI mass spectra of the monoterpene product profile of SscTPS1, using GPP as the substrate and an acid (HCl) as the hydrolysis agent, showing two cyclic, two olefinic and two oxygenated cyclic and two oxygenated olefinic products. TIC's shown in the Appendix (NG_GPP08v2).

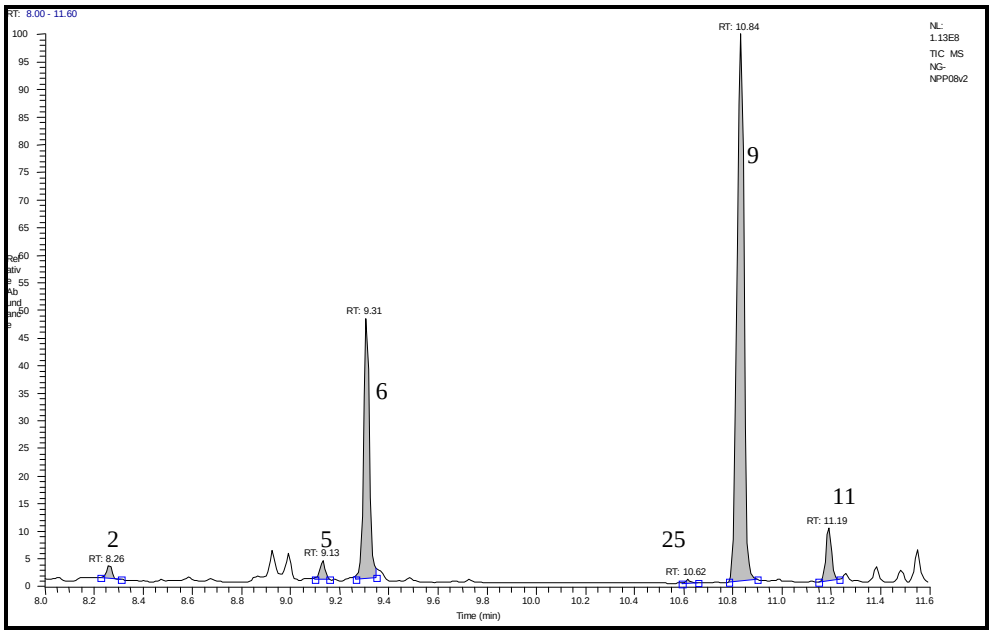


Figure 30: GC-MS analysis (R_t = 8 - 11.6 min) of the acid (HCl) hydrolized NPP monoterpene products of SscTPS1: limonene (**2**), terpinolene (**5**), β-linalool (**6**), (-)-terpinen-4-ol (**25**), α-terpineol (**9**), cis-geraniol (**11**), Appendix (NG_NPP08v2).

No.	Compound	R_t (min)	EI-MS (*m/z*, rel int.)
2	limonene	8.26	136 (M^+, 22), 121 ($[M\text{-}Me]^+$, 22), 93 ($[M\text{-}C_3H_7]^+$, 67), 79 (34), 69 (7), 68 (100),67 (84), 53 (24)
5	terpinolene	9.13	136 (M^+, 74), 121 ($[M\text{-}Me]^+$, 84), 93 ($[M\text{-}C_3H_7]^+$, 100), 79 (36), 67 (9), 53 (10)
6	β-linalool	9.31	154 (M^+, 1), 139 ($[M\text{-}Me]^+$, 3), 136 ($[M\text{-}H_2O]^+$, 13), 93 ($[M\text{-}H_2O\text{-}C_3H_7]^+$, 89), 79 (26), 70 (100), 69 (66), 67 (36), 53 (20)
25	(-)-terpinen -4-ol	10.62	154 (M^+, 13), 139 ($[M\text{-}Me]^+$, 2), 136 ($[M\text{-}H_2O]^+$, 14), 93 ($[M\text{-}H_2O\text{-}C_3H_7]^+$, 54), 79 (7), 71 (100), 69 (18), 67 (17), 53 (9)
9	α-terpineol	10.84	154 (M^+, -), 139 ($[M\text{-}Me]^+$, 27), 136 ($[M\text{-}H_2O]^+$, 65), 93 ($[M\text{-}H_2O\text{-}C_3H_7]^+$, 100), 79 (100), 69 (15), 67 (58), 53 (28)
11	cis-geraniol	11.19	154 (M^+, 1), 139 ($[M\text{-}Me]^+$, 4), 136 ($[M\text{-}H_2O]^+$, 4), 93 ($[M\text{-}H_2O\text{-}C_3H_7]^+$, 30), 79 (8), 69 (100), 67 (20), 53 (11)

Table 9: EI mass spectra of the monoterpene product profile of SscTPS1 using NPP as the substrate and an acid (HCl) as the hydrolysis agent, showing two cyclic three oxygenated cyclic and one oxygenated olefinic products. TIC's shown in the Appendix (NG_NPP08v2).

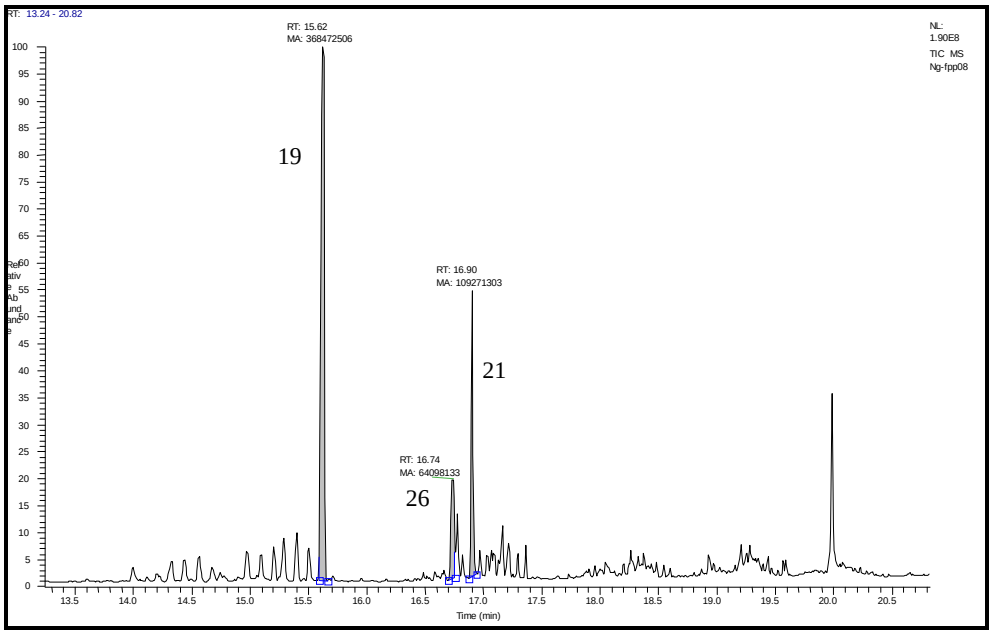


Figure 31: GC-MS analysis (R_t = 13.25 - 20.55 min) of the acid (HCl) hydrolized FPP sesquiterpene products of SscTPS1: nerolidol (**19**), α-bisabolol (**26**), farnesol (**21**), Appendix (NG_FPP08).

No.	Compound	R_t (min)	EI-MS (*m/z*, rel int.)
19	nerolidol	15.62	222 (M^+, -),207 ([M-Me]$^+$, 2), 204 ([M-H_2O]$^+$, 7), 189 (19), 161 (52), 148 (17) 136 (58), 135(32), 121 (45), 107 (71), 93 (100), 79 (28), 55 (61), 43 (67), 41 (71)
26	α-bisabolol	16.74	222 (M^+, -),207 ([M-Me]$^+$, 1), 204 ([M-H_2O]$^+$, 16), 189 (4), 161 (10), 148 (4) 136 (2), 135(4), 121 (23), 107 (16), 93 (47), 96 (100), 79 (23), 55 (28), 43 (82), 41 (66)
21	farnesol	16.90	222 (M^+, 1),207 ([M-Me]$^+$, 1), 204 ([M-H_2O]$^+$, 3), 189 (5), 161 (14), 148 (4) 136 (41), 135(14), 121 (36), 107 (46), 93 (68), 68 (100), 79 (38), 55 (36),43 (30), 41 (62)

Table 10: EI mass spectra of the sesquiterpene product profile of SscTPS1 using NPP as the substrate and an acid (HCl) as the hydrolysis agent showing one oxygenated cyclic and two oxygenated olefinic products. TIC's shown in the Appendix (NG_FPP08).

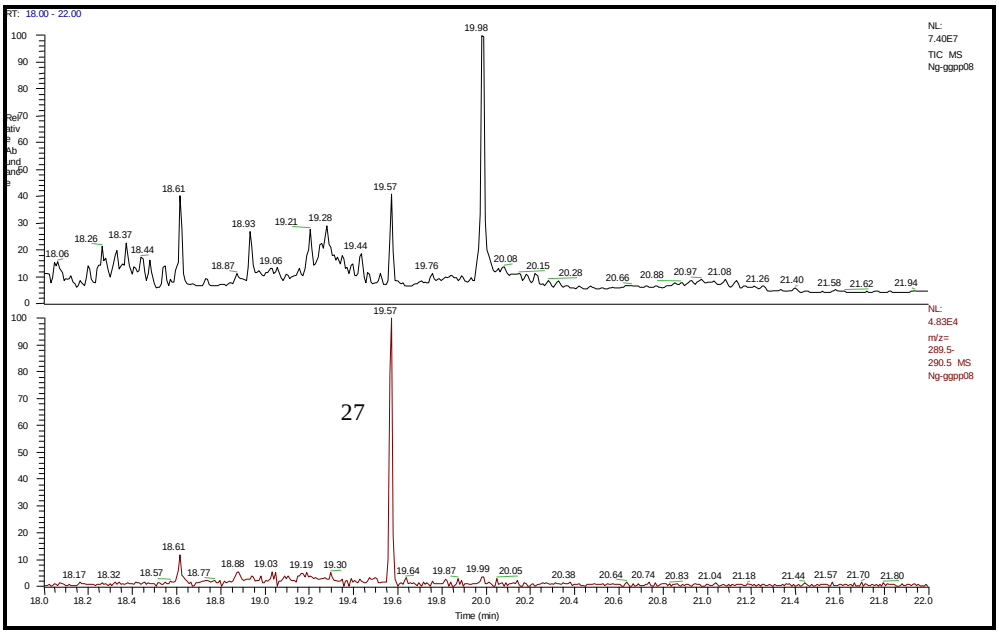


Figure 32: GC-MS analysis (R_t = 18 - 22 min) of the acid (HCl) hydrolized GGPP diterpene products of SscTPS1: sclareol (**27**), Appendix (NG_GGPP08).

No.	Compound	R_t (min)	EI-MS (*m/z*, rel int.)
27	sclareol	19.57	308 (M^+, -), 290 ($[M\text{-}H_2O]^+$, 3), 275 (2), 272 ($[M\text{-}2H_2O]^+$, 2), 257 ($[M\text{-}2H_2O\text{-}Me]^+$, 4), 204 (2), 191(13), 189 (6), 175 (4), 163 (8), 149 (11), 121 (24), 109 (44), 95 (60), 81 (70), 69 (74), 59 (10), 43 (100)

Table 11: EI mass spectra of the diterpene product profile of SscTPS1 using GGPP as the substrate and an acid (HCl) as the hydrolysis agent showing the cyclic diterpene sclareol as the only product. TIC shown in the Appendix (NG_GGPP08).

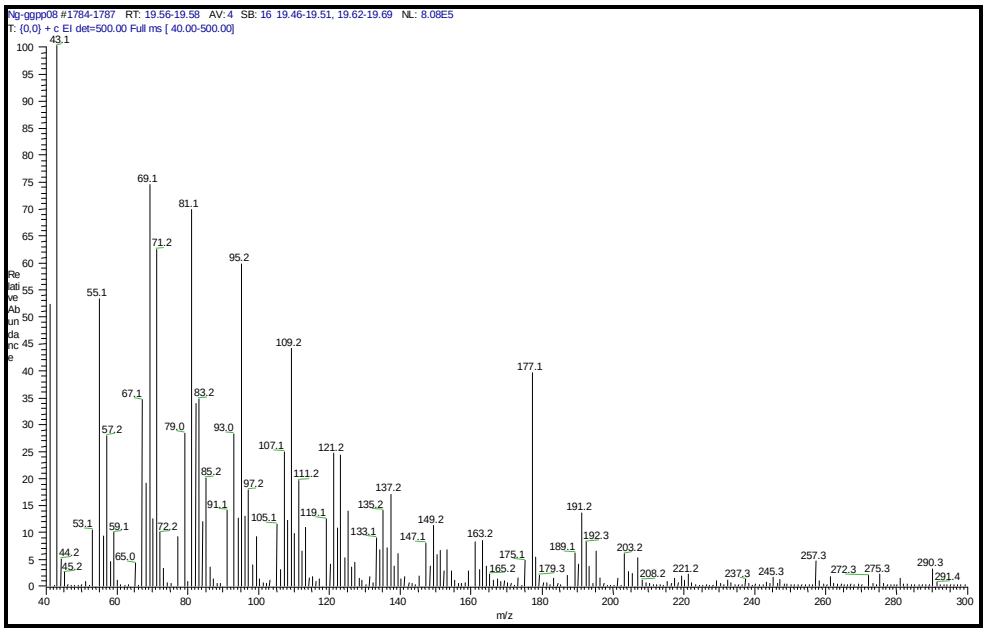


Figure 33: Mass spectrum of sclareol (**27**) at R_t = 19.57 min, formed by SscTPS1 using GGPP as substrate and HCl as the hydrolysis agent.

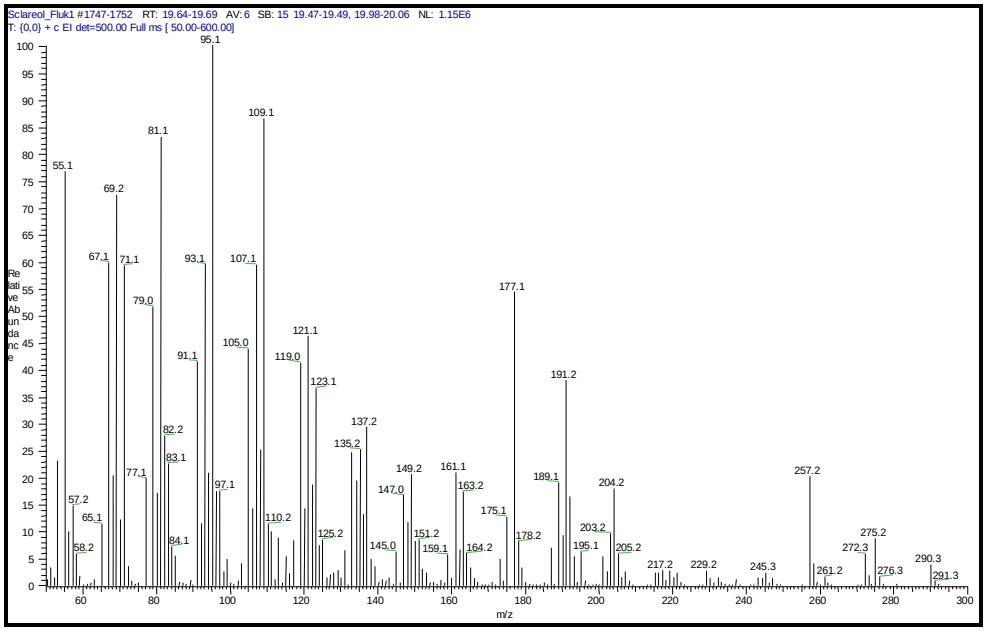


Figure 34: Mass spectrum from authentic sclareol (Fluka) which exhibits a similar total ion chromatogram and a R_t = 19.66 min in the range of the sclareol product formed by SscTPS1.

3.2.8. Phylogenetic (identity tree) analysis and sequence comparison of SscTPS1

SscTPS1 had a high homology (http://www.ncbi.nih.gov/ ; blastx) to copalyl diphosphate synthase of *Lycoperiscum esculentum* (now *Solanum lycopersicum*), *Lactuca sativa* and ent-kaurene synthase A of *Pisum sativum*. On the basis of this, an identity tree (**Figure 35**) was prepared with 31 aa sequences, including SscTPS1, CsTPS1 and CsTPS2, from gymnosperm and angiosperm monoterpene-; sesquiterpene-; and diterpene synthases. ClustalW (http://www.ebi.ac.uk/clustalw/) in fasta format and GeneBee in the Clastw.phb format file (http://www.genebee.msu.su/services/phtree_full.html) were used. The conditions were: scale-random; algorithm-cluster and topological; matrix-blosum 62; with bootstrap values. SscTPS1 placed within the *Tpsc* subfamily of terpene synthases alongside other diverse diterpene synthases from angiosperms [61].

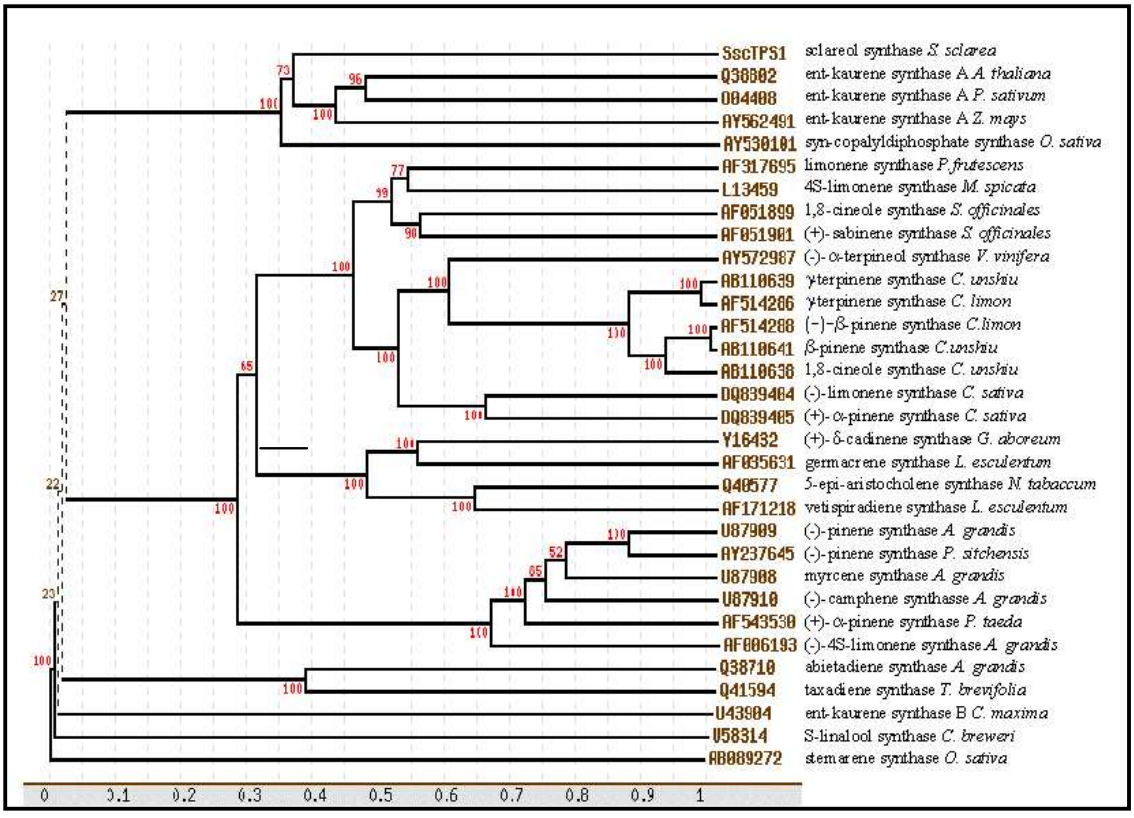


Figure 35: Identity tree analysis of 31 aa sequences of monoterpene-; sesquiterpene-; and diterpene synthases from gymnosperms and angiosperms using ClustalW and GeneBee (algorithm-cluster and topological; matrix-blosum 62) with bootstrap values

Further analysis of the full-length diterpene synthase showed *SscTPS1* bearing typical features of other diterpene synthases. As previously reported, the first 59 N-terminal aa are rich in serine and threonine (22% for SccTPS1), and the estimated pI of this region is close to 11 (10.89 for SccTPS1). These features of a diterpene synthase are common characteristics of transit peptides, which target the proteins to plastids [108, 109]. This information was also supported by the prediction for SscTPS1 to have an N-terminal transit peptide for plastidial targeting (http://hc.ims.u-tokyo.ac.jp/iPSORT/). Two kinds of terpene synthases/cyclases exist, with a different domain structure and different metal binding motifs. Class I terpene synthases contain an N-terminal and a C-terminal domain with DDXXD-motif for metal binding. Class II terpene synthases contain an insertional-, central and C-terminal domain whereas the central domain harbors the DXDD-motif for metal binding. The biosynthesis of labdane-related terpenes requires two cyclization steps. The first is the cyclization of GGPP to the characteristic bicyclic copalyldiphosphate, by stereo specific class II terpene synthases. Further modifications of CPP were done by stereoselective class I terpene synthases. Both terpene synthases classes catalyzed reactions involving electrophilic cyclization/rearrangement, and were biochemically distinct. Class I enzymes facilitated, the metal ion assisted allylic diphosphate ionization reactions commonly associated with terpene synthases, whereas class II enzymes mediated protonation of a carbon-carbon double bond in an acid/base catalyzed reaction. SscTPS1 contained the DXDD-motif and is a class II terpene synthase. The sequence comparison also showed that most diterpene synthases contained the highly conserved SAYDTAW, although the first A in the motif of SscTPS1 is replaced by other unpolar aa P like in OsDTC1 (ent-cassa-12,15-diene synthase; *Oryza sativa*) and just like the W, which is replaced by V (SPYDTAV). The second known conserved motif, the QXXDGSW-motif, was fully represented in SscTPS1 (**Figure 36**) [78]. The roles of the motifs remain unknown, whereas the latter one was suggested in bacterial squalene hopene synthases to be involved in the stabilization of the whole enzyme [109].

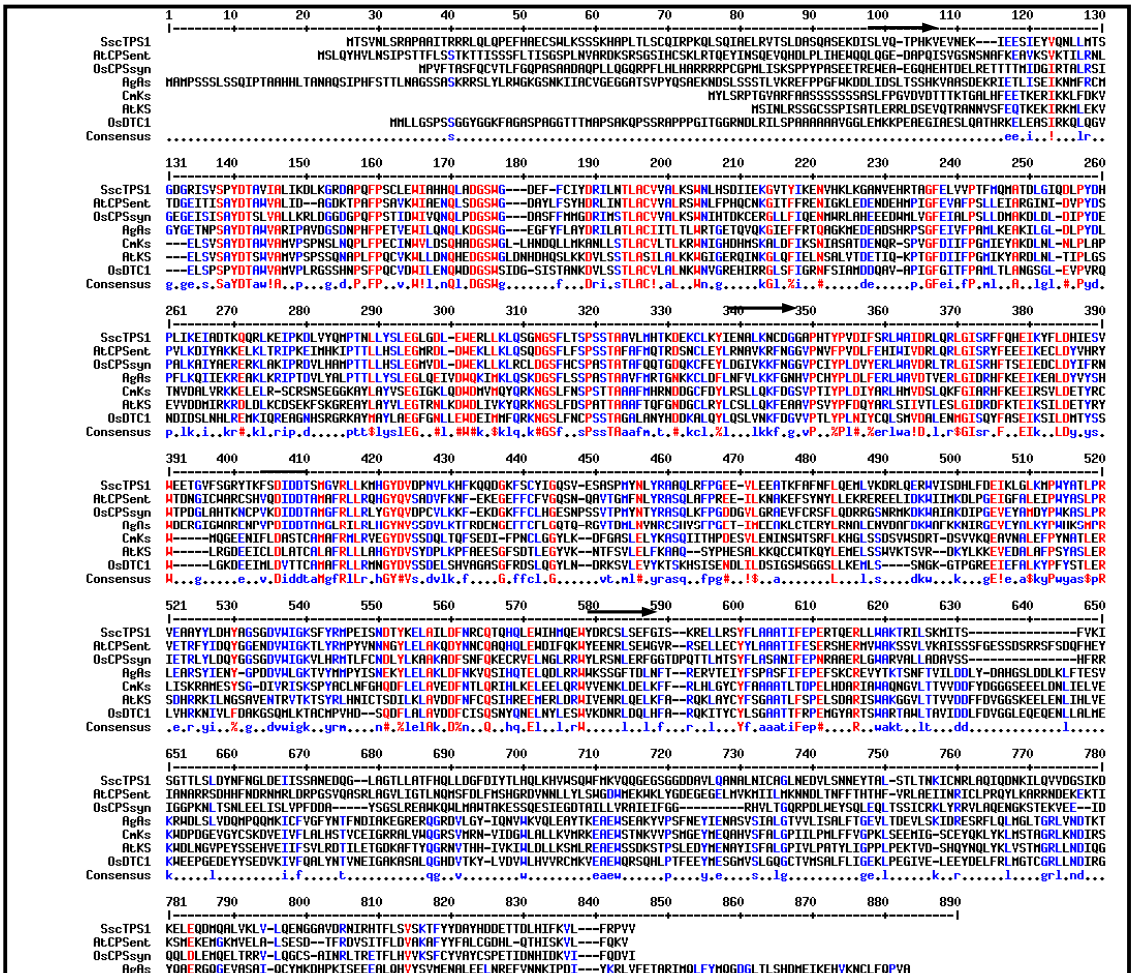


Figure 36: Sequence comparison of SscTPS1 with other diterpene synthases. This alignment represents class II terpene synthases (AtCPSent; NP_192187) and 'insertional' element containing class I terpene synthases, kaurene synthases (AtKs; Q9SAK2), from *Arabidopsis*, as well as bifunctional class II/I synthase, abietadiene synthase from *Abies grandis* (AgAs; Q38710). The boundaries used for domain comparisons are marked with arrows, and the class II DXDD motif marked with an overhead line. *Oryza sativa* syn-copalyldiphophate synthase (OsCPSsyn; AAS98158) and ent-cassa-12,15-diene synthase (OsDTC1; BAC56714), and *Cucurbita maxima* kaurene synthase (CmKs; AAB39482).

3.2.9. Size exclusion chromatography (SEC) of SscTPS1

A SEC was also used for SscTPS1, to estimate the molecular mass of a protein and to distinguish between monomeric and multimeric enzymes, which is already mentioned for the CsTPS's. ÄKTA-FPLC-system with a HiLoad Superdex 200 16/60 (GE Healthcare) column was equilibrated with the DSB-Buffer containing 150 mM NaCl and 5% glycerol instead of 10% glycerol. SscTPS1 has a calculated molecular weight of 92.94 kDa (DNAstar Inc.). SscTPS1 eluted at ~56.5 ml (**Figure 37**), checked by SDS-PAGE (**Figure 38**; the fractions 15.-17. corresponds to the elution volume of ~56.5 ml), according to the calibration curve, would have a molecular mass of ~476.5 kDa which led to the speculation that in vitro and under the buffer conditions used here, the enzymes seemed to be most likely a pentamer or at least a oligomer; the stokes radius was determined at 6.32 nm (**Table 12**; **2.11.5.)**.

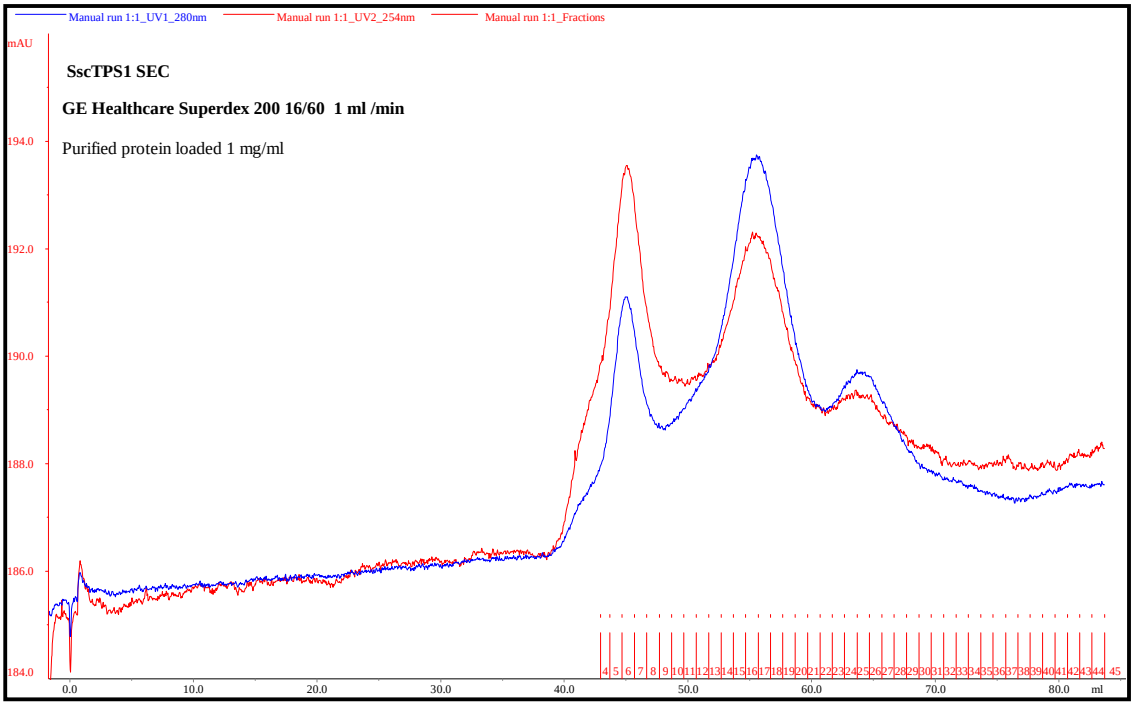


Figure 37: SEC of SscTPS1

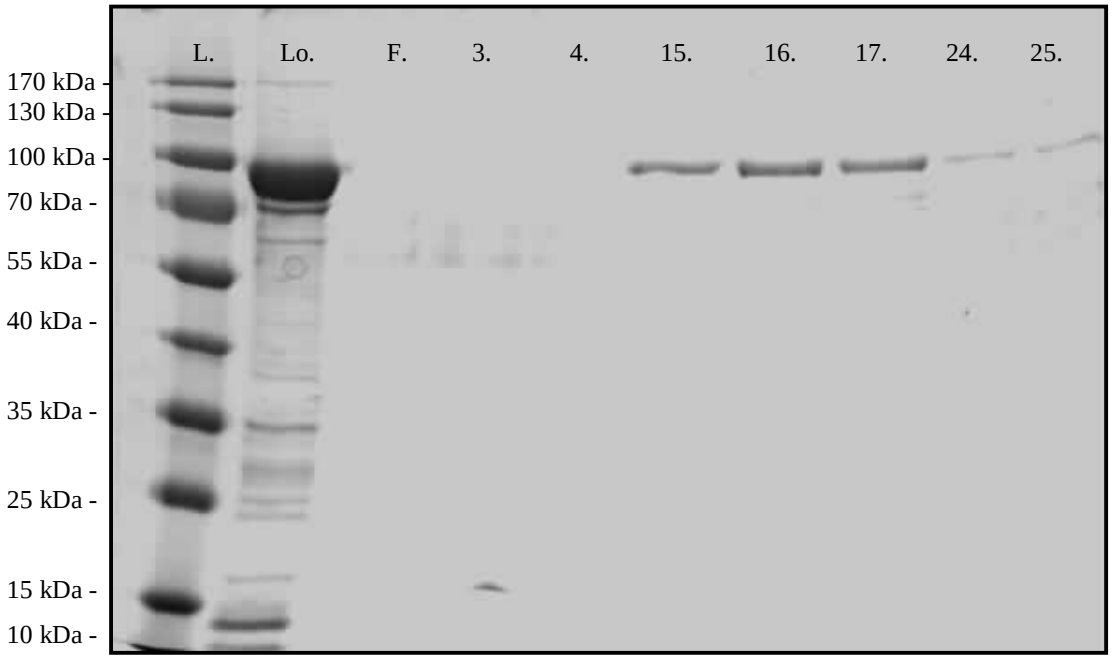


Figure 38: [L]adder, [Lo]ad, Amicon^{TM} [F]low thru and fractions 3.+4., 15.-17. and 24.+25. of the SscTPS1 SEC (Figure).

Protein	M_r [kDa]	Stokes radius [nm]	V_e [ml]	V_e/V_o	$(-\log k_{av})^{1/2}$
SscTPS1	476.5	6.32	56.5	1.361	0.850

Table 12: Determination of the calculated stokes radius and molecular weight of *S. sclarea* L. SscTPS1, according to the monoterpene synthases of *C. sativa* L. (**3.1.2.4.**)

3.2.10. Graphical overview of SscTPS1

Parameter	SscTPS1
size (nt)/truncated version	2.358 kb/2.328 kb
size (M_w)/truncated version	89.758 kDa/88,929 kDa
size (aa)/truncated version	785/776
isoelectric point (pI)	5.708
3D structure	most likely a pentamer
stokes radius	6.32 nm
phylogenetic subfamily	*Tpsc*
diterpene class	II

Table 13: Graphical overview of SscTPS1.

4. Discussion

4.1. Discussion of *C. sativa* L. monoterpene synthases

During the course of a trichome-specific secondary metabolism *Cannabis sativa* EST-project within the research group of Prof. Dr. Toni M. Kutchan and under the guidance of Dr. Jonathan E. Page, putative prenyltransferases and terpenoid synthases were found. Due to the ecological importance of terpenoids in the plant kingdom, their uses as fragrances in perfumery, as flavoring agents in the food industry, and also in pharmaceutical applications, an attempt was made to characterize these enzymes. Two putative monoterpene synthases were investigated that might be involved in the scent and defense of the *Cannabis* plant [12, 13, 14, 110]. During my diploma thesis both were cloned and expressed in *E. coli*, and afterwards the product profile was determinated. One monoterpene synthase was identified to be a (-)-limonene synthase. (-)-limonene is broadly used by industry as a flavour and fragrance compound (turpentine, orange scent), whereas its enantiomer (+)-limonene is known to have antineoplastic activities [111]. The second monoterpene synthase found was a (+)-α-pinene synthase; which is known in plants as a defense compound against predators [12]. It is shown that each CsTPS produces, in addition to a major monoterpene product, several minor products. CsTPS1 produced, besides limonene, minor amounts of α-pinene, β-pinene, β-myrcene and α-terpinolene. CsTPS2 produced α-pinene and minor amounts of β-pinene, β-myrcene and limonene, which are known from other monoterpene synthases as well. *Mentha piperita* and *Mentha spicata* limonene synthases produced α-pinene, β-pinene and myrcene as minor products and a (-)-α-terpineol synthase from *Vitis vinifera* produced in total fourteen different monoterpenes. The biosynthesis of multiple products by monoterpene synthases can be explained by the formation of the minor products from carbocationic intermediates that are formed in the course of the conversion of GPP to the various monoterpenes [112, 113, 114]. The minor products could all have the same stereochemistry, as proposed for the gymnosperm mirror image pathway of monoterpene biosynthesis [66] or be enantiomeric mixtures, as shown for the angiosperm (-)-α-terpineol synthase from *V. vinifera* [114]. The accumulation of terpenoids in the plant family of the *Cannabaceae* was localized to the glandular trichomes. This tissue specificity was previously reported for angiosperms from other species, for example the glandular trichomes of *Mentha x piperita* [115], the trichomes of

Nicotiana tabacum [116], the scent producing floral tissues of *Clarkia breweri* [117] and the flower anthesis and leaves of *Citrus unshiu* [118]. A *de novo* formation of traumatic resin ducts in the wood and the specific production of terpenoids in these ducts were reported for gymnosperms such as *Picea abies* [12]. Within the tissue, the terpene synthase expression was organelle-specific; the monoterpenoids and diterpenoids were produced in the plastids [119, 120]. This is similar to all other published monoterpene synthases that contain an N-terminal transit peptide. The nuclear-encoded monoterpene synthases are preproteins that are proteolytically processed into their mature form upon transport into the plastid. All terpene synthase transit peptides are rich in serine and threonine, low in acidic and basic aa, and they are about 45-70 aa long. Both of the characterized *C. sativa* terpene synthase N-terminal transit peptides are consistent with these characteristics, but show little aa sequence identity to other transit peptides [85]. During my PhD thesis, I have reported the first characterization of enzymes involved in terpenoid metabolism of *C. sativa* L. with respect to basic kinetic characteristics, molecular mass and the Stokes-, or hydrodynamic radius. The phylogenetic (identity tree) integration of these enzymes in the plant kingdom is also represented. The kinetic parameters of previously characterized terpene synthases varied between the native form, full-length preprotein and truncated versions as has been shown for a limonene synthase of *Mentha spicata*. The concentrations of divalent metal cations in the buffer can also affect the enzymatic activities, as shown with various monoterpene synthases of *Citrus limon* [70, 85]. The kinetic parameters of limonene synthases varied between 0.7 μM and 32.4 μM, which lies within the range of monoterpene synthases reported here [70, 85, 112, 121]. The applied pH was in the range of previously characterized monoterpene synthases from other angiosperm species such as *Thymus vulgaris*, *Mentha piperita*, *Ricciocarpos natans*, *Citrus limon*, *Mentha citrata* [122, 85, 112, 121, 123]. The pH optima of gymnosperms such as *Pinus taeda* are generally around pH 7.5 [71]. The predicted isoelectric points are within 6.0-7.0, somewhat higher than known monoterpene synthases, which have mostly isoelectric points in the range of 4.0-6.0 [112, 118, 124]. Little has been done to characterize temperature optima of terpene synthases. Protein extracts of leaves of *Picea abies* and *Quercus ilex* demonstrated monoterpene production at 40°C, determined in a range of 10°C – 60°C [125]. A temperature optimum of a pinene cyclase from *Abies grandis* has

been reported at 42°C [126]. The temperature optimum of CsTPS1c was in the range of those previously reported, whereas the temperature optimum of CsTPS2c was determined at a lower temperature. Past 50°C, pipetting of the hexane was problematic, because of evaporation of the organic phase. After this project was ended, the advice of Prof. Dr. M.H. Zenk to avoid this problem by adding mineral oil came unfortunately too late. The results above including an identity tree, discussed below, are published in Günnewich *et al.* 2007 [95]. Monoterpene synthases could exist either in monomers or homodimers. Known monoterpene synthases with molecular weights between 80 kDa – 100 kDa appear to be dimers of two identical subunits, whereas synthases, such as limonene synthases with molecular weights of approximately 40 kDa – 60 kDa, are apparently monomers [61, 127]. Due to the fact that both CsTPS's are around ~70 kDa in molecular weight, a size exclusion chromatography (SEC) was done to analyze the oligomeric state. It could be shown that both enzymes are homodimers under the buffer conditions used. The stokes- or hydrodynamic radiuses were determined at 4.47 nm according to Andrews 1965 [101] and Laurent & Killander 1964 [102]. As a result of the SEC the kinetic data of the dimers was rechecked on the basis of a sigmoid behavior. The kinetic properties change slightly, but they are still in the range of data known from other monoterpene synthases already discussed, but GPP seems to be a slightly better substrate for CsTPS2c than for CsTPS1c. Previous studies of the identity tree origin and sub-grouping of plant terpene synthases were based on a minimum sequence identity of 40 % along these groups; it was shown here that the terpene synthases were grouped into the *Tpsb* family alongside other angiosperm monoterpene synthases. The dendrogram analysis in **Figure 18** showed that abietadiene synthase (*A. grandis*) and a taxadiene synthase (*T. brevifolia*) grouped within the *Tpsd* family, contrary to other comparisons that placed them alongside other gymnosperm terpene synthases [61]. As shown previously, the clustering of both terpene synthases in the sequence analysis was according to their plant family and not with limonene and pinene synthases from other angiosperm species [85]. This clustering based on plant family has been reported for *Citrus*, *Arabidopsis*, *Mentha*, *Artemisia* and *Salvia* monoterpenes synthases and has been described as an interesting parallel molecular evolution for these plants [118]. This evolution of the monoterpene synthases and clustering is described to have occurred in ancient times. For example, four monoterpenes

synthase lemon (*Citrus*) genes exist whereas only two at a time cluster together [85]. Two cluster with monoterpenes synthase genes of *Artemisia* and two with a monoterpene synthase of *Quercus*; it is supposed that divergence occurred before *Quercus* and *Artemisia* separated from *Citrus*. One can see that the identities of both lemon gene clusters with respect to their N-terminal target sequence and the gene coding sequence is relatively low. This tight clustering leads to very specific functional roles in different plants [85, 118, 128]. Further study should address the functional identification of any remaining *C. sativa* monoterpene synthases. In our research, only seven different monoterpenes have been found in *C. sativa* and CsTPS1 and CsTPS2 are likely to produce five of these compounds [95] this leads to the conclusion that at least one additional monoterpene synthase is present in the plant. The unravelling of monoterpene biosynthesis will contribute to our understanding of the complex fragrance components produced by *C. sativa* [110]. For other projects both enzymes were provided to the working group of Prof. Dr. Ludger A. Wessjohann (department head of Bioorganic Chemistry; IPB). A modelling study and site-directed-mutagenesis with CsTPS1 and CsTPS2 was done by Lars Bräuer (Dept. PD Dr. W. Brandt, IPB) during his PhD thesis [98].For the computer modelling, a known bornyl diphosphate synthase X-ray 3D structure from *Salvia officinalis* L., with ~44% sequence identity to CsTPS1 and CsTPS2, was used. The C-terminus is known to be the catalytic site, whereas the function of the N-terminus is not unravelled. The active site was determined by a comparison of the crystal structure and the DDXXD-motif, the binding site of the divalent-metal-ions (Mg^{2+}, Mn^{2+}, etc.). The active site pocket of the CsTPSs was calculated by docking programs, using the GPP substrate. The calculated complexation of the GPP was done over the conserved aa arginine and aspartate, beside the DDXXD-motif. The unpolar moiety of the GPP was stabilizing over the π-electron-system of tryptophan. Beside this, other aa site-chains were determined to be involved in the catalytic mechanism. It could be shown that the active site of CsTPS1 compared to CsTPS2 is differing only at one position (CsTPS1: C343, CsTPS2: A337); this alanine might be responsible for the formation of pinene instead of limonene. The template differs in 3 positions (F578, S320 and I344) (**Figure 39**). Aim of the project was to verify the computer model by site-directed-mutagenesis of the putative in the catalysis involved amino acids. For this the active site of CsTPS1 was mutated by single-,

double- and triple-mutations to mimic the one of CsTPS2 or the bornyl diphosphate synthase. The changes were detected by the alteration of the product formation. During this research traces of more side products of CsTPS1 compared to the already known products, were discovered. CsTPS1 produce traces of linalool, fenchol, *trans*-pinan-2-ol, *cis*-pinan-2-ol, α-terpineol, geraniol, camphen, menthenol, and *trans*-citral. After treatment of the assay with wheat germ acid phosphatase, bornyl diphosphate (borneol) was discovered, too. The conclusion of remaining monoterpene synthases above still persists, although we found new monoterpenes here, dominant products are still missing. Site-directed-mutagenesis constructs could show that the product profile changed from limonene to linalool, which is a sign of involvement of the researched aa of the putative active site. A change in the direction of the product profile to CsTPS2 or the bornyl diphosphate synthase was not determined; it is most likely that more aa are involved in product formation. The catalytic mechanism of the cyclization reaction from monoterpene synthases could be more enlighten. A putative catalytic diade, of an aspartate and a histidine, is involved by increasing the histidine basicity. This leads to a deprotonation to a cation and the release of the diphosphate-moiety, complexed to an arginine. The postulated allyl-rearrangement of the GPP leads now to the linalyl-diphosphate. The second elimination of the diphosphate moiety by the arginine resulted in a cyclization to the terpinyl-cation, and a subsequent deprotonation by the histidine leads to the end product limonene. The lack of a histidine (having instead phenylalanine; F578; Figure) by the bornyl diphosphate synthase tends to result in no diphosphate elimination and deprotonation. Due to later results of a different subunit architecture, *C. sativa* L. monoterpene synthases seem to be dimeric, the model is currently recalculated (Diana Schulz; Dept. PD Dr. W. Brandt, IPB). Another study is a substrate specificity study to investigate the capability of both enzymes to convert artificial GPP like substrates (Roman Weber; Dept. Prof. L. A. Wessjohann, IPB).

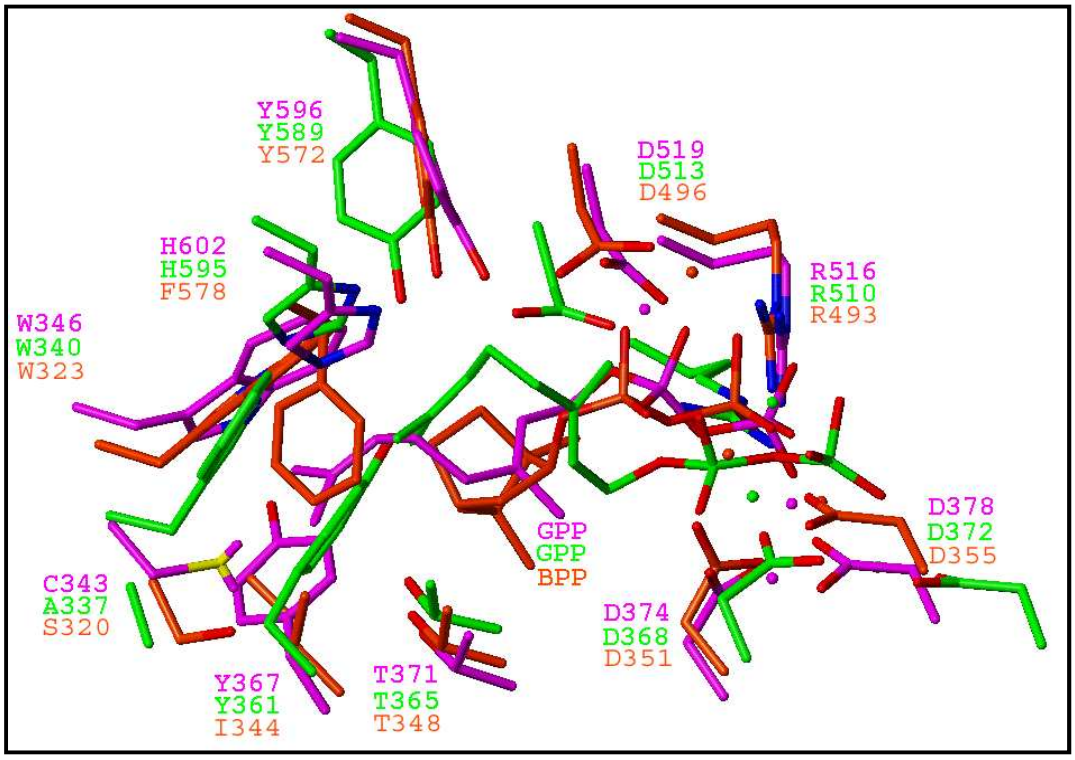


Figure 39: Calculated active site with the amino acids, of CsTPS1 (limonene synthase), CsTPS2 (pinene synthase) and the bornyl-diphosphate synthase (*Salvia officinalis*)

4.2. Discussion of *S. sclarea* L. diterpene synthase

During another course of a trichome-specific, secondary metabolism *S. sclarea* L. EST-project, started by Dr. Kum-Boo Choi, putative terpenoid synthases were found, too. Due to the economical importance of sclareol as a fragrance in perfumery, an attempt was made to characterize its biosynthetic pathway. This project was continued. Sclareol is widely used in the industry as a fragrance and flavor starting compound (ambrox®) it resembles the odor of the legendary amber or ambergris (33 tons of sclareol are needed per year), which is a pathological metabolite of the sperm whale intestines, a compound which has had a high value in perfumery since antiquity [129, 130, 131, 132]. Sclareol and its derivatives have antibacterial properties [133, 134, 135] and have been shown more and more to be active against several types of cancer [136, 137, 138]. Herein, we have reported the first cloning and characterization of an enzyme involved in terpenoid metabolism of *S. sclarea* L. with respect to tissue specificity, sequence comparisons, and identity tree integration of these enzyme in the plant kingdom. The full-length clone of the putative diterpene synthase EST, which had a high homology to known copalyl pyro-

phosphate synthases, was obtained by RACE-PCR. Copalyl pyrophosphate is a possible intermediate in the sclareol biosynthesis. In the past, mainly the expression of truncated terpenoid synthase full-length clones were successful [70, 85]. This was confirmed here also for the putative diterpene synthase of *S. sclarea* L.. The truncation was done preceding an internal R_3-motif, which is not a characterisic motif of diterpene synthases compared to those of the monoterpene synthases. Some diterpene synthases contain RR_X-motifs, some not (**Figure 36**). Until know this issue is not discussed, and a conclusion about this cannot yet be reached. The putative diterpene synthase named SscTPS1, shows the same characteristics (nucleotide and amino acid sequence length, molecular weight and pI) such as known diterpene synthases [109, 139, 140, 141, 142, 143]. Recombinant SscTPS1was assayed using the general terpene synthase substrates (GPP, NPP, FPP and GGPP). In general, NPP and linalyl pyrophosphate (LPP; not available, not used), isomers of GPP and intermediates in the monoterpene biosynthesis, were often used for the determination of stereochemical distinct cyclization reactions of monoterpene synthases. NPP was taken here to show the capability of SscTPS1 to use different substrates, not for studies about different stereochemistry [144]. It could be shown that the enzyme has a promiscuous behavior with respect to the substrates; mono-, sesqui- and diterpenes are formed. All the terpenoid products formed here are known from wild growing *Salvia sclarea* L. etheric oil extracts [145, 146, 147]. Although nerol-acetate (**13**) is known from the etheric oil, it could be here an artefact incurred by the extraction method. The acetate of the sodium acetate buffer, used for the wheat germ acid phosphatase extraction, could have formed an ester with cis-geraniol (**11**). Just as well α-pyronene (**8**) could be an artefact too, it is known from the etheric oil too, but in general it is known to be a thermolysis product of α-pinene. This raises the question whether is it formed by lower temperature during the GC-MS than described in the literature [107]. A thermolytic isomerization is also possible for α-phellandrene (**14**), α-terpinene (**15**) and γ-terpinene (**16**). That terpene synthases could use different substrates is already known from a sesquiterpene synthase from *A. thaliana* (At5g44630) that uses GPP to form in addition to sesquiterpenes, seven different monoterpenes like myrcene, limonene, (Z)-β-ocimene, (E)-β-ocimene, terpinolene, α-terpinolene and one unidentified compound. The formation of theses compounds by the enzyme *in vivo* is rather unlikely, as the protein lacks an N-terminal transit peptide

and therefore is not expected to be present in plastids, where GPP is thought to be produced. This conclusion is supported by analysis of an At5g44630 T-DNA insertion line, that showed no difference in monoterpene emission profiles in comparison to the wild type [148]. Contrary to this, in *Fragaria* terpene synthases are located in the cytosol/plastids/mitochondria depending on the isoform (containing or not containing N-terminal transit sequences) and synthesize a monoterpenoid (linalool) and a sesquiterpenoid (nerolidol), already mentioned in chapter **1.5.**. A transformation of *A. thaliana* with the *Fragaria* synthase (*FaNES1*), which is located to the cytosol, leads to mono- and sesquiterpenoid formation *in vivo,* and feeding labelled 1-deoxy-D-xylulose did not result in labelled linalool in *Fragaria* [84]. It could be suggested, that we should be careful to give statements about compartmentation of terpenoid production *in vivo;* it could be that one enzyme like SscTPS1 is responsible for the mono-, sesqui-, and diterpene production in *S. sclarea* L.. All needed substrates could be synthesized in the plastid where SscTPS1 is located, GPP and GGPP are endogenous, known FPP synthases are located into the plastid as already mentioned in chapter **1.5.**, too [81]. It is possible that FPP is freely available from GGPP synthases located in the plastids, based on product-substrate changes in the equilibrium and competition of the active site with the IPP and DMAPP substrates. The extraction methods used here are common for terpenoid synthases; the resulting solvolysis products of any allylic diphosphate esters are present (i.e. geraniol and farnesol) [78, 79]. Both were used here to investigate differences between the acid and the phosphatase procedure. One of the methods used here is necessary, because unsolvolized radioactive assays resulted in almost no radioactivity in the organic phase (data not shown). Other methods do not work with solvolysis, here only the hydrophobic products should be trapped in an organic phase. $MgSO_4$ columns are often use to trap the remaining aqueous components or hydroxylated substrates (geranylgeraniol i. e.). A problem of this procedure is that the $MgSO_4$ matrices could lead to rearrangements [67]. In general, we can see that the phosphatase treatment yielded a wider distribution in the product formation. It is assumed that (13R)- and (13S)-manoyl oxide (**22** and **23**) are derived from sclareol (**27**). The OH group C13 of sclareol might be activated to an anion, most likely by the buffer, and a following nucleophilic attack of it at C8, under water elimination, leads to both manoyl oxides (**22**, **23**). An enzymatic reaction should be excluded due to the fact

that we have a consistent stereocenter at C8 (personal communication Prof. B. Westermann; Dept. Prof. L. Wessjohann, IPB), and that SscTPS1 belongs to the class II terpene synthases. For further cyclizations, to manoyl oxide i. e., a class I terpene synthase like ent-kaurene synthases or a bifunctional class I/II terpene synthase like an abietadiene synthase is required [78]. In the beginning of our research the biosynthetic pathway to sclareol was proposed to proceed through the formation of copalyl diphosphate and its following allyl rearrangement and a dephosphorylation and hydroxylation by CytP450 enzymes. Nevertheless other possibilities where discussed in the literature. The biosynthesis of sclareol from mevalonate, acetate, and CO_2 in sage leaves has been reported by Nicholas 1964 [149], and the direct formation of manool and sclareol via cyclization of geranylgeraniol has been proposed by Geissman and Crout 1969 [150] and Ruzicka et al. 1994 [151], but the speculative scheme lacked experimental evidence. The biosynthesis of cis-abienol was shown to be done by a direct cyclization of GGPP [152], the same group showed in cell free extracts of trichome bearing tobacco tissue the direct cyclization of GGPP to sclareol [116]. We can now verify their research for the first time by showing direct cyclization with a recombinant sclareol synthase, without any CytP450 enzymes involved in this process. That involvement of phosphatases could not be excluded as discussed above. An exclusively trichome specific expression could not be shown, we found a slightly expression in other tissues as well. The RNA-blot (**Figure 20**) was supported by an RT-PCR. That this low expression could be come from remnant trichomes on the analyzed tissues could be excluded, as shown for C. sativa monoterpene synthases [95]. Contrary to the findings here, Schmiderer *et al.* 2008 reported an exclusively trichome specific incurrence of sclareol. It could be shown that capitate trichomes contain sclareol beside linalool and linalyl acetate and peltate ones accumulate noticeable concentrations of sesquiterpenoids [153]. The commercial aspect of this thesis is undeniable. In the late nineties the *S. sclarea* market, as a volatile oil producing plant, had a value of 5.4 million $US [1]. When this project started, there was a need of 33 tons sclareol per anno to yield 20 tons Ambrox® [132]. In the past, the biotechnological production of plant secondary metabolites was based on tissue and cell culture engineering. Later modern approaches to the genetic manipulation of volatile and medicinal plants where done, by *Agrobacterium rizhogenes* and *tumefaciens* transformation for example

[1]. In the future, after the kinetic characterization of SscTPS1, it is possible to establish a fermentative sclareol production. This could lead to a decrease in the waste production during the synthesis of Ambrox®. From the planting of S. sclarea L. to the end product 206 tons waste are produced to yield 1 ton of Ambrox® [132]. That sclareol is a natural product in *S. sclarea* is undeniable [153] but the final question is, is a phosphatase involved in the biosynthetic pathway to sclareol or not. If not we have to search for the occurrence of non-enzymatically dephosphorylations in the cell. If yes we have to search for a phosphatase to unravel the biosynthetic pathway of sclareol fully. But for the technical production this doesn't matter, with the addition of hydrochloric acid to the SscTPS1 reaction, this problem is circumvented.

The aim of my thesis was to unravel the biosynthetic pathway of sclareol in *Salvia sclarea* L., this is hereby done. And I can report here that we found a novel diterpene synthase the first sclareol synthase, so far I know.

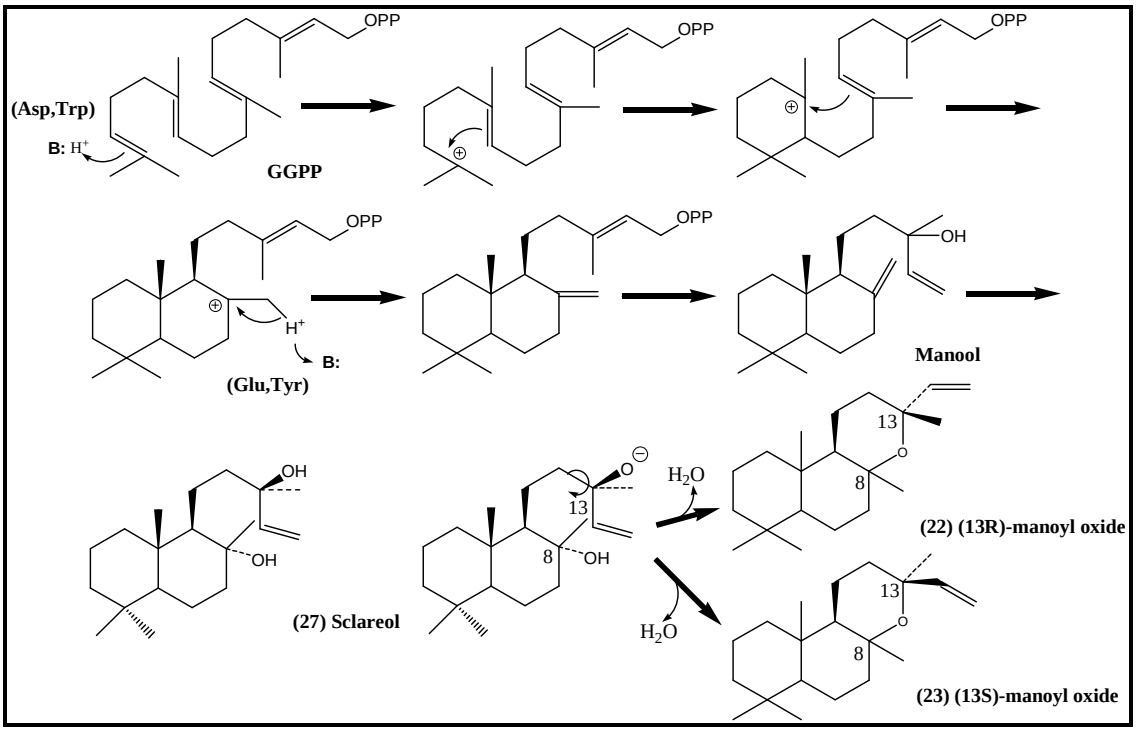


Figure 40: Proposed biosynthetic pathway of sclareol and its derivatives of manoyloxide.

5. Summary

5.1. Summary of *C. sativa* L. monoterpene synthases

Two recombinant, stereospecific and dimeric monoterpene synthases, a (-)-limonene synthase (*CsTPS1*) and a (+)-α-pinene synthase (*CsTPS2*), encoded by *Cannabis sativa* L. cv. Skunk trichome mRNA, have been isolated and characterized. Recombinant CsTPS1 shows a K_m value at 6.25 µM and a K_{cat} at 0.09 s^{-1}, the pH optimum was determined at pH 6.5, and a temperature optimum at 40°C. Recombinant CsTPS2 shows a K_m values at 4.96 µM and a K_{cat} at 0.14 s^{-1}, the pH optimum was determined at pH 7.0, and a temperature optimum at 30°C. The Stokes radius was determined at 4.47 nm. Phylogenetic analysis showed that both CsTPSs group within the angiosperm family and belong to the *Tpsb* subgroup of monoterpene synthases. The enzymatic products (-)-limonene and (+)-α-pinene were detected as natural products in *C. sativa* trichomes.

5.2. Summary of *S. sclarea* L. diterpene synthase

One recombinant oligomeric (most likely a pentamer) diterpene synthase, a sclareol synthase (*SscTPS1*), encoded by Salvia sclarea L. cv Trakystra trichome mRNA, has been isolated and partially characterized for the first time. It could be shown that SscTPS1 has a promiscous behavior in its ability to use different substrates. Beside GGPP as the substrate for diterpene synthases, SscTPS1 is capable to produce various mono- and sesquiterpenoids, out of GPP, NPP and FPP. The stokes radius was determined at 6.32 nm. Phylogenetic analysis showed that SscTPS1 group within the angiosperm family and belongs to the *Tpsc* subgroup and is a Class II diterpene synthase. The enzymatic products are known as natural products of *S. sclarea L.*

6. Literature (cited by numbers)

[1] Hay RKM, Waterman PG. **(1993)** Volatile oil crops: their biology, biochemistry and production. Longman Scientific & Technical. ISBN 0-582-00557-4.

[2] Croteau R. **(1984)** Biosynthesis and catabolism of monoterpenes. In Ness W.D., Fuller G. and Tsai L.S. (eds.). Isopentenoids in plants biochemistry and function. New York, Marcel Dekker, 31-64.

[3] Flemming MP, Clarke RC. **(1998)** Physical evidence for the antiquity of *Cannabis sativa* L. (Cannabaceae). *Journal of the International Hemp Association*. **5**, 80-92.

[4] Clarke RC. (2000) Haschisch: Geschichte, Kultur, Inhaltsstoffe, Genuss, Heilkunde, Herstellung. AT Verlag, Aarau, Schweiz. ISBN 3-85502-669-6

[5] Mechoulam R**.** Marihuana chemistry. **(1970)** *Science,* **168,** 1159-166.

[6] Di Marzo V, De Petrocellis L, Bisogno T, Maurelli S. **(1995)** Pharmacology and physiology of the endogenous cannabimimetic mediator anandamide. *Journal of Drug Development and Clinical Practice* **7**, 199-219.

[7] Devane WA, Dysarz FA, Johnson MR, Melvin LS, Howlett AC. **(1988)** Determination and characterization of a cannabinoid receptor in rat brain. *Molecular Pharmacology,* **34,** 605-613.

[8] Baker D, Pryce G, Croxford JL, Brown P, Pertwee RG, Huffman JW, Layward L. **(2000)** Cannabinoids control spasticity and tremor in a multiple sclerosis model. *Nature,* **404**, 84-87.

[9] Clebsch B. **(2003)** The New Book of Salvias; sages for every garden. Timber Press, Inc. ISBN 0-88192-560-8 (hb).

[10] Croteau R, Kutchan TM, Lewis N. **(2000)** Natural products (secondary metabolism). In BB Bucchanan, W Gruissem, RL Jones, eds, Biochemistry and Molecular Biology of Plants. American Society of Plant Biologists, Rockville, MD, pp 1250-1318.

[11] R.A. Hill Dictionary of natural products on CD-ROM, (**2002**) Ed version **10:2**, Chapman & Hall/CRC.

[12] Martin D, Tholl D, Gershenzon J, Bohlmann J. **(2002)** Methyl jasmonate induces traumatic resin ducts, terpenoid resin biosynthesis, and terpenoid accumulation in developing xylem of Norway spruce stems. *Plant Physiology*, **129,** 1003-1018.

[13] Schnee C, Köllner TG, Held M, Turlings TCJ, Gershenzon J, Degenhardt J. **(2006)** The products of a single maize sesquiterpenes synthase from a volatile defense signal that attracts natural enemies of maize herbivores. *Proceedings of the National Academy of Sciences* USA, **103(4),**1129-1134.

[14] Kessler A, Bladwin IT. **(2002)** Plant responses to insect herbivory: The emerging molecular analysis. *Annual Review of Plant Biology,* **53,** 299-328.

[15] McGarvey DJ, Croteau R. **(1995)** Terpenoid metabolism. *Plant Cell,* **7,** 1015-1026.

[16] Zimmerman PR, Chatfield RB, Fishman J, Crutzen PJ, Hanst PL. **(1978)** Estimates on the production of CO and H_2 from the oxidation of hydrocarbon emissions from vegetation. *Geophysical Research Letters,* **5,** 679-682.

[17] Qureshi N, Porter JW. **(1981)** Conversion of acetyl-coenzyme A to isopentenyl pyrophosphate. In Biosynthesis of Isoprenoid Compounds, **Vol. 1**, Porter JW. and Spurgeon SL. (eds), New York: Jon Wiley, 47-94.

[18] Chappel J. **(1995)** Biochemistry and molecular biology of isoprenoid biosynthetic pathway in plants. *Annual Review of Plant Physiology Plant Molecular Biology,* **46,** 521-547.

[19] Banthorpe DV, Charlwood BV, Francis MJO. **(1972)** The biosynthesis of monoterpenes, *Chemical Reviews,* **72,** 115-155.

[20] Beyer P, Kreuz K, Kleinig H. **(1980)** β-Carotene synthesis in isolated chromoplasts from Narcissus pseudonarcissus. *Planta,* **150,** 435-438.

[21] Schulze-Siebert D, Schultz G. **(1987)** Full autonomy in isoprenoid synthesis in spinach chloroplasts. *Plant Physiology Biochemistry,* **25,** 145-153.

[22] Flesch G, Rohmer M. **(1988)** Prokaryotic hopanoids: the biosynthesis of the bacteriohopane skeleton. *European Journal of Biochemistry*, **175,** 405-411.

[23] Rohmer M, Knani M, Simonin P, Sutter B, Sahm H. **(1993)** Isoprenoid biosynthesis in bacteria: a novel pathway for the early steps leading to isopentenyl diphosphate. *Biochemical Journal*, **295,** 517-524.

[24] Broers STJ. **(1994).** Über die frühen Stufen der Biosynthese von Isoprenoiden in *Escherichia coli*. Dissertation **Nr. 10978**, ETH Zürich, Schweiz.

[25] Schwarz M. **(1994)** Terpene-Biosynthese in *Ginkgo biloba*: Eine überraschende Geschichte. Dissertation **Nr. 10951**, ETH Zürich, Schweiz.

[26] White RH. **(1978)** Stable isotope studies on the biosynthesis of the thiazole moiety of thiamin in *Escherichia coli*. *Biochemistry*, **17,** 3833-3840.

[27] Hill RE, Himmeldirk K, Kennedy IA, Panloski RM, Sayer BG, Wolf E, Spenser ID. **(1996)** The biogenetic anatomy of vitamin B_6. A ^{13}C NMR investigation of the biosynthesis of pyridoxol in *Escherichia coli*. *Journal of Biological Chemistry*, **271,** 30426-30435.

[28] Disch A, Rohmer M. **(1998)** On the absence of the glyceraldehyde 3-phosphate/pyruvate pathway for isoprenoid biosynthesis in fungi and yeasts. *FEMS Microbiology Letters*., **168,** 201-208.

[29] Eisenreich W, Schwarz M, Cartayrade A, Arigoni D, Zenk MH, Bacher A. **(1998)** The deoxyxylulose phosphate pathway of terpenoid biosynthesis in plants and microorganisms. *Chemistry & Biology*, **5,** R221-R233.

[30] Proteau PJ. **(1998)** Biosynthesis of phytol in the cyanobacterium *Synechocystis* sp. UTEX 2470: utilization of the non-mevalonate pathway. *Journal of Natural Products*, **61,** 841-843.

[31] Disch A, Hemmerlin A, Bach TJ, Rohmer M. **(1998)** Mevalonate-derived isopentenyl diphosphate is the biosynthetic precursor of ubiquinone prenyl side chain in tobacca BY-2 cells. *Biochemical Journal.*, **331,** 615-621.

[32] McCaskill D, Croteu R. **(1995)** Monoterpene and sesquiterpene biosynthesis in glandular trichomes of peppermint (*Mentha x piperita*) rely exclusively on plastid-derived isopentenyl diphosphate. *Planta*, **197,** 49-56.

[33] Lichtenthaler HK. **(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.

[34] López-Garcia P, Moreira D. **(1999)** Metabolic symbiosis at origin of eukaryotes. *Trends in Biochemical Science*, **24,** 88-93.

[35] Adam KP, Thiel R, Zapp J. **(1999)** Incorporation of 1-[1-^{13}C]-Deoxy-D-xylulose in Chamomile sesquiterpenes. *Archives of Biochemistry and Biophysics*, **369,** 127-132.

[36] Piel J, Donath J, Bandemer K, Boland W. **(1998)** Mevalonate-independent biosynthesis of terpenoid volatiles in plants - induced and constitutive emission of volatiles. *Angewandte Chemie International Edition*, **37,** 2478-2481.

[37] Laule O, Fürholz A, Chang H-S, Wang X, Heifetz PB, Gruissem W, Lange MB. **(2003)** Crosstalk between cytosolic and plastidial pathways of isoprenoid

biosynthesis in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences* USA, **100,** 6866-6871.

[38] Nagata N, Suzuki M, Yoshida S, Muranaka T. **(2002)** Mevalonic acid partially restores chloroplast and etioplast development in Arabidopsis lacking the non-mevalonate pathway. *Planta,* **216,** 345–350.

[39] Bach TJ, Boronat A, Caelles C, Ferrer A, Weber T, Wettstein A. **(1991).** Aspects related to mevalonate biosynthesis in plants. *Lipids* **26**, 637-648.

[40] Weber T, Bach TJ. **(1994)** Conversation of acetyl-coenzyme A into 3-hydroxy-3-methylglutaryl-conenzyme A in radish seedlings: Evidence of a single monomeric protein catalyzing a Fe^{II}/quinine-stimulated double condensation reaction. *Biochimica et Biophysica Acta,* **1211**, 85-96.

[41] Lehninger AL, Nelson DL, Cox MM. **(1998)** Prinzipien der Biochemie. Spektrum Akademischer Verlag, Heidelberg • Berlin • Oxford

[42] Voet & Voet. **(1995)** Biochemistry, second Edition, John Wiley & Sons, Inc.

[43] Walter MH, Floß DS, Hans J, Fester T, Strack D. **(2007)** Apocarotenoid biosynthesis in arbuscular mycorrhizal roots: Contributions from methylerythritol phosphate pathway isogenes and tools for its manipulation. *Phytochemistry,* **68,** 130–138.

[44] Rodriguez-Concépcion M, Boronat A. **(2002)** Elucidation of the methylerythritol phosphate pathway for isoprenoid biosynthesis in bacteria and plastids. A metabolic milestone achieved through genomics. *Plant Physiology,* **130**, 1079-1089.

[45] Tholl D, Croteau R, Gershenzon J. **(2001)** Partial purification and characterization of the short-chain prenyltransferases, geranyl diphosphate synthase and farnesyl diphosphate synthase, from Abies grandis (Grand Fir*). Archives of Biochemistry and Biophysics,* **386,** 233-242.

[46] Poulter CD, Rilling HC. **(1978)** The prenyl transfer reaction. Enzymatic and mechanistic studies of the 1'-4 coupling reaction in the terpene biosynthetic pathway. *Accounts of Chemical Research,* **11,** 307-313.

[47] Koyama T. **(1999)** Molecular analysis of prenyl chain elongation enzymes. *Bioscience, Biotechnology, and Biochemistry,* **63,** 1671-1676.

[48] Burke CC, Wildung MR, Croteau R. **(1999)** Geranyl diphosphate synthase: Cloning, expression and characterization of this prenyltransferase as a heterodimer. *Proceedings of the National Academy of Sciences* USA, **96,** 13062-13067.

[49] Burke C, Croteau R. **(2002)** Interaction with the small subunit of geranyl diphosphate synthase modifies the chain length specifity of geranylgeranyl diphosphate synthase to produce geranyl diphosphate. *The Journal of Biological Chemistry*, **277,** 3141-3149.

[50] Tholl D, Kish CM, Orlova I, Sherman D, Gershenzon J, Pichersky E, Dudareva N. **(2004)** Formation of monoterpenes in Anthirrum majus and Clarkia breweri flowers involves heterodimeric geranyl diphosphate synthases. *The Plant Cell*, **16,** 977-992.

[51] Okada K, Suzuki K, Kamiya Y, Zhu X-F, Fujisaki S, Nishimura Y, Nishino T, Nakagawa T, Kawamukai M, Matsuda H. **(1996)** Polyprenyl diphophate synthase essentially defines the length of the side chain of ubiquinone. *Biochimica et Biophysica Acta*, **1302,** 217-223.

[52] Dewick PM. **(1999)** The biosynthesis of C5-C25 terpenoid compounds. *Natural Products Reports*, **16,** 97-130.

[53] Lohmann A, Schöttler MA, Bréhélin C, Kessler F, Bock R, Cahoon EB, Dörmann P. **(2006)** Deficiency in phylloquinone (Vitamin K_1) methylation affects prenyl quinone distribution, photosystem I abundance, and anthocyanin accumulation in the *Arabidopsis AtmenG* mutant. *The Journal of Biological Chemistry*, **281,** 40461-40472.

[54] Dogbo O, Bardat F, Quennemet J, Camara B. **(1987)** Metabolism of plastid terpenoids: in vitro inhibition of phytoene synthesis by phenetyl pyrophosphate derivatives. *FEBS Letters*, **210,** 211-215.

[55] Hunter SC, Cahoon EB. **(2007)** Enhancing Vitamin E in Oilseeds: Unravelling tocopherol and tocotrienol biosynthesis. *Lipids*, **42,** 97-108.

[56] Cheng Z, Sattler S, Maeda H, Sakuragi Y, Bryant DA, DellaPenna D. **(2003)** Highly divergent methyltransferases catalyze a conserved reaction in tocopherol and plastoquinone synthesis in cyanobacteria and photosynthetic eukaryotes. *The Plant Cell*, **15,** 2343–2356.

[57] Michal G. **(1999)** Biochemical Pathways. Spektrum Akademischer Verlag, Berlin, ISBN 3-86025-239-9.

[58] Fellermeier M, Zenk MH. **(1998)** Prenylation of olivetolate by a hemp transferase yields cannabigerolic acid, the precursor of tetrahydrocannabinol. *FEBS Letters*, **427,** 283-285.

[59] Zuurbier KWM, Fung SY, Scheffer JJC, Verpoorte R. **(1998)** *In-vitro* prenylation of aromatic intermediates in the biosynthesis of bitter acids in *Humulus lupulus*. *Phytochemistry*, **49,** 2315-2322.

[60] Maurer-Stroh S, Washietl S, Eisenhaber F. **(2003)** Protein prenyltransferases. *Genome Biology* **4**, 212

[61] Bohlmann J, Mayer-Gauen G, Croteau R. **(1998b)** Plant terpenoid synthases: Molecular biology and phylogenetic analysis *Proceedings of the National Academy of Sciences USA*, **95,** 4126-4133.

[62] Cane DE. **(1990)** Enzymatic formation of sesquiterpenes. *Chemical Reviews*, **90**, 1089-1103.

[63] Haudenschild C, Schalk M, Karp F, Croteau R. **(2000)** Functional expression of regiospecific cytochrome P450 limonene hydroxylases from mint (*Mentha* spp.) in *Escherichia coli* and *Saccharomyces cerevisiae*. *Archives of Biochemistry and Biophysics*, **379,** 127-136.

[64] Nualkaew N, Guennewich N, Springob K, De-Eknamkul W, Zenk MH, Kutchan TM. **(2007)** cDNA cloning of prenyl diphosphate phosphatase from *Croton stellatopilosus* Ohba. *Planta Medica*, **73,** 1019-1020.

[65] Jasiński M, Stukkens Y, Degand H, Purnelle B, Marchand-Brynaert J, Boutry M. **(2001)** A plant plasma membrane ATP binding cassette-type transporter is involved in antifungal terpenoid secretion. *The Plant Cell*, **13,** 1095-1107.

[66] Hyatt DC, Croteau R. **(2005)** Mutantional analysis of monoterpene synthase reaction: altered catalysis through direct mutagenesis of (-)-pinene synthase from *Abies grandis*. *Archives of Biochemistry and Biophysics,* **439,** 222-223.

[67] Peters RJ, Flory JE, Jetter R, Ravn MM, Lee HJ, Coates RM, Croteau RB. **(2000)** Abietadiene synthase from grand fir (*Abies grandis*): Characterization and mechanism of action of the "pseudomature" recombinant enzyme. *Biochemistry*, **39,** 15592-15602

[68] Bohlmann J, Steele CL, 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.

[69] Wise ML, Croteau R. **(1999)** Comprehensive Natural Products Chemistry: Isoprenoids Including Carotenoids and Steroids. D.E. Cane, editor. 2. Elsevier Science; Oxford 155-200.

[70] Williams DC, McGarvey DJ, Katahira EJ, 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.

[71] Phillips MA, Wildung MR, Williams DC, Hyatt DC, Croteau R. **(2003)** cDNA isolation, functional expression, and characterization of (+)-α-pinene synthase and (-)-α-pinene synthase from loblolly pine (*Pinus taeda*): stereocontrol in pinene biosynthesis. *Archives of Biochemistry and Biophysics*, **411,** 267-276.

[72] Rivera SB, Swedlund BD, King GJ, Bell RN , Hussey Jr CE, Shattuck-Eidens DM, Wrobel WM, Peiser GD, Poulter CD. **(2001)** Chrysanthemyl diphosphate synthase: Isolation of the gene and characterization of the recombinant non-head-to-tail monoterpene synthase from Chrysanthemum cinerariaefolium. *Proceedings of the National Academy of Sciences* USA, **98**, 4373–4378.

[73] Crock J, Wildung M, Croteau R. **(1997)** Isolation and bacterial expression of a sesquiterpene synthase cDNA clone from peppermint (*Mentha x piperita*, L.) that produces the aphid alarm pheromone (E)-beta-farnesene. *Proceedings of the National Academy of Sciences* USA, **94**, 12833-12838.

[74] Cane DE, Tsantrizos YS. **(1996)** Aristolochene synthase. Elucidation of the cryptic germacrene a synthase activity using the anomalous substrate dihydrofarnesyl diphosphate. *Journal of the American Chemical Society* . **118,** 10037-10040.

[75] Cano-Camacho H, Lopez-Romero E, Lozoya-Gloria E. **(1997)** Partial purification and characterization of an elicitor stimulated sesquiterpene cyclase from chili pepper (*Capsicum annuum* L.) fruits. *Plant Science.* **124,** 23-31.

[76] Back K, Chappell J. **(1996)** Identifying functional domains within terpene cyclases using a domain-swapping strategy. *Proceedings of the National Academy of Sciences* USA, **93,** 6481-6485.

[77] Starks CM, Back K, Chappell J, Noel JP. **(1997)** Structural basis for cyclic terpene biosynthesis by tobacco 5-epi-aristolochene synthase. *Science*, **277,** 1815-1820.

[78] Xu M, Hillwig ML, Prisic S, Coates RM, Peters RJ. **(2004)** Functional identification of rice syn-copalyl diphosphate synthase and its role in initiating biosynthesis of diterpenoid phytoalexin/allelopathic natural products. *The Plant Journal*, **39,** 309-318.

[79] Peters RJ, Carter OA, Zhang Y, Matthews BW, Croteau RB. **(2003)** Bifunctional abietadiene synthase: Mutual structural dependence of the active sites for protonation-initiated and ionization-initiated cyclizations. *Biochemistry*, **42,** 2700-2707.

[80] Cunillera N, Boronat A, Ferrer A. **(1997)** The Arabidopsis thaliana FPS1 gene generates a novel mRNA that encodes a mitochondrial farnesyl-diphosphate synthase isoform. *Journal of Biological Chemistry*, **272,** 15381-15388.

[81] Sanmiya K, Ueno O, Matsuoka M, Yamamoto N. **(1999)** Localization of farnesyl diphosphate synthase in chloroplast. *Plant & Cell Physiology*, **40,** 348-354.

[82] Okada K, Saito T, Nakagawa T, Kawamukai M, Kamiya Y. **(2000)** Five geranylgeranyl diphosphate synthases expressed in different organs are localized into three subcellular compartments in Arabidopsis. *Plant Physiology*, **122,** 1045-1056.

[83] Cervantes-Cervantes M, Gallagher CE, Zhu C, Wurtzel ET. **(2006)** Maize cDNAs expressed in endosperm encode functional farnesyl diphosphate synthase with geranylgeranyl diphosphate synthase activity. *Plant Physiology*, **141**, 220-231.

[84] Aharoni A, Giri AP, Verstappen FWA, Bertea CM, Sevenier R, Sun Z, Jongsma MA, Schwab W, Bouwmeester HJ. **(2004)** Gain and loss of fruit flavor compounds produced by wild and cultivated strawberry species. *The Plant Cell*, **16,** 3110-3131.

[85] Lücker J, El Tamer MK, Schwab W, Verstappen FWA, van der Plas LHW, Bouwmeester HJ, Verhoeven HA. **(2002)** Monoterpene biosynthesis in lemon (*Citrus limon*). cDNA isolation and functional analysis of four monoterpene synthases. *European Journal of Biochemistry*, **269,** 3160-3171.

[86] Maniatis T, Fritsch EF, Sambrook J. **(1982)** Molecular cloning: A laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

[87] Hanahan D. **(1985)** Techniques for transformation of E. coli. DNA cloning: A pratical approach I. Gover DM, ed., p 109-135, IRL-Press, Oxford, Washington D.C..

[88] Laemmli UK. **(1970)** Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227,** 680-685.

[89] Chomczynski P, Sacchi N. **(1987)** A single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Analytical Biochemistry*, **162,** 156-159.

[90] Bradford MM. **(1976)** A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry*, **72,** 248-54.

[91] Mullis KB, Fallona FA. **(1987)** Specific synthesis of DNA *in vitro* via a polymerase-catalyzed chain reaction. *Methods in Enzymology*, **155**, 335-350.

[92] Sanger F, Micklen S, Coulsten AR. **(1977)** DNA sequencing with chain termination inhibitors. *Proceedings of the National Academy of Sciences* USA, **74**, 5463-5467.

[93] Gershenzon J, McCaskill D, Rajaonarivony JI, Mihaliak C, Karp F, Croteau R. **(1992)** Isolation of secretory cells from plant glandular trichomes and their use in biosynthetic studies of monoterpenes and other gland products. *Analytical Biochemistry*, **200,** 130-138.

[94] Bjerrum OJ, Schafer-Nielsen C. **(1986)** Analytical Electrophoresis, M. J. Dunn, ed., pp. 315; Verlag Chemie, Weinheim.

[95] Günnewich N, Page JE, Köllner TG, Degenhardt J, Kutchan TM. **(2007)** Functional expression and characterization of trichome specific (-)-limonene synthase and (+)-α-pinene synthase from *Cannabis sativa*. *Natural Product Communications*, **2,** 223-232.

[96] Wong KKW, Whilton NT, Cölfen H, Douglas T, Mann S. **(1998)** Hydrophobic proteins: synthesis and characterisation of organic-soluble alkylated ferritins. *Chemical Communications*, 1621-1622.

[97] Schnee C, Köllner TG, Gershenzon J, Degenhardt J. **(2002)** The maize gene *terpene synthase 1* encodes a sesquiterpene synthase catalyzing the formation of (E)-β-farnesene, (E)-nerolidol, and (E,E)-farnesol after herbivore damage. *Plant Physiology*, **130,** 2049-2060.

[98] Bräuer L. Modelling- und Mutationsstudien an ausgewählten prenylierenden Enzymen. (2006) Dissertation, Mathematisch-Naturwissenschaftlich-Technische Fakultät der Martin-Luther- Universität Halle-Wittenberg.

[99] Siegel LM, Monty KJ. **(1966)** Determination of molecular weight and frictional ratios of proteins in impure systems by use of gel filtration and density gradient centrifugation. Application to crude preparations of sulfite and hydroxylamine reductases. *Biochimica et Biophysica Acta*, **112,** 346-362.

[100] Wyss M, Pasamontes L, Friedlein A, Re´My R, Tessier M, Kronenberger A, Middendorf A, Lehmann M, Schnoebelen L, Röthlisberger U, Kusznir E, Wahl G, Müller F, Lahm HW, Vogel K, van Loon APGM. **(1999)** Biophysical characterization of fungal phytases (*myo*-inositol hexakisphosphate phosphohydrolases): Molecular size, glycosylation pattern, and engineering of proteolytic resistance. *Applied and Environmental Microbiology*, **65,** 359-366.

[101] Andrews P. **(1965)** The gel-filtration behaviour of proteins related to their molecular weights over a wide range. *Biochemical Journal*, **96,** 595-606.

[102] Laurent TC, Killander J. **(1964)** A theory of gel filtration and its experimental verification. *Journal of Chromatography*, **14,** 317-330.

[103] Dudareva N, Martin D, Kish CM, Kolosova N, Gorenstein N, Fäldt J, Miller B, Bohlmann J. **(2003)** (*E*)-β-ocimene and myrcene synthase genes of floral scent biosynthesis in snapdragon: Function and expression of three terpene synthase genes of a new terpene synthase subfamily. *The Plant Cell*, **15,** 1227-1241.

[104] Hieda T, Mikami Y, Obi Y, Kisaki T. **(1983)** Microbial tranformation of labdanes, cis-abienol and sclareol. *Agricultural Biology and Chemistry*, **47,** 243-250.

[105] Beyer & Walter. **(1998)** Lehrbuch der Organischen Chemie, S.Hirzel Verlag Stuttgart, **23. Auflage,** 714-724, ISBN 3-7776-0808-4.

[106] Connolly JD, Hill RA. **(1991)** Dictionary of Terpenoids, Chapman & Hall, **Vol 1.** + **2.,** ISBN 0-412-25770 X Three-volume set.

[107] Chibiryaev AM, Anikeev VI, Yermakova A, Mikenin PE, Koshevnikov IV, Sal'nikova OI. **(2006)** Thermolysis of α-pinene in supercritical lower alcohols. *Russian Chemical Bulletin; International Edition*, **5,** 987-992.

[108] Keegstra K, Olsen LJ, Theg SM. **(1989)** Chloroplastic precursors and their transport across the envelope membranes. *Annual Reviews of Plant Physiology and Plant Molecular Biology***. 40,** 471-501.

[109] Cho E, Okada A, Kenmoku H, Otomo K, Toyomasu T, Mitsuhashi W, Sassa T, Yajima A, Yabuta G, Mori K, Oikawa H, Toshima H, Shibuya N, Nojiri H, Omori T, Nishiyama M, Yamane H. **(2004)** Molecular cloning and characterization of a cDNA encoding ent-cassa-12,15-diene synthase, a putative diterpenoid phytoalexin biosynthetic enzyme, from suspension-cultured rice cells treated with a chitin elicitor. *The Plant Journal*, **37,** 1-8.

[110] Weiss EA. **(1997)** *Essential Oil Crops*, CAB International, Wallingford, UK.

[111] Crowell PL, Gould MN. **(1994)** Chemoprevention and therapy of cancer by d-limonene. *CRC Critical Reviews in Oncogenesis*, **5,** 1-22.

[112] Rajaonarivony JIM, Gershenzon J, Croteau R. **(1992)** Characterization and mechanism of (4S)-limonene synthase, a monoterpenes cyclase from the glandular trichomes of peppermint (*Mentha* x *piperita*). *Archives of Biochemistry and Biophysics*, **296,** 49-57.

[113] Colby SM, Alonso WR, Katahira E, McGarvey DJ, Croteau R. **(1993)** 4S-limonene synthase from the oil glands of spearmint *(Mentha spicata)* cDNA isolation, characterization, and bacterial expression of the catalytically active monoterpenes cyclase. *Journal of Biological Chemistry* **268,** 23016-23024.

[114] Martin MD, Bohlmann J. **(2004)** Identification of *Vitis vinifera* (-)-□-terpineol synthase by in silico screening of full length cDNA ESTs and functional characterization of recombinant terpene synthase. *Phytochemistry*, **65,** 1223-1229.

[115] McConkey ME, Gershenzon J, Croteau R. **(2000)** Developmental regulation of monoterpene biosynthesis in the glandular trichomes of peppermint. *Plant Physiology*, **122,** 215-223.

[116] Guo Z, Wagner GJ. **(1995)** Biosynthesis of cembratrienols in cell-free extracts from trichomes of *Nicotiana tabacum*. *Plant Science*, **110,** 1-10.

[117] Dudareva N, Ceske L, Blank VM, Pichersky E. **(1996)** Evolution of floral scent in *Clarkia:* nove1 patterns of S-linalool synthase gene expression in the *C. breweri* flower. *Plant Cell,* **8,** 1137-1148.

[118] Shimada T, Endo T, Fujii H, Hara M, Omura M. **(2005)** Isolation and characterization of (E)-beta-ocimene and 1,8 cineole synthases in *Citrus unshiu* Marc. *Plant Science*, **168,** 987-995.

[119] Turner GW, Croteau R. **(2004)** Organization of monoterpenes biosynthesis in menthe. Immunocytochemical localizations of geranyl diphosphate synthase, limonene-6-hydroxylase, isopiperitenol dehydrogenase, and pulegone reductase. *Plant Physiology*, **136,** 4215-4227.

[120] Helliwell CA, Sullivan JA, Mould RM, Gray JC, Peacock WJ, Dennis ES. **(2001)** A plastid envelope location of Arabidopsis ent-kaurene oxidase links the plastid and endoplasmic reticulum steps of the gibberellin biosynthesis pathway. *The Plant Journal*, **28,** 201-208.

[121] Adam KP, Crock J, Croteau R. **(1996).** Partial purification and characterization of a monoterpene cyclase, limonene synthase, from the liverwort *Ricciocarpos natans*. *Archives of Biochemistry and Biophysics*, **332,** 352-356.

[122] Alonso RW, Croteau R. **(1991)** Purification and characterization of the monoterpene cyclase γ-terpinene synthase from *Thymus vulgaris*. *Archives of Biochemistry and Biophysics,* **2,** 511-517.

[123] Crowell AL, Williams DC, Davis EM, Wildung MR, Croteau R. **(2002)** Molecular cloning and characterization of a new linalool synthase. *Archives of Biochemistry and Biophysics,* **405,** 112-121.

[124] Colby SM, Crock J, Dowdle-Rizzo B, Lemaux PG, Croteau R. **(1998)** Germacrene C synthase from *Lycopersicon esculentum* cv. VFNT cherry tomato: cDNA isolation, characterization, and bacterial expression of the multiple product

sesquiterpene cyclase *Proceedings of the National Academy of Sciences USA.* **95** 2216–2221.

[125] Fischbach RJ, Zimmer I, Steinbrecher R, Pfichner A, Schnitzler JP. **(2000)** Monoterpene synthase activities in leaves of *Picea abies* (L.) Karst. and *Quercus ilex* L. *Phytochemistry,* **54(3),** 257-265.

[126] Lewinsohn E, Gijzen M, Croteau R. **(1992)** Wound-inducible pinene cyclase from grand fir: purification, characterization, and renaturation after SDS-PAGE. *Archives of Biochemistry and Biophysics*, **293,** 167-173.

[127] Alonso WR, Rajaonarivony JI, Gershenzon J, 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.

[128] Lu S, Xu R, Jia JW, Pang J, Matsuda SPT, Chen XY. **(2002)** Cloning and functional characterization of a β-pinene synthase from *Artemisia annua* that shows a circadian pattern of expression. *Plant Physiology.,* **30,** 477–486.

[129] Barrero AF, Alvarez-Manzaneda EJ, Altarejos J, Salido S, Ramos JM. **(1993)** Synthesis of Ambrox® from (-)-sclareol and (+)-cis-abienol. *Tetrahedron,* **49,** 10405-10412.

[130] Müller PM, Lamparsky D. **(1991)** Perfumes: Art, Science & Technology. Amsterdam, New York, Elsevier.

[131] Cheetham PSJ. **(1993)** The use of biotranformations for the production of flavours and fragrances. *Trends in Biotechnology*, **11,** 478-488.

[132] Choi K, Gerth A, Henrich J, Kutchan TM, Wilken D. (2004) Erarbeitung von Grundlagen für ein nachhaltiges biotechnologisches Verfahren zur Herstellung von Ambrariechstoffen, NAROSSA 2004, 10. Internationaler Fachkongress für nachwachsende Rohstoffe, Magdeburg, Germany.

[133] Hayet E, Fatma B, Souhir I, Waheb FA, Abderaouf K, Mahjoub A, Maha M. **(2007)** Antibacterial and cytotoxic activity of the aceton extract of the flowers of *Salvia sclarea* and some natural products. *Pakistan Journal of Pharmaceutical Sciences* **20,** 146-148.

[134] Choudhary MI, Siddiqiu ZA, Hussain S, Atta-ur-Rahman. **(2006)** Structure elucidation and antibacterial activity of new fungal metabolites of sclareol. *Chemistry & Biodiversity,* **3,** 54-61.

[135] Jassabi AR, Zamanizadehnajari S, Azar PA, Tahara S. **(2002)** Antibacterial diterpenoids from *Astragalus brachystachys*. *Zeitschrift für Naturforschung C*, **57,** 1016-1021.

[136] Dimas K, Demetzos C, Vaos V, Ioannidis P, Trangas T. **(2001)** Labdane type diterpenes down-regulate the expression of c-Myc protein, but not Bcl-2, in human leukemia T-cells undergoing apoptosis. *Leukemia research,* **25,** 449-454.

[137] Paradissis A, Hatziantoniou S, Georgopoulos A, Psarra AM, Dimas K, Demetzos C. **(2007)** Liposomes modify the subcellular distribution of sclareol uptake in HCT-116 cancer cell lines. *Biomedicine & pharmacotherapy,* **61,** 120-124.

[138] Sashidhara KV, Rosaiah JN, Kumar A, Bid HK, Konwar R, Chattopadhyay N. **(2007)** Cell growth inhibitory action of an unusual labdane diterpene, 13-epi-sclareol in breast and uterine cancers in vitro. *Phytotherapy research,* **21,** 1105-8.

[139] Sun T, Kamiya J. **(1994)** The arabidopsis *GA1* locus encodes the cyclase ent-kaurene synthetase A of gibberelin biosynthesis. *The Plant Cell,* **6,** 1509-1518.

[140] Bensen RJ, Johal GS, Crane VC, Tossberg JT, Schnable PS, Meeley RB, Briggs SP. **(1995)** Cloning and characterization of the maize *An1* gene. *The Plant Cell,* **7,** 75-84.

[141] Yamaguchi S, Saito T, Abe H, Yamane H, Murofushi N, Kamiya Y. **(1996)** Molecular cloning and characterization of a cDNA encoding the gibberellin biosynthetic enzyme *ent*-kaurene synthase B from pumpkin (*Cucurbita maxima* L.). *The Plant Journal,* **10,** 203-213.

[142] Stofer-Vogel B, Wildung MR, Vogel G, Croteau R. **(1996)** Abietadiene synthase from grand fir (*Abies grandis*). *The Journal of Biological Chemistry,* **271,** 23262-23268.

[143] Richman AS, Gijzen M, Starratt AN, Yang Z, Brandle JE. **(1999)** Diterpene synthesis in *Stevia rebaudiana*: recruitment and up-regulation of key enzymes from the gibberellin biosynthetic pathway. *The Plant Journal,* **19,** 411-421.

[144] Croteau R, Felton MN, Wheeler CJ. **(1985)** Stereochemistry at C-1 of geranyl pyrophosphate and neryl pyrophosphate in the cyclization to (+)- and (-)-bornyl pyrophosphate. *The Journal of Biological Chemistry,* **10,** 5956-5962.

[145] Souleles Chr, Argyriadou N. **(1997)** Constituents of the essential oil of *Salvia sclarea* growing wild in greece. *International Journal of Pharmacognosy,* **35,** 218-220.

[146] Ulubelen A, Topcu G. **(1998)** Chemical and biological investigations of *Salvia* species growing in turkey. *Studies in Natural Products Chemistry,* **20,** 659-718.

[147] Hudaib M, Bellardi MG, Rubies-Autonell C, Fiori J, Cavrini V. **(2001)** Chromatographic (GC-MS, HPLC) and virological evaluations of *Salvia sclarea* infected by BBWV-I. *Il Farmaco*, **56,** 219-227.

[148] Tholl D, Chen F, Petri J, Gershenzon J, Pichersky E. **(2005)** Two sesquiterpene synthases are responsible for the complex mixture emitted from *Arabidopsis* flowers. *The Plant Journal*, **42,** 757-771.

[149] Nicholas HJ. **(1964)** Biosynthesis and metabolism of (^{14}C)sclareol. *Biochimica et Biophysica Acta*, **84,** 80-90.

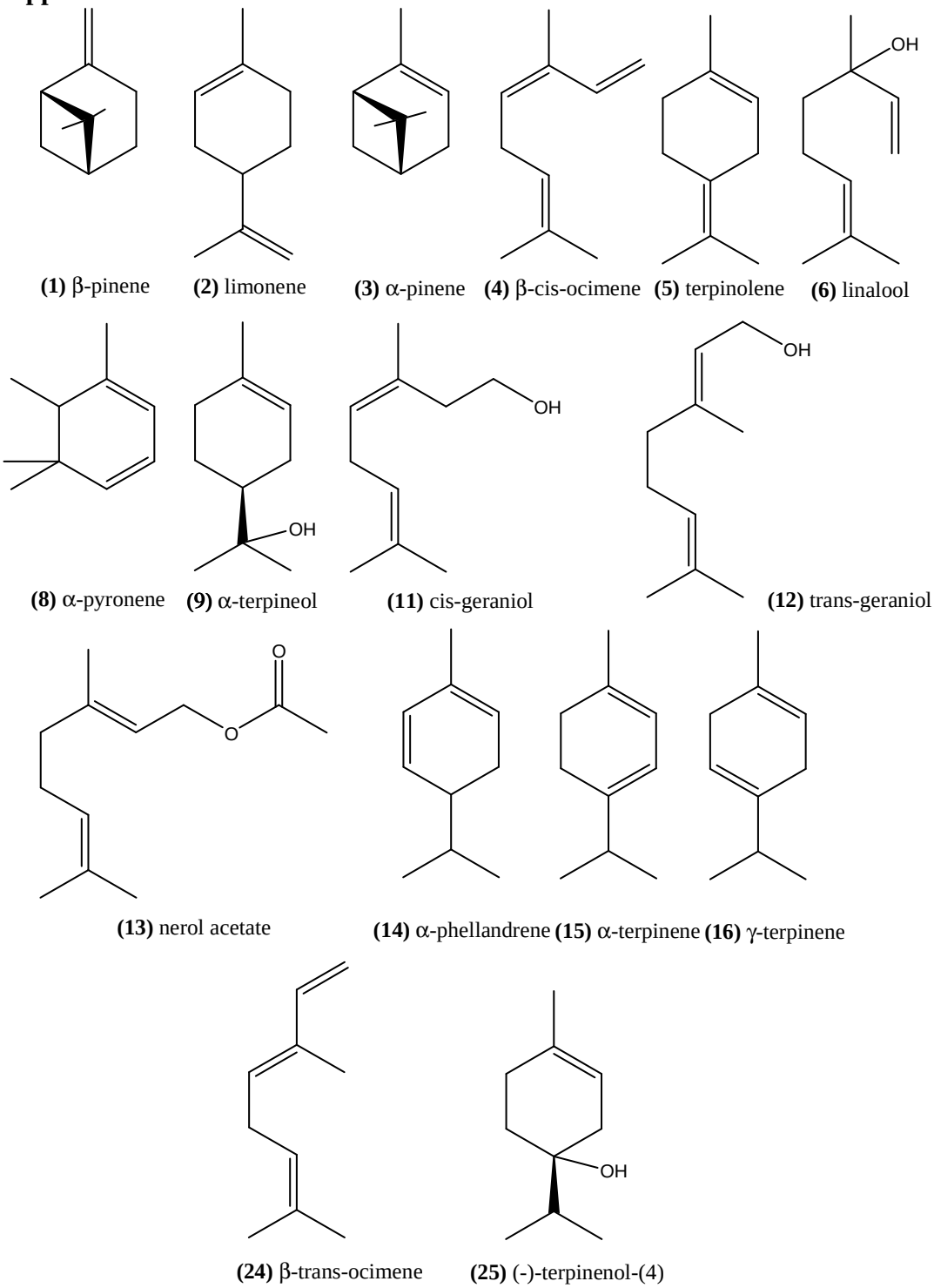
[150] Geissman TA, Crout DGH, (1969) Organic chemistry of secondary plant metabolism, Freeman, Cooper and Co. San Francisco, California, pp 290–311.

[151] Ruzicka L. **(1994)** The isoprene rule and the biogenesis of terpenic compounds. 1953. *Experientia*, **50,** 395-405.

[152] Guo Z, Severson RF, Wagner GJ. **(1994)** Biosynthesis of the diterpene cis-abienol in cell free extracts of tobacco trichomes. *Archives of Biochemistry and Biophysics*, **308,** 103-108.

[153] Schmiderer C, Grassi P, Novak J, Weber M, Franz C. **(2008)** Diversity of essential oil glands of clary sage (Salvia sclarea L., Lamiaceae). *Plant Biology*, **10,** 443-440.

7. Appendix



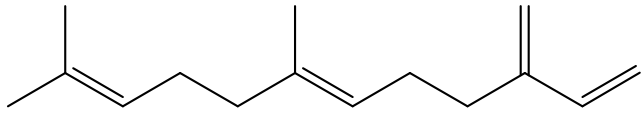
(1) β-pinene **(2)** limonene **(3)** α-pinene **(4)** β-cis-ocimene **(5)** terpinolene **(6)** linalool

(8) α-pyronene **(9)** α-terpineol **(11)** cis-geraniol **(12)** trans-geraniol

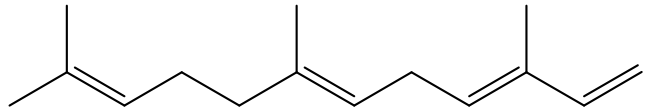
(13) nerol acetate **(14)** α-phellandrene **(15)** α-terpinene **(16)** γ-terpinene

(24) β-trans-ocimene **(25)** (-)-terpinenol-(4)

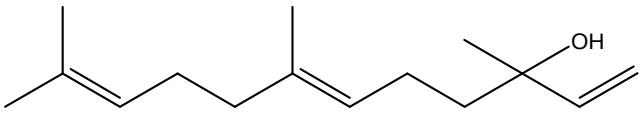
Figure 41: SscTPS1 monoterpenoid products using GPP and NPP as substrates



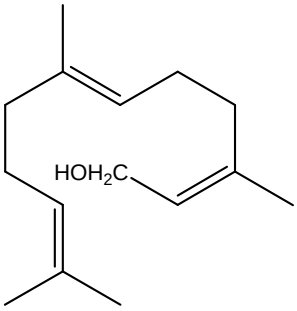
(17) β-farnesene



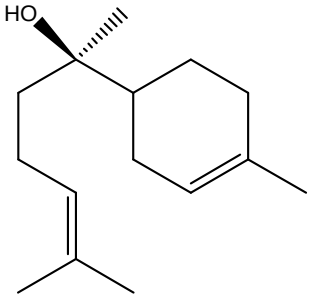
(18) α-farnesene



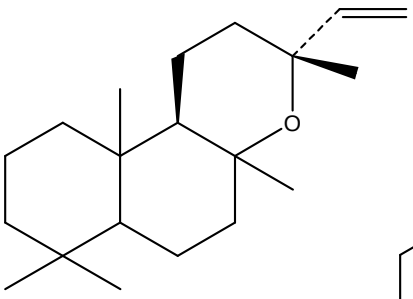
(19) nerolidol



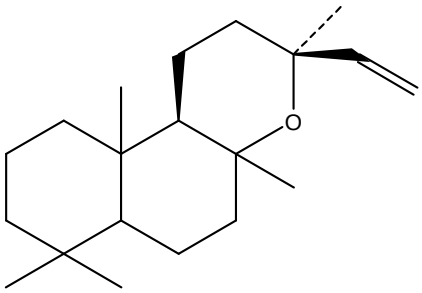
(21) farnesol



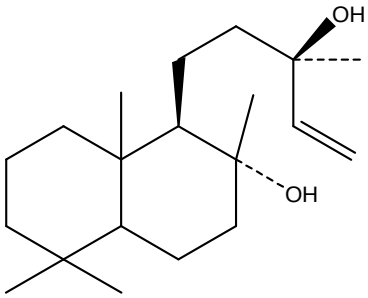
(26) α-bisabolol



(22) (13R)-manoyl oxide



(23) (13S)-manoyl oxide



(27) sclareol

Figure 42: SscTPS1 sesqui- and diterpenoid products using FPP and GGPP as substrates

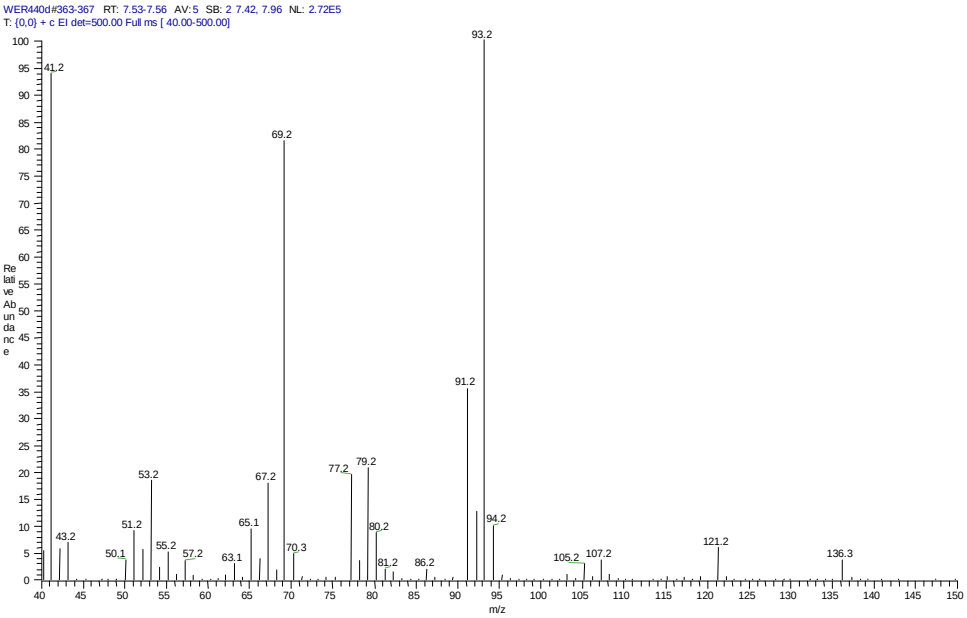


Figure 43 (WER440d): MS-spectrum of compound β-pinene (**1**) R_t = 7.53 min obtained by GC-MS.

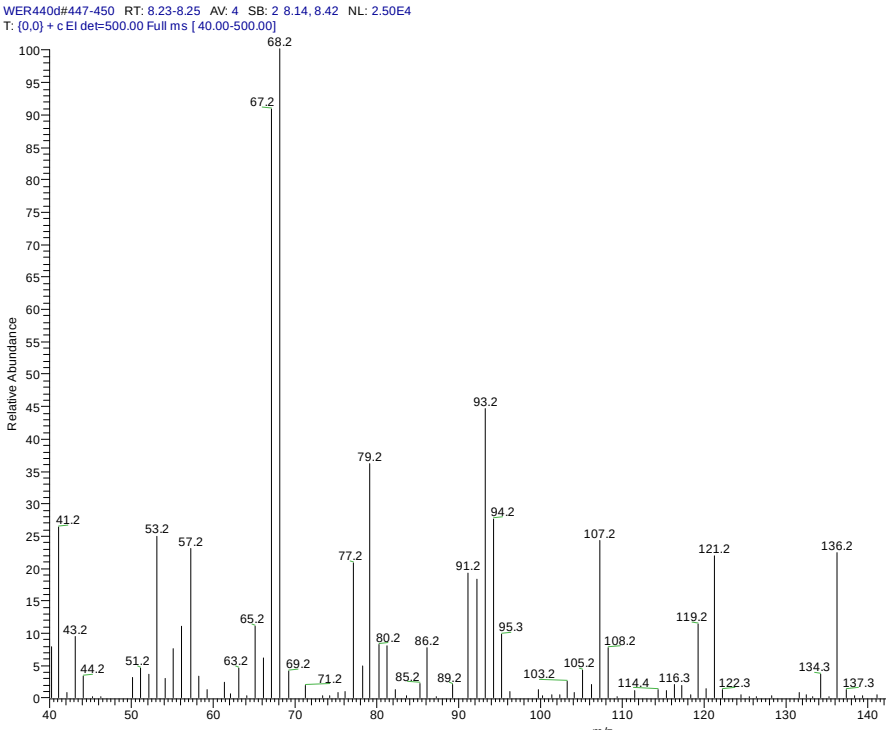


Figure 44 (WER440d): MS- spectrum of compound limonene (**2**) R_t = 8.24 min obtained by GC-MS.

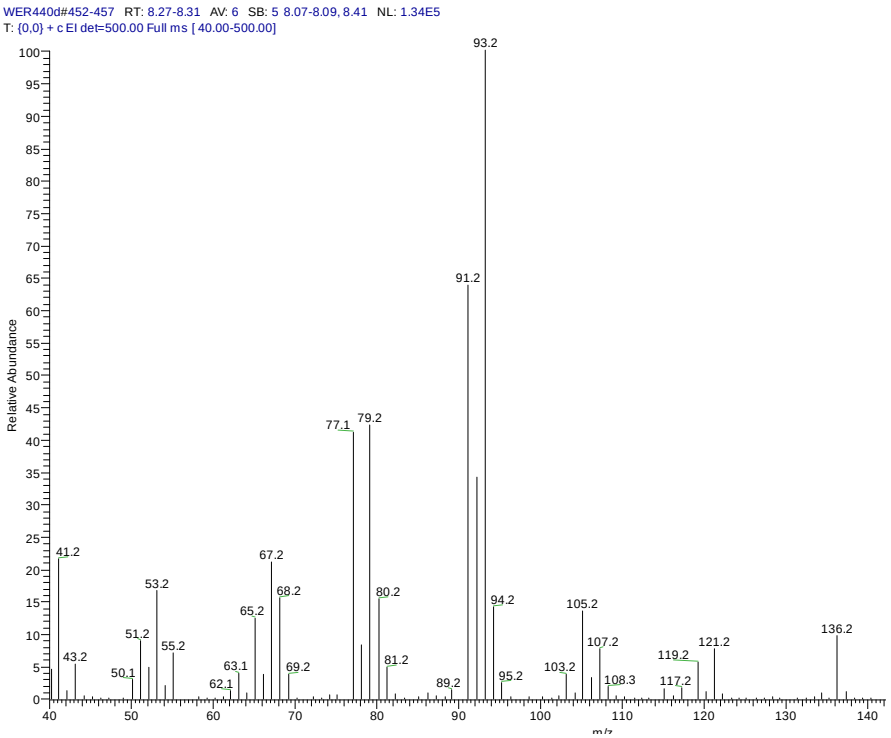


Figure 45 (WER440d): MS- spectrum of compound α-pinene (**3**) R_t = 8.28 min obtained by GC-MS.

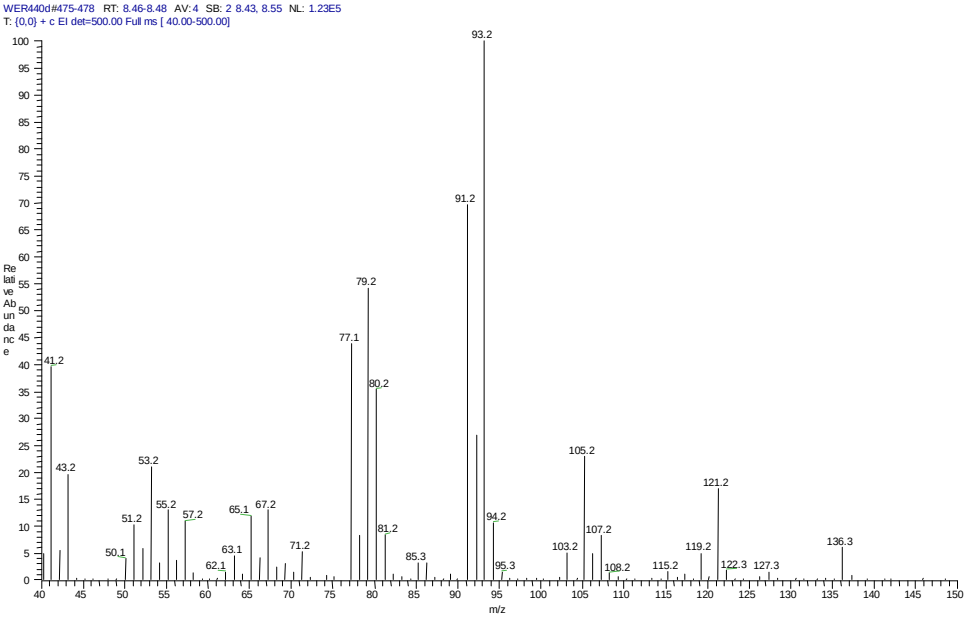


Figure 46 (WER440d): MS- spectrum of compound β-cis-ocimene (**4**) R_t = 8.47 min obtained by GC-MS.

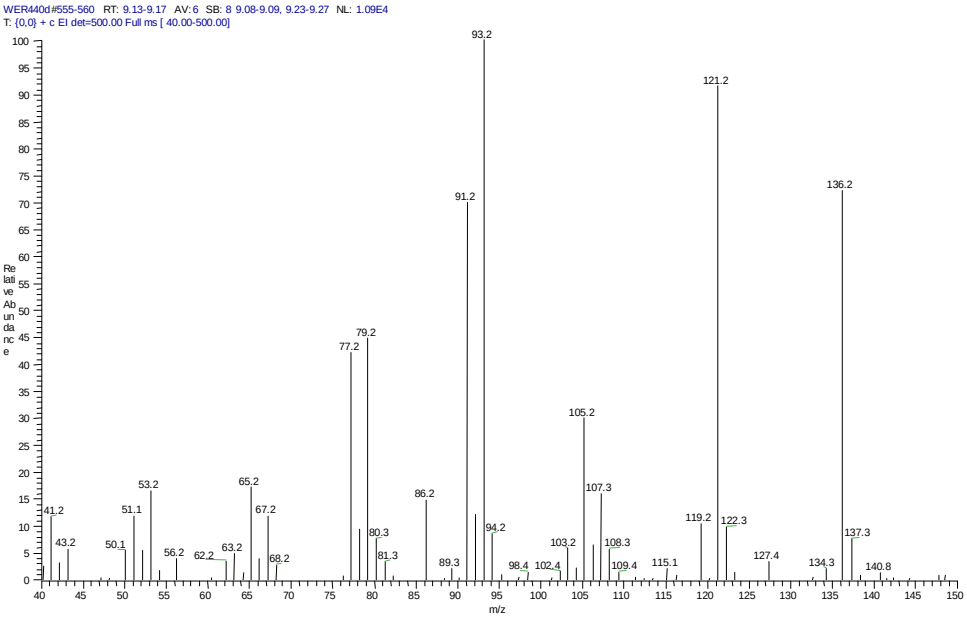


Figure 47 (WER440d): MS- spectrum of compound terpinolene (**5**) R_t = 9.13 min obtained by GC-MS.

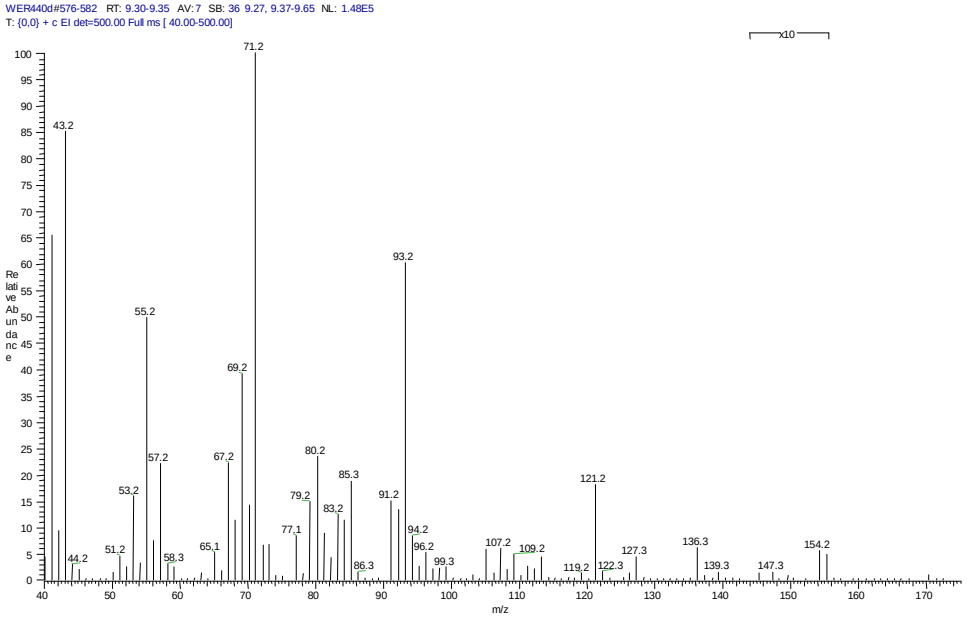


Figure 48 (WER440d): MS- spectrum of compound β-linalool (**6**) R_t = 9.32 min obtained by GC-MS.

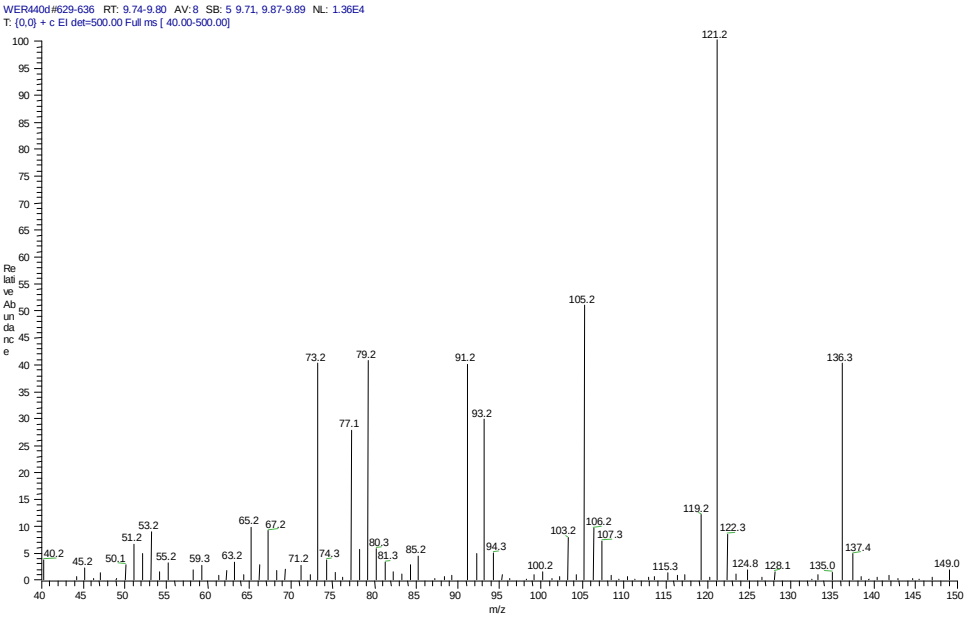


Figure 49 (WER440d): MS- spectrum of compound not-identified monterpene (**7**) R_t = 9.77 min obtained by GC-MS.

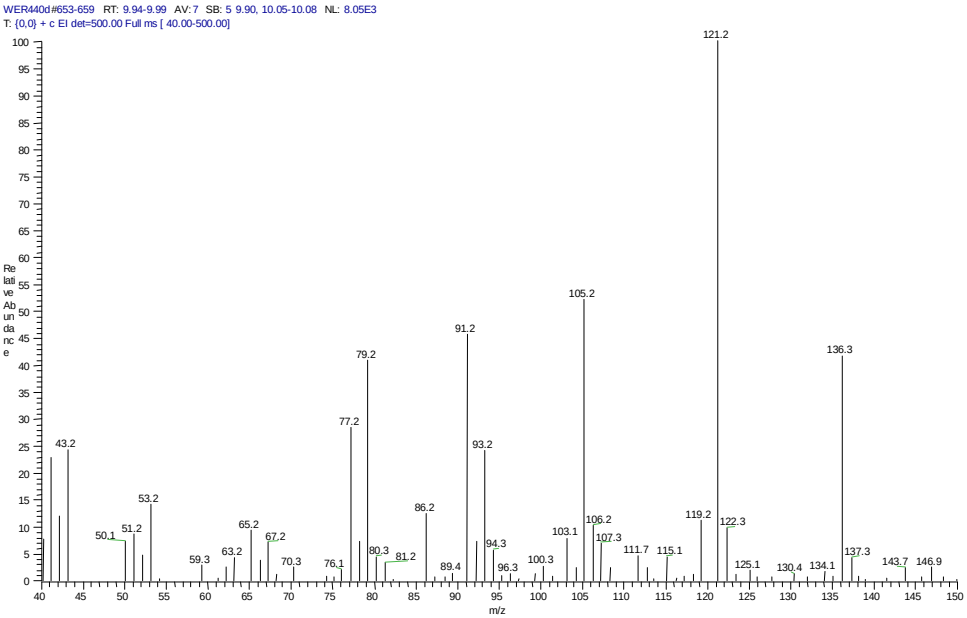


Figure 50 (WER440d): MS- spectrum of compound α-pyronene (**8**) R_t = 9.96 min obtained by GC-MS.

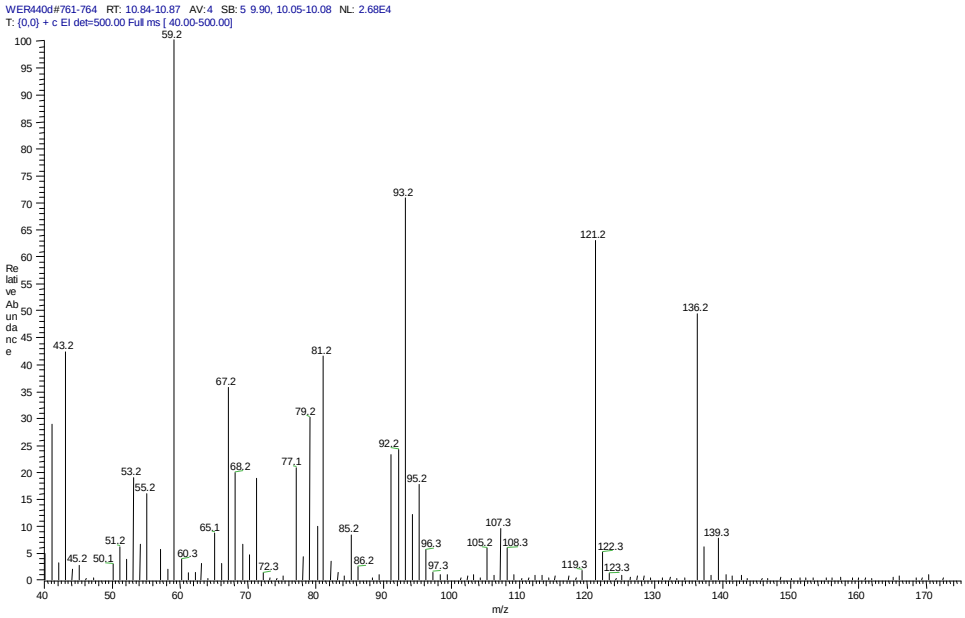


Figure 51 (WER440d): MS- spectrum of compound α-terpineol (**9**) R_t = 10.85 min obtained by GC-MS.

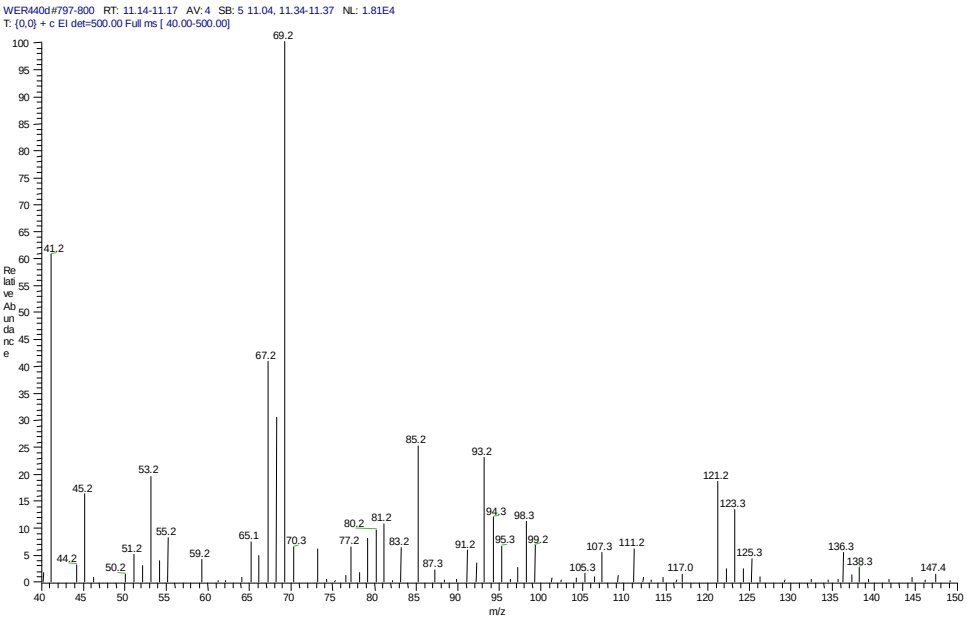


Figure 52 (WER440d): MS- spectrum of compound not-identified monoterpene (**10**) R_t = 11.15 min obtained by GC-MS.

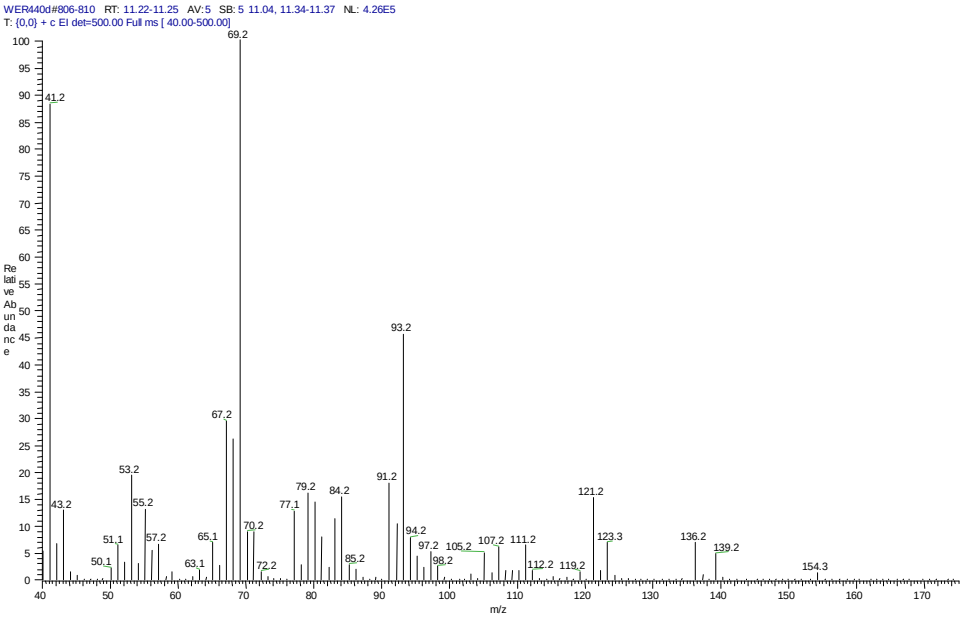


Figure 53 (WER440d): MS-spectrum of compound cis-geraniol (**11**) R_t = 11.23 min obtained by GC-MS.

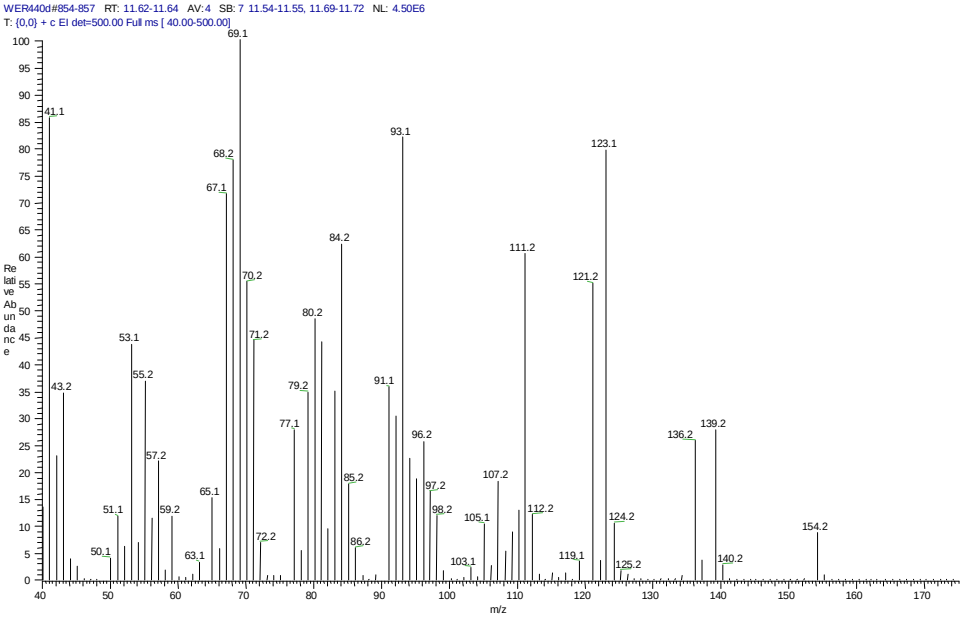


Figure 54 (WER440d): MS-spectrum of compound trans-geraniol (**12**) R_t = 11.64 min obtained by GC-MS.

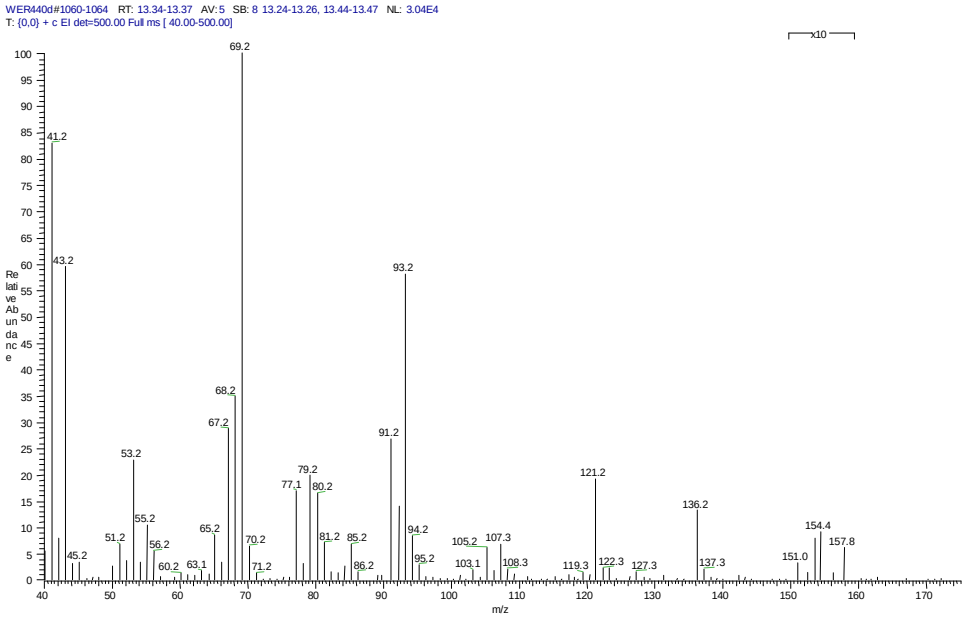


Figure 55 (WER440d): MS-spectrum of compound neryl acetate (**13**) R_t = 13.34 min obtained by GC-MS.

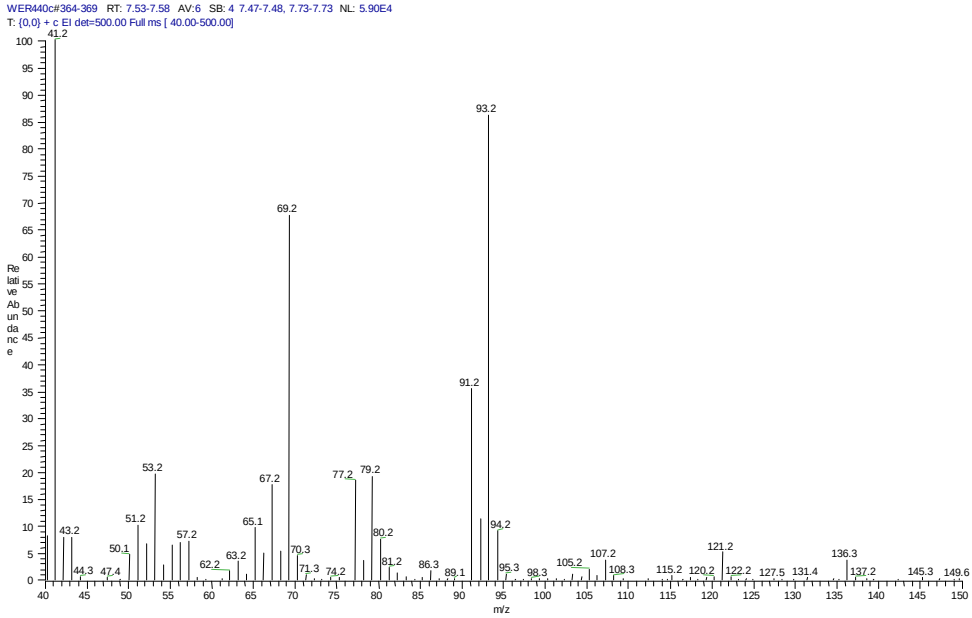


Figure 56 (WER440c): MS-spectrum of compound β-pinene (**1**) R_t = 7.55 min obtained by GC-MS.

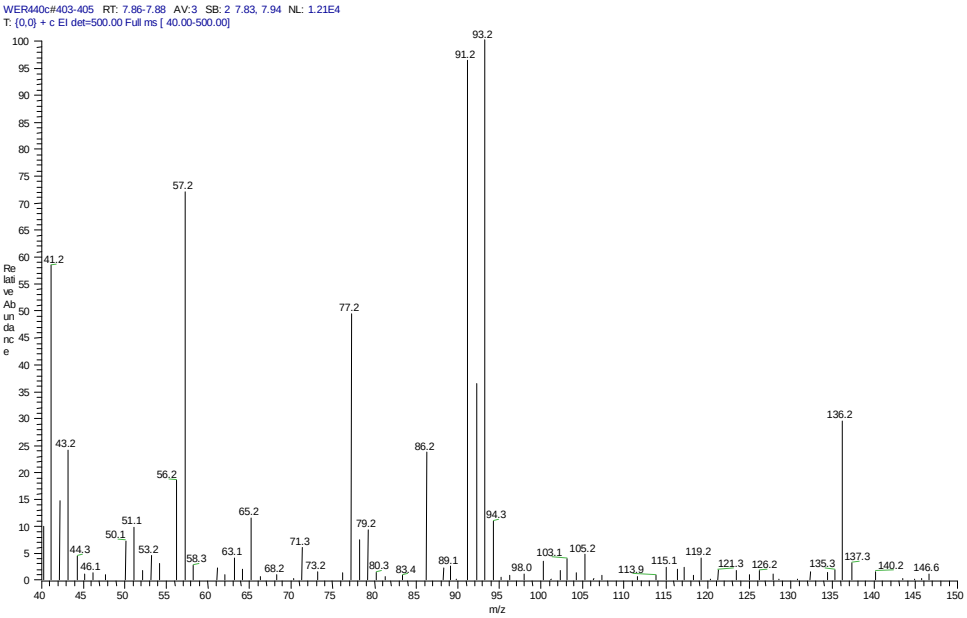


Figure 57 (WER440c): MS-spectrum of compound α-phellandrene (**14**) R_t = 7.88 min obtained by GC-MS.

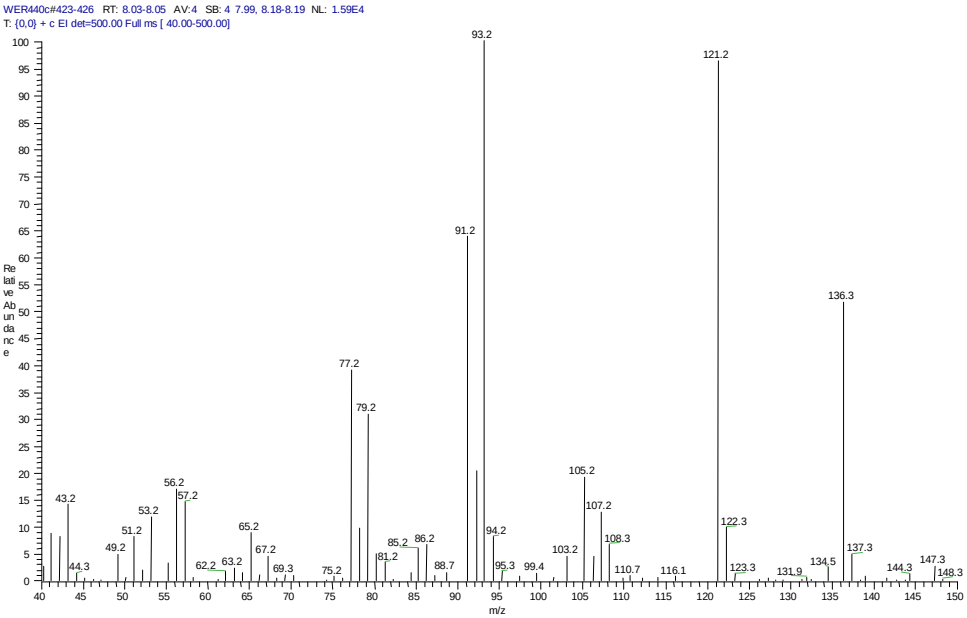


Figure 58 (WER440c): MS-spectrum of compound α-terpinene (**15**) R_t = 8.05 min obtained by GC-MS.

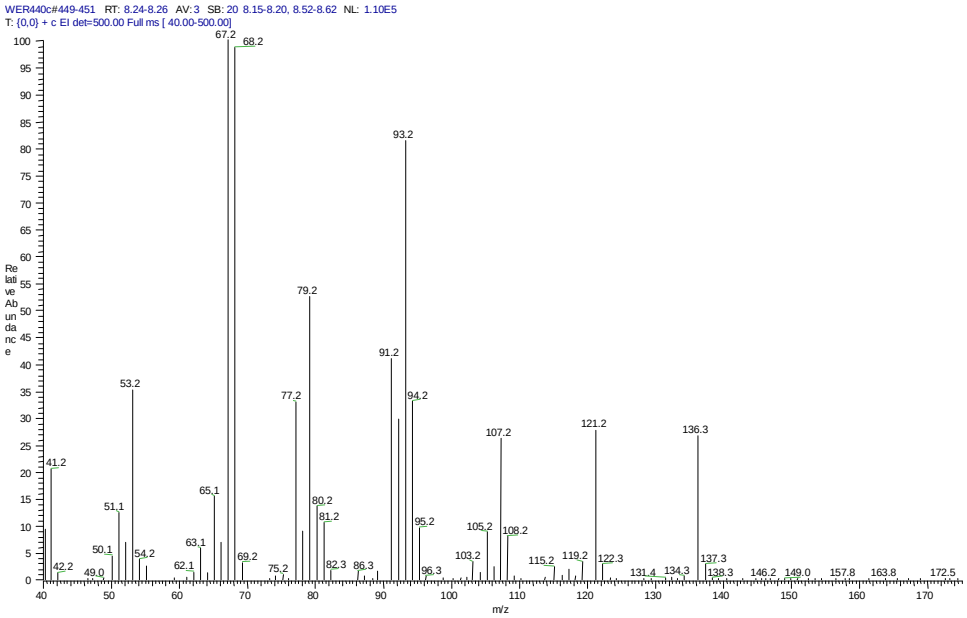


Figure 59 (WER440c): MS-spectrum of compound limonene (**2**) R_t = 8.26 min obtained by GC-MS.

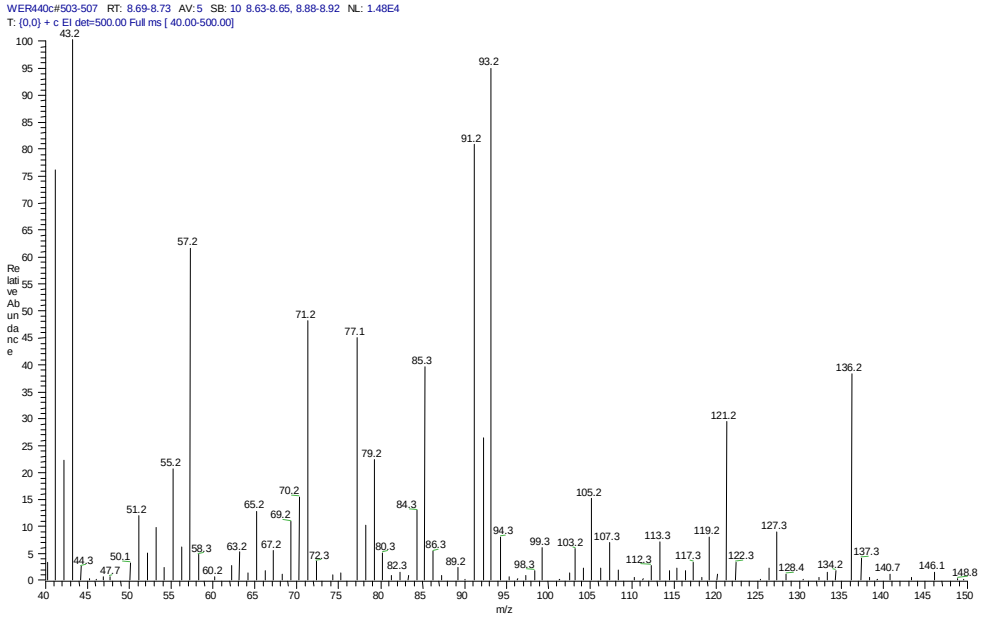


Figure 60 (WER440c): MS-spectrum of compound γ-terpinene (**16**) R_t = 8.71 min obtained by GC-MS.

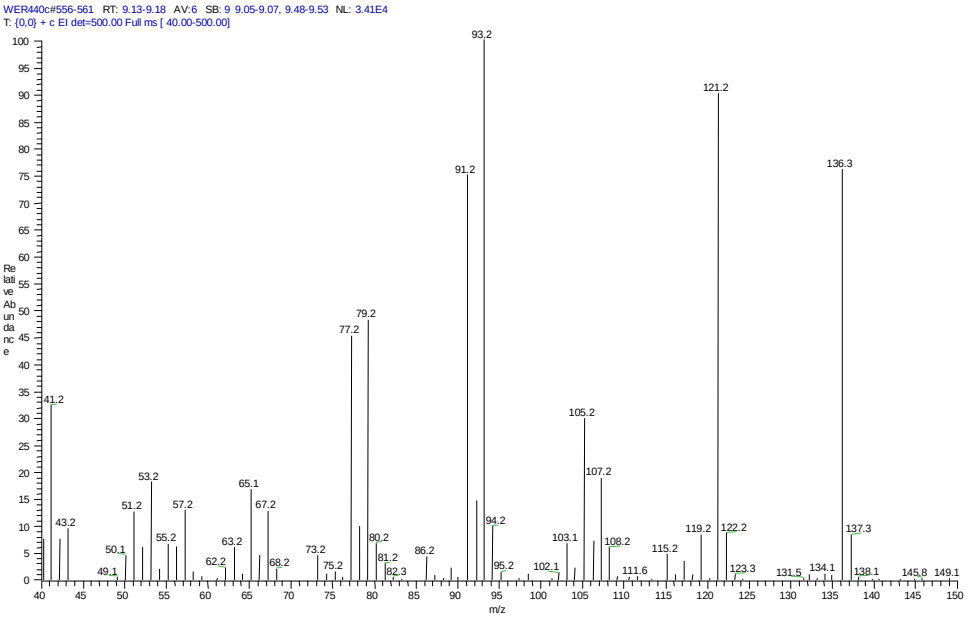


Figure 61 (WER440c): MS-spectrum of compound terpinolene (**5**) R_t = 9.15 min obtained by GC-MS.

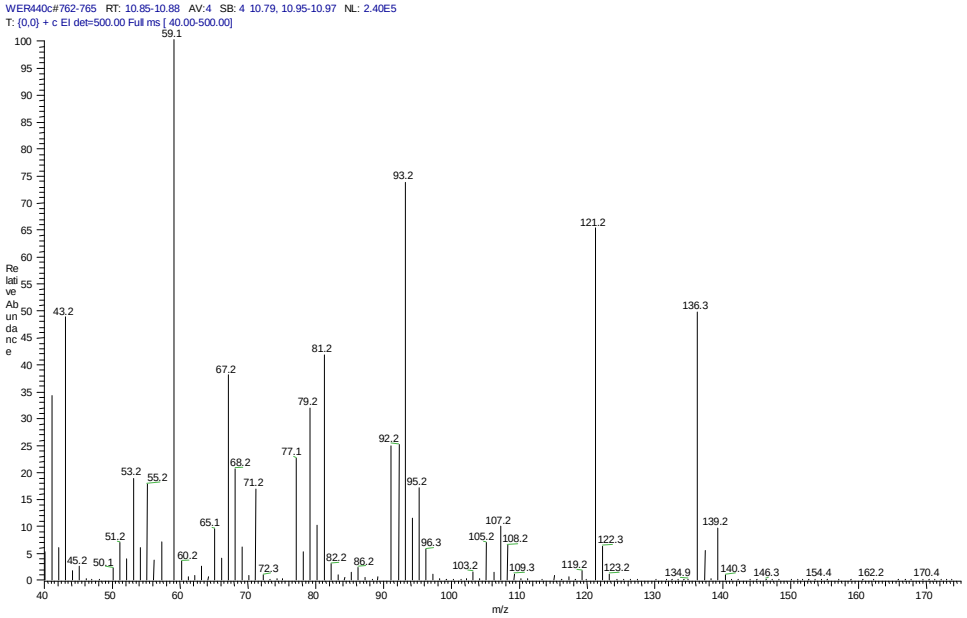


Figure 62 (WER440c): MS-spectrum of compound α-terpineol (**9**) R_t = 10.86 min obtained by GC-MS.

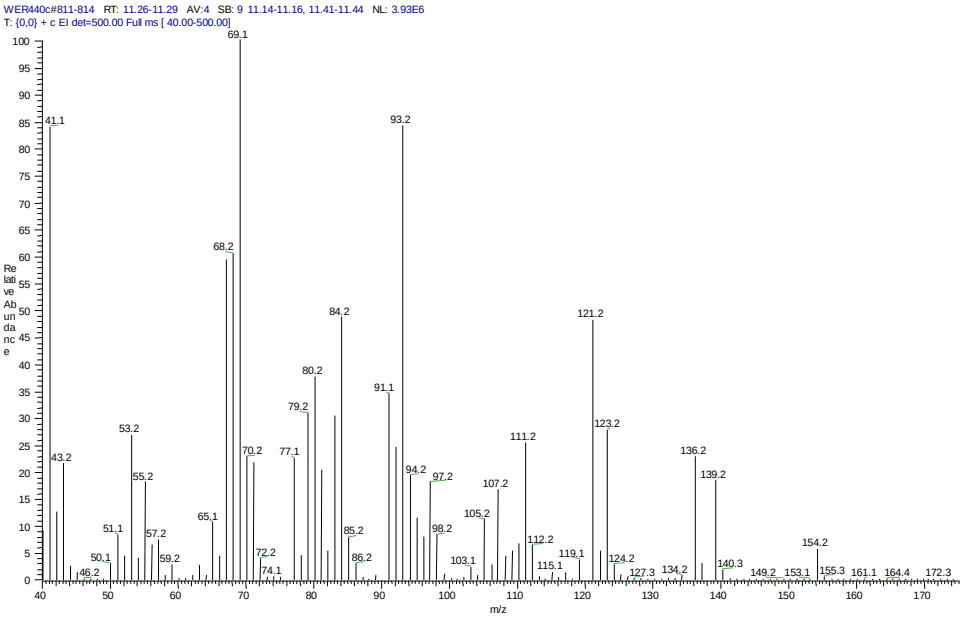


Figure 63 (WER440c): MS-spectrum of compound cis-geraniol (**11**) R_t = 11.28 min obtained by GC-MS.

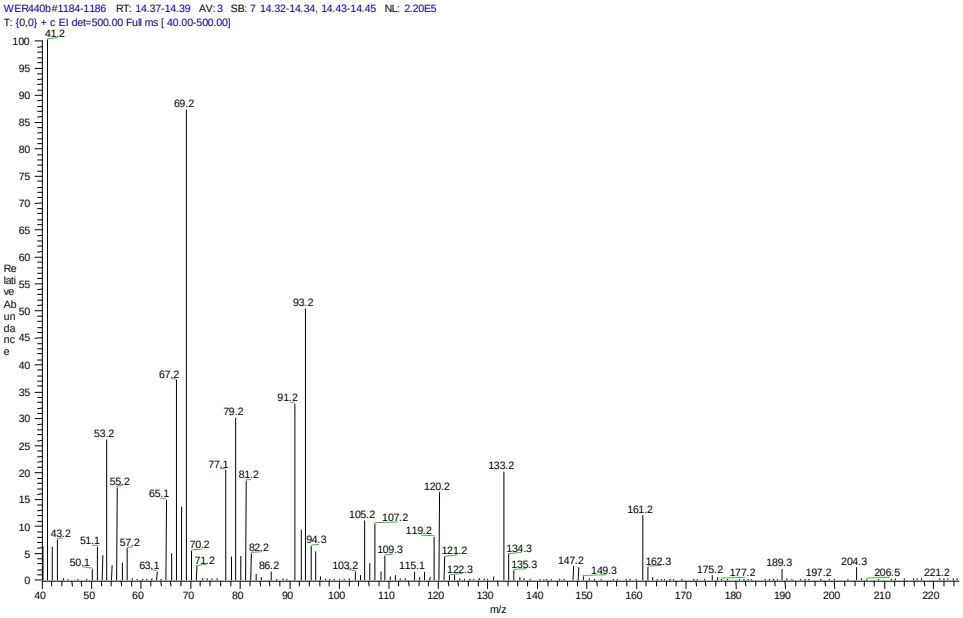


Figure 64 (WER440b): MS-spectrum of compound β-farnesene (**17**) R_t = 14.38 min obtained by GC-MS.

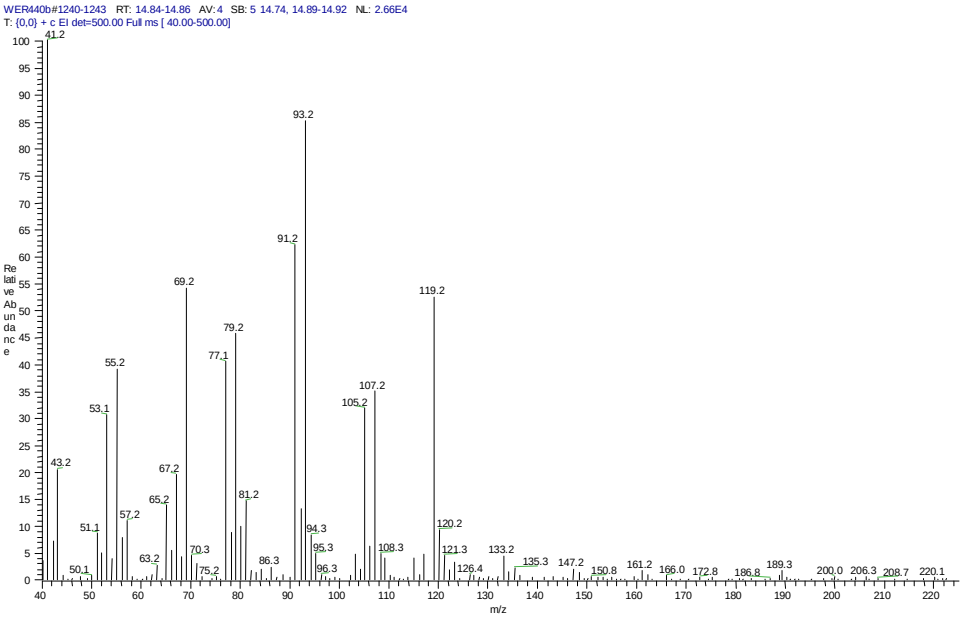


Figure 65 (WER440b): MS-spectrum of compound (Z,E)-α-farnesene (**18**) R_t = 14.85 min obtained by GC-MS.

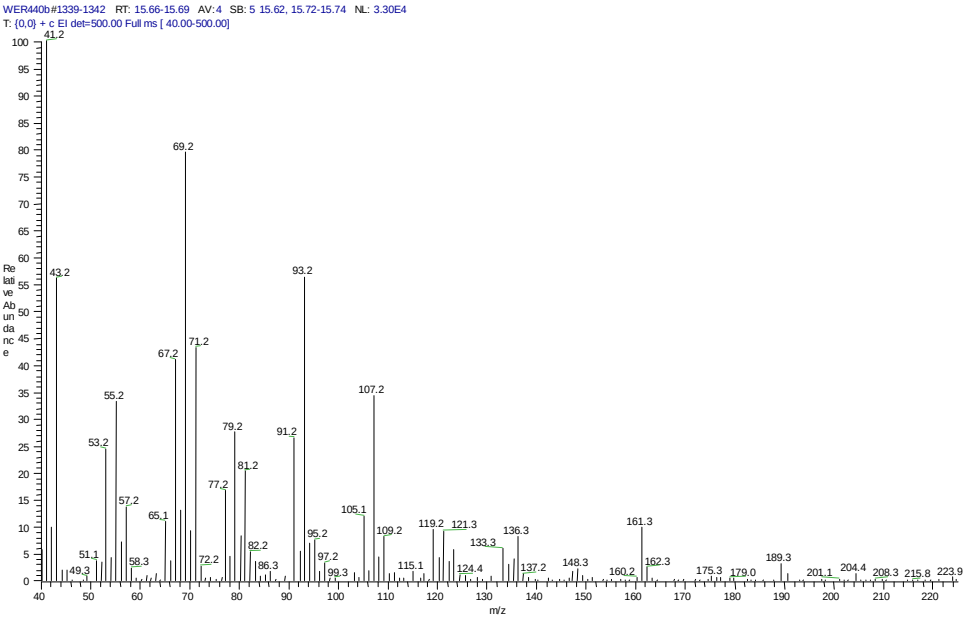


Figure 66 (WER440b): MS-spectrum of compound nerolidol (**19**) R_t = 15.67 min obtained by GC-MS.

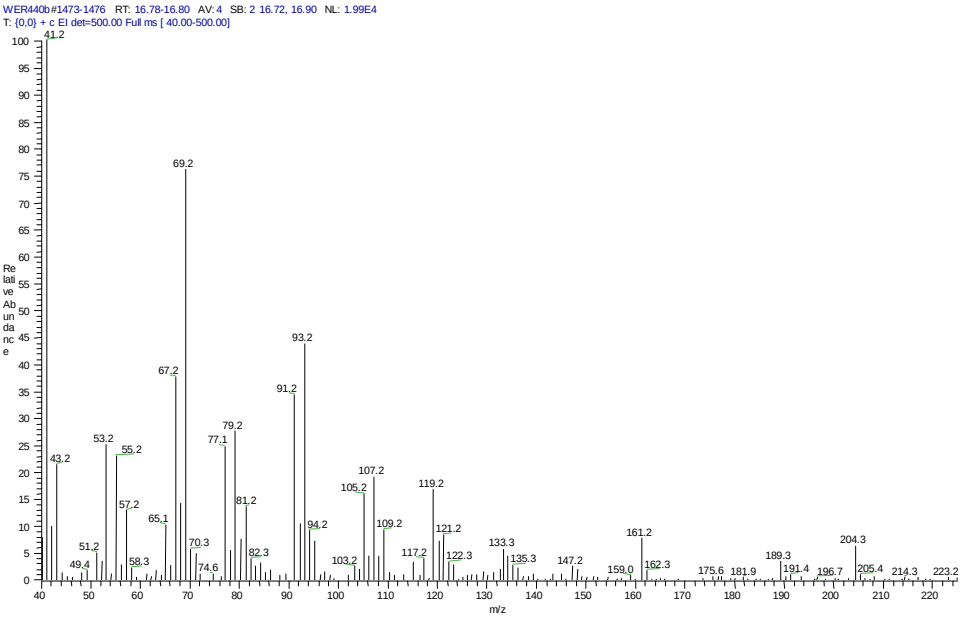


Figure 67 (WER440b): MS-spectrum of compound not-identified (**20**) R_t = 16.80 min obtained by GC-MS.

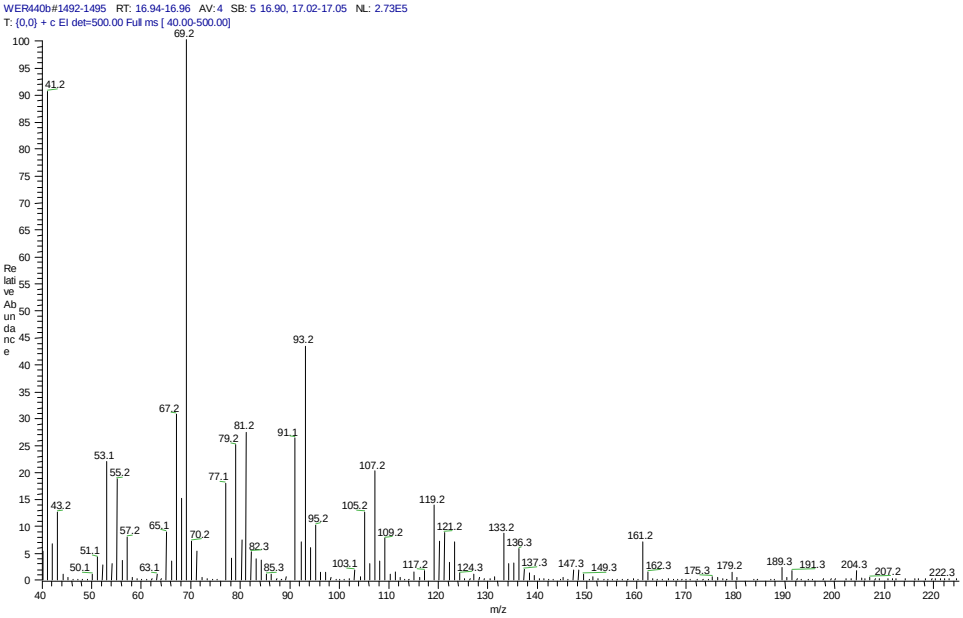


Figure 68 (WER440b): MS-spectrum of compound farnesol (2**1**) R_t = 16.95 min obtained by GC-MS.

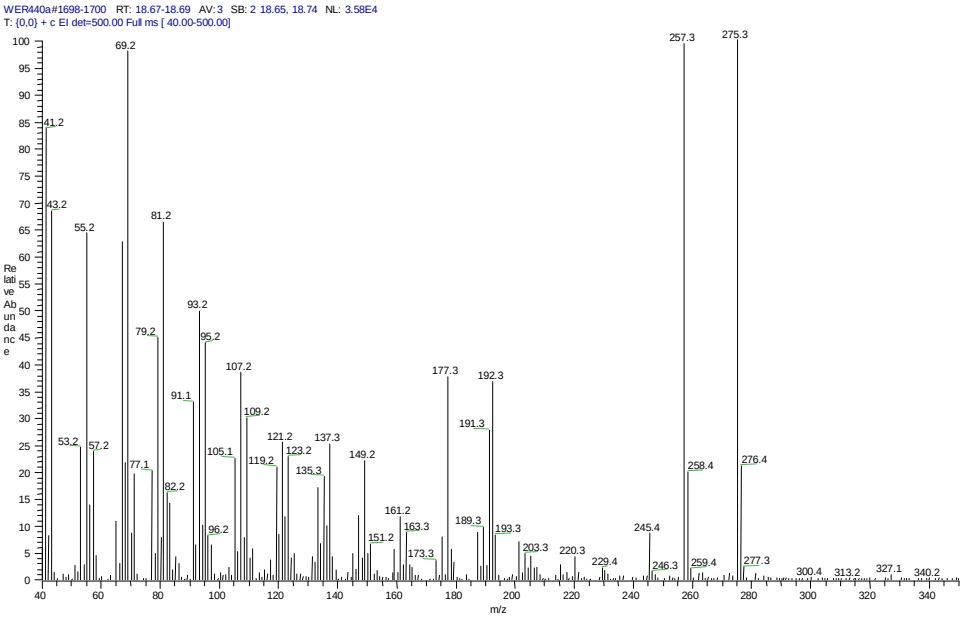


Figure 69 (WER440a): MS-spectrum of compound (13R)-manoyl-oxide (**22**) R_t = 18.69 min obtained by GC-MS.

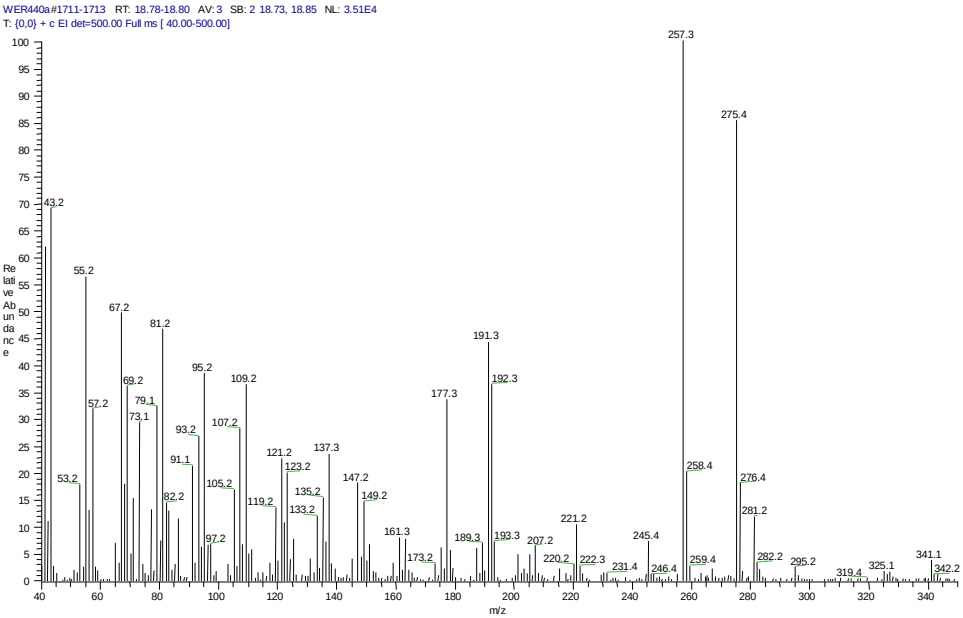


Figure 70 (WER440a): MS-spectrum of compound (13S)-manoyl-oxide (**23**) R_t = 18.79 min obtained by GC-MS.

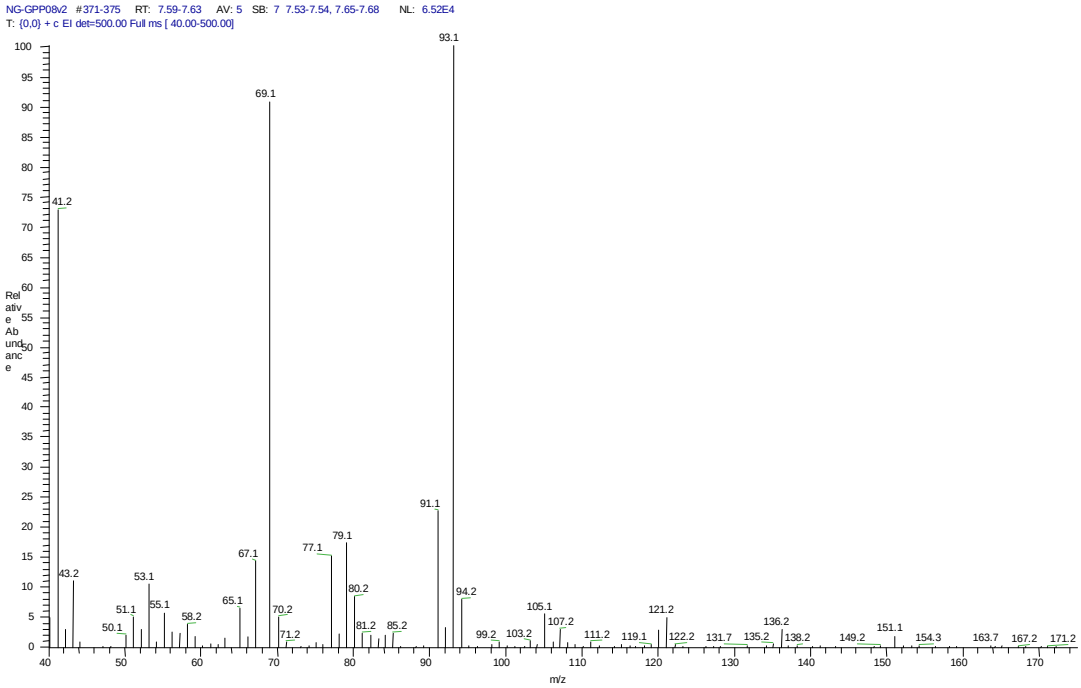


Figure 71 (NG_GPP08v2): MS-spectrum of compound β-pinene (**1**) R_t = 7.61 min obtained by GC-MS.

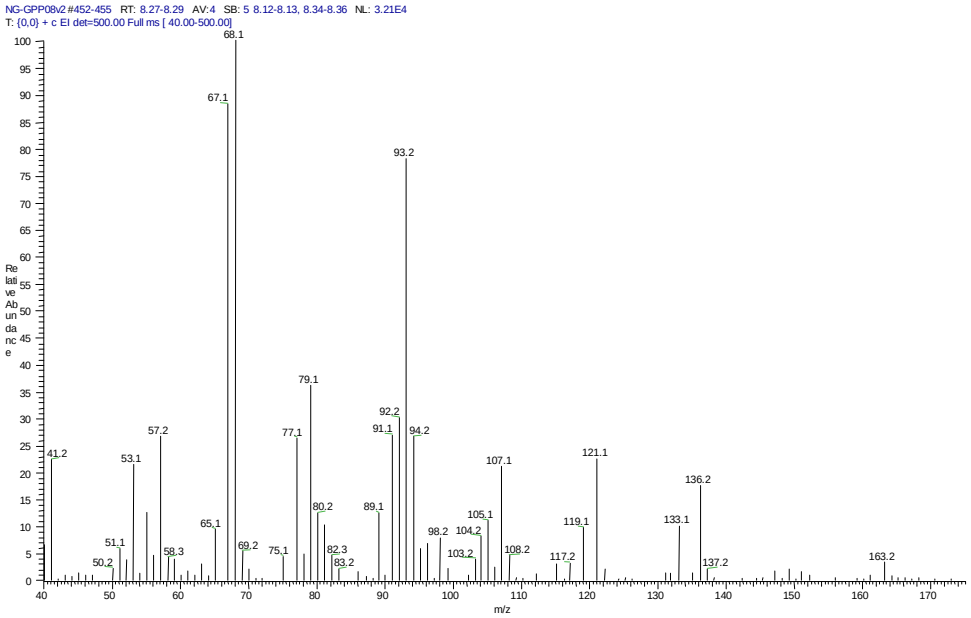


Figure 72 (NG_GPP08v2): MS-spectrum of compound limonene (**2**) R_t = 8.28 min obtained by GC-MS.

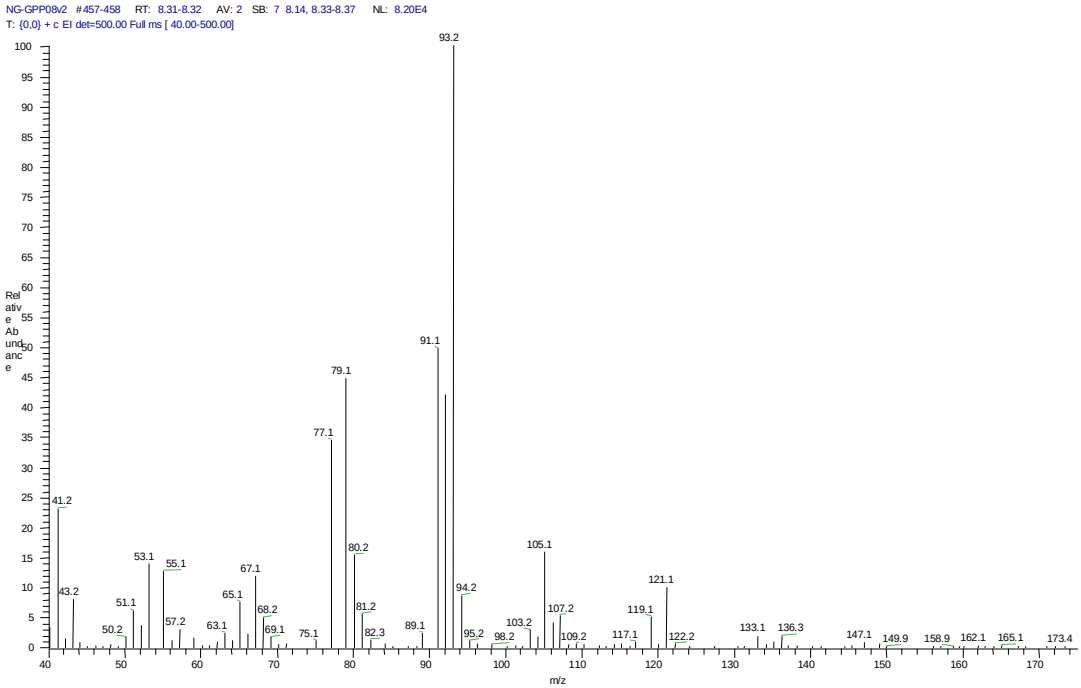


Figure 73 (NG_GPP08v2): MS-spectrum of compound β-trans-ocimene (**24**) R_t = 8.31 min obtained by GC-MS.

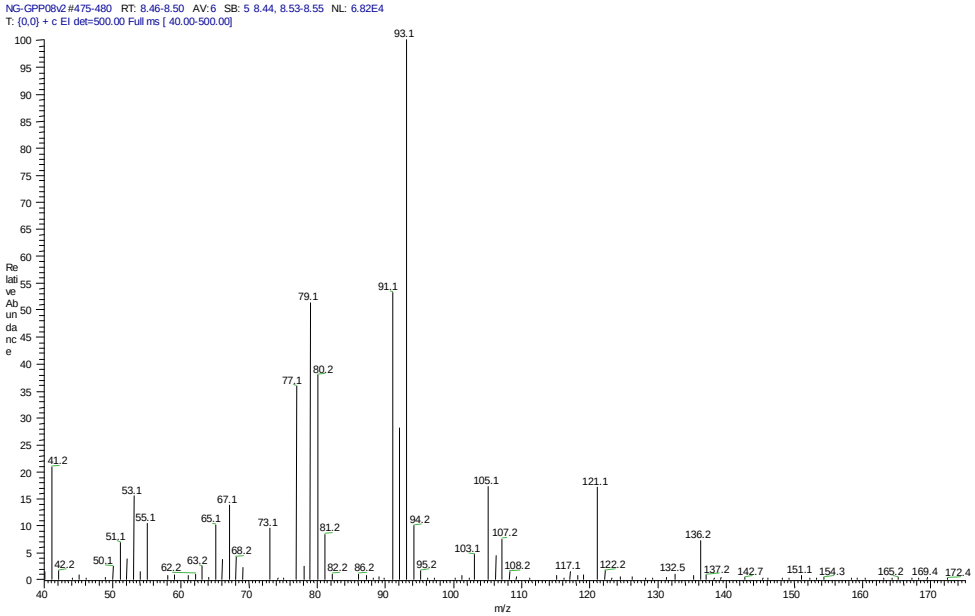


Figure 74 (NG_GPP08v2): MS-spectrum of compound β-cis-ocimene (**4**) R_t = 8.48 min obtained by GC-MS.

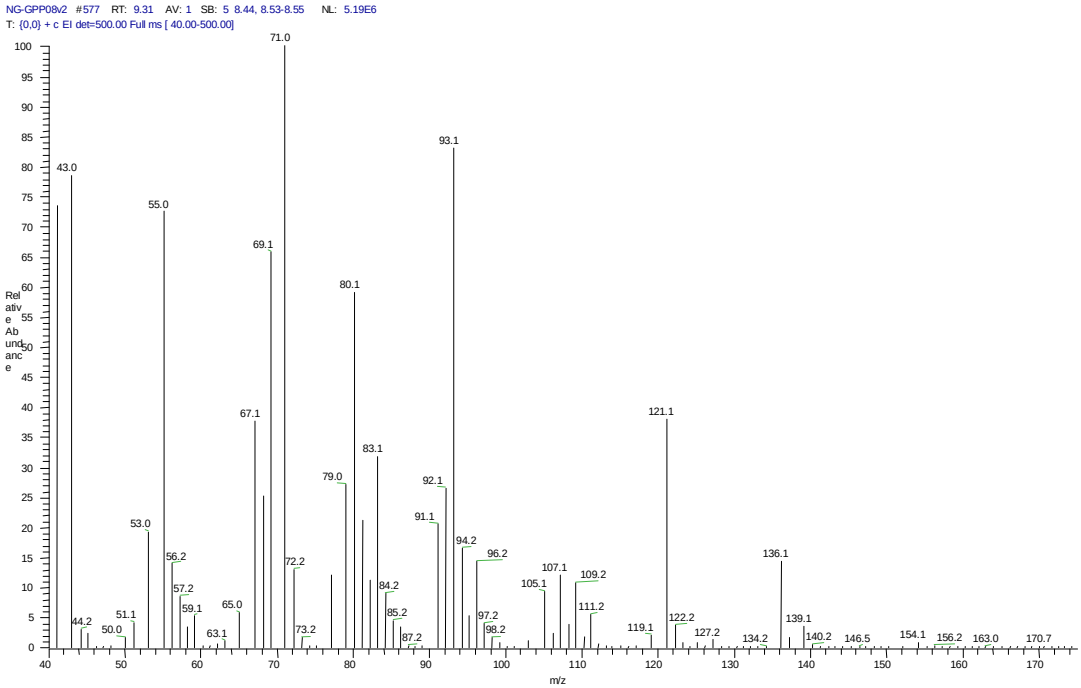


Figure 75 (NG_GPP08v2): MS-spectrum of compound β-linalool (**6**) R_t = 9.31 min obtained by GC-MS.

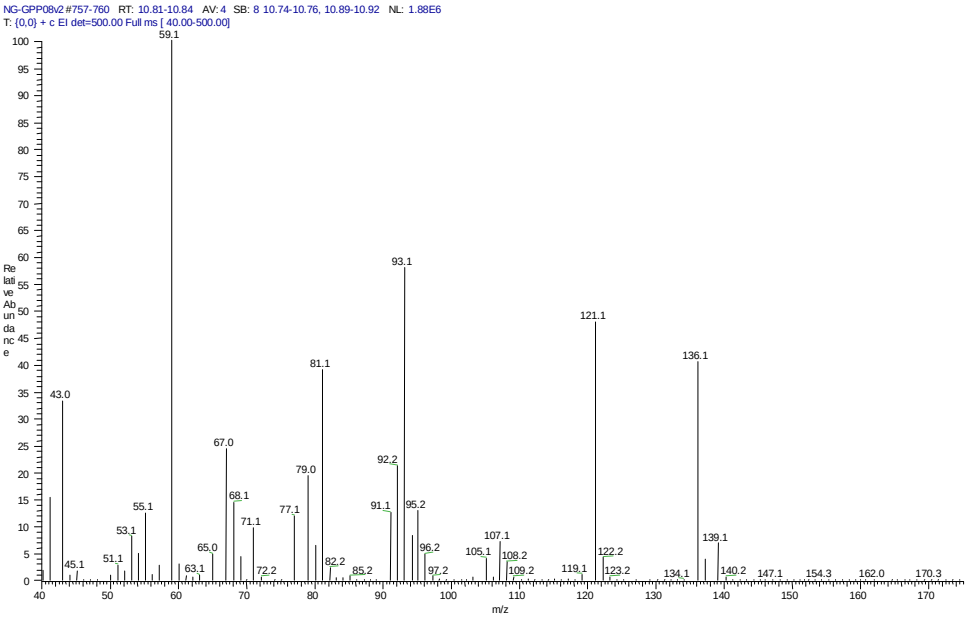


Figure 76 (NG_GPP08v2): MS-spectrum of compound α-terpineol (**9**) R_t = 10.83 min obtained by GC-MS.

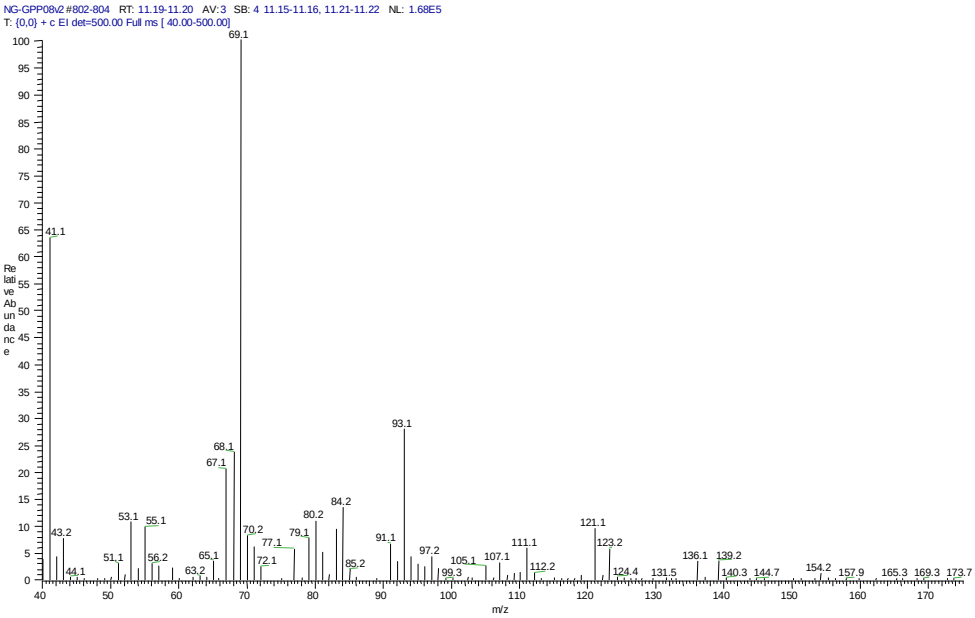


Figure 77 (NG_GPP08v2): MS-spectrum of compound cis-geraniol (**11**) R_t = 11.19 min obtained by GC-MS.

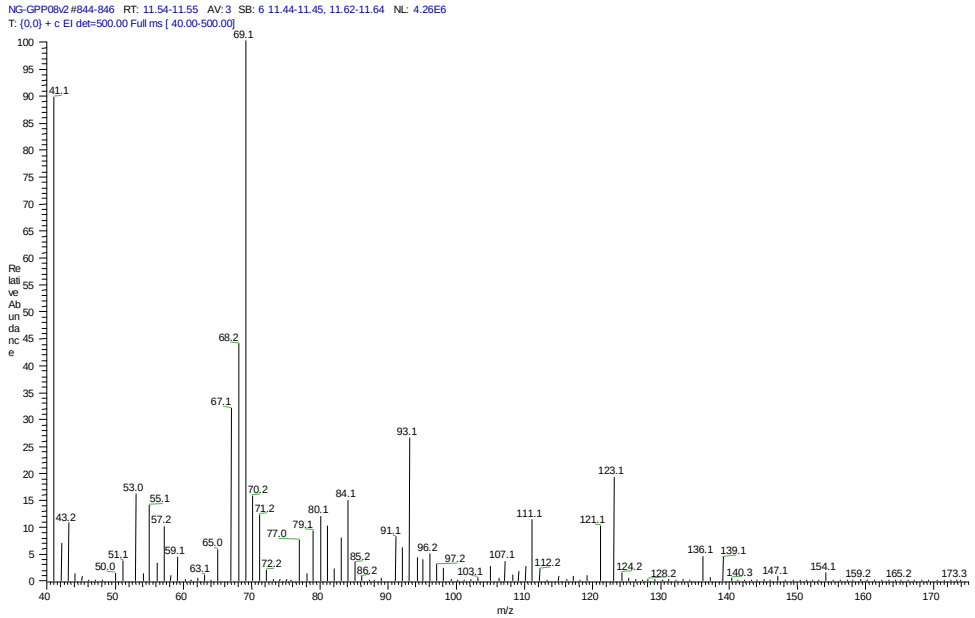


Figure 78 (NG_GPP08v2): MS-spectrum of compound trans-geraniol (**12**) R_t = 11.54 min obtained by GC-MS.

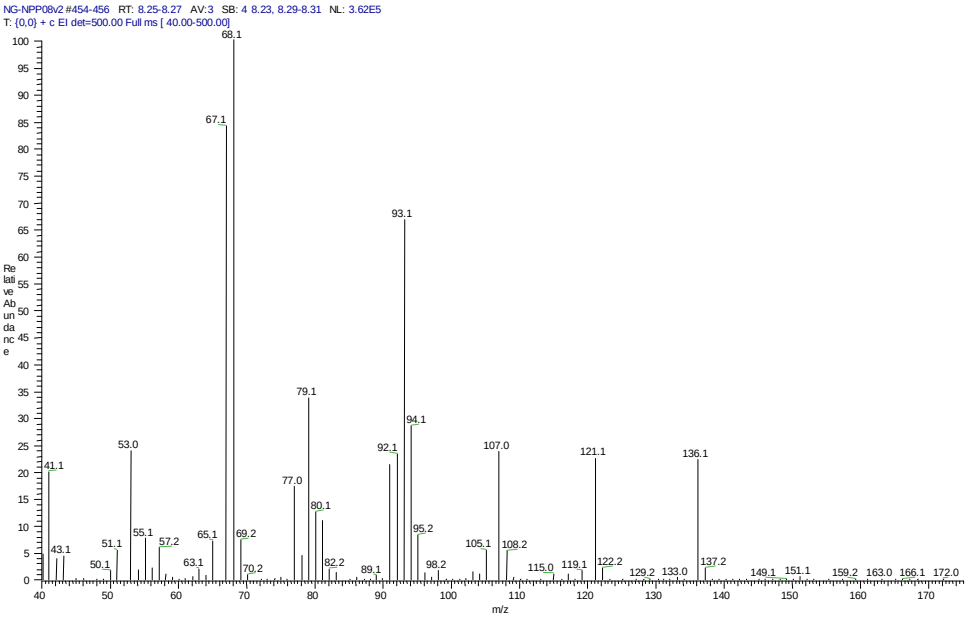


Figure 79 (NG_NPP08v2): MS-spectrum of compound limonene (**2**) R_t = 8.26 min obtained by GC-MS.

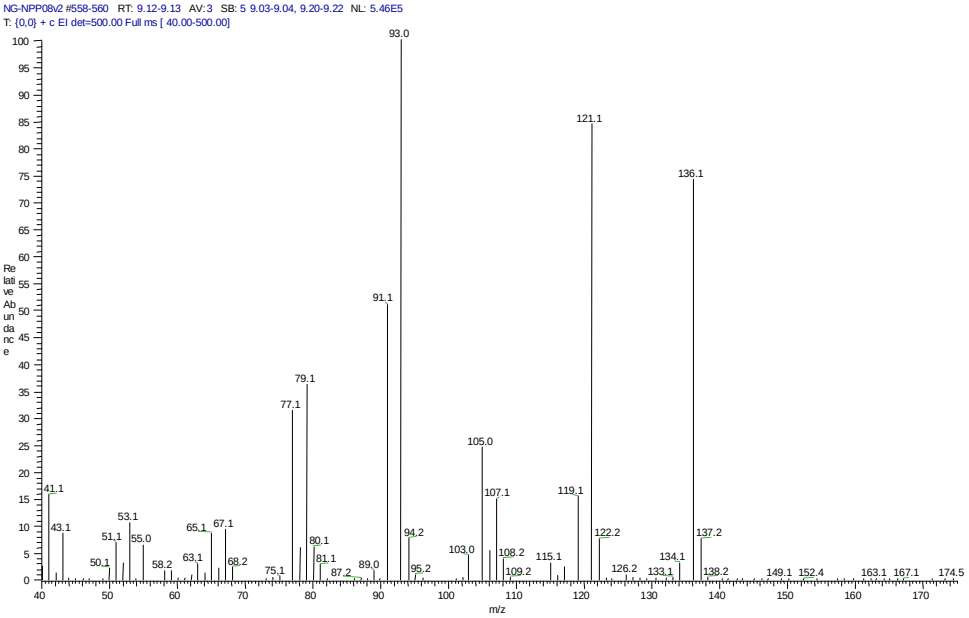


Figure 80 (NG_NPP08v2): MS-spectrum of compound terpinolene (**5**) R_t = 9.13 min obtained by GC-MS.

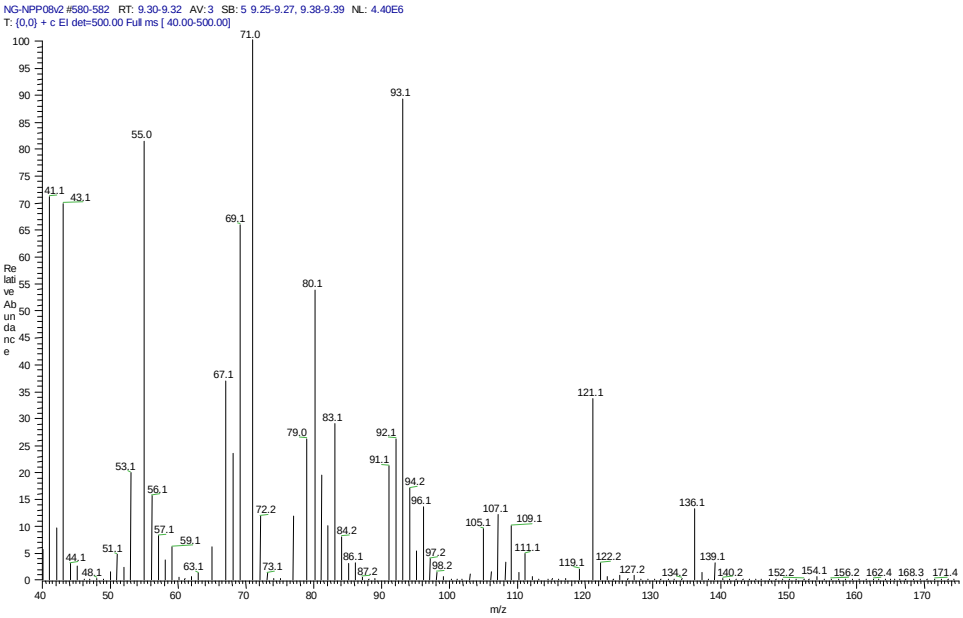


Figure 81 (NG_NPP08v2): MS-spectrum of compound β-linalool (**6**) R_t = 9.31 min obtained by GC-MS.

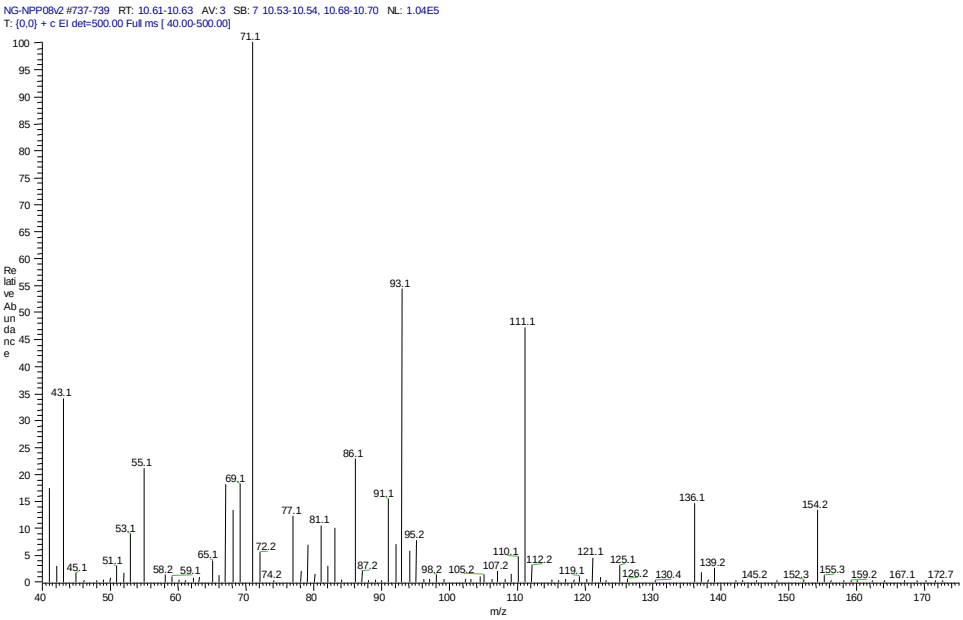


Figure 82 (NG_NPP08v2): MS-spectrum of compound (-)-terpine-4-ol (**25**) R_t = 10.62 min obtained by GC-MS.

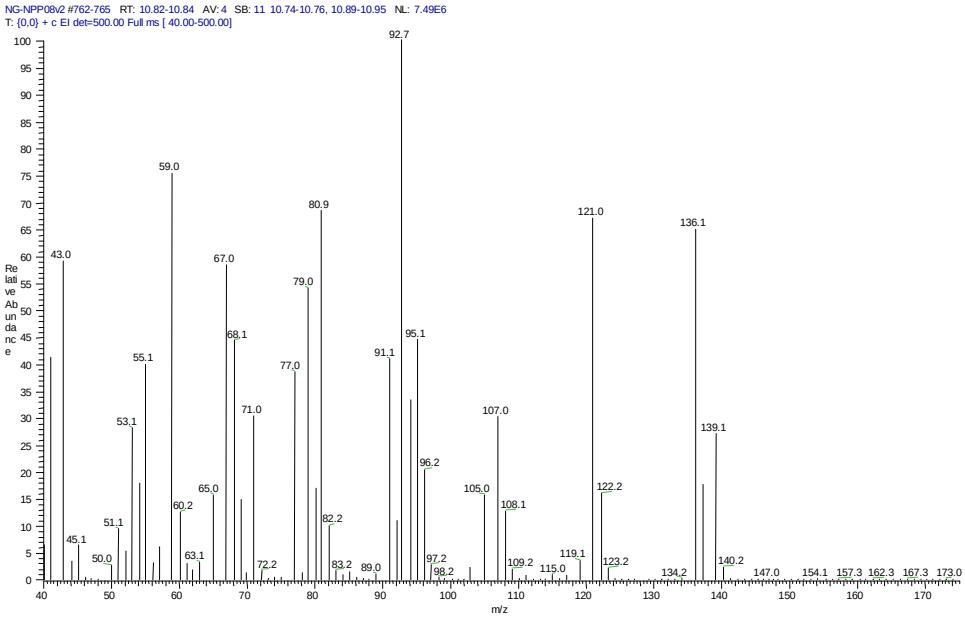


Figure 83 (NG_NPP08v2): MS-spectrum of compound α-terpineol (**9**) R_t = 10.84 min obtained by GC-MS.

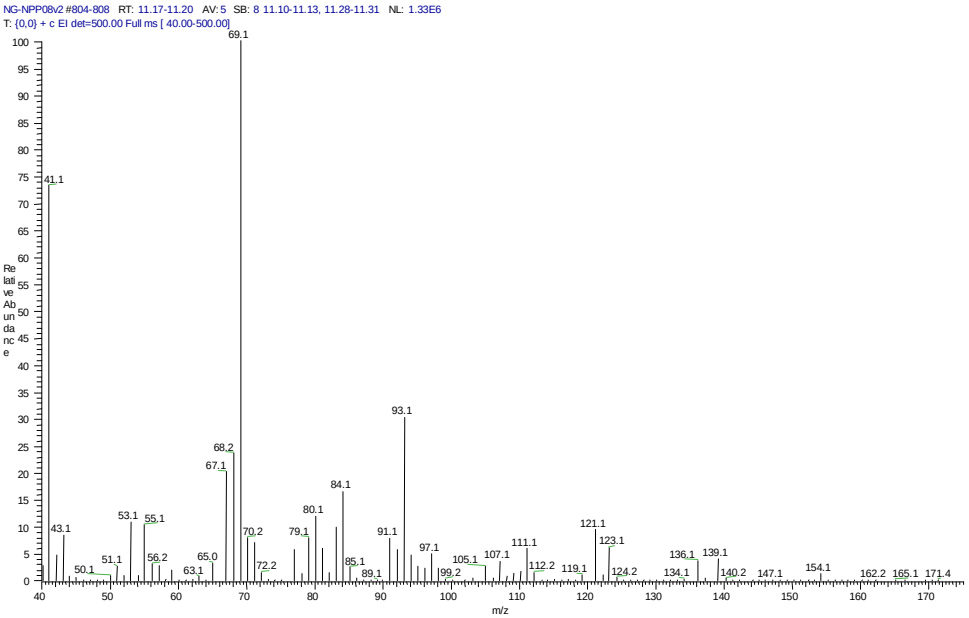


Figure 84 (NG_NPP08v2): MS-spectrum of compound cis-geraniol (**11**) R_t = 11.19 min obtained by GC-MS.

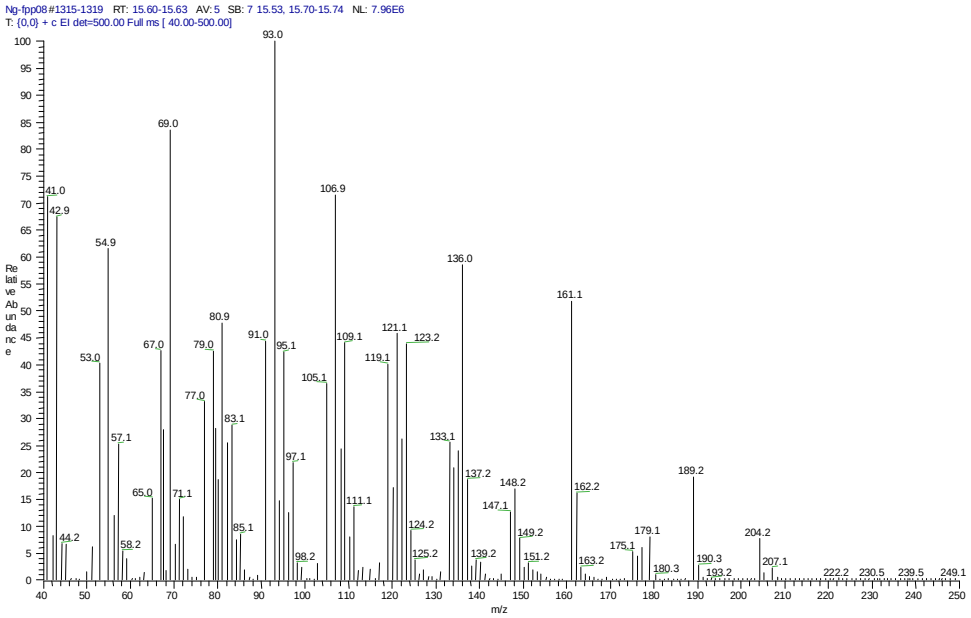


Figure 85 (NG_FPP08): MS-spectrum of compound nerolidol (**19**) R_t = 15.62 min obtained by GC-MS.

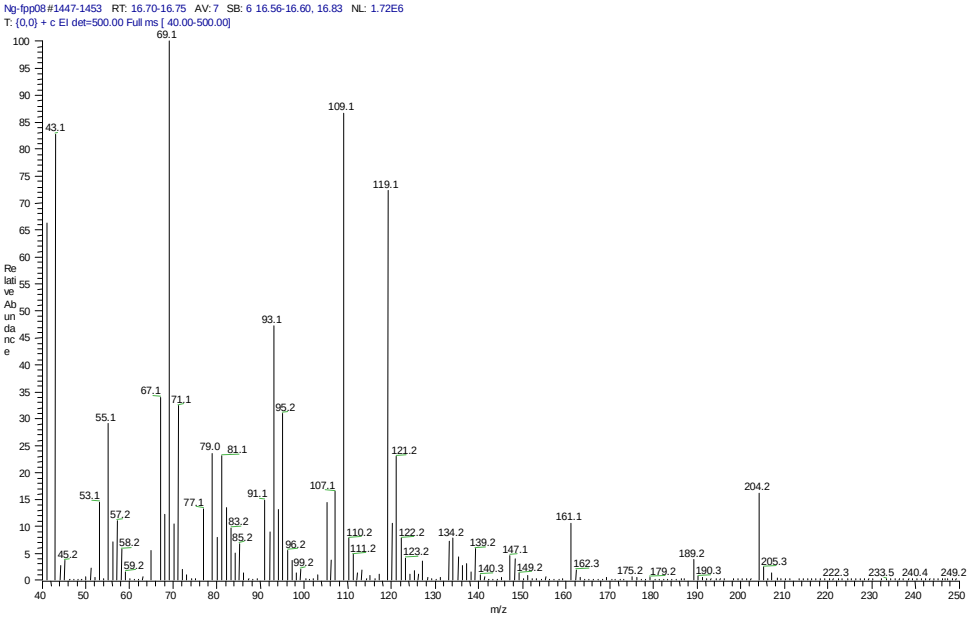


Figure 86 (NG_FPP08): MS-spectrum of compound α-bisabolol (**26**) R_t = 16.74 min obtained by GC-MS.

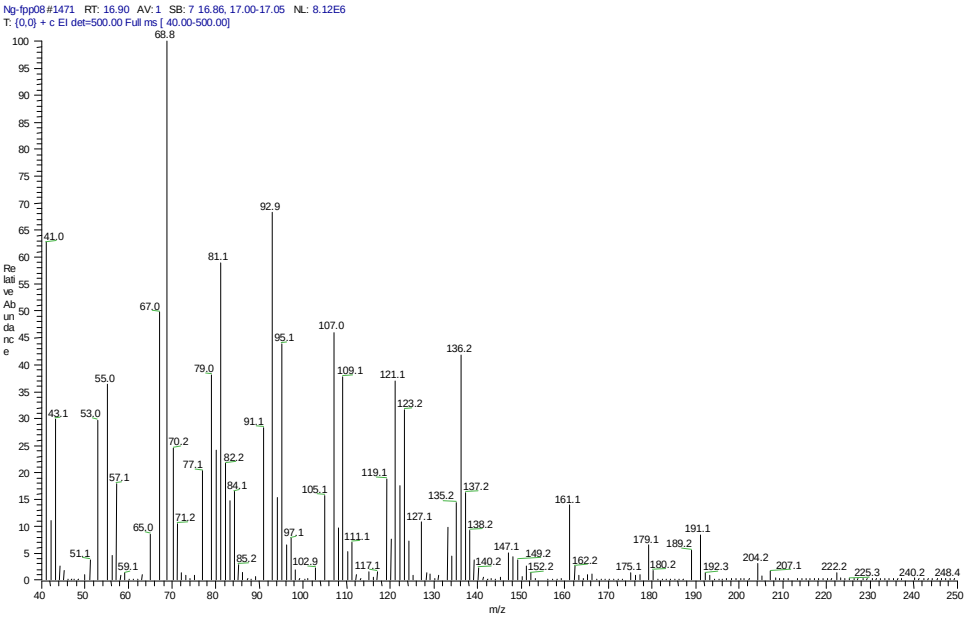


Figure 87 (NG_FPP08): MS-spectrum of compound farnesol (**21**) R_t = 16.90 min obtained by GC-MS.

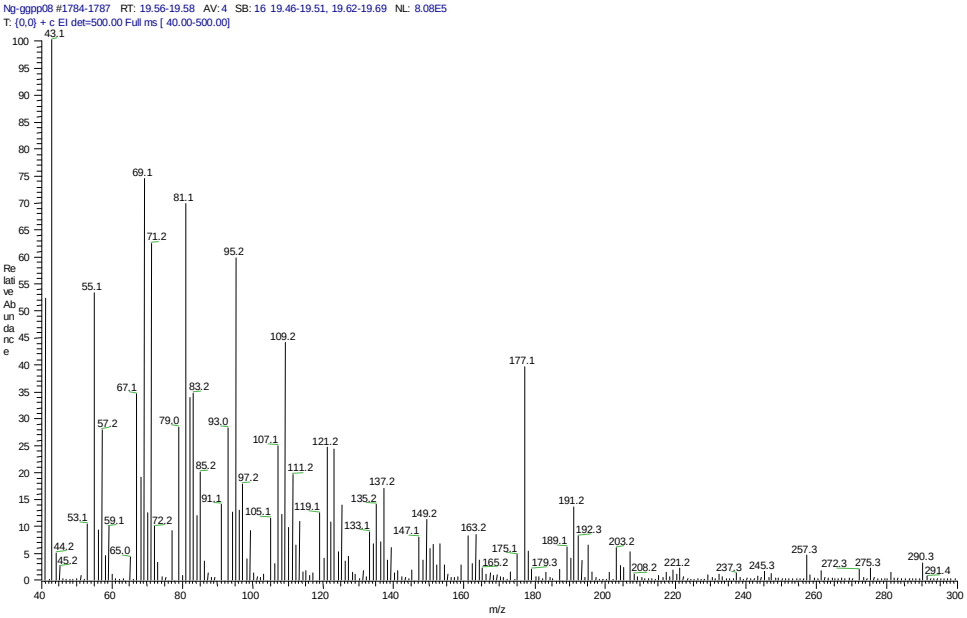


Figure 88 (NG_GGPP08): MS-spectrum of compound sclareol (**27**) R_t = 19.57 min obtained by GC-MS.

8. Acknowledgement

I would like to thank…

Prof. Dr. Toni M. Kutchan for giving me the opportunity to work on such an interessting subject, for the scientific support during my work and for the experience to work at the Donald Danforth Plant Science Center in St. Louis, Missouri. Prof. Dr. Meinhart H. Zenk for the usefull fights and discussions we had, for the critics and impulses he gaves me.

Judy Coyle for the correction of my thesis and my paper and a little bit to be our mother at the Kutchan's lab.

Dr. Kum-Boo Choi for starting the *Salvia sclarea* L. project, especially for the clones.

Prof. Klaus Wasternack for the financial support after my return home from the USA and Prof. Ludger A. Wessjohann for the opportunity to work on different terpenoid projects.

Dr. Susi Frick for beeing a friend you support me in hard times. Dr. Jonathan E. Page and Verona Dietel for giving me all the skills I needed for a Lab. Dr. Jörg Ziegler for being sometimes a fair bully to me, this helps me a lot to learn how to work independent.

and Sylvia, Marco, Otto, Robert, Birgit, Andreas and Maria, Katja, Nadja, Khaled, Alfonso, Jana Na., Jana Ne., Carolin, Matti, Julia, Chotima and Christian, Melanie, Heike, Claudia, Swanhild, Karin and Thomas, Stef to being nice colleques.

Jim, Howard, Dean, Christian, Pete, Chuck, Leslie, Jeanny, Ed, Joe, Andrew, Vince, Tim, Becky and Ed, Sara, Joe, Jang-Jang, Ban, Chuck, Mary, Coumar, Vicky and Billie, and all the other nice colleques from the DDPSC.

Dr. J.Schmidt and Mrs. Kuhnt for help with the GC-MS analysis, Dr. W. Brandt and R. Weber for cooperation on the *Cannabis* project and Dr. T. Vogt for help with the SEC.

Dr. Holger Neef and Dr. Ralph Golbik from the MLU-University for supporting me during my study and beyond.

my friends Anne (reading my thesis), Knut (helping me with the characterization) and Lars for their deep friendship and support. Our "Stammtisch": Ernst and Diana, Ute (thanks for the taxonomy) and Schorsch, Susi and Michael, Katharina and Heike. My friends Danny, Charly, Bill and Cesar (both from the USA).

my brother Benjamin and his wife Sandra for keeping me based. My grandma Sophie for her love and financial support. My family and friends in Minden: Karin, Julia and Marc, Dieter and Elke, Ingo and Sönke, Arndt, Thomas, Lutz and Christina, Reinhardt and Renate, Willi and Jutta.

9. *Curriculum Vitae*/Lebenslauf

Personal detail	
date of birth 05. 23.1977, Minden/Westphalia	
Education	
05/2007 - 02/2008	„Wissenschaftlicher Mitarbeiter“ at Leibniz institute of plant biochemistry (Halle/Saale)
04/2006 – 03/2007	continuation of PhD thesis at Donald Danforth Plant Science Center (St. Louis, MO, USA)
12/2003 - 03/2006	PhD thesis at Leibniz institute of plant biochemistry (Halle/Saale)
10/1997 - 10/2003	study of biochemistry at the Martin-Luther-University Halle-Wittenberg, diploma thesis at Leibniz institute of plant biochemistry (Halle/Saale)
09/1996 – 07/1997	civil service at the neurology unit II, hospital Minden
07/1996 – 08/1996	navy aidman basic training in List/Sylt
08/1987 - 05/1996	Kurt-Tucholsky comprehensive school Minden Abitur degree
Other qualifications	
teaching activities:	presentation of PowerPoint-talks, mentoring of undergraduate and graduate students
Safety Coordinator:	Donald Danforth Plant Science Center (St. Louis, MO, USA),
language:	German (native speaker) English (fluently writing and speaking skills) Spanish (basic skills)
DP:	MS Word; MS PowerPoint; Corel Photo-Paint; Corel Draw; Adobe Photoshop, ChemDraw; DNAstar Software; MS Excel
driver license:	class 3/F

10. Statement/Erklärung

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

Nils Günnewich

Halle an der Saale, November 2008