

# Determination of amino-acidic positions important for Ocimum basilicum geraniol synthase activity

Marc Fischer, Sophie Meyer, Patricia Claudel, Damien Steyer, Marc Bergdoll,

Philippe Hugueney

### ▶ To cite this version:

Marc Fischer, Sophie Meyer, Patricia Claudel, Damien Steyer, Marc Bergdoll, et al.. Determination of amino-acidic positions important for Ocimum basilicum geraniol synthase activity. Advances in Bioscience and Biotechnology, 2013, 2013 (4), pp.242-249. 10.4236/abb.2013.42033 . hal-02644032

## HAL Id: hal-02644032 https://hal.inrae.fr/hal-02644032v1

Submitted on 28 May 2020  $\,$ 

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Determination of amino-acidic positions important for *Ocimum basilicum* geraniol synthase activity

Marc J. C. Fischer<sup>1\*</sup>, Sophie Meyer<sup>1</sup>, Patricia Claudel<sup>1</sup>, Damien Steyer<sup>1</sup>, Marc Bergdoll<sup>2</sup>, Philippe Hugueney<sup>1</sup>

<sup>1</sup>Unit Mixte Recherche Santé Vigne & Qual Vins, Métab Second Vigne, Institut National de Recherche Agronomique, Université de Strasbourg, Colmar, France

<sup>2</sup>Institut de Biologie Moléculaire des Plantes, Université de Strasbourg, Strasbourg, France Email: <u>fischer@colmar.inra.fr</u>

Received 3 December 2012; revised 5 January 2013; accepted 13 January 2013

#### ABSTRACT

Terpenes are one of the largest and most diversified families of natural compounds. Although they have found numerous industrial applications, the molecular basis of their synthesis in plants has, until now, not been fully understood. Plant genomes have been shown to contain dozens of terpene synthase (TPS) genes, however knowledge of their amino-acidic protein sequence in not sufficient to predict which terpene(s) will be produced by a particular enzyme. In order to investigate the structural basis of a TPS specificity, we performed site directed mutations in the geraniol synthase from Ocimum basilicum. The results obtained suggest that a specific region on the catalytic site plays an important role in GPP transformation, either by stabilizing the GPP substrate on the catalytic site, or by enabling its transformation into a monoterpenol via an intermediate carbocation.

**Keywords:** Geraniol Synthase; Heterologous Expression; Monoterpenols; Point Mutations; Yeast

#### **1. INTRODUCTION**

The immense variety of chemical compositions encountered in aromatic plants has always attracted human interest, and such plants are believed to have been cultivated since the Neolithic for their fragrance, food improvement abilities, and traditional medicinal properties. Each of these plants can contain unusually high concentrations of chemicals and several hundreds of volatile compounds, with one or two of them being major compounds. Monoterpene aroma are amongst the most basic and prevalent fragrance representatives. They belong to the family of plant isoprenoids, which are amongst the most diversified compounds, with over 40,000 molecules having been described [1]. Their diversification started

with ancestral plants, which were almost devoid of terpenoids, as illustrated by the fact that the Physcomitrella patens moss has only one terpene synthase gene, encoding a diterpene precursor from the synthesis of the gibberellic acid hormone [2]. Amplification of the terpenoid synthase gene family occurred via gene duplication and is still an ongoing process [3]. The economic importance of some isoprenoids is based on their perfume, food flavor improvement properties, and pharmaceutical or industrial properties. Recent plant genome analyses have highlighted hundreds of putative monoterpenoid synthase genes. The functional characterization of monoterpenoid synthases has therefore generated a large amount of data, and shown that a large proportion of the characterized enzymes catalyze the synthesis of several end products [4,5]. However, it has until now been impossible to predict which monoterpenoid(s) will be produced by a particular enzyme. The tremendous range of possible variations in the carbocationic reactions (cyclizations, hydride shifts, rearrangements, and termination steps) catalyzed by terpenoid synthases explains the wide range of possible products [6]. Such studies have usually been carried out using heterologous truncated cDNA expression in Escherichia coli (since the native plant enzyme targets the chloroplast and bacteria cannot process introns), followed by enzymatic assays with crude cell extracts.

Some plants, such as basils, have been more extensively studied for their aromatic properties and can be considered as model plants for monoterpenoid synthase studies. The basil genus comprises approximately 150 species, which are found mainly in the tropics, and which were probably first cultivated in India [7]. Many varieties of these aromatic plants are now grown worldwide, particularly in Southeast Asia, Mediterranean countries and California. Sweet basil (*Ocimum basilicum*) has various cultivars, varying in shape, color and fragrance. Differences in fragrance have been linked to variations in compounds such as geranial, neral, methylchavicol, eugenol [8,9]. The functional characterization of sweet basil mo-



noterpenoid synthases expressed in *E. coli* has enabled the identification of a geraniol synthase (GES) [9] and a (*R*)-linalool synthase ((*R*)-LIS) [10]. Geraniol and linalool are the only products synthesized after heterologous expression by *O. basilicum* GES and (*R*)-LIS. Both genes are very similar, with a 78% identity on the nucleotide sequence and a 82% identity on the aminoacid sequence. Iijima *et al.* [9] made the first functional/structural analysis of these genes by investigating whether a hybrid protein could be affected in terms of either its activity or its specificity. A GES-(*R*)-LIS construct exhibited a dual geraniol/linalool synthase activity, whereas a (*R*)-LIS-GES construct was inactive.

An alternative approach to the investigation of terpene synthase specificities is to use 3D modeling, combined with site directed mutagenesis. This method has been used only once before, on two cyclic monoterpene synthases, the 1,8 cineole and sabinene synthases from Salvia fruticosa [11]. In the present study, our aim was to characterize the geraniol synthase catalytic site via a similar site-directed mutation approach, based on the specificities of geraniol or linalool synthases. We first used an array of functionally characterized GES and LIS from several plants, to determine the amino-acid signatures on the catalytic site that would correlate with the GES or LIS activity (Figure 1). Both GES and LIS have been functionally characterized in O. basilicum and Chinese basil (Perilla frutescens) [12]. We therefore used an alignment of O. basilicum, P. frutescens GES and (R)-LIS cDNA, combined with a 3D model construction, to define putative GES/LIS-specific amino-acids for sitedirected mutagenesis. In a previous study we engineered Saccharomyces cerevisiae to convert it into a suitable heterologous expression system for truncated O. basilicum GES cDNA [13]. We used this system to test whether any O. basilicum GES modifications would result in a modification of the GES enzyme specificity.

#### 2. MATERIALS AND METHODS

#### 2.1. Directed Mutagenesis of GES

Plasmid pGB7/PMA1 was constructed by integrating a 700 pb PMA1 promoter in the *NotI-SacI* digested plasmid pGB7 [14]. The pSM5 plasmid was constructed through integration of *O. basilicum* geraniol synthase (GES), by means of recombination in the pGB7/PMA1 plasmid, in fusion with the green fluorescent protein (GFP) from pGB7. The pSM5 plasmid was then used for site-directed mutations of GES. The plasmids carrying the mutations were created using the QuikChange Site-Directed Mutagenesis Kit from Stratagene, in accordance with the manufacturer's instructions. The complementary Forward and Reverse primers were designed as described in **Table 1**. All of the plasmids were verified by sequencing.

#### 2.2. Yeast Transformation

The haploid G strain (*Mata*, *his* $3\Delta$ , *leu* $2\Delta$ 0, *ura*3-, *trp* $1\Delta$ 63, *YJL*167*W::kanMX*4 [*pFL*44*erg*20*K*197*G*]) [13] was transformed with a pSM5 plasmid series expressing wild type and mutated GES. A 1.5 mL aliquot of G cells, grown to the stationary phase YNB medium supplemented with the appropriate auxotrophic amino acids, was centrifuged (1 min, 10,000 g) and 10 µL of mutated plasmid pSM5 (125 ng) was added to the yeast pellet. Transformed cells were selected for uracil prototrophy and G418 resistance on a minimal medium.

#### 2.3. Monoterpenoid Extraction and Quantification

The cells from a stationary phase culture were harvested by centrifugation (5000 g for 5 min). Octanol (4  $\mu$ g) and Ethylheptanoate (4  $\mu$ g) were added as an internal standard to 20 mL culture medium supernatant. Monoterpenoids were extracted using a stir bar (Twister; Gerstel, Mülheim a/d Ruhr; Germany) sorptive extraction liquid desorption gas chromatograhy mass spectrometry (SBSE-LD-GC-MS) [15] with the following modifications: acetonitrile as solvent and a 1  $\mu$ L injection volume.

Monoterpenol extracts were analyzed by GC-MS using an Agilent 6890N gas chromatograph equipped with a Gerstel MPS2 sampler, coupled to an Agilent 5975B inert MSD (Agilent Technologies). The gas chromatograph was fitted with a DB-Wax capillary column (60 m  $\times$  0.32 mm i.d.  $\times$  0.50 µm film thickness, J&W Scientific) and helium was used as the carrier gas  $(1 \text{ mL} \cdot \text{min}^{-1} \text{ con-}$ stant flow). The GC oven temperature was programmed to increase from 45°C to 82°C at a rate of 20°C·min<sup>-1</sup>, and then to increase to  $235^{\circ}$ C at a rate of  $2.7^{\circ}$ C·min<sup>-1</sup> (hold for 15 min). The injector was set to 250°C and used in pulsed splitless mode (25 psi for 0.50 min). The temperatures of the interface, MS ion source and quadrupole were 270°C, 230°C and 150°C, respectively. The mass spectrometer was operated in the electron impact ionization mode (EI, 70 eV), and the masses were scanned over a m/z range of 29 - 300 amu. Agilent MSD ChemStation software (G1701DA, Rev D.03.00) was used for instrument control and data processing. The mass spectra were compared with the Wiley's library reference spectral bank. Total amounts of geraniol, linalool, citronellol, nerol and  $\alpha$ -terpineol were determined using linear calibration curves, leading to a value of 0.98 for  $R^2$ , over the range of concentrations from 0 to 200 mg/mL.

#### 2.4. GES Model Building

Homology modeling was implemented using Modeller [16] with *Mentha spicata* (4S)-limonene synthase as template [17]. The *O. basilicum* GES sequence was aligned with *Mentha spicata* (4S)-limonene using ClustalX

		1				50		381			420	)	
	AAR11765.1		MSCARITVTL	PYRSA	KTSIQRGITH	YPALIRPRFS	/ /	LPEPKYSSVF	R IELAKAISI	LVIDDIPD	Y GEMDDLILF	r //	
	Q5SBP3.1		MSCARITVTL	PYRSA	KTSIQRGITH	CPALLRPRFS	/ /	LPHPKYSSSF	IESAKAAALI	. YVIDDIFD	Y GKMDELILF	г //	
	ADK62524.1		MASARSTISL	SSQSSHHGFS	KNSFPWQLRH	SRFVMGSRAR	/ /	LPEPKYSACF	IELAKTIAI	L LVIDDIFD	Y GKMEELALF	г //	
NP_(	ADR74217.1		MSRFVTM	PSHVLPSSFV	APSLQVSS	SPCSWRTRPS	/ /	FPDPRHSSCF	IELTKAIAI	LVIDDIYD	Y GSLDELALF	г //	
	ADR74218.1	M	AFNMSRFVTM	PSHVLPSSFV	APSLQVSS	SPCSWRTRPS	/ /	FPDPRHSSCF	IELTKAIAI	LVIDDIYD	SY GSLDELALF	г //	
	ADR66822.1		MSRFVTM	PSHVRPSSFV	APSLQVSS	SPCSWRTRPS	/ /	FPDPRHSSCF	IELTKAIAI	LVXDDVYD	SY GSLDELALF	г //	
	CAD29734.2		.MALQMIAPF	LSSFLPN.PR	HSLAAHGLTH	QKCVSKHISC	/ /	TPEPQFGRCQ	EVVAKVAQL	I III DDIYD	YY GTVDELELF	г //	
	ADR74209.1	М	ELTLTSLSPL	AYGALNC.RK	NFAMASPRMR	IKQGRSELPN	11	APEPHFSKSF	IGLTKFICI	. TAIDDMYD	Y GSPDELRRF	г //	
	BAG82825.1	M	ALPALFGSSL	PSSIRHNOPS	LFSFRHPRFC	SSSSSASFSS	11	VFEPOYTDFF	ELNTRIACM	A TLIDDVYD	Y GTPEELELL	r //	
	001233805.1			MVSTLSNIGM	MVVTFKRPSL	FTSLERRSAN	/ /	LEEPOHSYEE	RIJTTKVIVE	STIDDIYD	Y GTLDELELF	т //	
	09SPN1.1			MASISLEPYS	TLKOTSPLAR	GTAYNELYST		SYLPNESHGE	RTTTKVAAM	TTLDDVYD	F GTLGELEOF	- , , т //	
	ABB30218.1	MSSTSO	KVVTGLNKAA	ANNNLONLDR	RGFKTRCVSS	SKAASCLEAS		FEPHOYGYOF	RVAAKTTTL	TSTDDVYD	Y GTLDELOLF	- , , т //	
	AAY88965 1	MSSISO	KVVIGLNKAA	ANNNLONLDR	RGFKTRCVSS	SKAASCLEAS		FEPHOYGYOF	RVAAKIITI.	TSIDDVYD	Y GTLDELOLF	т //	
	ACN42012	MCSISO	KAALCINKYY	ANNILONLOR	RGENTROVSS	SKAASCIRAS	, ,	FERHOVOVOR	DVAAKTTTL	TSIDDVID	Y GTIDELQUE	т //	
	ACN42010 1	MCSISQ	KAALCINKYY	ANNOLONI DR	POFKTPRUSS	SEAASCLEAS	, ,	FERHOVOVOR	DVAAKTTTL	TSIDDVID	Y GTIDELQUE	т //	
	ACN42011 1	MVSI	BIANT INKKD	SAKHADMADK	KYCKDCMBAC	LSSSACLRAS		FEDHOVCVHE	KVAAKTITI.	TSUDDVYD	Y GTIDELQUE	т //	
	ACN42011.1	MCCM	DTVUATMEED	QUEUUDNUDV	VACKDOMICUO	TO ACTRO	, ,	FEDUOVCVUE	VINANTITI.	TELDDVID	V CTIDEIQUE	T / /	
	ACN42005.1	MCCT	DIVUNTMEED	SVEIVDNUDV	VACUT OWDUC	CONTACIDAC	, ,	FEDUOVCVUE	VINAVITID.	TELDDVID	V CTIDEIQUE	T / /	
	09U2D/ 1	MOTTOMUU	NTI VATNIKE	VIEUTVNIVDA	ACTNI DWCT C	DCCCAACDDT		FEAUVECVED	NTAANTITI.	TALDDVID	V CTIDEIQUE	T //	
	Q00204.1	MCIIISVNHH	NVAILSKPKV	KLEHIKNKKS	ASINLPWSLS	MDAAAULDDE		FEARAGIEF	KIAAKIIIL.	I TALDDVID.	.I GILDELQLF	I //	
	ABB/3045.1				MSININ	MPAAAVLRPF		FEPHQIGIQF	KVATKIITL.	I TSLDDVYD.	Y GTLDELQLF	r //	
	ACM92062.1				MALISSWRVN	MPAAAVLRPF	//	FEPHQIGIEF	K KVAAKIITM/	A TSLDDVYD	Y GTLGELQLF	r //	
D1/1	AA085533.1	• • • • • • • • • • •	• • • • • • • • • • •		MALIATKISS	RSCFVSAYPN	/ /	LQDPTLTEQF	LDLTKPISL	/ YVIDDIFD	Y GELEELTIF	r //	
DM1 DM2								VP	LAKATST	_ YVIDDIPD1			
DMZ									DANA131	<u> </u>			
		461			500					550			580
	AAR11765.1	GRYVLLNLKS	TWIDMIEGFM	EEAKWFNGGS	APKLEEYIEN	GVSTAGAYMA	FAHIFFLIGE	GVTHQNSQLF	TQKPYPKVFS	AAGRILRLW	DLGTAKEEQE	RGDLASCVQL	FMKEKS
	Q5SBP3.1	GRIALPYLKS	VWIETIEAYM	VEVKWFSGGS	APKLEEYIEN	GASTVGAYMV	LVHLFFLIGE	GLTHQNVLFF	KOKPYHKPFS	AAGRIFRLW	DLGTSQEEEE	RGDMASSIRL	FMKEYK
	ADK62524.1	GWSVLPYLRY	TWMDMIEGEM	VEAKWENGGS	APNLEEYTEN	GVSTAGAYMA	LVHLEFLIGE	GVSAONAOTI.	LKKPYPKLES	AAGRILRIW	DIGTAKEEEG	RGDLASSIRL	FMKEKN
	ADR74217.1	GWSVIEDLKR	TWMDIFGAFL	AEAYCEKGGH	VPSLEEYLTN	AVTTGGTYMA	LVHAFFLMGO	GVTRENMAML	KPYPNIES	CSGKTLBLW	DIGTAREEOE	RGDNASSIEC	YKRERE MI
	ADR74218 1	CWSVIEHIKR	TWMDIFGAEL	AFAVOFKOCH	VESLEEVITN	AVTTCCTVMA	LVHAFFIMGO	CUTRENMAMI.	KDVDNIES	CSCKTLPLW	DIGTAREFOR	RCDNASSIEC	AKDEDE WI
	ADR66822 1	GWSVIEDIKR	TWMDIFGAEL	AFAHCEKCCH	VESLEEVINN	AVTTGGTYMA	LVHAFFLTRO	GVTRENMAMI.	KDVDNIES	CSCKTLRINE	DIGTAREEOE	RCDNASSIEC	VKDEDEDEMI
	CAD29734 2	CRNVIDVIRN	TWTELCKAEL	VEAKWYSSCY	TOTLEEVIOT	SWISTOSLDM	OTVVFALLCK	NLAPESSDEA	EKISDILP	LCCMMIRI	DIGTSTOFIK	REDURKSTOC	VMHEAC
	ADD74200 1	DIVIIDVING	OWINI CTOVO	MELOWEVNCY	VDOTDEVION	AMTOLCODAA	MULINEIMCC	ATKONI NNCI	DNA CNI L V	WCCITTDIC	DICTRIARIA	RODVINGIQC	VMTEKC
	DAC02025 1	CINDIDUIDU	INCDECKADM	VENUMENNOT	KP31DE1LSN	AWISVGGEAA	TINCVETTED	VITERCINVU	OVTDOLMU	CONCEPTION	DIGISLADIA	RODVARSIQC	VMNETC
ND	DAG02023.1	GINPIPILKK	CWADI CKOVI	OFFICENTIAL	KPILKEIMDV	AVDSIGGLIL	LUNGIFLIID	OTEKENTO	SKIPSVMH	NONTEDINE	DISISSHELA	RGDNSKALEC	IMINEIG
NP_	001233805.1	DINVLPILIK	SWADLCKSYL	QEARWIHNGI	KPNLEEIMDN	ARISIGVPMV	LVHSLFLVTN	QITKEALDSL	TNYPDIIR	WSATIFRINI	DIGTSSDELK	RGDVSKSIQC	IMNERG
	Q9SPNI.I	GFLILPYIKK	AWADLCKSYL	VEAQWYHRGH	I PTLNEYLDN	ACVSISGPVA	LMHVHFLTSV	SSTKEIHHCI	ERTQNIVR	YVSLIFRLT	DLGTSLGEME	RGDTLKSIQL	YMHETG
	ABB30218.1	GFIVIPYLQR	SWVDLAESFL	KEANWYYSGY	TPSLEEYIDN	GSISIGAVAV	LSQVYFTLAN	SIEKPKIESM	YKYHHILR	LSGLLVRLHL	DLGTSLFEKK	RGDVPKAVEI	CMKERN
	AA188965.1	GFIVIPYLQR	SWVDLAESFL	KEANWYYSGY	TPSLEEYIDN	GSISIGAVAV	LSQVYFTLAN	SIEKPKIESM	YKYHHILR	LSGLLVRLHL	DLGTSLFEKK	RGDVPKAVEI	CMKERN
	ACN42012	GFIVIPYLQR	SWIDLAESFL	KEANWYYSGY	TPSLEEYIDN	GSVSIGAVAV	LSQVYFTLAN	SIEKPKIESM	YKYHHILR	LSGLLVRLH	DLGTSLFEKK	RGDVPKAVEI	CMKERN
	ACN42010.1	GFIAIPYLQR	SWVDLAESFL	KEANWYYSGY	TPSLEEYIDN	GSISIGAVAV	LSQVYFTLAN	SIEKPKIESM	YKYHHILR	LSGLLVRLH	DLGTSLFEKK	RGDVPKAVEI	CMKERN
	ACN42011.1	GFITIPYLQR	SWADLVEAYL	KEAKWFYNGY	VPSMEEYLNN	AYISIGATPV	ISQVFFTLAT	SIDKPVIDSL	YEYHRILR	LSGMLVRLPI	DLGTSPFEMK	RGDVPKAIQL	YMKERN
	ACN42009.1	GFITIPYLQR	SWADLVEAYL	KEAKWYYNGY	TPSMEEYLNN	AYISIGATPV	ISQVFFTLAT	SIDKPVIDSL	YEYHRILR	LSGILVRLPI	DLGTSPFEMK	RGDVPKAIQL	YMKERN
	AAX16075.1	GFI.IPYLQR	SWADLVVAYL	KEAKWFYNGY	TPSMEEYLNN	AYISIGATPV	ISQVFFTLAT	SIDKPVIESL	YEYHRILR	LSGMLVRLPI	DLGTSSFEMK	RGDVPKTIEL	YMKERN
	Q8H2B4.1	GFISIPFLHR	AWVDLVEGYL	QEAKWYYTKY	TPTMEEYLNY	ASITIGAPAV	ISQIYFMLAK	SKEKPVIESF	YEYDEIIR	LSGMLVRLP	DLGTLPFEMK	RGDVAKSIQI	YMKEQN
	ABB73045.1	GFTSIVYLQR	SWVDLLKGYL	KEAKWYNSGY	TPSLEEYFDN	AFMTIGAPPV	LSQAYFTLGS	SMEKPIIESM	YEYDNILR	VSGMLVRLP	DLGTSSFEME	RGDVPKSVQL	YMKETN
	ACM92062.1	GFI.LPYLKR	SWEDLIDSYL	KEAQWINNGY	TPSLEEYLNN	ACISFGATPV	IMHVFFTLSV	SIDKPVIECL	YRTHNILR	YVGMLVRLT	DLSTSSGEME	RGDELKTIEL	YMKERG
	AA085533.1	GWNPTYALRQ	SWASLCKAFL	VEAKWFNSGY	LPTTEEYMKN	GVVSSGVHLV	MLHAYILLGE	ELTKEKVELI	ESNPGIVS	SAATILRLW	DLGSAKDENQ	DGTDGSYVEC	YLNEYK
DM3		S	TWIDMIEAYM	<u>VEV</u> KWFNGGS	A								
DM4			WIDMIEGYM	EEAKWF									
DM5					YIEN	GVSTVGAYMA	FAHI						
DM0 DM7						AGAY PA	FAHIFFLI		FC	AACDTEDIWE	DT		
DM8									10	AAGKITKIWI	EEOE	RGDMASCVOL	F
2110											1191	10012100121	-
		581	600					650				684	
	AAR11765.1	.L.TEEEARS	RILEEIKGLW	RDLNGELVYN	KNLPLSII	KVALNMARAS	QVVYKHDQDT	YFSS	VDNYVDALFF	TQ			
	Q5SBP3.1	.LSTVEEARS	CVLEEISRLW	KDLNEGLISI	KD.ALPLTIV	KVALNIARTS	QVVYKHEQHT	YMLS	VDNYVEALFF	TPLLSS			
	ADK62524.1	.LTTEEEGRN	GIQEEIYSLW	KDLNGELIS.	KG.RMPLAII	KVALNMARAS	QVVYKHDEDS	YFSC	VDNYVEALFF	TPLL			
	ADR74217.1	TVLEDEACRK	HIRQMIQSLW	VELNGELVA.	SS.ALPLSII	KAAFNLSRTA	QVIYQHGDDN	KTSS	VEDHVQALLF	RPVSSNGHAC	ITMH		
	ADR74218.1	TVLEDEACRK	HIRQMIQSLW	VELNGELVA.	SS.ALPLSII	KAAFNLSRTA	QVIYQHGDDN	KTSS	VEDHVQALLF	RPVSSNGHAQ	ITMH		
	ADK00822.1	IVLUDEACRK	HINGLEOFTW	VELNGELVA.	SS.ALPLSII	DHAMNLOPUS	VCTYKHGDCE	SDGFGDPGPO	ARDHAÖYTTL ARDHAÖYTTL	REVSSNGHAQ	TSEVUDGCCA		
	ADR74209 1	TSEEOARD	OVEKLIBYSW	KKLNEASTDS	SLPKSLT	NSSLNMARSA	OCIFOEGDGI	GTSVG V	TKDRLTSFIT	KPILIEPS	INPYLDGMKM	SNRR	
	BAG82825.1	ASEEIARE	HIRHLVRETW	KKMNKEVFED	YPFSGFGPFL	SACLNLARAS	HCFYEYGDGY	GLPDHQ	TRDHLASTIF	ESVSLD			
NP	001233805.1	ASEEEAIE	HIEFLIQETW	EAMNTAQS.K	NS.PLSETFI	EVAKNITKAS	HFMYLHSD	VKSSI	SKILFEPIII	SNVAFALK			
_	Q9SPN1.1	ATEPEARS	YIKSLIDKTW	KKLNKERAIV	SS.ESSREFI	DYATNLARMA	HFMYGEGDED	FRLDVIKSHV	SSLLFTPIQG	I			
	ABB30218.1	VTEEEAEE	HVKYLIREAW	KEMNTATTAA	GC.PFMDELN	VAAANLGRAA	QFVYLDGDGH	GVQHSK	IHQQMGGLMF	EPYV			
	AAY88965.1	VTEEEAEE	HVKYLIREAW	KEMNTATTAA	GC.PFMDELN	VAAANLGRAA	QFVYLDGDGH	GVQHSK	IHQQMGGLMF	EPYV			
	ACN42012	VTEEEAEE	HVKYLIREAW	KEMNTATAAA	GC.PFMDELN	VAAANLGRAA	QFVYLDGDGH	GVQHSK	1HQQMGGLMF	EPYL			
	ACN42010.1	DIEEEAEE	HVDFT.TDF7M	KEMNTATAAA	GC.PFMDELN	VAAANLGRAA	QEVILDGDGH	GVQHSK	THORTACTTE	RPIV			
	ACN42011.1	. ATELEAGE	HVRFLIREAW	KEMNTATAAV	DC.PFTDDLV	TAAANI,GRAA	OFMYLDGDGN	нзо	THORITACITE	EPYA			
	AAX16075.1	ATEIEAOE	HVRFLIREAW	KEMNTATAAA	DC.PFTDDLV	AAAANLGRAA	QFMYLDGDGN	HSO	LHQRIASLLF	EQYA			
	Q8H2B4.1	ATREEAEE	HVRFMIREAW	KEMNT.TMAA	NS.DLRGDVV	MAAANLGRDA	QFMYLDGDGN	HSQ	LQHRIANLLF	KPYV			
	ABB73045.1	ATEEEAVE	HVRFLNREAW	KKMNTAEAAG	DS.PLVSDVV	AVAANLGRAA	QFMYFDGDGN	QSS	LQQWIVSMLF	EPYA			
	ACM92062.1	ATEIEAQE	HIRFLINKTW	KKMNKEVAIA	DC.PPF	TLATNLGRMA	HFMYVDGDGN	GNRHSQ	IHQRIMSLLF	TQYALI			
DMC	AA085533.1	.GSTVDEART	HVAQKISRAW	KRLNRECLNP	CPFSRSFS	KACLNIARTV	PLMYSYDDD.	QRL	PDEYLKSLM.				
DM10							AAAKHDOGM AAAYHDOGM	IFSS VFCC	VDNYVDA				
DUIT U							* * TIUDÃGN	ırəə	* DIVI V DIN				

**Figure 1.** Partial multi-alignment from functionally characterized GES and LIS, made with Multalin (http://multalin.toulouse.inra.fr/multalin/). GenBank accession code for GES: *O. basilicum* (AAR11765.1), *Vitis vinifera* (ADR74217.1, ADR66822.1, ADR74218.1, ADR74217.1), *Cinnamomum tenuipile* (CAD29734.2), *P. Frutescens* (ABB30218.1), *P. Frutescens var. hirtella GES* (GenBank: ACN42012), *Perilla setoyensis* (ACN42010.1), *Perilla citriodora* (AAY88965.1), *Phyla dulcis* (ADK62524.1). GenBank accession code for LIS: *P. Frutescens var. hirtella* (ACN42011.1), *Lavandula angustifolia* (ABB73045.1), *Perilla setoyensis* (ACN42009.1), *Perilla citriodora* (AAX16075.1), *Osmanthus fragrans var. thunbergii* (ACM92062.1), *Solanum lycopersicum* (NP\_001233805.1), *Ocimum basilicum* (Q5SBP3.1), *Artemisia annua* (Q9SPN1.1), *Arabidopsis thaliana* (AAO85533.1), *Backhousia citriodora* (BAG82825.1), *Mentha aquatic* (Q8H2B4.1), *Vitis vinifera* (ADR74209.1). FARM (first Asp-rich motif) and SARM (second Asp-rich motif) domains are surrounded. Directed mutations made on *O. basilicum* GES (AAR11765.1) are specified (DM1-10) and the amino acids modified in each directed mutation are underlined.

Table 1. Specificities of forward oligonucleotides used to modify the O. Basilicum GES. Mutated codons are underlined.

Mutation	FW primer sequence
DM1	GTT AGA ATA GAG TTG GCG AAA GCC GCC GCT CTT CTC TAC GTG ATC GAT GAT ATT TTC GAT ACC
DM2	GTT GGC GAA AGC CAT CTC TAT T <u>AT CAC A</u> GT GAT CGA TGA TAT TTT CG
DM3	TCT ACG TGG ATA GAC ATG ATT GAA GCT TAC ATG GTG GAA GTA AAA TGG TTC AAT GGT GGA AGT GCA
DM4	TGG ATA GAC ATG ATT GAA GGT <u>TAC</u> ATG GAG GAA GCA AAA TGG TTC
DM5	TAT ATA GAG AAT GGA GTG TCC ACG <u>GTA GGA GCA TAC ATG GCT TTT GCA CAC ATC</u>
DM6	GCA GGA GCA TAC <u>CCG</u> GCT TTT GCA CAC ATC TTC TTT CTC ATA GG
DM7	TTC TCC GCC GCC GGC CGC ATT TTT CGC CTC TGG GAT GAT CTC GG
DM8	GAA GAG CAA GAG CGA GGA GAT <u>ATG</u> GCT TCG TGC GTG CAG TTA TTT
DM9	GTT GTG TAC AAG CAC GAT CAA <u>CAC</u> ACT TAT TTT TCA AGC GTA GAC
DM10	GTG TAC AAG CAC GAT CAA GAC <u>GGT AAT</u> TTT TCA AGC GTA GAC AAT TAT GTG GAT GCC CTC

[18]. Visual analysis of the model and the illus- trations were made using PyMOL [19].

#### **3. RESULTS AND DISCUSSIONS**

#### 3.1. Monoterpenoids Quantification

Although the product specificities of monoterpenol synthases have been extensively studied using heterologous expression, the manner in which their amino acidic sequence relates to the specific enzymatic function still remains largely unknown, with the notable exception of the sabinene and 1,8 cineole synthases from S. fruticosa [11]. Although classical sequence comparisons led to the identification of seven subfamilies [20], further analysis of their enzymatic activity was not possible, due to the history of this protein's family. The original plant terpenoid synthase was a di-terpene synthase, and it is thought that monoterpenoid synthases originated via the loss of the domain of approximately 200 amino-acids [21]. This loss did not take place in a single ancestral event, but occurred several times during plant evolution, leading to an ongoing process of mono and sesqui-TPS generation [22]. In the present study, to enable catalytic site analysis, we positioned each residue differing from O. basilicum and P. frutescens GES and LIS within the predicted catalytic site, and used an in vivo characterization approach to investigate the functional consequences of amino-acidic substitutions at the putative catalytic site of O. basilicum GES. We discovered that 3 mutations resulted in an abolished function of GES and one mutation resulted in a partial loss of function, whereas 5 mutations did not affect the original function.

Expression of GES in the erg20K197G yeast strain resulted in the biosynthesis of a number of monoterpenols (geraniol, linalool, citronellol, nerol) in a YNB medium (**Table 2**), whereas almost no production of  $\alpha$ -terpineol was detected. Untransformed control erg20K197G cells exhibited a much lower production of monoterpenols. The production level of geraniol was 16 times higher arising from expression of GES. An increase in linalool, citronellol and nerol was also detected, but their overall amounts represented only 4% - 12% of the geraniol detected. Expression of 4 mutated GES K197G DM1, K197G DM2, K197G DM5 and K197G DM7 resulted in a monoterpenol profile very similar to that observed with the untransformed control. The expression of 5 other GES mutants (K197G DM3 - K197G DM4 -K197G DM6 - K197G DM8 - K197G DM9) resulted in a monoterpenol profile, very similar to that observed with erg20K197G yeast cells transformed with wild type GES. Only one strain (K197G DM10) exhibited an intermediate monoterpenol profile, between that of erg20K197G transformed with wild type GES and that of the untransformed control.

#### 3.2. Correlation between Model Prediction and Experimental Results

It is of interest to compare our results with previous structure function studies of *O. basilicum* GES. Using a GES-(*R*)-LIS protein chimeric approach, the function was evaluated by Iijima *et al.* [10]. Replacement of the first 360 aa on GES by their (*R*)-LIS homologues resulted in a loss of function. Our catalytic site modeling indicates that only one area in the first 360 aa on GES could potentially be involved in the catalytic site, suggesting a significant role for this region in the proposed GES function (**Figure 2**(A) and (B)). The site-directed modification of this area corresponds to DM1, and the resulting K197G DM1 strain indeed exhibits a dramatic loss of GES activity (**Table 2**). This result prompted us to test another mutation, DM2, replacing the 2 amino-acids

**Table 2.** Monoterpenol content ( $\mu g \cdot L^{-1}$ ) of FPPS mutated yeast strains, untransformed or transformed by control or mutated GES plasmid. Terpenoids were extracted from a minimal medium at their stationary growth phase. The results are the mean of 3 experiments  $\pm$  standard deviation.

Yeast strain	Geraniol	Linalool	Citronellol	Nerol	$\alpha$ -terpineol
K197G	$527 \pm -22$	$400 \pm 23$	$38 \pm 1$	ND	ND
K197G GES	$8423\pm297$	$1031 \pm 9$	$419\pm20$	321 ± 3	ND
K197G DM1	$703\pm286$	$301\pm131$	$24\pm28$	ND	ND
K197G DM2	$736\pm234$	$486\pm216$	$41 \pm 40$	$151\pm138$	ND
K197G DM3	$6624\pm2019$	$590\pm362$	$189\pm147$	$157\pm181$	ND
K197G DM4	$10715\pm366$	$832\pm63$	$252 \pm 27$	$285 \pm 3$	ND
K197G DM5	$566 \pm 10$	$431 \pm 17$	ND	ND	ND
K197G DM6	$6782\pm758$	$897\pm52$	$285\pm27$	$290 \pm 3$	ND
K197G DM7	$736\pm210$	$353\pm124$	$24\pm28$	ND	ND
K197G DM8	$9389\pm2058$	$509 \pm 63$	$177 \pm 79$	$273\pm12$	ND
K197G DM9	$10,635 \pm 240$	$1187\pm211$	$484\pm73$	$329 \pm 24$	$3 \pm 3$
K197G DM10	$2790\pm370$	$270\pm30$	$50 \pm 1$	$260 \pm 2$	ND

ND: Not Detected.



**Figure 2.** GES catalytic site, showing the amino acids in the catalytic site pocket. Native amino-acids are color-coded in grey and cyan (A, B), and the amino-acids modified by directed mutagenesis are color-coded in purple for each construct (DM1-DM10).

corresponding to the difference between P. frutescens GES and (R)-LIS in O. basilicum GES. GES activity in the K197G DM2 strain is also suppressed. It is also noteworthy that *P. frutescens* belongs to the same family (Lamiaceae) as sweet basil. From this comparison it appears that there is a good correlation between DM1, DM2 and the GES-(R)-LIS inactive chimeric construct. One can infer that the loss of GES function obtained by Iijima et al. [10] was due to a modification of this area, defined by DM1 and DM2. The same area also appeared to be of interest in the sabinene and 1,8 cineole synthases analysis of S. fruticosa made by Kampranis et al. [11], since it corresponds to their Region 1. Kampranis et al. [11] suggested that these residues, located on helix  $\alpha$  14 and aligned along the pore of the catalytic centre, could be involved in the deprotonation of the water molecule, facilitating the attack on the intermediate  $\alpha$ -terpinyl carbocation. The multiple changes they introduced in Region 1 resulted in a progressive shift from a sabinene or 1.8 cineole synthase activity to a wider range of detected products. On the basis of our results and those reported by Iijima et al. [10], we infer that any modification of size (Y) or charge (T) of the amino-acids in this region is likely to substantially alter the GES enzymatic function on the intermediate geranyl carbocation, whereas a simple change from a GES→LIS signature in this area may not be sufficient to affect the original GES enzymatic specificity.

Following replacement of the aa after position 360 on the GES, by their (R)-LIS homologues, Iijima et al. [10] found a dual geraniol/linalool synthase function. As monoterpenol biosynthesis proceeds via an intermediate carbocation [9], one could therefore infer that they affected the intermediate geranyl carbocation, thereby shifting the charges towards an intermediate potential geranyl or linalyl carbocation. One could therefore consider that mutations in the second half of the GES protein may either inactivate the enzyme or shift the specificity of GES towards mixed geraniol and linalool production. A comparison of O. basilicum GES, (R)-LIS, P. frutescens GES and (R)-LIS (Figure 1), combined with O basilicum GES modeling (Figure 2), suggested 6 O basilicum GES modeling (Figure 2), suggested 6 areas of potential interest.

DM3 and DM4 are in the middle of the catalytic site, and in the area where the 3 Mg<sup>2+</sup> are thought to be located [17]. DM3 was intended to modify this area, making it more similar to *O. basilicum* (*R*)-LIS, and DM4 was used to induce a  $F \rightarrow Y$  change since both *O. basilicum* and *P. frutescens* GES have an F at this position, whereas their (*R*)-LIS have a Y. The DM3 and DM4 modifications did not affect GES activity or specificity, suggesting that this area has only a minor *in vivo* role in directing the enzymatic function.

DM5 and DM7 are located in the region, which is predicted to interact with the first and middle parts of GPP, respectively [17]. Although both the V and F residues, introduced respectively into DM5 and DM7, are larger than the original A and L amino-acids, they remain hydrophobic. DM7 is targeted to the motif in order to achieve metal dependent ionization of the prenvl diphosphate substrates (N,D)DXX(S,T)XXXE. The modification of these two GPP-interacting residues tested in our experiments gave similar results, leading to a loss of GES function. These results indicate that conformational changes in the processing of the GPP substrate are needed for GES function. It is likely that any mutation of these residues affects the conformation, and either induces the release of GPP out of the catalytic region before it is transformed into a monoterpenol, or even prevents its access to the catalytic region. These observations are corroborated by the fact that any disturbance of the substrate stability in the catalytic pocket (i.e. GPP from yeast FPPS) [13,23] results in a partial or total loss of enzymatic function. One could thus infer that the loss of GES specificity by GPP destabilization would not result in a shift towards a geraniol/linalool monoterpenol synthase function, but rather in a loss of activity.

DM6 is located at the very bottom of the pocket, and affects an amino-acid pointing to a region complementtary to the area tested in DM1 and DM2. This drastic  $M \rightarrow P$  change observed in *P. frutescens* (*R*)-LIS, and recreated in DM6, had the potential to affect the nearby helix and thus to produce constraints on the pocket. However, it resulted in a GES activity very similar to that observed with the wild type GES protein (**Table 2**). This area is therefore likely to play no, or only a minor, role in the GES activity and specificity.

DM8 is on a loop at the very bottom of the pocket site, and is therefore located outside the helix parts of the catalytic site. It is likely that this residue forms part of the entrance of the GES pore. Amino acids located on a loop are not considered to be critical in monoterpene synthase activity, but they have been shown to be important in diterpene synthases [24]. DM8 was tested because the L $\rightarrow$ M modification could destabilize the neighboring metallic cations from the catalytic pocket. This might affect the enzymatic reaction via substrate wobbling at the main catalytic site, and release different reaction products. However, quite expectedly, this minor modification did not affect GES efficiency or specificity, indicating that its presence in *O. basilicum* (*R*)-LIS may be fortuitous.

DM9 corresponds to a single amino-acidic difference observed between *O. basilicum* GES and LIS, at the end of the protein sequence, but within the catalytic site. DM9 had no effect on GES protein activity. This was very surprising, since changing the negatively charged D into a positively charged H at this position, facing SARM and in contact with  $Mg^{2+}$ , was expected to substantially affect the enzyme activity. This outcome, together with the results observed with DM3, DM4 and DM8, suggests that these residues, which are involved in the coordination of the divalent cations, act strongly and that the link is very difficult to disturb with a single directed mutagenesis.

DM10 is also at the end of the protein sequence and corresponds to a single amino-acid difference, observed between P. frutescens GES and LIS. Albeit being next to the amino-acidic position described in DM9, in P. fru*tescens* this  $H \rightarrow N$  mutation acts on the carbonate part of the substrate. In DM10, this area was modified to make this amino acid and the previous one look similar to the amino acid observed in P. frutescens. Strain K197G DM10 exhibited a monoterpenol profile, intermediate between the profile observed with the untransformed erg20K197G yeast strain, and that corresponding to the same strain expressing the wild-type GES (Table 2). This area therefore participates in the formation of the catalytic site. More interestingly, the enzyme with DM10 was altered in its efficiency in behaving as GES, but maintained its entire specificity, with no shift of production towards other terpenes. This type of behavior is clearly different to that observed by Kampranis et al. [11] for S. fruticosa sabinene and 1,8 cineole synthases.

#### 4. CONCLUSION

The data from this study provides initial insight into the location and composition of the GES catalytic site. This catalytic site has the shape of a pocket surrounded by the cyan amino-acids from Figure 2(A) and (B) that were mutated in our study. It is possible to draw an imaginary line leaving F 355 and D507 on the left side and L277, A381 and L442 on the right side from the catalytic site. All introduced mutations affecting the enzymatic efficiency were located on the right half of the catalytic site, as shown in Figure 2, suggesting that the most critical residues for the enzymatic function are located in this area. Conversely, none of the mutations located on its left half resulted in a reduced enzymatic efficiency. The residues from the right half of the catalytic site are involved in holding the carbon chain of GPP, whereas the residues from the left half of the catalytic site are involved in interactions with the 3  $Mg^{2+}$ , which holds the pyrophosphate O<sup>-</sup> charges. Destabilization of the Mg<sup>2+</sup>-GPP link was impossible to achieve, contrary to our observations of the yeast farnesyl diphosphate synthase site, where single amino-acidic mutations in K197 or K254 residues could induce a release of intermediate GPP out of the catalytic site [13,23]. Furthermore, none of the GES $\rightarrow$ LIS modifications we tested was able to induce the formation of new products, contrary to our observations of sabinene or 1,8 cineole synthase [11]. We expected that at least one of the tested DM would be able to induce not only a change in the amount of detected geraniol, but also a modification in the monoterpenol's specificity, with the detection of linalool or other monoterpenols. This was clearly not the case, suggesting that the enzymatic specificity of GES is more difficult to modify than its enzymatic efficiency. The transformation from the intermediate carbocation into products other than geraniol might therefore be due to a coordinated interaction produced by several critical amino-acids, such as the amino-acid observed in the present study, or could be induced by specific action from an area which has not vet be identified by 3D predictions and functionally directed mutagenesis experiments. Further biochemical and structure-function studies are needed, to determine the specificities of the GES catalytic site with greater accuracy.

#### 5. ACKNOWLEDGEMENTS

The authors thank professor Eran Pichersky for providing the *O. basilicum* GES cDNA.

#### REFERENCES

- Withers, S.T. and Keasling, J.D. (2007) Biosynthesis and engineering of isoprenoid small molecules. *Applied Microbiology and Biotechnology*, **73**, 980-990. doi:10.1007/s00253-006-0593-1
- [2] Hayashi, K., Kawaide, H., Notomi, M., Sakigi, Y., Matsuo, A. and Nosaki, H. (2006) Identification and functional analysis of bifunctional ent-kaurene synthase from the moss *Physcomitrella patens*. *FEBS Letters*, **580**, 6175-6181. doi:10.1016/j.febslet.2006.10.018
- [3] Dittmar, K. and Liberles, D. (2010) Evolution after gene duplication. Wiley-Blackwell, New York.
- [4] Trapp, S.C. and R.B. (2001) Genomic organization of plant terpene synthases and molecular evolutionary implications. *Genetics*, **158**, 811-832.
- [5] Martin, D.M., Aubourg, S., Schouwey, M.B., Daviet, L., Schalk, M., Toub, O., Lund, S.T. and Bohlmann, J. (2010) Functional annotation, genome organization and phylogeny of the grapevine (*Vitis vinifera*) terpene synthase gene family based on genome assembly, FLcDNA cloning, and enzyme assays. *BMC Plant Biology*, **10**, 226. doi:10.1186/1471-2229-10-226
- [6] Dickschat, J.S. (2011) Isoprenoids in three-dimensional space: The stereochemistry of terpene biosynthesis. *Natural Product Reports*, 28, 1917-1936. doi:10.1039/c1np00063b
- [7] Mukherjee, M. and Datta, A.K. (2007) The basils—A review. *Plant Arches*, 7, 473-483.
- [8] Morales, M.R. and Simon, J.E. (1997) "Sweet Dani": A new culinary and ornamental lemon basil. *HortScience*, 32, 148-149.

- [9] Iijima, Y., Gang, D.R., Fridman, E., Lewinsohn, E. and Pichersky, E. (2004) Characterization of geraniol synthase from the peltate glands of sweet basil. *Plant Physiology*, **134**, 370-379. <u>doi:10.1104/pp.103.032946</u>
- [10] Iijima, Y., Davidovich-Rikanati, R., Fridman, E., Gang, D., Bar, E., Lewinsohn, E. and Pichersky, E. (2004) The biochemical and molecular basis for the divergent patterns in the biosynthesis of terpenes and phenylpropenes in the peltate glands of three cultivars of basil. *Plant Physiology*, **136**, 3724-3736. doi:10.1104/pp.104.051318
- [11] Kampranis, S.C., Ioannidis, D., Purvis, A., Mahrez, W., Ninga, N., Katerelos, N.A., Ansour, S., Dunwell, J.M., Degenhardt, J., Makris, A.M., Goodenough, P.W. and Johnson, C.B. (2007) Rational conversion of substrate and product specificity in a *Salvia* monoterpene synthase: Structural insights into the evolution of terpene synthase function. *Plant Cell*, **19**, 1994-2005. doi:10.1105/tpc.106.047779
- [12] Masumoto, N., Korin, M. and Ito, M. (2010) Geraniol and linalool synthases from wild species of perilla. *Phytochemistry*, **71**, 1068-1075. doi:10.1016/j.phytochem.2010.04.006
- [13] Fischer, M.J.C., Meyer, S., Claudel, P., Bergdoll, M. and Karst, F. (2011) Metabolic engineering of monoterpene synthesis in yeast. *Biotechnology and Bioengineering*, **108**, 1883-1892. doi:10.1002/bit.23129
- [14] Sagot, I., Bonneu, M., Balguerie, A. and Aigle, M. (1999) Imaging fluorescence resonance energy transfer between two green fluorescent proteins in living yeast. *FEBS Letters*, **447**, 53-57. <u>doi:10.1016/S0014-5793(99)00258-6</u>
- [15] Coelho, E., Perestrelo, R., Neng, N.R., Câmara, J.S., Coimbra, M.A., Nogueira, J.M.F. and Rocha, S.M. (2008) Optimisation of stir bar sorptive extraction and liquid desorption combined with large volume injection-gas chromatography-quadrupole mass spectrometry for the determination of volatile compounds in wines. *Analytica Chimica Acta*, **624**, 79-89. doi:10.1016/j.aca.2008.06.032
- [16] Sali, A. and Blundell, T.L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *Journal of Molecular Biology*, 234, 779-815.

doi:10.1006/jmbi.1993.1626

- [17] Hyatt, D.C., Youn, B.Y., Zhao, Y.X., Santhamma, B., Coates, R.M., Croteau, R.B. and Kang, C.H. (2007) Structure of limonene synthase, a simple model for terpenoid cyclase catalysis. *Proceedings of National Academy of Sciences USA*, **104**, 5360-5365. doi:10.1073/pnas.0700915104
- [18] Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. and Higgins, D.G. (1997) The CLUSTAL\_X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research*, 25, 4876-4882. doi:10.1093/nar/25.24.4876
- [19] DeLano, W.L. (2002) The PyMOL molecular graphics system on world wide web. <u>http://www.pymol.org</u>
- [20] Chen, F., Tholl, D., Bohlmann, J. and Pichersky, E. (2011) he family of terpene synthases in plants: A mid-size family of genes for specialized metabolism that is highly diversified throughout the kingdom. *Plant Journal*, 66, 212-229. doi:10.1111/j.1365-313X.2011.04520.x
- [21] Köksal, M., Jin, Y., Coates, R.M., Croteau, R. and Christianson, D.W. (2011) Taxadiene synthase structure and evolution of modular architecture in terpene biosynthesis. *Nature*, **469**, 116-120. doi:10.1038/nature09628
- [22] Hillwig, M.L., Xu, M., Toyomasu, T., Tiernan, M.S., Wei, G., Cui, G., Huang, L. and Peters, R.J. (2011) Domain loss has independently occurred multiple times in plant terpene synthase evolution. *Plant Journal*, 68, 1051-1060. <u>doi:10.1111/j.1365-313X.2011.04756.x</u>
- [23] Fischer, M.J.C., Meyer, S., Claudel, P., Bergdoll, M. and Karst, F. (2011) Identification of a lysine residue important for the catalytic activity of yeast farnesyl diphosphate synthase. *Protein Journal*, **30**, 334-339. doi:10.1007/s10930-011-9336-y
- [24] Zhou, K., Gao, Y., Hoy, J.A., Mann, F.M., Honzatko, R.B. and Peters, R.B. (2012) Insights into diterpene cyclization from structure of bifunctional abietadiene synthase from *Abies grandis*. *Journal of Biological Chemistry*, 287, 6840-6850. doi:10.1074/jbc.M111.337592