Molecular Cloning, Functional Expression and Characterization of *d*-Limonene Synthase from *Schizonepeta tenuifolia*

Takuro MARUYAMA, Michiho ITO, Fumiyuki KIUCHI, and Gisho HONDA*

Department of Pharmacognosy, Graduate School of Pharmaceutical Sciences, Kyoto University, 46–29, Yoshida-Shimoadachi, Sakyo, Kyoto 606–8501, Japan. Received November 27, 2000; accepted January 10, 2001

Limonene is one of the most simple cyclic monoterpenes, and two enantiomers, *d*- and *l*-limonene occur due to the chiral carbon at 4-position. Cyclization of GPP into limonene is catalyzed by the limonene synthase, and some *l*-limonene synthase cDNAs have already been cloned from several species of plants, mainly from Labiatae family. However, the *d*-limonene synthase gene has not yet been obtained, therefore, no information is available on the molecular mechanism of stereochemical regulation in limonene formation. To resolve this, we cloned the *d*-limonene synthase gene (dLMS) from *Schizonepeta tenuifolia* (Labiatae) by a reverse genetic approach, and we found that both *d*- and *l*-limonene synthase share similar features such as a transit peptide, an arginine rich domain, and a metal cation binding site in their structures. Here, we report on the cloning of dLMS, and the putative stereochemical regulation mechanism is discussed based on the comparison of the deduced amino acid sequence of dLMS with those of known *l*-limonene synthases.

Key words *d*-limonene synthase; monoterpene biosynthesis; *Schizonepeta tenuifolia* (Labiatae); cDNA cloning

Terpenoids are members of the large family of natural products including alkaloids, hydrocarbons, flavonoids and fatty acids and so on. The terpenoid group includes many useful bioactive compounds, like steroids, saponins and essential oils. Recent development of molecular biological techniques allowed the isolation of biosynthetic enzymes for these compounds by a reverse genetic approach.[1)] As many sequences of biosynthetic enzymes are analyzed, and X-ray crystallography analysis of purified enzymes are performed, additional molecular mechanisms of biosynthetic reactions are revealed. As for terpenoid synthases, sequence comparison of mono-, sesqui- and diterpene synthases showed that they share a similar structure.[1)] Limonene, one of the most simple cyclic monoterpenes among plant secondary metabolites would be a good example for the analysis of the cyclization mechanism. This compound is contained in significant amount in mint leaves and orange fruits, and is widely distributed in many plant species as an intermediate of monoterpene biosynthesis. Limonene has a chiral center in its structure, and therefore, *d*- and *l*-enantiomers occur. The enantiomer ratio is different among species; mint and orange contain almost pure *d*- or *l*-limonene, but several Umbelliferae plants show different enantiomer ratios according to the growth stage or in different organ.[2—7)] Such differences are interesting with respect to the role of limonene for plant physiology. As a defense against harmful insects and as a pollinator inductor are two known roles for this compound in plants.[8)]

Biosynthesis of limonene is thought to proceed from geranyl pyrophosphate (GPP), which is a common precursor for all monoterpenes, to limonene through linalyl pyrophosphate (LPP)[9)] (Fig. 1). Limonene synthases, which promote the cyclization of GPP into limonene, are already cloned from *Perilla frutescens*, *P. citriodora*, *Mentha spicata*, *Abies grandis*,[10—13)] but all of these cloned enzymes produce the *l*-form, however, the *d*-limonene synthase gene has not yet been obtained. Therefore, the molecular mechanism of regulating the stereochemistry in limonene biosynthesis remains unclear. To clarify this, we carried out the cloning of the *d*-limonene synthase gene (dLMS) from *Schizonepeta tenuifolia* (Labiatae), which contains ubiquitous *d*-limonene along with its derivatives (*l*-pulegone, *d*-menthone). The putative stereochemical regulation mechanisms in this enzyme are also discussed.

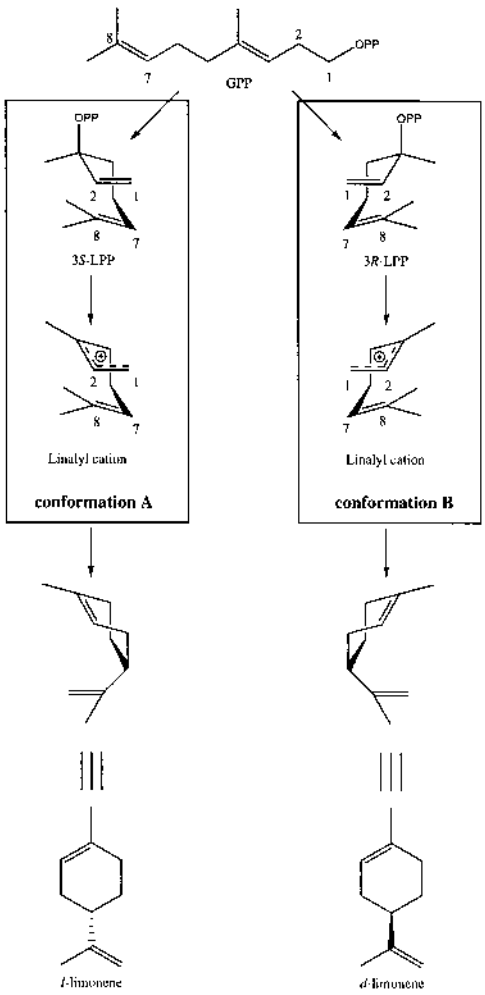

Fig. 1. Biosynthetic Route from GPP to Limonene

* To whom correspondence should be addressed. e-mail: ghonda@pharm.kyoto-u.ac.jp

© 2001 Pharmaceutical Society of Japan

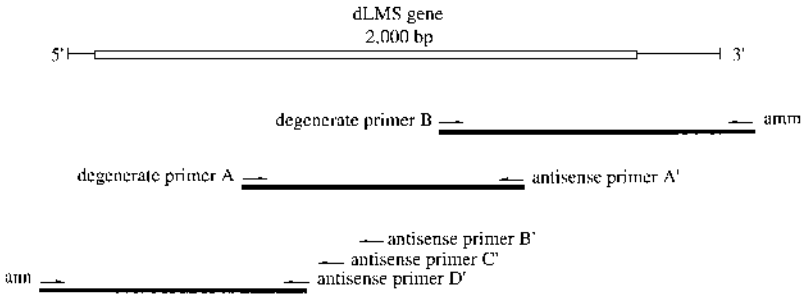

Fig. 2. Strategy for dLMS Gene Cloning Based on PCR Amplification

The open box on the top indicates the translational region, and solid lines represent the 5′- and 3′-flanking regions. Each arrow shows relative position and direction of the primers. cDNA fragments amplified with each primer are indicated by solid boxes.

MATERIALS AND METHODS

***S. tenuifolia* cDNA Preparation** Total RNA was extracted from the early stage of the *S. tenuifolia* flower using RNeasy Plant Mini Kit (Qiagen), and reverse transcribed with oligo (dT) primer designed to have an adaptor sequence at the 5′-end to obtain the cDNA for the 3′-RACE method. The cDNA for the 5′-RACE method was prepared by reverse-transcription with antisense primer B′ (Fig. 2) described below, followed by removing excess primer with the Suprec-2 column (Takara Shuzo), and then the oligo (dC) chain was added to the 5′-end of resulting cDNA by terminal deoxynucleotidyl transferase (Toyobo).

cDNA Cloning Based on PCR Amplification Based on the highly conserved amino acid sequence among several plant terpene synthases, the two degenerated sense primers were designed {A: GGA (C/T) T (A/G) (C/T) TG (I) A (I) (C/T) T (I) TA (C/T) GAAGC (A/T) TC, B: GA (C/T) GA (C/T) AT (A/C/T) TA (C/T) GA (C/T) GT (A/C/G/T) TA (C/T) GG (A/C/G/T)}. First, the polymerase chain reaction (PCR) for 3′-RACE was carried out with degenerate primer A and an adaptor primer ("amm" in Fig. 2). The PCR was performed by recombinant Taq DNA polymerase (Toyobo) and dNTP (0.1 mM) on the PCR Thermal Cycler Personal (Takara Shuzo) with the following program (94 °C 45 s, 50 °C 45 s, 72 °C 60 s, 35 cycles). The subsequent nested PCR with degenerate primer B and primer amm (temperature program 94 °C 45 s, 45 °C 45 s, 72 °C 60 s, 35 cycles) gave a 900 bp PCR product, which was electrophoresed in 0.7% agarose gel and purified using the GENECLEAN II Kit (BIO 101). The resulting DNA fragment was cloned into plasmid vector pCR 2.1 using the TOPO TA cloning kit (Invitrogen) and the plasmid DNA was sequenced on a GeneRapid system (Amersham Pharmacia Biotech). All PCR products described below were cloned and sequenced following the same procedure as above.

Specific antisense primer A′ designed on the basis of the sequence obtained from 3′-RACE, was applied to the PCR together with degenerate primer A (temperature program 94 °C 45 s, 50 °C 45 s, 72 °C 60 s, 35 cycles). The sequence of amplified DNA was used as a reference to design three additional antisense primers for the 5′-RACE procedure (B′: ACTGGATTCATGTCTGGTC, C′: GTTTTCGCCTTCCTTGAAC, D′: TCACGCTCCAGTAAGTTGGG). The first PCR for 5′-RACE was performed with antisense primer C′ and the adaptor primer with oligo (dG) ("ann" in Fig. 2), followed by the second PCR with antisense primer D′ and the primer ann to afford a DNA fragment with a complete 5′-end sequence.

Expression of dLMS The DNA fragment corresponding to the translational region except for the transit peptide within the dLMS gene was amplified by PCR with primers, dLMS-S and dLMS-A (dLMS-S: TCTAGAATGCGACGATCTGGAAAC, dLMS-A: GAGCTCGGCCACGCGTCGACT) using Ex Taq DNA polymerase (Takara Shuzo). To compensate for the lost start codon, the additional nucleotide ATG was attached to the 5′-end of dLMS-S. The pCR 2.1 vector containing the resultant DNA fragment downstream of the T7 promotor was constructed using the TOPO TA cloning kit (Invitrogen), and applied to the TNT T7 Quick Coupled Transcription/Translation System (Promega) to obtain the dLMS recombinant protein according to the manufacturer's protocol. However, biotinylated recombinant protein was prepared by the same system combined with Transcend Non-Radioactive Translation Detection Systems (Promega) in order to analyze molecular weight of the protein on SDS-PAGE. Prestained Protein Marker (Bio Labs) was used as a protein marker. As for the positive control in protein expression and subsequent enzyme assay, the same procedure was performed on the *l*-limonene synthase gene of *Perilla frutescens*, whose function was already confirmed by the expression system with *E. coli*.[10]

Enzyme assay Fifty microliters dLMS recombinant protein solution (accurate protein amount was not determined) was incubated with 50 μM GPP in 500 μl of 25 mM potassium phosphate buffer (pH 7.6, containing 1 mM 1, 4-dithiothreitol and 15 mM $MgCl_2$) for 12 h at 31 °C. The incubation mixture was overlaid with 500 μl diethyl ether to trap volatile products. After incubation, the reaction mixture was extracted with 500 μl of diethyl ether twice, and the combined extract was dried on $MgSO_4$ for GC-MS analysis.

Identification of the Products of Enzymatic Reaction GC-MS analysis was performed on the G9000/3DQMS (HITACHI) equipped with a DB-5 column (30 m×5.0 mm internal diameter, J & W), or chiral column, Cyclodex-B (30 m×5.0 mm i.d., J & W) using helium as carrier gas (1.0 ml/min). The analysis using the DB-5 column was done with an oven program started at 40 °C (1 min hold), increasing 10 °C/min to 160 °C, and then 15 °C/min to 240 °C (5 min hold). For the chiral phase separation with Cyclodex-B, the same conditions as reported by Bohlmann[14] were employed.

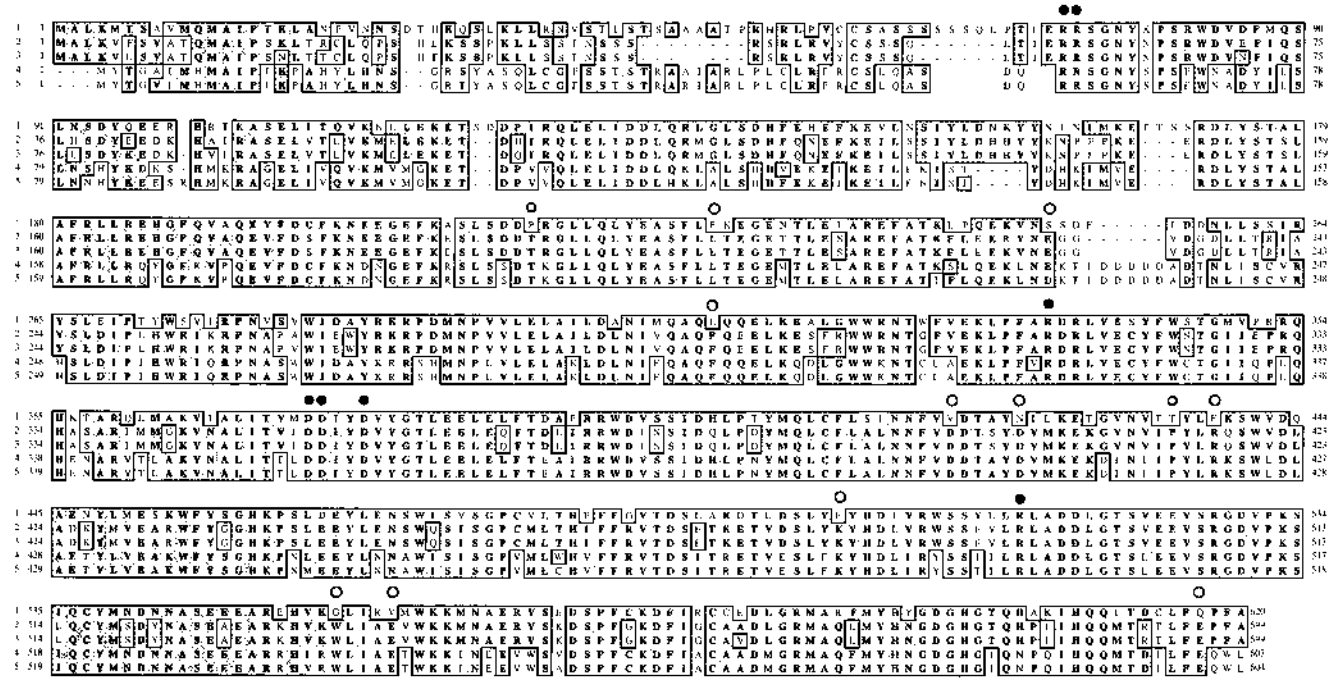

Fig. 3. Deduced Amino Acid Sequence of Several Limonene Synthases

1) *d*-Limonene synthase from *Schizonepeta tenuifolia* (dLMS), 2—5) *l*-limonene synthase from *Mentha longifolia*, *M. spicata*, *Perilla frutescens*, *P. citriodora*, respectively. Solid circles indicate highly conserved amino acid residues among many terpene synthases. Open circles indicate aromatic and acidic amino acids mutated in dLMS compared with other *l*-limonene synthases. Gray circles show mutations with respect to important threonine and proline residues.

RESULTS

The PCR with degenerate primers designed with reference to the conserved amino acid sequence among several plant terpene synthases, and following sequence analysis, using the cDNA prepared from the spike of *S. tenuifolia* as a template, revealed the core sequence containing the 3′-flanking region of limonene synthase. By clarifying the remaining 5′-region using the 5′-RACE technique, the full length of limonene synthase cDNA (dLMS) was cloned (Figs. 2, 3). The complete nucleotide sequence was registered at GenBank accession number AF282875. dLMS is 2108 nucleotides in length and contained a 1860 bp translational region encoding 620 amino acids. The deduced amino acid sequence of dLMS showed the presence of a plastidial transit-peptide-like domain in its *N* terminal (1-72 a.a. in Fig. 3). Transit peptide is a targeting sequence for transporting the protein synthesized in endoplasmic reticulum to the plastid, the organelle responsible for monoterpene biosynthesis, and it is considered to be eliminated after migration to produce the active mature protein.[15] In general, the transit peptide of monoterpene synthase is 50—70 a.a. in length, rich in serine and threonine and poor in acidic residues,[16] and the dLMS reported here also shared these features. dLMS possesses another important domain, the DDxxD motif, which appeared to be the binding site for substrate-divalent metal cation (Mn^{2+}, Mg^{2+} etc.) complex and its sequence of xx was IY, the same observed in all known limonene synthases.

Subsequently, recombinant protein was prepared for the functional analysis of dLMS with an *in vitro* transcription/translation system. Recombinant protein was detected on SDS-PAGE followed by Western blotting on a nitrocellulose membrane and by chemical staining. As a result, the molecular weight of both the recombinant protein from dLMS and from the *l*-limonene synthase gene of *Perilla frutescens* (positive control) were 62 kDa whereas their calculated weight was 64 kDa (Fig. 4).

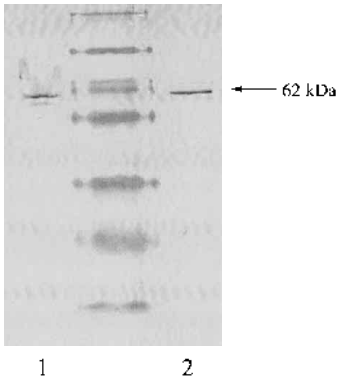

Fig. 4. SDS-PAGE Analysis

1) Recombinant limonene synthase from *Perilla frutescens*. 2) Recombinant limonene synthase from *Schizonepeta tenuifolia* (dLMS).

These recombinant proteins were applied to the following enzyme assays and the enzymatic products were analyzed on GC-MS. According to the analysis with the DB-5 column, the dLMS recombinant protein yielded limonene (peak 1 in Fig. 5) as a principal product together with a small amount of by-product (peak 2 in Fig. 5). The mass spectrum of peak 2 (Fig. 6C) suggested that this by-product was a cyclic monoterpene with the same MW of 136 as limonene, but this was not determined. Stereochemistry of the resulting limonene was determined using the Cyclodex-B column; the retention time of limonene from the dLMS recombinant protein was identical to that of authentic *d*-limonene (Figs. 7A,

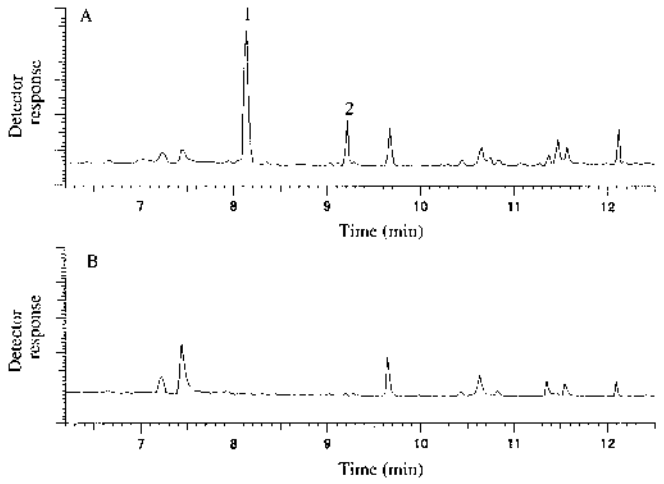

Fig. 5. GC-MS Chromatogram

A) Olefine products generated from GPP by recombinant dLMS protein. Peak 1: limonene, 2: unknown. B) Blank: in enzyme assay, distilled water was used instead of recombinat dLMS protein.

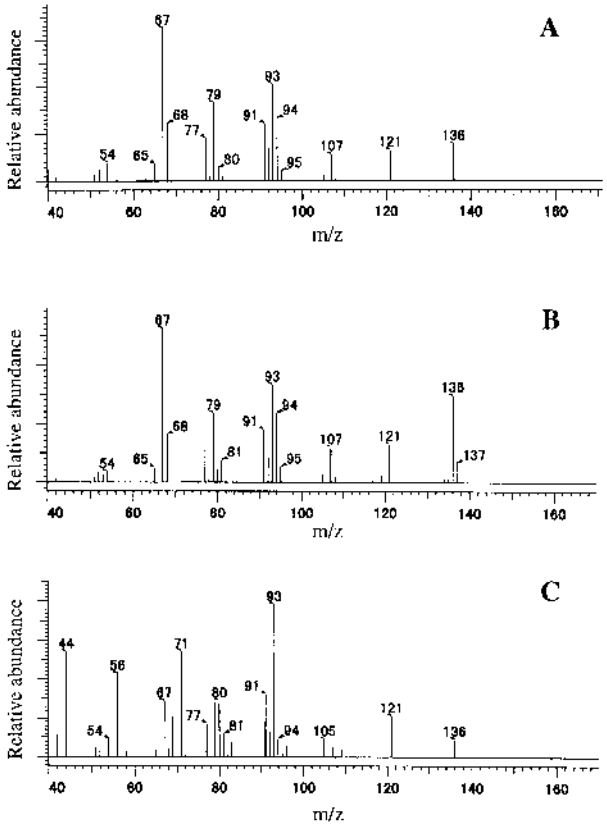

Fig. 6. MS Spectra on GC-MS Analysis

A) Peak 1 in Fig. 5, B) authentic *d*-limonene, C) peak 2 in Fig. 5.

B). However, limonene from *Perilla frutescens* *l*-limonene synthase (positive control) was shown to be *l*-limonene (Fig. 7C).

Based on these findings, we confirmed that dLMS encoded *d*-limonene synthase. This study was the first to report the cloning of *d*-limonene synthase gene from a plant.

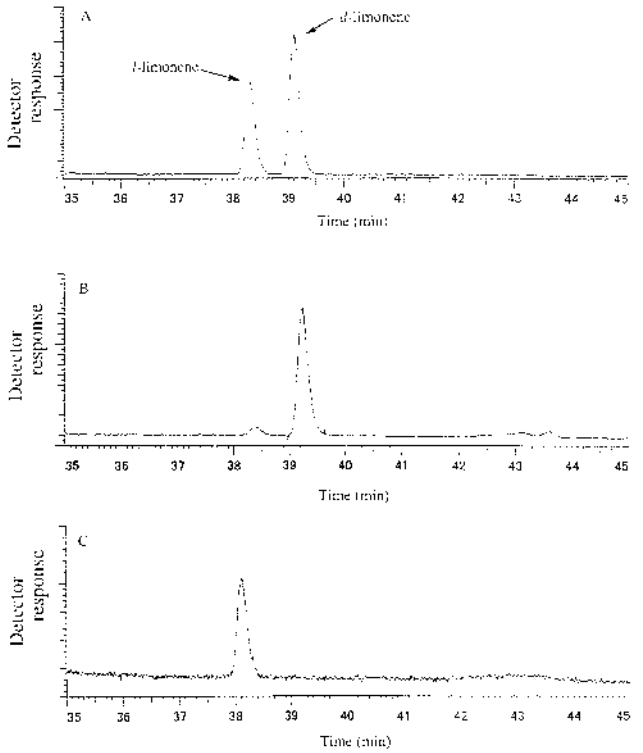

Fig. 7. Chiral GC-MS Chromatogram

A) Authentic *d*, *l*-limonene, B) Olefine products generated from GPP by recombinant dLMS protein, C) Olefine products generated from GPP by recombinant limonene synthase from *Perilla frutescens*.

DISCUSSION

The deduced amino acid sequence of dLMS indicates a close identity with the *l*-limonene synthase genes from Labiatae plants throughout its whole region (68.5, 67.7, 61.4, 61.3% with *Mentha longifolia*, *M. spicata*, *Perilla frutescens*, *P. citriodora*, respectively) except for the transit peptide. Furthermore, two important domains exist at the same position as all known *l*-limonene synthases; the first one is the RR motif (position 73, 74 in Fig. 3) which is thought to play an important role for isomerization of GPP into LPP.[17] The second is the DDxxD motif (position 373 to 377 in Fig. 3) and 2 arginines (position 336 and 514 in Fig. 3) suggested to contribute to stabilize the negative charge of the diphosphate group released from LPP during the generation of the important intermediate, linalyl cation.[18] These findings suggest that both *l*- and *d*-limonene synthases share a similar structure in the active pocket site regardless of the difference in stereochemistry. Therefore, the regulation of the stereochemical structure in limonene biosynthesis probably relies on a very minor mutation of the limonene synthase gene.

Using native *l*-limonene synthase from *Mentha spicata* and *d*-limonene synthase from *Citrus unshiu*, Hiraga *et al.* showed that the cyclization from linalyl cation to limonene was proceeded by an electrophilic attack from the 2*Si*-face through the conformation shown in Fig. 1A or B.[19] With respect to these findings, the diphosphate group and 7,8-olefine of LPP or the linalyl cation occupied similar parts of the active pocket site in each enzyme, respectively, which suggested that these two functional groups appeared not to relate to the stereochemical regulation. So, it was suggested that the

stereochemistry of limonene biosynthesis depended on the regulation of the position of another functional group, the allylic cation in the core enzyme pocket; amino acid residues, possible to control the movement of the allylic cation, are likely to be the stereochemical regulation factor. We paid attention to aromatic and acidic residues because these residues were suggested to be able to stabilize allylic cation *via* the π-orbital to cation interaction and formation of ion pairs, respectively. In fact, stabilization of the oleanyl cation *via* the π-orbital of tryptophan in β-amyrin synthase was suggested to be a main factor responsible for the conversion of the lupenyl cation into the oleanyl cation.[20] Among the aromatic and acidic residues in dLMS, F227, 392, Y596, E438, 501 (Fig. 3) are different from the residues conserved in the corresponding position of *l*-limonene synthases from Labiatae plants. However, several aromatic and acidic residues, conserved in all *l*-limonene synthases from Labiatae plants, changed to non-aromatic and acidic residues (L312, G555, S251, V419, 559, N424) in dLMS. It is noteworthy that E501 and S251 can be found at positions of equal distance from the DDIYD motif in opposite direction.

Based on the X-ray crystallography analysis of *epi*-aristolochene synthase from *Nicotiana tabacum*, Joseph *et al*. suggested the contribution of the peptide carbonyl and hydroxyl groups of the threonine residues as the factor stabilizing the allylic cation.[18] Comparing mutations of the threonine residue with *l*-limonene synthases, we found that highly conserved threonine and proline (T194, P414 in *l*-limonene synthase from *M. longifolia*; Fig. 3) changed to P214 and T435 in dLMS. Further studies of these mutations in *d*- and *l*-limonene synthase may help clarify a method for regulating the stereochemical structures in terpenoid biosynthesis. Presently, a domain swapping experiment by constructing the chimera gene with *d*- and *l*-limonene synthases is on going to clarify the relationship between these amino acid residues and stereochemical regulation.

Acknowledgements We are grateful to Tochimoto Tenkaido Co., Ltd. for providing seed of *Schizonepeta tenuifolia*.

REFERENCES

1) Bohlmann J., Meyer-Guene G., Croeau R., *Proc. Natl. Acad. Sci. U.S.A.*, **95**, 4126 (1998).
2) Shaw P. E., *J. Agric. Food Chem*., **27**, 246 (1979).
3) Umemoto K., Nagasawa T., *Nogeikagakukaishi*, **53**, 269 (1979).
4) Fujita S., Fujita Y., *Yakugaku Zasshi*, **93**, 1622 (1973).
5) Fujita S., Fujita Y., *Yakugaku Zasshi*, **93**, 1679 (1973).
6) Faber B., Bangert K., Mosandl A., *Flavour and Fragrance Journal*, **12**, 305 (1997).
7) Bouwmeester H. J., Davies J. A. R., Toxopeus H., *J. Agric. Food Chem*., **43**, 3057 (1995).
8) Harborne J. B., *in Ecological Chemistry and Biochemistry of Plant Terpenoids*, pp. 399—426 (1991).
9) Croteau R., *Chem. Rev*., **87**, 929 (1987).
10) Yuba A., Yazaki K., Tabata M., Honda G., Croteau R., *Arch. Biochem. Biophys*., **332**, 280 (1996).
11) Ito M., Kiuchi F., Yang L. L., Honda G., *Biol. Pharm. Bull*., **23**, 359 (2000).
12) Colby S. M., Alonso W. R., Katahira E. J., McGarvey D. J., Croteau R., *J. Biol. Chem*., **268**, 23016 (1993).
13) Bohlmann J., Steels C. L., Croteau R., *J. Biol. Chem*., **272**, 21784 (1997).
14) Bohlmann J., Martin D., Oldham N. J., Gershenzon J., *Arch. Biochem. Biophys*., **375**, 261 (2000).
15) Keegstra K., Olsen L. J., *Annu. Rev. Plant Physiol. Mol. Biol*., **40**, 471 (1989).
16) von Heijne G., Steppuhn J., Herrmann R. G., *Eur. J. Biochem*., **180**, 535 (1989).
17) Williams D. C., McGarvey D. J., Katahira E. J., Croteau R., *Biochemistry*, **37**, 12213 (1998).
18) Starks C. M., Back K., Chappell J., Nuel J. P., *Science*, **277**, 1815 (1997).
19) Hiraga Y., Shi W., Ito D. I., Ohta S., Suga T., *J. Chem. Soc. Chem. Commun*., **1993**, 1370 (1993).
20) Kushiro T., Shibuya M., Masuda K., Ebizuka Y., *J. Am. Chem. Soc*., **122**, 6816 (2000).