Note

Plant Biotechnology, **16** (4), 311-314 (1999)

Molecular Cloning and Sequencing of an *Allium macrostemon* cDNA Probably Encoding Oxidosqualene Cyclase

Song YOU[1,2], Noriaki KAWANO[1], Koji ICHINOSE[1], Xin-Sheng YAO[2] and Yutaka EBIZUKA*[1]

[1] *Graduate School of Pharmaceutical Sciences, The University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan.* E-mail: yebiz@mol.f.u-tokyo.ac.jp
[2] *Department of Natural Product, Shenyang Pharmaceutical University, Wen-hua Road, Shenyang 110015, P. R. C.*

Received 7 April 1999; accepted 4 June 1999

Abstract

Plant oxidosqualene cyclase (OSC) is classified into two types depeneding on its catalytic mechanisms and products: cycloartenol synthase and triterpene synthase. A cDNA probably encoding OSC was cloned from bulbs of *Allium macrostemon* by RT-PCR. The cDNA contains a 2289-bp open reading frame, encoding 762 amino acids. A higher degree of homology (73 %) was found for the gene product with the known cycloartenol synthases than with the triterpene synthases (51-55 %), indicating its most likely function as cycloartenol synthase. This is the first cloning example of a putative cDNA encoding OSC from a monocotyledonous plant.

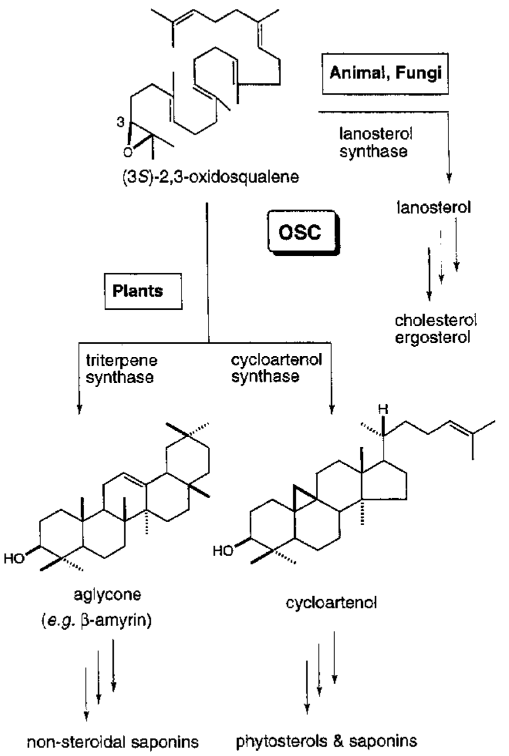

Fig. 1 Reactions catalyzed by oxidosqualene cyclase (OSC) converting 2,3-oxidosqualene to various cyclized products

Phytosterols, such as sitosterol, campesterol, and stigmasterol, are known to function as the essential components of cell membranes. They are biosynthesized from a common isoprenoid precursor, cycloartenol, the first cyclic intermediate in the pathways. A step affording cycloartenol is one of the most complex reactions found in nature: a cyclization of acylic 2, 3-oxidosqualene to a tetracylic cation intermediate followed by a series of rearrangements of hydrogen atoms and methyl groups. 2, 3-oxidosqualene cyclase (OSC, EC 5.4.99.-) involved in this remarkable reaction is designated as cycloartenol synthase (**Fig. 1**). Unlike animal, fungi, and yeast, plant OSCs provide a number of options in their catalytic control, giving rise to a vast variety of non-steroidal triterpenes leading to saponins. Distinct cDNAs encoding OSCs for cycloartenol (Corey *et al.*, 1993; Morita *et al.*, 1997) and triterpenes (Herrera *et al.*, 1998; Kushiro *et al.*, 1998) have been characterized in plants, indicating that a marked metabolic branching point exists at the stage of 2, 3-oxidosqualene (**Fig. 1**). All the OSC-cDNAs so far characterized are from the triterpene-producing dicotyledonous (dicot) plants, and virtually nothing has been studied on OSCs from monocotyledonous (monocot) ones. Instead of triterpene saponins, some monocot plants produce steroidal glycosides, mostly occurring as furostanol glycosides (bisdesmosidic, sugars at C3 and C26) as well as spirostanol glycosides (monodesmosidic, sugars at C3). It was proved (Inoue *et al.*, 1996a; 1996b) that a specific β-glucosidase converts a furostanol glycoside to its corresponding spirostanol glycoside. Spirostanol glycosides are a major class of steroid saponins with various biological activities including antifungal and hemolytic ones (Shvets *et al.*, 1996). Concerning the biosynthesis of aglycones leading to steroid saponins and their related metabolites, very little is known except for some indications: 1) cholesterol and sitosterol are its precursors (Stohs *et al.*, 1969; Stohs *et al.*, 1974); 2) precursors are derived from cycloartenol (Bergenstråhle, *et al.*, 1996). A preliminary attempt (You *et al.*, 1998) was made to clone a cDNA encoding cylcoartenol synthase from *Allium macrostemon* Bge. (Liliaceae) which is a monocot plant known as an original plant for traditional Chinese medicine "Xiè Bái". The bulb is a rich source of steroid saponins (Peng *et al.*, 1996), whereas no cyclic triterpenes have been reported from this plant. We report here a full length of cDNA probably encoding OSC from this plant as the first example from a monocot plant.

Cloning strategy is based on reverse transcription polymerase chain reaction (RT-PCR) using degenerate primers designed from the highly conserved regions of the known OSCs from *Arabidopsis thaliana* (GenBank accession no. U02555), *Rattus rattus* (GenBank accession no. D45252), and *Saccharomyces cerevisiae* (GenBank accession no. U23488). Following the amplification of an initial fragment (CORE) using the degenerate primers (162S, 623A, 467S, and 556A), a "rapid amplification of cDNA ends" (RACE) (Frohman *et al.*, 1988) method was applied to obtain sequence information on 5'- and 3'- ends of a target cDNA. Overall strategy is summarized in **Fig. 2**, and detailed experimental procedures including preparation of RNA from fresh bulbs are as decribed (You *et al.*, 1998) except for an additional purification step of messenger RNA (mRNA) using an oligo-dT column (Quick Prep^{TM} mRNA Purification Kit, Amersham Pharmacia Biotech) for an effcient RT reaction.

PCR amplification of a full length of cDNA was performed based on a nested method (Mullis *et al.*, 1987) using a primer combination of AL-OSC-5'-END and AL-OSC-3'-END followed by K-ALL-N and B-ALL-C. The reactions were carried out in a final volume of 100 μl for 30 cycles with *Ex Taq*^{TM} (Takara Shuzo, Japan) using a step program (1min at 94 ℃, 2 min at 58 ℃ and 3 min at 72 ℃). The 2.3-kb fragment obtained was gel-purified and subcloned into pT7Blue (R) T-vector (Novagen) for DNA sequencing using Thermo Sequenase^{TM} cyclesequencing kit (Amersham Pharmacia Biotech).

Sequencing (an automated DNA sequencer model 4000L, Licor) of the several independent clones identified a 2289-bp open reading frame, *All-OSC1*, encoding 762 amino acids (The nucleotide sequence reported here has been deposited in the DDBJ/EMBL/GenBank data bases under the accession number AB025353). Sequence analysis (the DNASIS programs, Hitachi Software Engineering, Japan) showed that the deduced protein shares 73 %

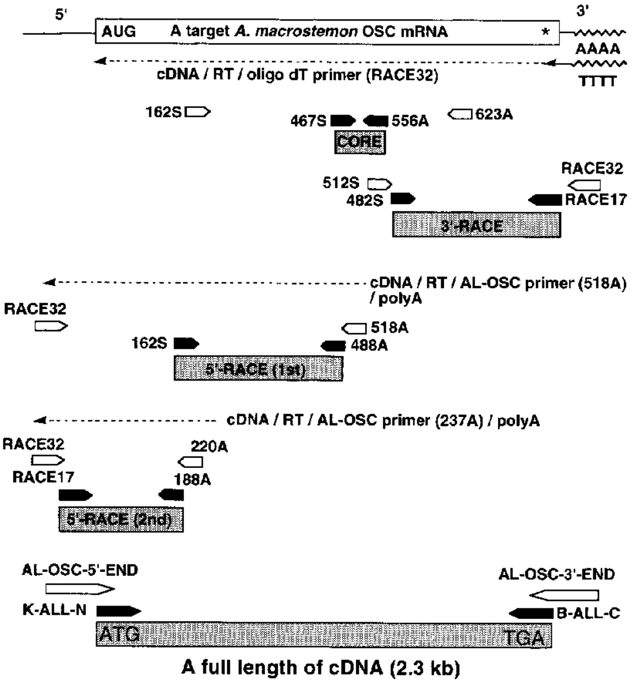

Fig. 2 PCR-based strategy for cloning of cDNA encoding OSC from A. *macrostemon*. All the PCR amplifications were performed by a nested method (Mullis *et al.*, 1987): the first and second primer sets are indicated by open and filled arrows, respectively. Primer sequences used in this study are as follows (one-letter abbreviations are based on the rule of International Union of Biochemistry, IUB): RACE17 (5'-GACTCGAGTCGACATCG-3'); RACE32 (5'-GACTCGAGTCGACATCGATTTTTTTTTTTTTT-3'); 162S (5'-GAYGGIGGITGGGGIYTICA-3', designed from DGGWGLH); 623A (5'-CCCAISWICC ITMCCAISWICCRTC-3', designed from DGSWYGCW); 467S (5'-AARGGIGCITGGCCITTYWSIAC-3', designed from KGAWPFST); 556A (5'-GTRCAYTCIACRTAIGGRTAITC-3', designed from D(E)YPYVECT); 482S (5'-TCTGATTGTACAGCTGAAGGA-3'); 488A (5'-TCCTTCAGCTGTACAATCAGA-3'); 512S (5'-CTTTATGATGCTGTAAATGTG-3'); 518A (5'-CACATTTACAGCATCATAAAG-3'); 188A (5'-AAGCAATCTCAGAGTAACATA-3'); 220A (5'-ATTTTCCCCATGATGTTATTG-3'); 237A (5'-TAGGGGATTGTTGCCAGACCA-3'); AL-OSC-5'-END (5'-CAATATCCAAGCTACTTACAGGGC-3'); AL-OSC-3'-END (5'-TACAGCAAT AACCATGTACGGTTC-3'); K-ALL-N (5'-GGATGGTACC**ATG** TGGAAGCTGCTGAA G-3'); B-ALL-C (5'-GTCAGGATCC**TCA**ATGACCTGCAGAGGA-3'); the engineering sites, *Kpn*I and *Bam*HI, are underlined; the start and stop (complement) codons are indicated in boldface.

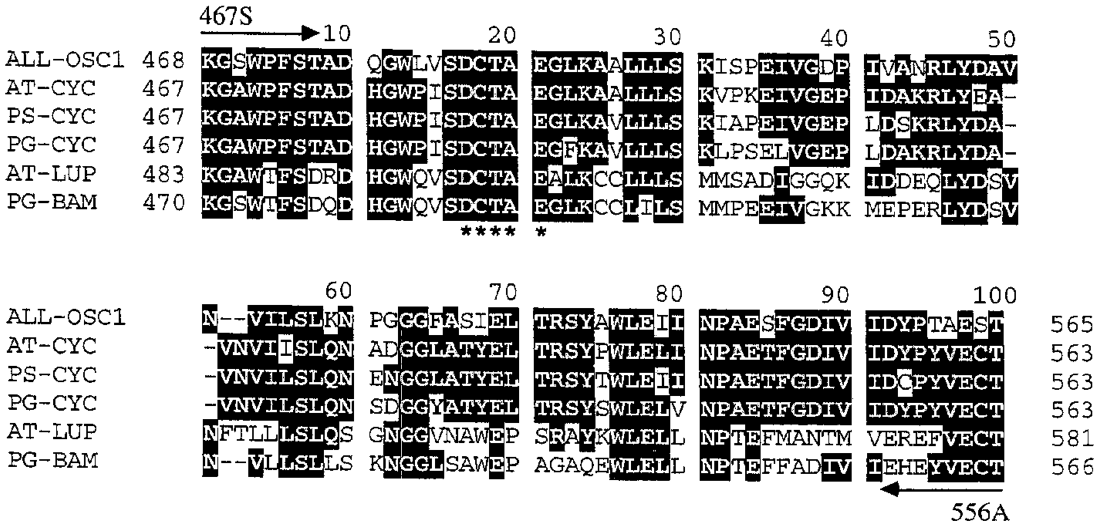

Fig. 3 Alignment of amino acid sequences of the CORE region of ALL-OSC1 and the known plant OSCs. Highly conserved residues are outlined in black. The DCTAE motif (see text) is asterisked. Residue numbers at line starts and ends refer to those of each sequence. The numbered arrows indicate the regions corresponding to the PCR primers. The entries are: *Arabidopsis thaliana* cycloartenol synthase [AT-CYC / 763 aa (Corey *et al.*, 1993)]; *Pisum sativum* cycloartenol synthase [PS-CYC / 756 aa (Morita *et al.*, 1997)]; *Panax ginseng* cycloartenol synthase [PG-CYC / 758 aa (Kushiro *et al.*, 1998)]; A. *thaliana* lupeol synthase [AT-LUP / 783 aa (Herrera *et al.*, 1998)]; *P. ginseng* β-amyrin synthase [PG-BAM / 762 aa (Kushiro *et al.*, 1998)].

identity with the known cycloartenol synthases from *Arabidopsis thaliana*, *Pisum sativum*, and *Panax ginseng*. A lower degree of identity was found with the triterpene synthases: A. *thaliana* lupeol synthase (51 %); *P. ginseng* β-amyrin synthase (55 %). The DCTAE motif (Abe *et al.*, 1994), which is highly conserved among the known OSCs and claimed to be in an active site of lanosterol synthase, was also found in the CORE sequence. Alignment (**Fig. 3**) of the CORE regions revealed perfect identity of this motif for ALL-OSC1. Based on the foregoing results, it is suggested that ALL-OSC1 functions an OSC, most likely as a cycloartenol synthase. This is the first cDNA cloning of a putative OSC from monocot plants, providing useful information for further studies on regulation of steroid saponin biosynthesis and molecular evolution of plant OSCs as well.

Acknowledgements

We thank Dr. M. Shibuya and Ms. K. Hayashi for their technical advice and assistance. S. Y. was a recipient of Short-term Student Exchange Promotion Program Scholarship of Association of International Education, Japan for his study at the University of Tokyo.

References

Abe, I., Prestwich, G. D. 1994. Active site mapping of affinity-labeled rat oxido-squalene cyclase. J. Biol. Chem., **269**: 802-804.

Bergenstråhle, A., Borgå, P., Jonsson, L. M. V. 1996. Sterol composition and synthesis in potato tuber discs in relation to glycoalkaloids synthesis. Phytochemistry, **41**: 155-161.

Corey, E. J., Matsuda, S. P. T., Bartel , B. 1993. Isolation of an *Arabidopsis thaliana* gene encoding cycloartenol synthase by functional expression in a yeast mutant lacking lanosterol synthase by the use of a chromatographic screen. Proc. Natl. Acad. Sci. USA, **90**: 11628-11632.

Frohman, M. A., Dush, M. K., Martin, M. K. 1988. Rapid production of a full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. Proc. Natl. Acad. Sci. USA, **85**: 8998-9002.

Herrera, J. B. R., Bartel, Wilson, W. K., Matsuda, S. P. T. 1998. Cloning and characterization of the *Arabidopsis thaliana* lupeol synthase gene. Phytochemistry. **49**: 1905-1911.

Inoue, K., Ebizuka, Y. 1996a. Purification and characterization of furostanol glycoside 26-*O*-β-glucosidase from *Costus speciosus* rhizomes. FEBS Letters, **378**: 157-160.

Inoue, K., Ebizuka, Y. 1996b. Purification and characterization of a β-glucosidase which converts furostanol glycosides to spirostanol glycoside from *Costus speciosus*. Adv. Exp. Med. Biol., **404**: 57-69.

Kushiro, T., Shibuya, M., Ebizuka, Y. 1998. β-Amyrin synthase. Cloning of oxidosqualene cyclase that catalyzes the formation of the most popular triterpene among higher plants. Eur. J. Biochem., **256**: 238-244.

Morita, M., Shibuya, M., Lee, M.-S., Sankawa, U. and Ebizuka, Y. 1997. Molecular cloning of pea cDNA encoding cycloartenol synthase and its functional expression in yeast. Biol. Pharm. Bull., **20**: 770-775.

Mullis, K. B., Faloona, F. A. 1987. Specific synthesis of DNA *in vitro* via a polymerase-catalysed chain reaction. Methods Enzymol., **155**: 335-350.

Peng, J.-P., Yao, X.-S. 1996. 19 new steroidal saponins from Allium plants: isolation, structural elucidation and effect on blood coagulability. Adv. Exp. Med. Biol., **404**: 511-526.

Shvets, S. A., Kintia P. K. and Naibi, M. A. 1996. Steroidal glycosides from *Petunia hybrida* L. seeds and their biological activity. Adv. Exp. Med. Biol. **404**: 251-260.

Stohs, S. J., Kaul, B., Staba, E. J. 1969. The metabolism of ^{14}C-cholesterol by *Dioscorea deltoidea* suspension cultures. Phytochemistry, **8**: 1679-1686.

Stohs, S. J. Sabatka, J. J., Rosenberg, H. 1974. Incorporation of 4-^{14}C-22, 23-^{3}H-sitosterol into diosgenin by *Dioscorea deltoidea* tissue suspension cultures. Phytochemistry, **13**: 2145-2148.

You S., Ichinose K., Inoue K., Shibuya M., Yao X.-S. and Ebizuka Y. 1998. Biosynthetic studies on steroid saponins. In : Ageta, H. *et al.* (Eds.) : Towards Natural Medicine Research in the 21st Century. Excerpta Medica International Congress Series **1157**: 429-441, Elsevier, Amsterdam.