Molecular Cloning and Characterization of Two cDNAs for *Glycyrrhiza glabra* Squalene Synthase¹⁾

Hiroaki HAYASHI,^{*,2)} Akiko HIROTA, Noboru HIRAOKA, and Yasumasa IKESHIRO

Niigata College of Pharmacy, 5–13–2 Kamishin-ei-cho, Niigata 950–2081, Japan.

Received March 1, 1999; accepted June 5, 1999

Two cDNAs (GgSQS1 and GgSQS2) encoding squalene synthase have been isolated from licorice, *Glycyrrhiza glabra* L., and characterized. The deduced amino acid sequence of GgSQS1 was 88%, 81%, 78%, 45–44%, and 45–41% identical to those of GgSQS2, *Nicotiana*, *Arabidopsis*, mammal and yeast squalene synthases, respectively. Squalene synthase activity was found in the cell-free extracts of *Escherichia coli* transformed with the recombinant plasmids for GgSQS1 and GgSQS2, respectively. Genomic Southern blot hybridization indicated that there are three squalene synthase genes in the licorice genome. Northern blot analysis showed that GgSQS2 mRNA is mainly expressed during the exponential growth phase of the cultured licorice cells.

Key words *Glycyrrhiza glabra* L.; Leguminosae; squalene synthase; biosynthesis; triterpenoid; cDNA

The roots and stolons of licorice, *Glycyrrhiza glabra* L., contain large amounts (3–10% of dry weight) of an oleanane-type triterpenoid saponin, glycyrrhizin, which is a well-known natural sweetener. In addition to glycyrrhizin, betulinic acid, a lupane-type triterpene, and soyasaponins, oleanane-type triterpenoid saponins, are localized in the different organs of *G. glabra*.^{3,4)} The cultured cells of *G. glabra* produce no detectable amount of glycyrrhizin but produce betulinic acid³⁾ and soyasaponins.⁵⁾ Our interest is focused on the regulation of sterol and triterpenoid biosynthesis in *G. glabra*.

Squalene synthase (SQS, EC 2.5.1.21) catalyzes the condensation of two molecules of farnesyl diphosphate into squalene, a common precursor of sterols and triterpenoids. It catalyzes the first pathway-specific step of the sterol and triterpenoid biosyntheses. cDNAs of squalene synthase have been cloned from yeast,^{6–8)} mammals,^{8–11)} and higher plants,^{12–16)} and squalene synthase has been shown to play an important role in the regulation of cholesterol biosynthesis in humans.^{10,17)} Since higher plants accumulate not only sterols but also various triterpenoids, the regulation of phytoosterol and triterpenoid biosynthesis should be different from that of cholesterol biosynthesis in mammals. Licorice is one of the model plants used to elucidate the regulation of a variety of triterpenoid biosyntheses in higher plants. Thus, we screened squalene synthase cDNAs from a cDNA library of cultured licorice cells, and two cDNAs (GgSQS1 and GgSQS2) encoding squalene synthase were isolated¹⁾ and characterized in the present study.

MATERIALS AND METHODS

Cell Culture Cell suspension cultures (strains RNS-1B and RDS-2A), derived from seedlings of *G. glabra*, have been maintained in Linsmaier Skoog (LS) medium¹⁸⁾ as reported previously.^{3,5)} For the triterpenoid production experiments, cultured cells (1 g fr. wt) were inoculated to 30 ml LS medium containing 10^{-6} M 1-naphthaleneacetic acid and 10^{-5} M 6-benzyladenine in 100 ml Erlenmeyer flasks, and incubated on a reciprocal shaker (100 strokes min⁻¹) at 25 °C in the dark.

Construction of a Licorice cDNA Library Total RNA was extracted from 14-d-old cultured cells by the Extract-A-

Plant^{TM} RNA Isolation Kit (Clontech, U.S.A.). Poly(A)-RNA was purified by Oligotex-dT30 (Super) (Takara, Japan), and used to construct a licorice cDNA library using the ZAP-cDNA Synthesis Kit (Stratagene, U. S. A.).

Screening of cDNA Library by Hybridization A 134-bp digoxigenin (DIG)-labelled DNA fragment was amplified by the polymerase chain reaction (PCR) using *Arabidopsis* SQS cDNA as a template, DIG-dNTP mixture (Boehringer Mannheim, Germany), *Taq*-DNA polymerase (Takara, Japan) and two primers, 5'-TACTGCCACTATGTTGC-3' and 5'-ATGTTTGTCTTCTGCAGAAA-3'. The DIG-labelled fragment was used as a probe to screen the cDNA library (ca. 200000 plaques) which was transferred to a colony-plaque screen (MEN Life Science Products, U.S.A.). The hybridization was carried out at 65 °C overnight in a solution recommended in the protocol of the DIG Nucleic Acid Detection Kit (Boehringer Mannheim, Germany). The membranes were washed twice for 5 min at room temperature in 2×SSC/0.1% sodium dodecyl sulfate (SDS), and then a further twice for 10 min at 60 °C in 0.5×SSC/0.1% SDS. The hybridized DIG-probe was detected using the DIG Nucleic Acid Detection Kit (Boehringer Mannheim, Germany) according to the manufacturer's protocol. Four positive clones were plaque purified, converted into plasmid form, and the inserts were sequenced by the dideoxy chain termination method using the ALFred DNA sequencer. The two longest different clones were designated as GgSQS1 and GgSQS2.¹⁹⁾ Sequence data were analyzed using Genetyx software (Software Development, Japan).

Construction of Expression Plasmid and Prokaryotic Expression A DNA fragment was amplified by PCR using GgSQS1 or GgSQS2 cDNA as a template, *Pfu*-DNA polymerase (Stratagene, U.S.A.) and two primers, 5'-CGGGATCCTATGGGGAGTTTGGGAGC-3' and 5'-TTTTTTTTT-TTTTTTTTTTTT-3'. The amplified fragment was digested by *Bam*HI and *Hind*III (GgSQS1), or *Bam*HI and *Nco*I (GgSQS2). The digested PCR fragment was substituted for the corresponding fragment of the GgSQS1 or GgSQS2 cDNA clone, so that the entire coding region was in frame with the N-terminal region of β -galactosidase in pBluescript SK (–) vector. The respective recombinant plasmid (designated pGgSQS1-exp and pGgSQS2-exp) was introduced into *E. coli* strain XL1-Blue (Stratagene, U.S.A.). The trans-

* To whom correspondence should be addressed.

formed *E. coli* ($OD_{600}=0.5$) was treated with isopropylthio- β -D-galactoside (IPTG), to a final concentration of 0.5 mM, for 3 h and the crude enzyme extract was prepared as reported previously.¹¹⁾

Assay of Squalene Synthase Activity The enzyme assay (total volume, 100 μ l) was performed in the presence of [³H]farnesyl diphosphate (60000 dpm, 15 Ci/mmol, ARC), 3 mM NADPH, 5 mM MgCl₂, 0.1 mM dithiothreitol, 10 mM KF, 50 mM Tris-HCl (pH 7.6) and the 30 μ g protein. The reaction mixture was incubated at 30 °C for 1 h. After the incubation, 10 μ g squalene (in 20 μ l of ethanol) was added as carrier, and half (60 μ l) the mixture was separated by silica gel TLC (solvent, ethylacetate-benzene (1:4)). The spot of squalene (visualized by I₂) was scraped off and extracted with ethanol, and the radioactivity of the extract was determined in a scintillation counter.

Genomic Southern Blot Analysis A 215-bp DIG-labelled PCR fragment, corresponding to amino acid residue number 1—65, was amplified by PCR using GgSQS1 or GgSQS2 cDNA as a template, DIG-dNTP mixture (Boehringer Mannheim, Germany), *Taq*-DNA polymerase and two primers, 5'-CGGGATCCTATGGGAGTTTGGGAGC-3' and 5'-AAWATGCAAACRGC GTTGCG-3'. Genomic DNA was extracted from 14-d-old cultured licorice cells (RDS-2A strain⁵⁾) by the cetyltrimethylammonium bromide (CTAB) method as reported previously.²⁰⁾ Twenty micrograms genomic DNA (per lane) was digested with *Eco*RI, *Hind*III or *Pst*I then separated on 0.8% agarose gel by electrophoresis. The DNA was transferred to a positively charged nylon membrane (Boehringer Mannheim, Germany), and hybridized with the DIG-labelled DNA probe for GgSQS2 at 65 °C. The membrane was washed twice for 5 min at room temperature in 2 $\times$ SSC/0.1% SDS, and then a further twice for 10 min at 65 °C in 0.2 $\times$ SSC/0.1% SDS. The hybridized DIG probe was detected using the DIG Nucleic Acid Detection Kit according to the manufacturer's manual.

Northern Blot Analysis Twenty micrograms total RNA (per lane) was separated on 1% agarose gel containing formaldehyde, blotted on to a positively charged nylon membrane (Boehringer Mannheim, Germany) and hybridized with the DIG labelled DNA probe for GgSQS2 at 50 °C. The membrane was washed twice for 5 min at room temperature in 2 $\times$ SSC/0.1% SDS, and then a further twice for 15 min at 55 °C in 0.2 $\times$ SSC/0.1% SDS. The hybridized DIG probe was detected using the DIG Nucleic Acid Detection Kit according to the manufacturer's manual.

Quantitative Analysis of Phytosterols and Triterpenoids A powdered sample (100 mg) of freeze-dried cultured cells was extracted with ethyl acetate (2 ml twice, reflux for 1 h). Cholesterol (1 mg) was added to the ethyl acetate extract as an internal standard, and the solvent was dried *in vacuo*. This dried sample was used for the GC analysis of betulinic acid and phytosterols. The residue of the ethyl acetate extraction was re-extracted with methanol (2 ml twice, reflux for 1 h), and the extract was dried *in vacuo*, and methanolized with 1 ml 9% HCl-methanol (reflux for 2 h). After addition of diethyl ether (4 ml) and cholesterol (1 mg), the solution was neutralized with Ag₂CO₃. The supernatant was dried, and used for the GC analysis of soyasaponins. All dried samples were dissolved in a mixture of pyridine (10 μ l) and *N,O*-bis(trimethylsilyl)acetamide (10 μ l), and incubated

at 40 °C for 1 h. An aliquot (2 μ l) of the solution was analyzed by GC as reported previously.⁵⁾ The triterpenoid and phytosterol content was calculated from the ratio of the peak area of the respective compound to that of the internal standard. The total soyasaponin content (estimated as soyasaponin II) was calculated from the quantity of soyasapogenol B. The phytosterol content was the total of the stigmasterol, β -sitosterol and campesterol content.

RESULTS AND DISCUSSION

cDNA Cloning of *Glycyrrhiza glabra* Squalene Synthase

A λ ZAP cDNA library was prepared from mRNA of 14-d-old cultured licorice cells (RNS-1B strain), which produced phytosterols, soyasaponins and betulinic acid.⁵⁾ The cDNA library was screened by a DIG-labelled PCR fragment of *Arabidopsis* squalene synthase cDNA (corresponding to amino acid residue number 168—212¹²⁾) under conditions of low stringency. Four positive clones were isolated from 200000 plaques of the cDNA library, and the cDNA inserts were subcloned into pBluescript SK (–) vector. Restriction enzyme digestion of these clones revealed that they are divided into two different clones. The longest clones, designated GgSQS1 and GgSQS2, were sequenced in both strands.¹⁹⁾ Their nucleotide sequences contained an open reading frame for a polypeptide of 413 amino acids (GgSQS1) and 412 amino acids (GgSQS2), respectively. The deduced amino acid sequence of GgSQS1 exhibits 88%, 78%, 81%, 45—44%, and 45—41% identity to those of the GgSQS2, *Arabidopsis thaliana*,¹²⁾ *Nicotiana* plants,^{13,16)} mammal^{8—11)} and yeast^{6—8)} squalene synthases, respectively. Figure 1 shows the alignment of the deduced amino acid sequences of GgSQS1 and GgSQS2. Although both proteins are highly conserved, the conservation is especially poor at the C-terminal region, which contains a highly hydrophobic sequence regarded as a transmembrane domain.^{8,14)}

Functional Expression of Two SQS cDNAs in *E. coli*

To confirm the enzyme activity, expression of the cDNAs was performed in *E. coli*. The entire coding region of both clones was subcloned into an expression vector (designated pGgSQS1-exp and pGgSQS2-exp, respectively), and the enzyme assay was performed as reported previously¹¹⁾ except for some minor modifications. As shown in Table 1, the cell-free extract of *E. coli* transformed with pGgSQS1-exp and pGgSQS2-exp converted [³H]farnesyl diphosphate into squalene. These results indicate that the cultured cells of *G. glabra* produce two functional isozymes for squalene synthase, which may play an important role in the regulation of the different triterpenoid metabolism pathways.

Genomic Southern Blot Analysis Although the squalene synthase genes of yeast^{6,8)} and humans⁸⁾ have been reported to be single-copy genes, two squalene synthase genes per genome were reported in *Arabidopsis thaliana*¹⁴⁾ and *Nicotiana tabacum*.¹⁶⁾ In addition, it was reported that the two genes are organized in a tandem array in *Arabidopsis*.¹⁴⁾ To estimate the copy number of the squalene synthase gene in licorice, genomic Southern blotting was performed using a DIG-labelled PCR fragment of GgSQS2, corresponding to the exon 1 of *Arabidopsis* and *Nicotiana* SQS genes,^{14,16)} as a probe. As shown in Fig. 2, one strong and two weak signals were observed in each lane, indicating that three squalene

GgSQS1	1	MGSLGATVRHPDEVYPLLKLMARHAEKQIPAEPHWAF	80
GgSQS2	1	MGSLGATILKHPDDLPLLLKLTIAARQAEKQIPREPHWAF	80
GgSQS1	81	DDTSIATDVKEPTLIAFHRHIYDRDWHFSCGTKEYKVLMDQFHHVSTAFLELEKNYQAATENITKEMGAGMAKFCICNEVE	160
GgSQS2	81	DDTSIATDVKEPTLIAFHRHIYDCDWHFSCGTKEYKVLMDQFHHVSTAFLELEKNYQAATENITKEMGAGMAKFCICKEVE	160
GgSQS1	161	TVDDYDEYCHYVAGLVGLGLSKLFHASGKEDLAPDHLNSMGLFLQKTNIIRDYLEDINEIPKSRMFWRQIWSKYVNKL	240
GgSQS2	161	TVDDYDEYCHYVAGLVGLGLSKLFHASGSEDLASDYLNSMGLFLQKTNIIRDYLEDINEIPKSRMFWRQIWSKYVNKL	240
GgSQS1	241	EDLKYEENSVKAVQCLNDMVTNALLHAEDCLKYMSALRDMSTFRFCAIPQIMAIGTLALCYNNIEVFRGVVKMRRGLTAK	320
GgSQS2	241	EDLKYEENSVKAVQCLNDMVTNALLHAEDCLKYMAALRDLATFRFCAIPQIMAIGTLALCYNNIEVFRGVVKMRRGLTAK	320
GgSQS1	321	VIDRTKTMAADVGAFFDFSCMLKSKVDKNDPNATKTLRLEAIQKTCRESGLLSKRKPYILRNESSTNSSTMVLILVILS	400
GgSQS2	321	VIDRTKTMGADVGAFFDFASMLESKVDKNDPNATKTLRLEAIQKTCRESGLLNKRKS-LLRNGNGYGSTMTLILVILFG	399
GgSQS1	401	TIIFAYLSAKRQDN	413
GgSQS2	400	TIIFAYLSANHHNN	412

Fig. 1. Alignment of the Deduced Amino Acid Sequences of Two Squalene Synthases from *Glycyrrhiza glabra*

Hyphens were inserted to maximize homology. The amino acid residues identical in two are boxed.

Table 1. Expression of *G. glabra* Squalene Synthase cDNAs in *E. coli*

Transformed vector	[³ H]Squalene formation (dpm)
pBluescript SK (-)	Not detected
pGgSQS1-exp	8650
pGgSQS2-exp	3070

The transformed *E. coli* (strain XL1-Blue) with pBluescript SK (-), pGgSQS1-exp or pGgSQS2-exp was cultured and each cell free extract was prepared 3 h after IPTG (0.5 mM) induction. The enzyme assay was performed in the presence of [³H]farnesyl diphosphate (60000 dpm) and 30 µg protein. After silica gel TLC separation, the radioactivity of the spot corresponding to squalene was determined.

Fig. 2. Genomic Southern Blot Analysis

Genomic DNA (20 µg) from licorice was digested by the restriction enzyme *EcoRI* (E), *HindIII* (H) or *PstI* (P), separated on 0.8% agarose gel by electrophoresis, and transferred to a positively charged nylon membrane. The filter was hybridized with the 215-bp DIG-labelled PCR fragment of GgSQS2. A λ -*HindIII* digest served as molecular markers.

synthase genes exist in the licorice genome. The DIG-labelled GgSQS1 DNA probe, corresponding to the same region, also showed the same three signals of different strength (data not shown).

Northern Blot Analysis Total RNA was isolated from the cultured licorice cells at different growth stages as shown in Fig.3, and Northern blot analysis was performed using the

Fig. 3. Time-Course of Cell Growth (●) and the Production of Soyasaponins (□), Betulinic Acid (○) and Total Phytosterols (△) in *Glycyrrhiza glabra* Cell Suspension Cultures (RNS-1B Strain)

Fig. 4. Northern Blot Analysis

Total RNA (20 µg) was separated on 1% agarose gel containing formaldehyde, blotted to a positively charged nylon membrane. The filter was hybridized with the 215-bp DIG-labelled PCR fragment of GgSQS2. The ethidium bromide staining of the gel before transfer is shown below.

same DIG-labelled GgSQS2 DNA probe. As shown in Fig. 4, a 1.6 kb SQS mRNA was detected at day 4 to 20, but a strong signal was observed on days 4 and 8. In preliminary Northern blot experiments using the GgSQS1 DNA probe, the expression of GgSQS1 mRNA was stronger on day 16 (data not

shown). This suggests that the two squalene synthase mRNAs are differentially regulated in the cultured licorice cells. However, to estimate the differential expression of squalene synthase genes in *G. glabra*, further experiments using the gene-specific probes are necessary. Experiments are underway to isolate the third squalene synthase gene, and to substantiate their differential expression in cultured cells and intact plants of *G. glabra*.

Acknowledgment The authors would like to thank Dr. S. Hata (Kyoto University) for providing the *Arabidopsis thaliana* squalene synthase cDNA. This research was partly supported by the Sasakawa Scientific Research Grant from the Japan Science Society, and by grant No. 07780500 from the Ministry of Education, Science, Sports and Culture, Japan.

REFERENCES AND NOTES

- 1) Preliminary communication on this subject: Hayashi H., Hiraoka N., Ikeshiro Y., *Biol. Pharm. Bull.*, **19**, 1387—1389 (1996).
- 2) Present address: Gifu Pharmaceutical University, 5-6-1 Mitahorahigashi, Gifu 502-8585, Japan.
- 3) Hayashi H., Fukui H., Tabata M., *Plant Cell Rep.*, **7**, 508—511 (1988).
- 4) Hayashi H., Fukui H., Tabata M., *Planta Med.*, **59**, 351—353 (1993).
- 5) Hayashi H., Sakai T., Fukui H., Tabata M., *Phytochemistry*, **29**, 3127—3129 (1990).
- 6) Jennings S. M., Tsay Y. H., Fisch T. M., Robinson G. W., *Proc. Natl. Acad. Sci. U.S.A.*, **88**, 6038—6042 (1991).
- 7) Fegueur M., Richard L., Charles A. D., Karst F., *Curr. Genet.*, **20**, 365—372 (1991).
- 8) Robinson G. W., Tsay Y. H., Kienzie B. K., Smith-Monroy C. A., Bishop R. W., *Mol. Cell. Biol.*, **13**, 2706—2717 (1993).
- 9) McKenzie T. L., Jiang G., Straubhaar J. R., Conrad D. G., Shechter I., *J. Biol. Chem.*, **267**, 21368—21374 (1992).
- 10) Jiang G., McKenzie T. L., Conrad D. G., Shechter I., *J. Biol. Chem.*, **268**, 12818—12824 (1993).
- 11) Inoue T., Osumi T., Hata S., *Biochim. Biophys. Acta*, **1260**, 49—54 (1995).
- 12) Nakashima T., Inoue T., Oka A., Nishino T., Osumi T., Hata S., *Proc. Natl. Acad. Sci. U.S.A.*, **92**, 2328—2332 (1995).
- 13) Hanley K. M., Nicolas O., Donaldson T. B., Smith-Monroy C., Robinson G. W., Hellmann G. M., *Plant Mol. Biol.*, **30**, 1139—1151 (1996).
- 14) Kribii R., Arro M., Del Arco A., Gonzalez V., Balcells L., Delourme D., Ferrer A., Karst F., Boronat A., *Eur. J. Biochem.*, **249**, 61—69 (1997).
- 15) Hata S., Sanmiya K., Kouchi H., Matsuoka M., Yamamoto N., Izui K., *Plant Cell Physiol.*, **38**, 1409—1413 (1997).
- 16) Devarenne T. P., Shin D. H., Back K., Yin S., Chappell J., *Arch. Biochem. Biophys.*, **349**, 205—215 (1998).
- 17) Guan G., Jian G., Koch R. L., Shechter I., *J. Biol. Chem.*, **270**, 21958—21965 (1995).
- 18) Linsmaier E. M., Skoog F., *Physiol. Plant.*, **18**, 100—127 (1965).
- 19) The sequence reported in this paper has been deposited in the DDBJ, EMBL, GenBank data base under the accession nos. D86409 (GgSQS1) and D86410 (GgSQS2).
- 20) Murray M. G., Thompson W. F., *Nucleic Acids Res.*, **8**, 4321—4325 (1980).