BIOLOGIA PLANTARUM 50 (4): 542-550, 2006

Characterization of 5-enolpyruvylshikimate 3-phosphate synthase gene from *Camptotheca acuminata*

Y. GONG*,***, Z. LIAO**, M. CHEN**, B. GUO*, H. JIN*, X. SUN** and K. TANG*,**[1]

*Shanghai Key Laboratory of Biotechnology, Plant Biotechnology Research Center, School of Agriculture and Biology, Shanghai Jiaotong University, Shanghai 200030, China**
*State Key Laboratory of Genetic Engineering, Fudan University, Shanghai 200433, China***
*Faculty of Life Science and Biotechnology, Ningbo University, Ningbo 315211, China****

Abstract

5-enolpyruvylshikimate 3-phosphate synthase (EPSPS; 3-phosphoshikimate 1-carboxyvinyl-transferase; EC 2.5.1.19) is a critical enzyme in the shikimate pathway. The full-length EPSPS cDNA sequence (*CaEPSPS*, GenBank accession number: AY639815) was cloned and characterized for the first time from woody plant, *Camptotheca acuminata*, using rapid amplification of cDNA ends (RACE) technique. The full-length cDNA of *CaEPSPS* was 1778 bp containing a 1557 bp ORF (open reading frame) encoding a polypeptide of 519 amino acids with a calculated molecular mass of 55.6 kDa and an isoelectric point of 8.22. Comparative and bioinformatic analyses revealed that CaEPSPS showed extensive homology with EPSPSs from other plant species. CaEPSPS contained two highly conserved motifs owned by plant and most bacteria EPSPSs in its N-terminal region. Phylogenetic analysis revealed that CaEPSPS belonged to dicotyledonous plant EPSPS group. Tissue expression pattern analysis indicated that *CaEPSPS* was constitutively expressed in leaves, stems and roots, with the lower expression being found in roots. The coding sequence of *CaEPSPS* gene was successfully subcloned in a plasmid-*Escherichia coli* system (pET-32a), and the cells containing the plasmid carrying the *CaEPSPS* gene exhibited enhanced tolerance to herbicide glyphosate, compared to the control.

Additional key words: RACE, shikimate pathway, tissue expression pattern.

Introduction

The enzyme 5-enolpyruvylshikimate 3-phosphate synthase (EPSPS; 3-phosphoshikimate 1-carboxyvinyl -transferase; EC2.5.1.19) is a critical enzyme in the shikimate biosynthesis pathway, which catalyzes the formation of 5-enolpyruvylshikimate 3-phosphate (EPSP) from shikimate-3-phosphate (S3P) and phosphoenolpyruvate (PEP) in the chloroplast (Herrmann and Weaver 1999). The products of this pathway are precursors to the synthesis of the aromatic amino acids as well as other essential aromatic amino acids. The shikimate pathway is an attractive target for herbicides, antibiotic and antimicrobial agent development because it is essential in bacteria, algae, fungi and higher plants, but absent from mammals (Bentley 1990). Interest in the characterization of plant EPSPS has been increased significantly since the enzyme has been identified as the predominant cellular target of the broad-spectrum, nonselective herbicide glyphosate (GLP; N-phosphonomethyl glycine). It has been shown that glyphosate can kill most weeds and crops by inhibiting EPSPS activity in a competitive manner with phosphoenolpyruvate (PEP). Glyphosate also blocks import of the cytoplasmically synthesized EPSPS pre-protein to chloroplast. *EPSPS* gene has been used to engineer glyphosate tolerance in transgenic plants either by the overproduction of the wild-type EPSPS or by the expression of a mutant gene (*aroA*) encoding glyphosate-resistant EPSPS (Ye *et al.* 2001, Howe *et al.* 2002, Wang *et al.* 2003). The genes encoding EPSPS have been isolated and sequenced from bacteria (Duncan *et al.* 1984, Garbe *et al.* 1990), fungi (Charles *et al.* 1986), dicotyledonous plants (Klee *et al.* 1987, Gasser *et al.* 1988) and monocotyledons plants

Received 22 January 2005, *accepted* 1 October 2005.
Abbreviations: CaEPSPS - *Camptotheca acuminata* 5-enolpyruvylshikimate 3-phosphate synthase; EPSPS - 5-enolpyruvylshikimate 3-phosphate synthase; ORF - open reading frame; PCR - polymerase chain reaction; RACE - rapid amplification of cDNA ends; RT-PCR - reverse transcriptase - polymerase chain reaction.
Acknowledgements: This work was funded by China National "863" High-Tech Program, China Ministry of Education and Shanghai Science and Technology Committee.
[1] Corresponding author; fax: (+86) 21 62824073, e-mail: kxtang1@yahoo.com

(Ream *et al*. 1988, Forlanni 1994, Xu *et al*. 2002). However, until now there are no reports on cloning of *EPSPS* gene from woody plants including Chinese happy tree, *Camptotheca acuminata*, which produces the important anti-cancer monoterpenoid indole alkaloid camptothecin (CPTP; Thomas *et al*. 2004).

In this paper, we report on the cloning and characterization of EPSP synthase gene from *C. acuminata* (*CaEPSPS*). The potential tolerance of *CaEPSPS* to herbicide glyphosate was also studied.

Materials and methods

Plants: Young leaves were collected from *Camptotheca acuminata* Decne. grown in the greenhouse at Shanghai Jiaotong University, Shanghai, China, and used as the starting material for RNA isolation. Total RNA was isolated by CTAB method and lithium chloride precipitation (Liao *et al*. 2004).

Cloning of CaEPSPS full-length cDNA by RACE: Single-strand cDNAs were synthesized from total RNA with an oligo$(dT)_{17}$ primer and reversely transcribed according to the manufacturer's protocol (PowerScript™, Clontech, Mountain View, USA). After RNase H treatment, the single-strand cDNA mixtures were used as templates for PCR amplification of the conserved region of *EPSPS* from *C. acuminata*. Two degenerate oligonucleotide primers, FEPSPS [5'-(A/C)A(T/C)C (A/G/T)A(G/C)(A/T/C)TA(T/C)(A/G)T(A/G)CTTGATG G-3'] and REPSPS [5'-GGCATTTTGTTCAT(A/G)TT (A/G/C)AC(A/G)TC-3'], were designed according to the conserved sequences of other EPSPS genes and used for the amplification of the core cDNA fragment of *CaEPSPS* by standard gradient PCR amplification (from 52 to 60 °C). The PCR products were purified and subcloned into pGEM T-easy vector (*Promega*, Madison, WI, USA) followed by sequencing. The core fragment was subsequently used to design the gene-specific primers for the cloning of full-length cDNA of *CaEPSPS* by RACE.

SMART™ RACE cDNA amplification kit (*Clontech*, Palo Alto, USA) was used to clone the 3'-end and 5'-end of *CaEPSPS* cDNA. The first-strand 3'-RACE-ready and 5'-RACE-ready cDNA samples from *C. acuminata* were prepared according to the manufacturer's protocol and used as templates for 3'-RACE and 5'-RACE respectively.

The 3'-end of *CaEPSPS* cDNA was amplified using two 3'-gene-specific primers and the universal primers provided by the kit. For the first PCR amplification of 3'-RACE, CaEPSPS3-1 (5'-GATGTTAAATTTGCCGA GGTTCTTG-3') and UPM (*Universal Primer A Mix*, 5'-CTAATACGACTCACTATAGGGCAAGCAGTGGT ATCAACGCAGAGT-3' and 5'-AAGCAGTGGTATC AACGCAGAGT-3') were used as the first PCR primers, and 3'-RACE-ready cDNA was used as the template. For the nested PCR amplification of 3'-RACE, CaEPSPS3-2 (5'-GAGCAAAAGTTACCTGGACAGAAAACAG-3') and NUP (*Nested Universal Primer A*, 5'-AAGCAGT GGTATCAACGCAGAGT-3') were used as the nested PCR primers, and the products of the first PCR amplification were used as templates. The 5'-end of *CaEPSPS* cDNA was amplified using two 5'-gene-specific primers and the universal primers (UPM and NUP) provided by the kit. For the first PCR amplification of 5'-RACE, CaEPSPS5-1 (5'-GATAGAG CCAGAGAGCTTCACCTTTCC-3') and UPM were used as the first PCR primers, and 5'-RACE-ready cDNA was used as the template. For the nested PCR amplification of 5'-RACE, CaEPSPS5-2 (5'-ACAATCAACATCTGCA CCAAGCTGC-3') and NUP were used as the nested PCR primers, and the products of the first PCR amplification were used as templates. The nested 3'-RACE and 5'-RACE products were purified and subcloned into pGEM T-easy vectors followed by sequencing. By aligning and assembling the sequences of 3'-RACE, 5'-RACE and the core fragment on *Contig Express* (*Vector NTI Suite 6.0*), the full-length cDNA sequence of *CaEPSPS* was deduced.

According to the deduced *CaEPSPS* cDNA sequence, two gene-specific primers, FCaEPSPS (5'-ATGGCGCAA GTTAGCAACATTCCTAATG-3') and RCaEPSPS (5'-ATGCTTTGCGAACCTCTGGAGAACTTC-3') were designed, synthesized and used to clone the coding sequence of *CaEPSPS* by RT-PCR using the 3'-RACE-ready cDNA as the template.

Comparative and bioinformatic analysis: Comparative and bioinformatic analyses of *CaEPSPS* were carried out online at the websites (http://www.ncbi.nlm.nih.gov and http://cn.expasy.org). The nucleotide sequence, deduced amino acid sequence and ORF (Open reading frame) encoded by *CaEPSPS* were analyzed and the sequence comparison was conducted through database search using *BLAST* program (*NCBI*, National Center for Biotechnology Services, http://www.ncbi.nlm.nih.gov).

The phylogenetic analyses of CaEPSPS and EPSPSs from other species were aligned with *Clustal W* (*1.82*) using default parameters. A phylogenetic tree was constructed using *MEGA* version *2.1* (Kumar *et al*. 2001) from *CLUSTAL W* alignments by neighbor-joining method (Saitou and Nei 1987). Two dimensional structural prediction of CaEPSPS was performed by the *SOMPA* (Comber *et al*. 2000) server (http://bip.weizmann.ac.il/bio-tools/faq.html). The homology-based 3-D structural modeling of CaEPSPS was accomplished by Swiss-Modeling (Schwede *et al*. 2003). *WebLab ViewerLite* was used for 3-D structure displaying (homology-based modeling by Swiss-Model).

Tissue expression pattern analysis: Semi-quantitative one-step RT-PCR was carried out to investigate the expression profile of *CaEPSPS* in different tissues

including leaves, stems and roots of *C. acuminata*. Aliquots of 0.5 µg total RNA extracted from leaves, stems and roots of *C. acuminata* were used as templates in one-step RT-PCR reaction with the forward primer fcaepsps (5'-AACGCGAGTTACGTGCTTGATGG-3') and the reverse primer rcaepsps (5'-TTGACA TCAACAGCACGCAGGTG-3') specific to the coding sequence of CaEPSPS using one-step RNA PCR kit (*Takara*, Shiga, Japan). Meanwhile, the RT-PCR reaction for the house-keeping gene (actin gene) using specific primers actF (5'-GTGACAATGGAACTGGAATGG-3') and actR (5'-AGACGGAGGATAGCGTGAGG-3') designed according to the conserved regions of plant actin genes was performed to estimate if equal amounts of RNA among samples were used in RT-PCR as an internal control. Amplifications were performed under the following condition: 50 °C for 30 min, 94 °C for 2 min followed by 25 cycles of amplification (94 °C for 50 s, 55 °C for 50 s and 72 °C for 120 s). The amplified products were separated on 1 % agarose gel and the densities of the target bands were measured using *Furi FR-200A* ultraviolet analyzer (*Furi Tech*., Shanghai, China).

Expression of *CaEPSPS* in *E. coli*: A PCR strategy was employed to subclone the coding sequence of *CaEPSPS* gene into the *E. coli* expression vector pET-32a(+) (*Novogen*). Two synthetic oligonucleotide primers were designed for the amplification of *CaEPSPS* gene based on the coding sequence of *CaEPSPS*. The forward primer fcaepsps-*Bgl*II (5'-CCCAGATCTATGGCGCAAGTTAGCAACATTCC-3') contained a *Bgl*II site (underlined) and the reverse primer rcaepsps-*Hind*III (5'-CCCAAGCTTTCAATG CTTTGCGAACCTCTGG-3') contained a *Hind*III site (underlined). These two primers were complementary to the amino-terminal and carboxyl-terminal coding strands of their respective structural genes containing *Bgl*II and *Hind*III restriction sites. The primers were used to amplify the *CaEPSPS* gene using *Pfu* DNA polymerase with 3'-RACE-ready cDNA as the template. The PCR products were purified, digested with *Bgl*II and *Hind*III, then repurified following by ligating into pET-32a(+) expression vector which was pre-digested with the same restriction enzymes to generate recombinant plasmid pET-32(a+)::*CaEPSPS*. *E. coli* strain BL21 (Novogen) was used as the host for the transformation and expression of the *caepsps* gene. The recombinant plasmid pET-32a(+)::*CaEPSPS* was transformed into *E. coli* strain BL21 cells and selected on LB agar plates containing 100 µg cm^{-3} carbenicillin. Single colony was inoculated in 1.5 dm^3 of Luria-Bertani (LB) broth with 100 µg cm^{-3} carbenicillin, grown at 37 °C until absorbance (A_{600}) reached 0.4 to 0.6. The plasmid DNA was isolated from selected colonies and its insert portion was sequenced to ensure correct sequence.

Measurement of growth in the presence of glyphosate: For comparative growth studies, fresh overnight cultures of *E. coli* strains BL21 harboring empty plasmid pET-32(a+) and recombinant plasmid pET-32(a+):: *CaEPSPS* were grown in LB medium with 100 µg cm^{-3} carbenicillin. The strain BL21 with empty pET-32a(+) was used as a control. The two transformants were cultured on solidified LB medium containing 100 µg cm^{-3} carbenicillin and different concentrations of glyphosate (0, 0.1, 0.5, 1, 2, 5, 10, 20, 40 and 80 mM). Cells were incubated at 37 °C for 24 h or 48 h.

Results and discussion

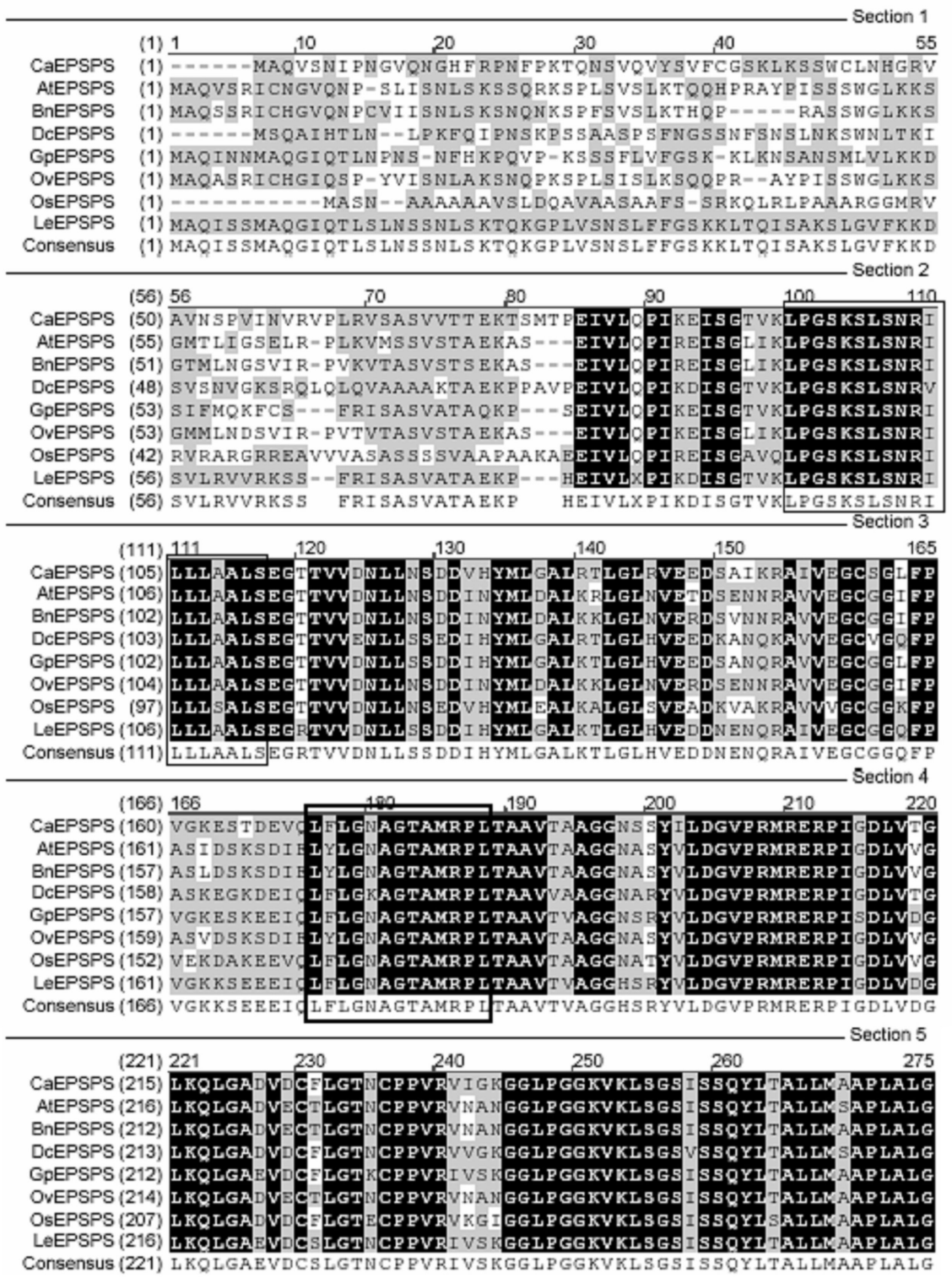
Cloning of the full-length cDNA of CAEPSPS: Based on the conserved regions of plant *EPSPS* sequences, two degenerate oligonucleotide primers (FEPSPS and REPSPS) were designed and used for gradient PCR-amplification of the core cDNA fragment of *EPSPS* from *C. acuminata*. Following PCR amplification, an approximately 650 bp product was amplified, subcloned and sequenced. The BLAST search result revealed a 642 bp cDNA fragment showing extensive homology to *EPSPS* genes from other plant species.

By 3'-RACE and 5'-RACE, the 500 bp and 850 bp nested PCR products were obtained, respectively. The products were subcloned into pGEM T-easy vector followed by sequencing and confirmed to be a 519 bp 3'-end and 846 bp 5'-end. The full-length cDNA sequence of *CaEPSPS* was 1778 bp, comprising 174 bp 5'-untranslated region, an ORF of 1557bp and 44 bp 3'-untranslated region. CaEPSPS encodes a peptide of 519 amino acids with a calculated molecular mass of 55.6 kDa and an isoelectric point of 8.22 (Fig. 1).

Comparative and bioinformatic analyses of CaEPSPS: PSI-BLAST of the deduced amino acid sequence of CaEPSPS revealed high homology with EPSPSs from other plant species, such as *Dicliptera chinensis* (77 % identities, 86 % positives), garden petunia (74 % identities, 83 % positives), *Nicotiana tabacum* (74 % identities, 84 % positives) and *Lycopersicon esculentum* (73 % identities, 82 % positives), indicating that CaEPSPS belonged to the EPSPS family. Two highly conserved motifs (LPGSKSLSNRILLLAAL and LFLGNAGTAMRPL), owned by all plants and most bacteria EPSPSs (Baerson *et al*. 2002), were also identified in CaEPSPS N-terminal region. The conserved residues may function as important catalytic domains of the enzyme. The amino acids in the first conserved motif are supposed to form a portion of a binding site for glyphosate. Mutation on amino acids especially lysine and arginine residues can alter the binding of glyphosate. Substitution of an alanine residue for the second glycine residue in the second conserved motif could produce a mutant EPSPS which exhibits

```
1       AAGCAGTGGTATCAACGCAGAGTACGCGGGGGTTGTTGGTGAACGCCTTCACTGTCAAAA
61      AAAAAACCCACCTTCCTTCCCACCAACCTCTTCCCCTCTCATAAAACATCAATTATAGAG
121     AGAGAAAAAGTTGGAAGAGTTTTTGGAGAAAGCAGGAGAAGGGAAAACGGAGAAATGGCG
                                                          M  A    2
181     CAAGTTAGCAACATTCCTAATGGAGTTCAAAACGGCCATTTTAGGCCCAATTTTCCTAAA
         Q  V  S  N  I  P  N  G  V  Q  N  G  H  F  R  P  N  F  P  K   22
241     ACCCAGAACTCCGTACAGGTGTATTCTGTATTCTGCGGATCAAAACTAAAGAGTTCATGG
         T  Q  N  S  V  Q  V  Y  S  V  F  C  G  S  K  L  K  S  S  W   42
301     TGTTTGAATCATGGGAGAGTTGCTGTCAACAGTCCCGTCATTAATGTTAGAGTTCCACTT
         C  L  N  H  G  R  V  A  V  N  S  P  V  I  N  V  R  V  P  L   62
361     AGGGTTTCAGCTTCGGTCGCCACGACCGAAAAGACGTCTATGACACCAGAGATTGTCTTG
        R  V  S  A  S  V  A  T  T  E  K  T  S  M  T  P  E  I  V  L    82
421     CAACCCATCAAAGAGATATCTGGTACCGTCAAATTACCGGGCTCTAAGTCGCTCTCGAAT
        Q  P  I  K  E  I  S  G  T  V  K  L  P  G  S  K  S  L  S  N   102
481     CGGATTCTCCTTCTTGCTGCCCTATCCGAGGGAACAACTGTTGTAGACAACCTGTTGGAC
        R  I  L  L  L  A  A  L  S  E  G  T  T  V  V  D  N  L  L  D   122
541     AGTGATGACGTCCATTACATGCTCGGAGCTTTGAGAACACTTGGGCTACGTGTGGAAGAA
        S  D  D  V  H  Y  M  L  G  A  L  R  T  L  G  L  R  V  E  E   142
601     GACAGCGCTATTAAGCGAGCAATTGTGGAAGGTTGCAGTGGTCCTTTCCCAGTTGGCAAA
        D  S  A  I  K  R  A  I  V  E  G  C  S  G  P  F  P  V  G  K   162
661     GAATCGACAGATGAAGTTCAACTTTTTCTGGGAAATGCAGGAACAGCAATGCGTCCATTG
        E  S  T  D  E  V  Q  L  F  L  G  N  A  G  T  A  M  R  P  L   182
721     ACAGCTGCTGTTACTGCTGCTGGAGGAAATTCAAGCTACATACTTGATGGGGTGCCCCGA
        T  A  A  V  T  A  A  G  G  N  S  S  Y  I  L  D  G  V  P  R   202
781     ATGAGAGAGAGACCAATTGGTGACTTGGTCACTGGTCTTAAGCAGCTTGGTGCAGATGTT
        M  R  E  R  P  I  G  D  L  V  T  G  L  K  Q  L  G  A  D  V   222
841     GATTGTTTTCTAGGTACAAACTGCCCCCCTGTACGTGTAATTGGAAAGGGAGGCCTTCCT
        D  C  F  L  G  T  N  C  P  P  V  R  V  I  G  K  G  G  L  P   242
901     GGGGGAAAGGTGAAGCTCTCTGGCTCTATCAGTAGTCAATATTTGACTGCTTTACTCATG
        G  G  K  V  K  L  S  G  S  I  S  S  Q  Y  L  T  A  L  L  M   262
961     GCAGCTCCATTGGCTCTGGGAGATGTGGAAATTGAGATTATTGATAAACTTATTTCCATA
        A  A  P  L  A  L  G  D  V  E  I  E  I  I  D  K  L  I  S  I   282
1021    CCGTATGTTGAGATGACCTTAAAGTTAATGAAACGCTTTGGGGTCACTGTAGGGCACAGT
        P  Y  V  E  M  T  L  K  L  M  K  R  F  G  V  T  V  G  H  S   302
1081    GATAACTGGGATAGGTTCTTAATCCAAGGAGGTCAAAAGTACAAGTCTCCTGGAAATTCT
        D  N  W  D  R  F  L  I  Q  G  G  Q  K  Y  K  S  P  G  N  S   322
1141    TATGTAGAAGGTGATGCTTCAAGTGCTAGTTACTTCCTAGCTGGTGCAGCTGTCACTGGT
        Y  V  E  G  D  A  S  S  A  S  Y  F  L  A  G  A  A  V  T  G   342
1201    GGGACCATCACTGTTGAAGGCTGTGGTTCAAGCAGTTTACAGGGAGATGTTAAATTTGCC
        G  T  I  T  V  E  G  C  G  S  S  S  L  Q  G  D  V  K  F  A   362
1261    GAGGTTCTTGAAAAAATGGGAGCAAAAGTTACCTGGACAGAAAACAGTGTAACCGTCACA
        E  V  L  E  K  M  G  A  K  V  T  W  T  E  N  S  V  T  V  T   382
1321    GGACCACCCCGCAATTCTTCTGGAAGGAAACACCTGCGTGCTGTTGATGTCAATATGAAC
        G  P  P  R  N  S  S  G  R  K  H  L  R  A  V  D  V  N  M  N   402
1381    AAAATGCCTGATGTTGCCATGACTCTTGCTGTTGTTGCCCTTTTTGCTGATGGTCCCACT
        K  M  P  D  V  A  M  T  L  A  V  V  A  L  F  A  D  G  P  T   422
1441    GCTATAAGAGACGTGGCTAGCTGGAGAGTGAAGGAAACAGAAAGGATGATTGCCATCTGC
        A  I  R  D  V  A  S  W  R  V  K  E  T  E  R  M  I  A  I  C   442
1501    ACAGAACTCAGAAAGTTGGGAGCAACAGTTGAAGAGGGGCCAGATTATTGTGTGATCACT
        T  E  L  R  K  L  G  A  T  V  E  E  G  P  D  Y  C  V  I  T   462
1561    CCACCAGAGAAATTAAATGTGACAGCCGTAGATACATATGATGATCACAGAATGGCAATG
        P  P  E  K  L  N  V  T  A  V  D  T  Y  D  D  H  R  M  A  M   482
1621    GCATTCTCTCTTGCTGCCTGTGCAAATGTTCCGGTTACCATCAAGGATCCTGGTTGCACT
        A  F  S  L  A  A  C  A  N  V  P  V  T  I  K  D  P  G  C  T   502
1681    CGGAAAACCTTTCCCGATTACTTTGAAGTTCTCCAGAGGTTCGCAAAGCATTGAACAACT
        R  K  T  F  P  D  Y  F  E  V  L  Q  R  F  A  K  H  *         519
1741    CTTTGACATAAAAATAAGAGGGAGAAAAAAAAAAAAAA
```

Fig. 1. The full-length cDNA sequence and the deduced amino acid sequence of *Camptotheca acuminata* 5-enolpyruvylshikimate 3-phosphate synthase (*CaEPSPS*).


Section 1
Section 2
Section 3
Section 4
Section 5
CaEPSPS
AtEPSPS
BnEPSPS
DcEPSPS
GpEPSPS
OvEPSPS
OsEPSPS
LeEPSPS
Consensus

(continued)

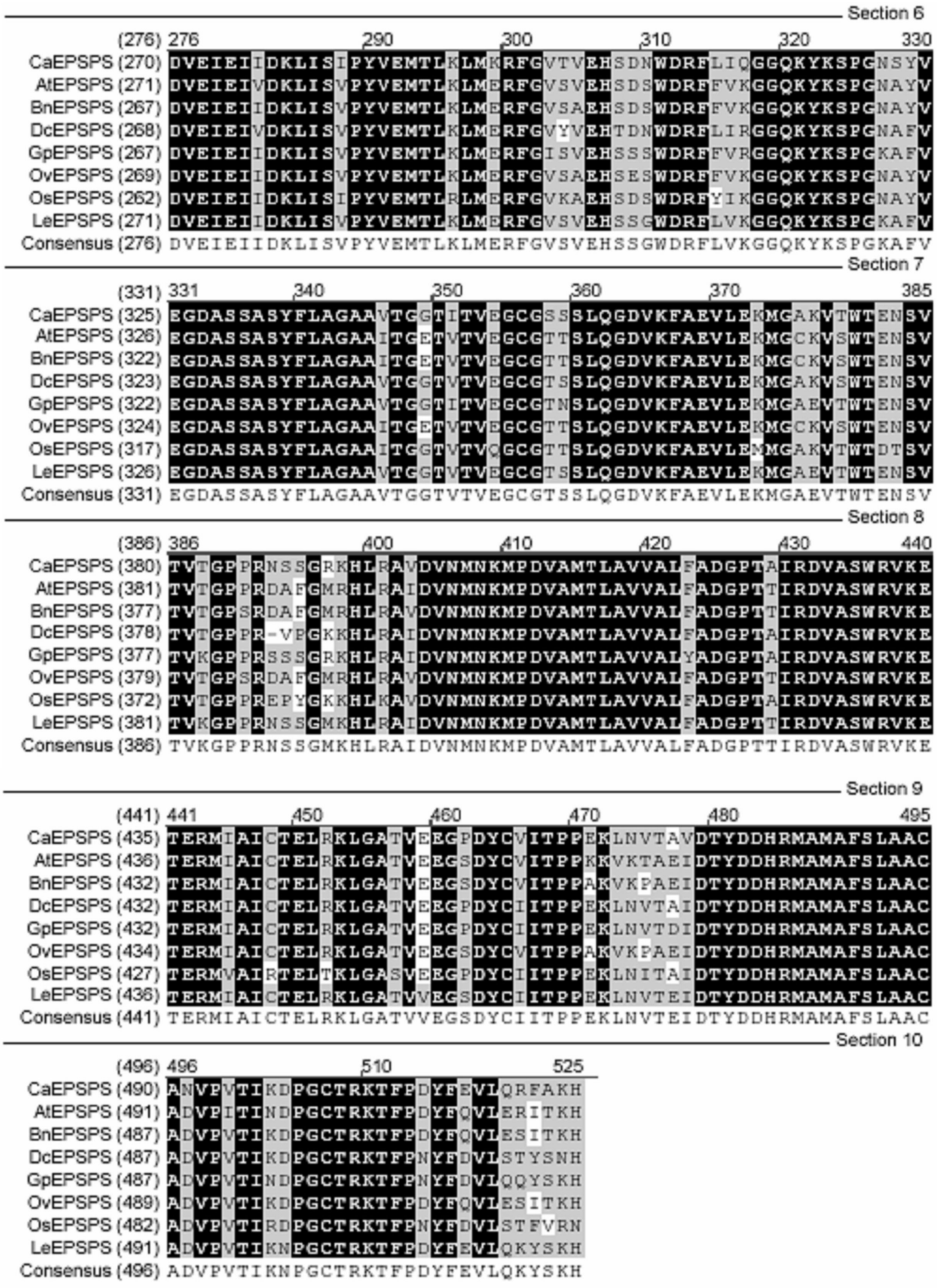


Fig. 2. Multiple-alignment of amino acid sequences of CaEPSPS and other plant EPSPSs. The identical amino acids are shown in white with black background and the conserved amino acids are shown in black with gray background. Two highly conserved motifs (LPGSKSLSNRILLLAAL and LFLGNAGTAMRPL) in all plants' and most bacteria' EPSPSs are boxed. The aligned EPSPSs are from *Camptotheca acuminata* (CaEPSPS, AY639815), *Arabidopsis thaliana* (AtEPSPS, AAB82633), *Brassica napus* (BnEPSPS, CAA35839), *Dicliptera chinensis* (DcEPSPS, AF371966), garden petunia (GpEPSPS, XUPJVS), *Orychophragmus violaceus* (OvEPSPS, AF440389), *Oryza sativa* (OsEPSPS, BAB61062) and *Lycopersicon esculentum* (LeEPSPS, XUTOVS).

a lower affinity for glyphosate while maintaining catalytic activity in an earlier study (Huynb *et al*. 1988).

Phylogenetic analysis of EPSPSs from *C. acuminata* and other plants and bacteria revealed two groups: plant

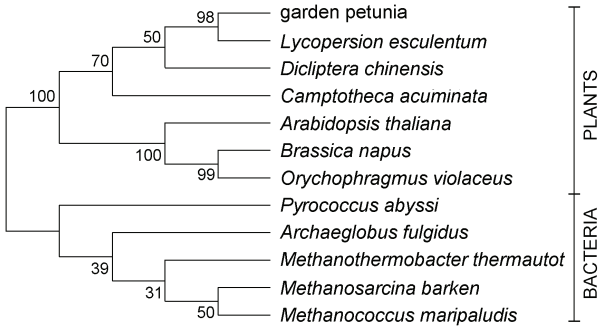


Fig. 3. The phylogenetic tree analysis of EPSPSs from plants and bacteria. The numbers on the branches represent bootstrap support for 1000 replicates. Sequences used were from plants and bacteria. The sequences used were listed bellow with GeneBank Accession number: garden petunia, XUPJVS; *Lycopersicon esculentum*, XUTOVS; *Dicliptera chinensis*, AF371966; *Camptotheca acuminata*, AY639815; *Arabidopsis thaliana*, AAB82633; *Brassica napus*, CAA35839; *Orychophragmus violaceus*, AF440389; *Pyrococcus abyssi*, CAB49378; *Archaeoglobus fulgidus*, AAB89746; *Methanothermobacter thermautot*, AAB85269; *Methanosarcina barken*, ZP_00079265; *Methanococcus maripaludis*, CAF30761.

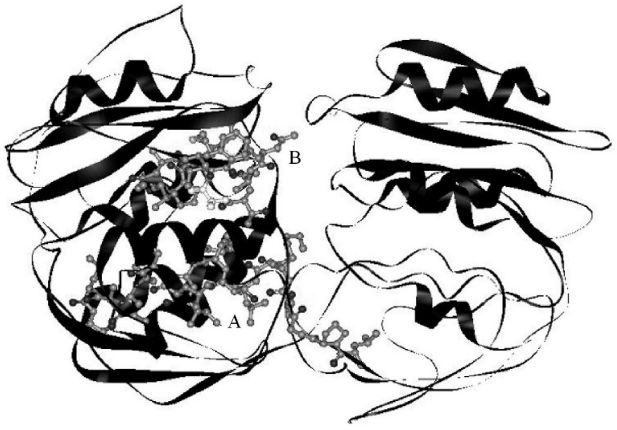


Fig. 4. The 3-D structure of CaEPSPS established by homology-based modeling. The α-helix is shown in black (helix-shaped), the β-sheet in grey (wide ribbon-shaped) and the random coil in grey (line-shaped). The highly conserved motifs A (LPGSKSLSNRILLLAAL) and B (LFLGNAGTAMRPL) in the N-terminal of the enzyme are also shown.

and bacterium EPSPS groups. According to the phylogenetic tree, CaEPSPS belonged to plant EPSPS group (Fig. 3). The analysis strongly suggested that CaEPSPS was a plant EPSPS protein involved in the shikimate biosynthesis.

Based on the Hierarchical Neural Network Analysis, CaEPSPS protein was composed of 30.25 % α-helix, 19.85 % extended strand, 7.51 % β-turn and 42.39 % random coil. The homology-based 3-D structural modeling of CaEPSPS was analyzed using the crystallographic structures of EPSP synthase from *E. coli* as template (Schonbrunn *et al*. 2001) by Swiss-Modeling and displayed by WebLab ViewerLite (Fig. 4). The overall folding of CaEPSPS, which was typically built from β-sheets connected by turns and loops, created very tight structural scaffold. The structure folded into two distinctive globular N-terminal and C-terminal domains of very similar size and symmetry connecting two crossover chains segments. Earlier studies showed that glyphosate formed a stable but noncovalent ternary with EPSPS and S3P (Marzabadi *et al*. 1996, McDowell *et al*. 1996). More complete structural studies of the EPSPS-S3P-glyphosate and EPSPS-EPSP-glyphosate ternary complexes will be needed to assist inhibitor design for the proposed allosteric glyphosate-binding domain (Sikorski *et al*. 1997). Recently the X-ray crystallographic structures of the EPSPS, and of the S3P-EPSPS-PEP and S3P-EPSPS -glyphosate complexes have been solved (Schonbrunn *et al*. 2001). However, the X-ray crystallographic structures have not sufficiently described the motion of domain closure of the enzyme. In addition, little is known about the precise structural data of the PEP binding site, the conformation changes induced by PEP binding, and the difference between the PEP and glyphosate binding site. Therefore studies on the domain-specific contribution of EPSPS to substrate binding and catalysis may give an overall explanation of the enzyme in action.

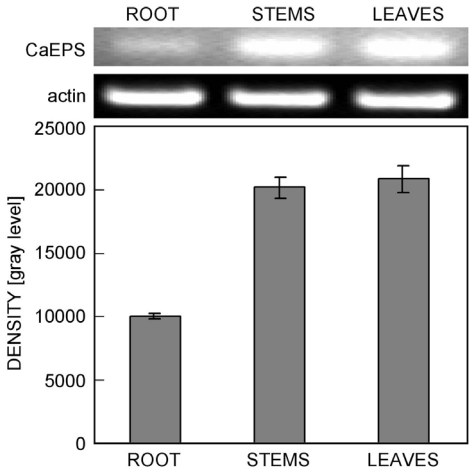


Fig. 5. Expression profile of *CaEPSPS* in different *C. acuminata* tissues by semi-quantitative one-step RT-PCR.

Tissue expression pattern analysis: Semi-quantitative one-step RT-PCR showed that *CaEPSPS* expression could be detected in all the tested tissues, with low expression being found in roots (Fig. 5).

Construction and characterization of pET32 carrying the EPSP synthase gene from C. acuminata: *E. coli* strain BL21 cells containing no expression vector

pET-32(a+) were sensitive to glyphosate and their growth was inhibited even in 0.1 mM glyphosate. In the absence of glyphosate, cells carrying empty pET-32(a+) and pET-32(a+)::*CaEPSPS* transformants could both grow well in LB medium containing carbenicillin within 24 h of incubation at 37 °C (Fig. 6). Inhibition to cell growth became more severe along with the increased concentrations of glyphosate. BL21/pET-32(a+) cells failed to grow in the presence of 10 mM glyphosate even after 48 h of incubation at 37 °C while the same glyphosate concentration showed little or no inhibitory effect on pET-32(a+)::*CaEPSPS* transformants (Fig. 6). BL21/pET-32(a+)::*CaEPSPS* could still grow on LB medium containing carbenicillin even in the presence of 20 mM glyphosate, a concentration twice the amount required to completely inhibit the growth of BL21/pET-32(a+) cells, but its growth was completely inhibited when exposed to 40 mM glyphosate. Earlier study showed that the transformants expressing goosegrass *EPSPS* gene were completely inhibited at only 5 mM glyphosate (Baerson *et al*. 2002). The glyphosate tolerance conferred by *CaEPSPS* gene can be demonstrated by colony formation on plates containing glyphosate. Therefore, glyphosate tolerance is a selectable marker for transformation of *E. coli* cells. This study showed that the cloned cDNA of *CaEPSPS* from *C. acuminata* is a functional gene conferring enhanced tolerance to glyphosate.

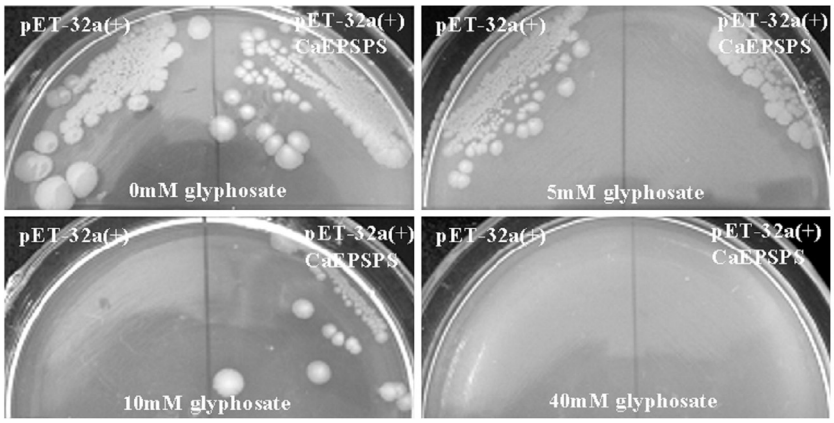


Fig. 6. Growth of *Escherichia coli* cells (BL21) carrying empty pET-32(a+) and pET-32(a+)::*CaEPSPS* in the presence of 0, 5, 10 and 40 mM glyphosate.

Conclusions: We have successfully presented the PCR amplification, cloning, sequencing and over-expression of a functional gene encoding 5-enolpyruvylshikimate 3-phosphate synthase, a committed-step enzyme involved in shikimate biosynthesis, from woody plant *C. acuminata*. The cloning and functional characterization of *CaEPSPS* will be helpful to understand more about the role of EPSPS at the molecular level. Of particular interest in our research is the potential use of cloned genes to increase plant tolerance to glyphosate by genetic engineering. Plant expression vector containing the *CaEPSPS* gene has been constructed and genetic transformation is being carried out.

References

Baerson, S.R., Rodriguez, D.J., Tran, M., Feng, Y.M., Biest, N.A., Dill, G.M.: Glyphosate-resistant goosegrass: identification of a mutation in the target enzyme 5-enolpyruvylshikimate-3-phosphate synthase. - Plant Physiol. **129**: 1265-1275, 2002.

Bentley, R.: The shikimate pathway – a metabolic tree with many branches. - Crit. Rev. Biochem. mol. Biol. **25**: 307-384, 1990.

Charles, I.J., Keyte, J.W., Brammar, W.J., Smith, M., Hawkins, A.R.: Structure and nucleotide sequence of the complex AROM locus of *Aspergillus nidulans*. - Nucl. Acids Res. **14**: 2201-2213, 1986.

Comber, C., Blanchet, C., Geourjon, C., Deleage, G.: NPS@:Network Protein Sequence Analysis. - Trends biochem. Sci. **25**: 147-150, 2000.

Duncan, K., Lewendon, A., Coggins, J.R.: The complete amino acid sequence of *Escherichia coli* 5-enol-pyruvyl-shikimate-3-phosphate synthase. - FEBS Lett. **170**: 59-63, 1984.

Forlanni, G., Parisi, B., Nielsen, E.: 5-enol-pyruvyl-shikimate -3-phosphate synthase from *Zea mays* cultured cells. - Plant Physiol. **105**: 1107-1114, 1994.

Garbe, T., Jones, C., Charles, I., Dougan, G., Young, D.: Cloning and characterization of the *aroA* gene from *Mycobacterium tumberculosis*. - J. Bacteriol. **172**: 6774-6782, 1990.

Gasser, C.S., Winter, J.A., Hironaka, C.M., Shah, D.M.: Structure, expression and devolution of the 5-enol-pyruvyl-shikimate-3-phosphate synthase genes of *Petunia* and tomato. - J. biol. Chem. **263**: 4280-4287, 1988.

Herrmann, K.M., Weaver, L.M.: The shikimate pathway. - Annu. Rev. Plant Physiol. Plant mol. Biol. **50**: 473-503, 1999.

Howe, A.R., Gasser, C.S., Brown, S.M., Padgette, S.R., Hart J., Parker, G.B., Fromm, M.E., Armstrong, C.L.: Glyphosate as a selective agent for the production of fertile transgenic

maize (*Zea mays* L.) plants. - Mol. Breed. **10**: 153-164, 2002.

Huynb, Q.K., Kisbore, G.M., Bild, G.S.: 5-enolpyruvyl shikimate-3-phosphate synthase from *Escherichia coli*: identification of Lys-22 as a potential active site residue. - J. biol. Chem. **263**: 735-739, 1988.

Klee, H.J., Muskopf, Y.M., Gasser, C.S.: Cloning of an *Arabidopsis thaliana* gene encoding 5-enol-pyruvyl-shikimate-3-phosphate synthase: sequence analysis and manipulation to obtain glyphosate-tolerant plant. - Mol. gen. Genet. **210**: 437-442, 1987.

Kumar, S.K., Tamura, K., Jakobsen, I.B., Nei, M.: MEGA2: molecular evolutionary genetics analysis software. - Bioinformatics **17**: 1244-1245, 2001.

Liao, Z., Chen, M., Guo, L., Gong, Y., Tan, F., Sun, X., Tang, K.: Rapid isolation of good-quality total RNA from *Taxus* and *Ginkgo*. - Preparative Biochem. Biotechnol. **34**: 209-214, 2004.

Marzabadi, M.R., Gruys, K.J., Pansegrau, P.D., Walker, M.C., Yuen, H.K., Sikorski, J.A.: An EPSP synthase inhibitor joining shikimate 3-phosphate with glyphosate: synthesis and ligand binding studies. - Biochemistry **35**: 4199-4210, 1996.

McDowell, L.M., Schmidt, A., Cohen, E.R., Studelska, D.R., Schaefer, J.: Structural constraints on the ternary complex of 5-enolpyruvylshikimate-3-phosphate synthase from rotational-echo double-resonance NMR. - J. mol. Biol. **256**: 160-171, 1996.

Ream, J.E., Steinrcken, H.C., Porter, C.A., Sikorsky, J.A.: Purification and properties of 5-enol-pyruvyl-shikimate-3-phosphate synthase from dark-grown seedlings of *Sorghum bicolor*. - Plant Physiol. **87**: 232-238, 1988.

Saitou, N., Nei, M.: The neighbor-joining method: a new method for reconstructing phylogenetic trees. - Mol. Biol. Evol. **4**: 406-425, 1987.

Schonbrunn, E., Eschenburg, S., Shuttleworth, W.A., Schloss, J.V., Amrhein, N., Evans, J.N.S., Kabsch, W.: Interaction of the herbicide glyphosate with its target enzyme 5-enolpyruvylshikimate 3-phosphate synthase in atomic detail. - Proc. nat. Acad. Sci. USA **98**: 1376-1380, 2001.

Schwede, T., Kopp, J., Guex, N., Peitsch, M.C.: SWISS-MODEL: an automated protein homology-modeling server. - Nucl. Acids Res. **31**: 3381-3385, 2003.

Sikorski, J.A., Gruys, K.J.: Understanding glyphosate's molecular mode of action with EPSP synthase: Evidence favoring an allosteric inhibitor model. - Assoc. chem. Res. **30**: 2-8, 1997.

Thomas, C.J., Rahier, N.J., Hecht, S.M.: Camptothecin: current perspectives. - Bioorgan. med. Chem. **12**: 1585-1604, 2004.

Wang, H.Y., Li, Y.F., Xie, L.X., Xu, P.: Expression of a bacterial *aroA-M1*, encoding 5-enolpyruvylshikimate-3-phosphate synthase for the production of glyphosate-resistant tobacco plants. - J. Plant Res. **116**: 455-460, 2003.

Xu, J.W., Feng, D.J., Li, X.G., Chang, T.J., Zhu, Z.: Cloning of genomic DNA of rice 5-enolpyruvylshikimate-3-phosphate synthase gene and chromosomal localization of the gene. - Sci. China (Ser. C) **45**: 251-259, 2002.

Ye, G.N., Hajdukiewicz, P.T., Broyles, D., Rodriguez, D., Xu, C.W., Nehra, N., Staub, J.M.: Plastid-expressed 5-enol-pyruvyl-shikimate-3-phosphate synthase genes provide high level glyphosate tolerance in tobacco. - Plant J. **25**: 261-270, 2001.