Cloning, expression, purification and functional characterization of full-length and truncated forms of *aroA* CP4 gene from genetically modified maize

MAHESH MADYAGOL – PAVEL UTEKAL – CSABA TÓTH – LUCIA PÁNČIOVÁ – VINH PHU HOANG – HANA DRAHOVSKÁ – JÁN TURŇA – STANISLAV STUCHLÍK

Summary

Development of genetically modified (GM) food has generated worldwide public controversy over food safety issues. To assess the food safety aspects and potential risk of horizontal gene transfer (HGT) from GM plants and food derived thereof, we have chosen *aroA* CP4 coding EPSPS (5-enolpyruvylshikimate-3-phosphate synthase) from GM maize. During the study of possible HGT of *aroA* CP4 gene from GM food, we have discovered and characterized truncated form of *aroA* CP4 within the cloning experiments. Comparison of properties of recombinant *E. coli* strains with both forms of CP4 EPSPS gene were performed on DNA and protein levels. We report on cloning, expression, purification and enzyme activity studies of both genes. The specific activities of both enzymes were indistinguishable indicating that the truncated CP4 EPSPS enzyme retained the activity of a full-length CP4 EPSPS. The in vitro growth test showed that both forms of *aroA* CP4 genes grew well at concentrations of 10–40 mmol·l^{-1} glyphosate and the in vivo complementation test revealed uniform growth. The results indicate that both full-length and truncated *aroA* CP4 have similar characteristics (full length CP4 EPSPS specific activity 0.063 ± 0.01 μmol·min^{-1}·mg^{-1}; truncated CP4 EPSPS specific activity 0.055 ± 0.01 μmol·min^{-1}·mg^{-1}). These results should also be taken into account in risk analysis of possible unintended effects from GM plants.

Keywords

horizontal gene transfer; truncated CP4 EPSPS; enzyme activity; food safety; risk analysis

In the last decade, recombinant DNA technologies have permitted the introduction of foreign genes into unrelated species, crossing species barriers at elevated frequencies [1]. Genetically modified (GM) foods have now been on the market for more than 12 years. Over 98% of all genetically modified organisms (GMO) that have been introduced into the environment worldwide are GM plants [2]. Prospects are that GM crops will continue to grow. The development of GM food has been a matter of considerable interest and worldwide public controversy. Food safety is of utmost concern within the European Union, being guaranteed by European Food Safety Authority (EFSA) [3]. Indeed, the European Union maintained a long "de facto" moratorium against the GM food because of the potential regards for health and environmental concerns [4]. The potential allergenicity of novel proteins expressed in GM crops is an important item in the food safety assessment. Likewise, expression of biologically active novel transgenic proteins in food derived from GM maize can potentially produce proteins with changed allergenic characteristics, which may have major effects on health of exposed individuals [3].

In this paper, we assess the food safety aspects of GM food and unintended gene transfer processes related to the use of GM crops. GMO contain transgenic genes from different species including those from different kingdoms of life [5]. The process of transmission of genes across species boundaries and/or mating barriers is known as horizontal gene transfer (HGT) [6]. This can

Mahesh Madyagol, Pavel Utekal, Csaba Tóth, Lucia Pánčiová, Vinh Phu Hoang, Hana Drahovská, Ján Turňa, Stanislav Stuchlík, Department of Molecular Biology, Faculty of Natural Sciences, Comenius University in Bratislava, Mlynská dolina, 842 15 Bratislava, Slovakia.

Correspondence author:
Mahesh Madyagol, e-mail: ing.mahesh@gmail.com, tel.: +421 2 60296 509, fax: +421 2 60296 508

increase genetic diversity and promote the spread of novel adaptations between organisms [7, 8]. Research demonstrates that transgenes can move beyond the intended organism and into the surrounding environment [9, 10] posing several risks, such as introgression into natural plant communities [11, 12] and genetic transformation into natural bacterial populations [13, 14]. Although GM crops have become a reality in agriculture, concerns have also been raised regarding the possibility that DNA from GM crops could be transferred into mammalian cells or into bacteria harboured in the gastrointestinal tract of the animals fed by these crops, and whether there might be any risk associated with such transfer [15, 16]. The health and environmental impact of potential unintended HGT from GMO is a debated concern and risk scenario [17]. Therefore, HGT is in the focus of the risk analysis of transgenic plants [18]. Thus the public and scientific concerns about the environmental and food safety of GM crops may overshadow the potential benefits offered by crop biotechnology to improve food quality.

The enzyme CP4 EPSPS (5-enolpyruvylshikimate-3-phosphate synthase) from *Agrobacterium tumefaciens* CP4, encoded by the *aroA* gene, has been used for the construction of GM crops resistant to the non-selective broad-spectrum herbicide glyphosate, marketed under the trade name Roundup. The CP4 EPSPS enzyme is a key enzyme in the biosynthesis of aromatic amino acids in microorganisms and plants via shikimate pathway [19]. The enzyme catalyses the reaction of phosphoenolpyruvate (P-E-P) with shikimate-3-phosphate (S-3-P) to form 5-enolpyruvylshikimate-3-phosphate (EPSP) and inorganic phosphate. It is the target of the herbicide glyphosate, which inhibits EPSP synthase [20]. Glyphosate specifically binds to and inactivates EPSPS, which is involved in the biosynthesis of the aromatic amino acids tyrosine, phenylalanine and tryptophan. This enzyme is present in all plants, bacteria and fungi, but not in animals, which do not synthesize their own aromatic amino acids. EPSPS is normally present in food derived from plant and microbial sources.

The main objective of this work was to assess the food safety and the potential risk of HGT from GM crops, and food derived thereof, to the bacterial microflora. For studying potential gene transfer, we have chosen *aroA* CP4 gene from Round Ready maize chromosome rendering herbicide glyphosate resistance with sequence of Monsanto Patent Seq No. 9. We report on cloning of this gene and construction of a system for high-level expression of the recombinant form of this enzyme by amplifying the *aroA* CP4 gene from the GM maize genomic DNA and subcloning into a vector suitable for expression (pMM-1). Furthermore, during the study of possible HGT of *aroA* CP4 gene from GM food in gastrointestinal tract to bacterial community living in the animal gut, we had earlier discovered and characterized the truncated form of *aroA* CP4 within the cloning experiments in *Escherichia coli.* The truncated form of *aroA* CP4 has been constructed by amplifying the *aroA* CP4 gene from the GM maize genomic DNA by specifically designed primers and subcloning into an expression vector (pMM-2). The expression, purification and initial characterization of enzyme activities for both forms of recombinant CP4-EPSP synthases were described. Growth studies in the presence of glyphosate in *aroA* auxotroph *E. coli* strain were carried out in order to compare in vitro glyphosate sensitivity between the full-length and truncated CP4 EPSPS enzymes. Comparison of properties of recombinant *E. coli* strains with both CP4 EPSPS enzyme forms was done on DNA and protein levels.

MATERIALS AND METHODS

Materials

Restriction enzymes, T4 DNA ligase, DNA polymerase (Maxima Hot Start Taq DNA Polymerase), DNA Ladder (GeneRuler 1 kb Plus), Unstained Protein Marker, GeneJET Plasmid Miniprep Kit and GeneJET Gel Extraction Kits were purchased from Fermentas (St. Leon-Rot, Germany). Phosphoenolpyruvate (P-E-P), shikimate-3-phosphate (S-3-P), HIS-Select Nickel Affinity Gel and IPTG (isopropyl-β-D-1-thiogalactopyranoside) were purchased from Sigma-Aldrich Chemie (Steinheim, Germany). pGEM-T Easy Vector System was purchased from Promega (Madison, Wisconsin, USA). BCA Protein Assay Kit and SpinPrep PCR Clean-up Kit were purchased from Novagen Merck (Darmstadt, Germany). Glyphosate (*N*-phosphonomethylglycine; Monsanto, St. Louis, Missouri, USA) was purchased locally.

Bacterial strains, vectors and plasmids

The strains, vectors and plasmids used in this study are listed in Tab. 1.

Amplification of genomic DNA

Chromosomal DNA was isolated from GM Roundup Ready maize (Monsanto), which has been used in the feeding experiments at the Animal Research Institute, Nitra (Slovakia). It car-

Tab. 1. Strains, vectors/plasmids, genotype/characteristics and their source.

E. coli strain	Relevant characteristic	Reference
XL1-Blue	[*recA1 supE44 endA1 hsdR17 gyrA46 relA1 thi lac- F'[proAB+lacIq lacZ M15 Tn10(tetr)*]	Stratagene
ER2799	*aroA* deficient strain (the *aroA* gene deleted from its chromosome)	New England Biolabs [21]
BL21(DE3)pLysS	[*F– ompT hsd SB (rB- mB-) gal dcm (DE3)*]	Novagen Merck
Rosetta-gami 2(DE3)	*Δ(ara-leu)7697 ΔlacX74 ΔphoA Pvu*II *phoR araD139 ahpC galE galK rpsL (*DE3*)* F'*[lac+ lacIq pro] gor522::*Tn*10 trxB* pRARE2 (Cam^R, Str^R, Tet^R)	Novagen Merck
Plasmid		
pGEM-T Easy	Vector for AT cloning; *ori* ColE1; Ap^R	Promega
pET-28b(+)	Expression vector; T7 promoter; Kan^R	Novagen Merck
pGEM full-length *aroA* CP4 (pGEM-1)	Full-length *aroA* CP4 gene amplified by genomic PCR and cloned into the pGEM-T Easy; Amp^R	This work
pGEM truncated *aroA* CP4 (pGEM-2)	Truncated *aroA* CP4 gene amplified by PCR and cloned into pGEM-T Easy; Amp^R	This work
pET-28b(+) full-length *aroA* CP4 (pMM-1)	Full-length *aroA* CP4 gene cloned into expression vector pET-28b(+); Kan^R	This work
pET-28b(+) truncated *aroA* CP4 (pMM-2)	Truncated *aroA* CP4 gene cloned into pET-28b(+) expression vector; Kan^R	This work

ried the synthetic *aroA* CP4 gene with sequence of Monsanto Patent Seq No. 9, US 5633435 [22] responsible for the glyphosate tolerance. PCR was performed using a BioRad Thermal Cycler (Hercules, California, USA). Amplification of the *aroA* CP4 was performed by PCR using two synthetic oligonucleotide primers following the procedure reported by [23]. Primer For1 (5'-GAC GCC ATG GCT CAC GGG TGC AAG CAG CCG TCC-3') incorporated a unique *Nco*I site whereas Rev (5'-CTC GAG AGC CTT CGT ATC GGA GAG TTC G-3') introduced a unique *Xho*I site at the 5' and 3' end, respectively. PCR amplifications were performed in a final volume 25 μl containing 2 μl of 10× reaction buffer (100 mmol·l^{-1} KCl, 50 mmol·l^{-1} $(NH_4)_2SO_4$, 200 mmol·l^{-1} Tris-HCl (pH 8.3 at 25 °C), 0.3 μl of *Taq* polymerase (5 U·μl^{-1}), 2 μl of DNA template (50–100 ng·μl^{-1} DNA), 0.5 μl of mixed deoxyribonucleotide solution (12.5 mmol·l^{-1}), 0.5 μl each of oligonucleotide primers (10 pmol·l^{-1} of each primer), and 2 μl of 50 mmol·l^{-1} $MgCl_2$. The following programme settings were used: incubation at 94 °C for 5 min followed by 30 cycles (94 °C, 30 s; 58 °C, 30 s; 72 °C, 2 min), linked to a final incubation at 72 °C for 5 min. By the amplification of DNA regions, a fragment of full length of 1 370 bp was obtained.

Generation of truncated *aroA* CP4 gene

The synthetic *aroA* CP4 gene with sequence of Monsanto Patent Seq No. 9, US 5633435, was used as a template in the amplification of the truncated *aroA* CP4 by PCR using two synthetic oligonucleotide primers. Primer For2 (5'-CCT CCA TGG TCG GCG GTC TCG CGA GCG GTG-3') incorporated a unique *Nco*I site, whereas Rev (5'-CTC GAG AGC CTT CGT ATC GGA GAG TTC G-3') introduced a unique *Xho*I site at the 5' and 3' end, respectively. The PCR amplification of truncated *aroA* CP4 was performed in 25 μl reaction mixture described above with the following programme settings: incubation at 95 °C for 2 min followed by 30 cycles (95 °C, 30 s; 58 °C, 30 s; 72 °C, 1 min), linked to a final incubation at 72 °C for 8 min. A fragment of 1 265 bp encoding the *aroA* CP4 was amplified.

Expression plasmid construction

The PCR-generated fragments encoding full-length and truncated *aroA* CP4 were purified by gel extraction using SpinPrep PCR Clean-up Kit (300 ng) and ligated into the pGEM-T Easy Vector (50 ng) with T4 DNA Ligase (3 Weiss units per·microlitre) in 2× Rapid Ligation Buffer (Fermentas, St. Leon-Rot, Germany) and incubated the reactions overnight at 4 °C to create plasmids pGEM-1 (pGEM-T Easy Vector + full-length *aroA* CP4) and pGEM-2 (pGEM-T Easy Vector + truncated *aroA* CP4). Initial clones were obtained by transformation of the ligation mixture into *E. coli* XL-1. Both pGEM-1 and pGEM-2 were digested with *Nco*I and *Xho*I. Digestion was accomplished by incubation of 3 μg of DNA with 2–5 U of *Nco*I and *Xho*I in Tango Buffer (Fermentas) at 37 °C for 2 h. Inserts were gel-purified using GeneJET Gel Extraction Kit and ligated into the

*Nco*I and *Xho*I sites of the IPTG-inducible expression vector pET-28b(+) to create plasmids pMM-1 (pET-28b(+) + full-length *aroA* CP4) and pMM-2 (pET-28b(+) + truncated *aroA* CP4). Ligation was carried out with restricted plasmid and insert DNA at a ratio of 1:3 in a volume of 10 μl. A volume of 1 μl of 10× reaction buffer (400 mmol·l^{-1} Tris-HCl, 100 mmol·l^{-1} $MgCl_2$, 100 mmol·l^{-1} DTT, 5 mmol·l^{-1} ATP, pH 7.8 at 25 °C) and 1 μl T4 DNA ligase (1 Weiss unit per microlitre) were added to the reaction mixture yielding a final volume of 10 μl. The ligation reaction was carried out overnight at 4 °C. Initial clones were obtained by transformation of the ligation mixture into *E. coli* XL-1 and then the constructs were transformed into Rosetta-gami 2(DE3) cells for expression.

DNA sequencing

Plasmid template DNA for sequencing was isolated and purified by GeneJET Plasmid Miniprep Kit according to the instructions supplied by the manufacturer. In order to determine whether the constructs contained the relevant *aroA* CP4 genes, the complete nucleotide sequence of the 1.37 kb and 1.26 kb insert from pMM-1 and pMM-2, respectively, were confirmed by sequencing of the open reading frame (*orf*) using standard relevant primers at Department of Molecular Biology sequencing facility on an automated sequencer Avant Genetic Analyser 3100 (Applied Biosystems, Foster City, California, USA) and analysed with Chromas v 2.33 software (Technelysium, Brisbane, Australia).

Heterologous expression of full-length and truncated CP4 EPSPS proteins

For the heterologous expression of the full-length and truncated *aroA* CP4 genes, the encoding genes were cloned using the T7 polymerase pET vector system with a polyhistidine tag at the C-terminus for the expression, purification and detection of the fusion protein. The recombinant plasmids (pMM-1 and pMM-2) encoding both forms of genes were transformed into electrocompetent *E. coli* Rosetta-gami 2(DE3) cells and selected on LB agar plates containing chloramphenicol 68 mg·ml^{-1}, kanamycine 50 mg·ml^{-1} and tetracycline 25 mg·ml^{-1}. Single colonies were used to inoculate 3 ml of LB media containing appropriate antibiotics and allowed to grow overnight at 37 °C. The following day, 1 ml of the culture was inoculated into 100 ml of fresh LB containing chloramphenicol 68 mg·ml^{-1}, kanamycine 50 mg·ml^{-1} and tetracycline 25 mg·ml^{-1} and grown at 37 °C. Cultures of *E. coli* cells bearing expression plasmids were grown until its absorbance at a wavelength of 600 nm (A_{600}) reached the value of 0.60. IPTG was added to a final concentration of 0.3 mmol·l^{-1} for induction of the expression of CP4 EPSPS and the cell cultures were further incubated for 3 h. The cells were harvested by ultracentrifugation on Andreas Hettich Universal 320R (Tuttlingen, Germany; 8 700 ×*g*, 10 min, 4 °C). Optimization of protein expression was carried out. For optimizing IPTG concentration, IPTG was added in a final concentration of 0.30 mmol·l^{-1}, 0.60 mmol·l^{-1} and 1 mmol·l^{-1} for induction of expression. For optimizing the induction time, periods of 2, 4, 6 h after IPTG addition were tested. After every 2 h, 1 ml of aliquots were withdrawn. The cells were harvested by centrifugation at 10 000 ×*g* for 2 min and the pellet was resuspended and heated in SDS-PAGE sample buffer and separated on 12% SDS-PAGE gel.

Preparation of crude CP4 EPSPS

The *E. coli* strain Rosetta-gami 2(DE3) containing either pMM-1 or pMM-2 plasmids were grown in 100 ml LB broth with appropriate antibiotics. When A_{600} reached 5.0, 0.3 mmol·l^{-1} IPTG was added to induce the expression. The cells were collected by centrifugation after 2 h and resuspended in 5 ml of lysis buffer (20 mmol·l^{-1} Tris; 150 mmol·l^{-1} NaCl; 5 mmol·l^{-1} imidazol, pH 8.0; 0.1% Triton X-100; 5 mmol·l^{-1} EDTA) and passed three times through a French pressure cell (Constant Systems, Daventry, United Kingdom) at 137.9 MPa. The crude homogenate was directly used for enzymatic activity assays to compare with enzyme assays of purified proteins.

Sodium dodecylsulfate polyacrylamide gel electrophoresis

For protein analysis and separation, denaturing sodium dodecylsulfate (SDS) polyacrylamide gel electrophoresis (PAGE) according to LAEMMLI [24] was used. Cell lysates and aliquots taken before and after induction were resuspended in 50 μl SDS-PAGE sample buffer (125 mmol·l^{-1} Tris-HCl pH 6.8; 10% (v/v) 2-mercaptoethanol; 0.01% (w/v) SDS; 10% (v/v) glycerol; 0.005% (w/v) bromophenol blue) prior to electrophoresis and heated for 3–5 min at 94 °C and separated on 12% SDS–PAGE gel. Gel runs were performed in an electrophoresis chamber for SDS-PAGE from Hoefer (Holliston, Massachusetts, USA) containing electrophoresis buffer (25 mmol·l^{-1} Tris-HCl; 190 mmol·l^{-1} glycine; 0.1% (v/v) SDS) at 120 V. Detection of the proteins was performed by staining with staining solution for protein gels (0.2% Coomassie brilliant blue R-250 in 10% acetic acid and 50% methanol) for 20 min at room tempera-

ture and de-stained in the same solvent without the dye. The gels were observed and analysed on sonic digitizer from LKB Bromma (Uppsala, Sweden).

Purification of His-tagged recombinant CP4 EPSPS proteins

Both forms of recombinant proteins expressed in *E. coli* Rosetta-gami 2(DE3) cells were centrifuged and resuspended in chilled lysis buffer (20 mmol·l^{-1} Tris; 150 mmol·l^{-1} NaCl; 5 mmol·l^{-1} imidazole, pH 8.0; 0.1% Triton X-100; 5 mmol·l^{-1} EDTA; 3 ml·g^{-1} cells) and passed three times through a French pressure cell at 137.9 MPa. Cell debris and unbroken cells were removed by centrifugation (8 700 ×g, 30 min, 4 °C). The supernatant and sediment were subsequently analysed by SDS-PAGE. The cell sediment was discarded and supernatant was used for purification.

Ni-NTA affinity chromatography was used to purify both forms of proteins. Appropriate amount of Nickel agarose (affinity gel) was transferred depending upon the amount of target histidine-containing proteins in the extract to a chromatography column. The affinity gel was washed with 2 volumes of deionized water, 3 volumes of strip buffer (100 mmol·l^{-1} EDTA, pH 7.0–8.8; 500 mmol·l^{-1} NaCl; 20 mmol·l^{-1} phosphate buffer, pH 7.0). Prior to application of the protein, column was equilibrated with 3 volumes of equilibration buffer (50 mmol·l^{-1} Tris-HCl, pH 8.0; 500 mmol·l^{-1} NaCl; 5 mmol·l^{-1} imidazole). Then, the clarified crude extract was loaded onto the column at a flow rate of 2–10 column volumes per hour. After all of the extract was loaded and flowed through, the column was equilibrated with 2 volumes of equilibration buffer and then washed with 3 volumes of wash buffer (50 mmol·l^{-1} Tris-HCl, pH 8.0; 500 mmol·l^{-1} NaCl; 20 mmol·l^{-1} imidazole). Elution of His-tagged proteins was done from the column using 3–10 column volumes of elution buffer (50 mmol·l^{-1} Tris-HCl, pH 8.0; 500 mmol·l^{-1} NaCl; 500 mmol·l^{-1} imidazole). Loading of the elution buffer was always 1 ml for each elution fraction. Single fractions of 1 ml were collected in each step and subsequently analysed using SDS-PAGE. Nickel agarose was cleaned by strip buffer after every run and recharged by $NiSO_4$ (100 mmol·l^{-1}).

CP4 EPSPS enzyme activity assay

CP4-EPSP synthase activity was measured in the forward direction by the production of inorganic phosphate (P_i) using the malachite green dye assay method [25]. The standard reaction was carried out at 30 °C in a final volume of 50 μl containing 50 mmol·l^{-1} HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) pH 7.0 (Sigma-Aldrich Chemie), 1 mmol·l^{-1} S-3-P, 1 mmol·l^{-1} P-E-P, 0.1 mmol·l^{-1} $(NH_4)_6Mo_7O_{24}.2H_2O$ and crude extracts (10 μl of diluted crude enzyme extraction)/purified proteins (10–25 pkat). After incubation for 20 min, 1 ml of malachite green-ammonium molybdate colorimetric solution was added and, 1 min later, 0.1 ml of 34% sodium citrate solution was added. After 15 min incubation at room temperature, absorbance of the samples was measured at 600 nm. The reaction solution without S-3-P was used as zero control. The activities of both forms of CP4 EPSPS enzymes were determined in both purified proteins and crude extracts of Rosetta-gami 2(DE3) harbouring pMM-1 and pMM-2 plasmids. The reactions were carried out at different intervals of time (0, 20, 30 and 45 min). The released phosphate was quantitatively analysed using Safire^2 Microplate Reader (Tecan, Grödig bei Salzburg, Austria).

Functional complementation and growth curves

To determine the activity of the full-length and truncated CP4-EPSP synthase in vivo, constructs pMM-1 and pMM-2 were transformed into *aroA* auxotrophic *E. coli* strain ER2799. Fresh overnight culture of *E. coli* ER2799 harbouring pMM-1 and pMM-2 plasmids was used to inoculate 20 ml of M9 minimal medium (6.8 g·l^{-1} Na_2HPO_4, 3 g·l^{-1} K_2HPO_4, 1 g·l^{-1} NH_4Cl, 0.5 g·l^{-1} NaCl, 0.12 g·l^{-1} $MgSO_4$, 0.4% w/v glucose) containing kanamycin (50 mg·ml^{-1}), 0.3 mmol·l^{-1} IPTG and different concentrations of glyphosate (0, 10, 20, 40 and 60 mmol·l^{-1}). The growth test was performed in duplicate at 37 °C for 32 h, aliquots were withdrawn at various points and the absorbance was measured spectrophotometrically at 600 nm. *E. coli* ER2799 without any construct was used as control. The cell growth was also tested on M9 minimal agar plates. The pMM-1 and pMM-2 transformed into *E. coli* ER2799 were cultured on solidified M9 medium containing 50 mg·ml^{-1} kanamycin, different concentrations of glyphosate (0, 5, 10, 20, and 40 mmol·l^{-1}) and 0.3 mmol·l^{-1} IPTG. Plates were incubated at 37 °C and the cell growth was observed after 72 h incubation.

RESULTS AND DISCUSSION

Generation of truncated coding sequence of *aroA* CP4 gene

Amplified full-length cDNA sequence was analysed for the presence of potential *orf* using ORF Finder by NCBI (National Center for

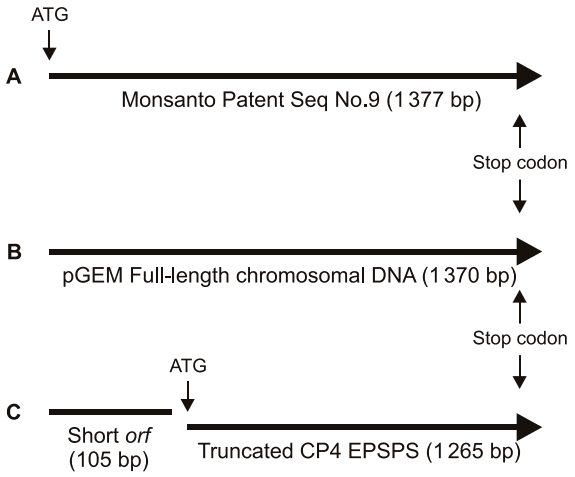

Fig. 1. Schematic diagram representing both forms of *aroA* CP4 genes and Monsanto patent Seq No.9.

A – patented Monsanto Sequence No.9, B – full-length *aroA* CP4 sequence, C – truncated *aroA* CP4 sequence with a short *orf* of 105 bp.

Biotechnology Information, Bethesda, Maryland, USA). A new *orf* within the sequence was noticed in the recombinant plasmid, which started at position 105 from the 5' end of the full-length sequence. This change in nucleotide level resulted in a new truncated *aroA* CP4 gene with an approximate length of 1 265 bp. Taken together, these findings led to speculate that cDNA has two open reading frames: one with approximate length of 105 bp and the second starting with a new initiation codon ATG at position 105 from 5' terminus (Fig. 1).

Bioinformatical analysis of both forms of *aroA* CP4 genes

The acquired nucleotide sequences of 1.37 kb and 1.26 kb inserts from pMM-1 and pMM-2, respectively, were compared in BLAST nucleotide sequence database (NCBI). It revealed that the recombinant gene *aroA* CP4 encoding the EPSPS protein was homological with other known *aroA* CP4 sequences.

However, the alignment of the cloned sequences of the full-length and truncated forms with the sequence Seq No. 9 of Monsanto, Patent US 5633435 [22], which is present in the GM Roundup Ready maize, showed only 82% similarity on nucleotide level, while on the amino acid sequence level, it showed revealed 98% similarity. The DNA sequences of both full-length and truncated forms of *aroA* CP4 genes obtained in this study showed maximum similarity of 99% when aligned with the synthetic construct of *aroA* CP4 gene sequence reported by Samsung Biomedical Research Institute (Seoul, Korea), while the alignment of the translated amino acid sequence revealed 98% identity.

The DNA sequence codes for a protein consiting of 462 amino acids in length for full-length CP4 EPSPS and 427 amino acids for truncated CP4 EPSPS. The predicted proteins were found to be acidic as the theoretical pI was 4.84 and 4.50 for full-length and truncated protein forms, respectively. The protein structure was predicted by SAM-T08 web server [26].

The alignment using ClustalW [27] of the acquired nucleotide sequences and translated amino acid sequences of recombinant truncated and full-length forms of *aroA* CP4 with the Monsanto Seq No. 9 CP4 EPSPS and synthetic CP4 EPSPS is shown in Fig. 2 and Fig. 3.

Protein expression analysis

The concentration of recombinant proteins can be increased by manipulations in bacterial cell culture conditions, such as time of induction or inducer. The effort was made to reduce the inclusion body formation, as this form of protein is inactive, and to express the proteins in a soluble form. One of the best ways to produce active enzymes is to decrease the rate of protein production. Among different factors that influence productivity, the

```
Truncated EPSPS      ------------------------------------------------------------
Full-length EPSPS    ATGGCTCACGGTGCAAGCAGCCGTCCCGCAACCGCCCGCAAATCCTCTGGCCTTTCCGGA 60
Synthetic EPSPS      ATGCTTCAGGGTGCAAGCAGCCGGCCCGCAACCGCCCGCAAATCCTCTGGCCTTTCCGGA 60
Monsanto Seq9 EPSPS  ATGGCTCACGGTGCAAGCAGCCGTCCAGCAACTGCTCGTAAGTCCTCTGGTCTTTCTGGA 60

Truncated EPSPS      ----------------------------------------------ATGGTCGGCGGTCTC 15
Full-length EPSPS    ACCGTCCGCATTCCCGGCGACAAGTCGATCTCCCACCGGTCCTTCATGTTCGGCGGTCTC 120
Synthetic EPSPS      ACCGTCCGCATTCCCGGCGACAAGTCGATCTCCCACCGGTCCTTCATGTTCGGCGGTCTC 120
Monsanto Seq9 EPSPS  ACCGTCCGCATTCCAGGTGACAAGTCTATCTCCCACAGGTCCTTCATGTTTGGAGGTCTC 120
                                                             *** * ** ******
```

Fig. 2. Alignment of the nucleotide sequences of full-length, truncated, Monsanto Patent Seq No. 9 and synthetic CP4-EPSPS.

The identical residues are signified by asterisks.

```
Monsanto Seq9 EPSPS  MAHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
Synthetic EPSPS      MAHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
Full-length EPSPS    MAHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGEARITGLLEGEDVINTG 60
Truncated EPSPS      ----------------------------------MVGGLASGEARITGLLEGEDVINTG 25
                                                        * ******* ***************

Monsanto Seq9 EPSPS  KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVZDFDSTFI 120
Synthetic EPSPS      KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVZDFDSTFI 120
Full-length EPSPS    KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVZDFDSTFI 120
Truncated EPSPS      KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVZDFDSTFI 85
                      ************************************************************
```

Fig. 3. Amino acid sequence comparison of full-length, truncated, Monsanto Patent Seq No. 9 and synthetic CP4-EPSPS.

Identical residues are marked by asterisks.

simplest one is to control the inducer concentration. In order to optimize the expression of both forms of *aroA* CP4 genes and to avoid formation of inclusion bodies, expression was performed with different induction times and with different concentrations of the inducer. The rate of production of both forms of recombinant proteins in Rosetta-gami 2(DE3) was optimized by decreasing the concentration of the inducer (IPTG) to a minimum level of 0.3 mmol·l^{-1}, as shown in Fig. 4A. Optimization of the induction time was also performed. Expression of the desired protein was achieved after 2 h of induction of the culture on a rotary shaker at 20 °C, as shown in Fig. 4B.

Protein analysis and quantitation

The determination of protein concentration was carried out using BCA protein assay method [28] following the instructions of the supplier. Bovine serum albumin (BSA; 2–10 μg·ml^{-1}) served as a standard. The concentration of full-length CP4 EPSPS was 1.14 mg·ml^{-1} and concentration of truncated CP4 EPSPS was 1.28 mg·ml^{-1}. Purified proteins analysed using SDS-PAGE are shown in Fig. 5. The molecular weight of the fusion protein (His tag + target gene) of full-length CP4 EPSPS protein was 48.53 kDa and 44.85 kDa for truncated CP4 EPSPS were determined.

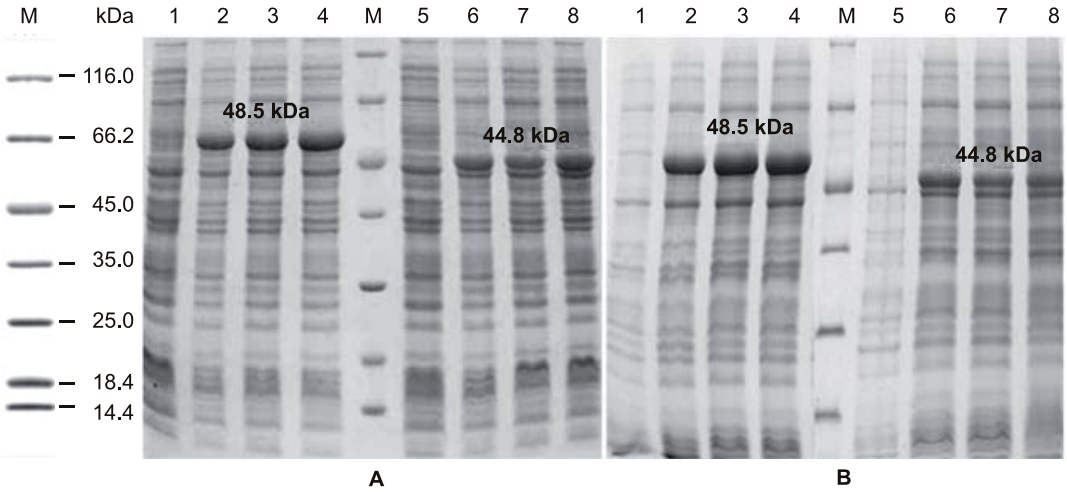

Fig. 4. Optimization of expression in *E. coli* strain Rosetta-gami 2(DE3).

Expressed products were separated by 12% SDS-PAGE.

A. Optimization of induction time: full-length CP4 EPSPS (lane 1 – lysate before induction; lanes 2–4 – lysates after 2, 4 and 6 h induction); lane M – protein marker; truncated CP4 EPSPS (lane 5 – lysate before induction; lanes 6–8 – lysates after 2, 4 and 6 h induction).

B. Optimization of IPTG concentration: full-length CP4 EPSPS (lane 1 – lysate before induction; lanes 2–4 – lysates after 2 h induction with 0.30, 0.60 and 1.0 mmol·l^{-1} of IPTG); lane M – protein marker; truncated CP4 EPSPS (lane 5 – lysate before induction; lanes 6–8 – lysates after 3 h induction with 0.30, 0.60 and 1.0 mmol·l^{-1} of IPTG).

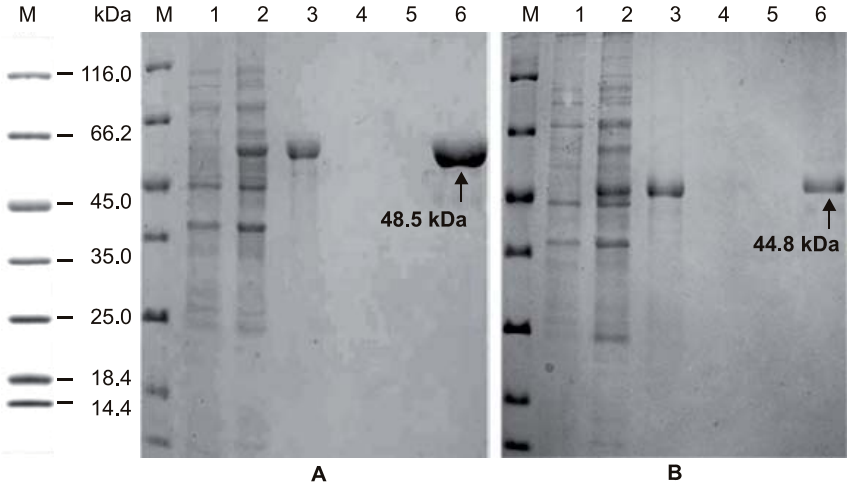

Fig. 5. 12% SDS-PAGE analysis.

A. Purification of His-tagged recombinant full-length CP4 EPSPS protein using Ni-NTA agarose chromatography: lane M – molecular weight standard, lane 1 – total proteins from cell culture before induction with IPTG, lane 2 – total proteins from cells after 2 h of induction, lane 3 – flow-through of the lysate after loading onto a column, lane 4 – equilibration wash, lane 5 – column wash, lane 6 – eluted protein from the column with elution buffer.
B. Purification of truncated CP4 EPSPS protein: lane M – molecular weight standard, lane 1 – total proteins from cell culture before induction with IPTG, lane 2 – total proteins from cells after 2 h of induction, lane 3 – flow-through of the lysate after loading onto a column, lane 4 – equilibration wash, lane 5 – column wash, lane 6 – eluted protein from the column with elution buffer.

Functional complementation and growth curves

E. coli strain ER2799 has the *aroA* gene deleted from its chromosome. Because ER2799 lacks the *aroA* gene, which is necessary for aromatic amino acid synthesis, it does not grow on M9 minimal medium. This strain is used to test the various *aroA* gene constructs to see if the new *aroA* gene can rescue the bacteria and allow growth on minimal media either in the presence or absence of glyphosate.

It was reported that amplification of EPSP synthase could confer tolerance to glyphosate [29]. Glyphosate was shown to disrupt the shikimate pathway, which is crucial for aromatic amino acids synthesis in bacteria, plants and microbial eukaryotes, by inhibiting EPSP synthase in the sixth step of the pathway [30]. The cells were grown in liquid M9 minimal medium containing IPTG (0.3 mmol·l^{-1}) and various concentrations of glyphosate (0, 10, 20, 40 and 60 mmol·l^{-1}). In order

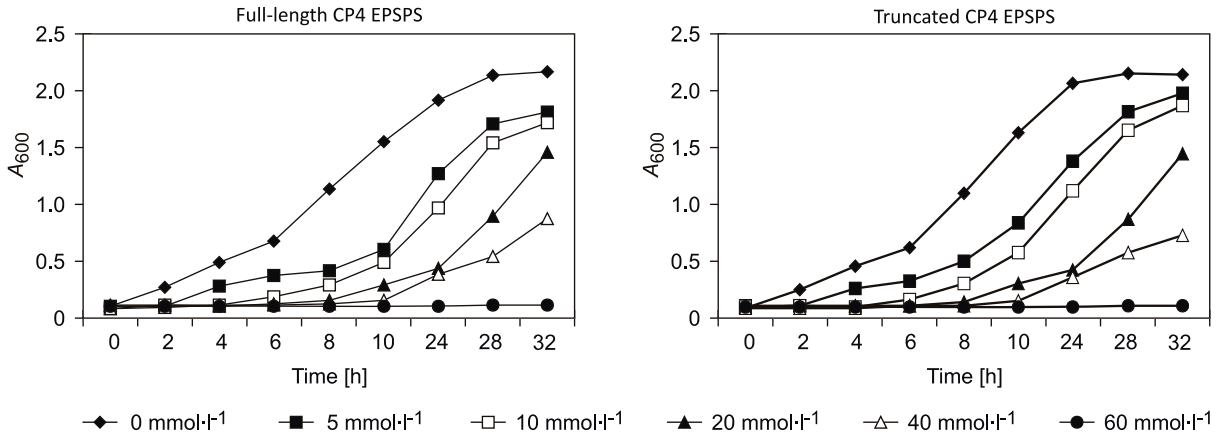

Fig. 6. Growth of cultures of ER2799 cells harbouring either pMM-1 (full-length CP4 EPSPS) or pMM-2 (truncated CP4 EPSPS) plasmids in liquid M9 minimal medium supplemented with 0.3 mmol·l^{-1} IPTG, 50 μg·ml^{-1} kanamycin and glyphosate at indicated concentrations.

Tab. 2. CP4 EPSPS enzyme activity studies of purified proteins, whole cell lysates of Rosetta-gami-2(DE3) and ER2799 strains expressing both forms of CP4 EPSPS recombinant genes.

Source of protein	Recombinant gene expressed	Specific activity [μmol·min^{-1}mg^{-1}]
Purified protein	Full-length CP4 EPSPS	0.063 ± 0.01
	Truncated CP4 EPSPS	0.055 ± 0.01
Crude lysate [Rosetta-gami 2(DE3)]	Full-length CP4 EPSPS	0.027 ± 0.008
	Truncated CP4 EPSPS	0.028 ± 0.014
Crude lysate (ER2799)	Full-length CP4 EPSPS	0.024 ± 0.004
	Truncated CP4 EPSPS	0.021 ± 0.003

Specific activity is expressed per milligram of protein.

to determine whether the over expression of the recombinant full-length and truncated CP4 EPSP synthase can bring glyphosate tolerance to *E. coli*, the growth curves of *E. coli* ER2799 harbouring pMM-1 and pMM-2 at various concentrations of glyphosate were constructed. Cultures of cells harbouring full-length and truncated *aroA* CP4 genes grew well at glyphosate concentrations of 10–40 mmol·l^{-1}, whereas ER2799 cells without the expression vectors were severely inhibited. *E. coli* ER2799 transformed with pMM-1 and pMM-2 had the same growth rate when no glyphosate was added to the culture. Inhibition became more severe with increasing concentration of glyphosate. At 40 mmol·l^{-1}, growth was barely detectable, being completely inhibited by 60 mmol·l^{-1} glyphosate. The results for growth of *E. coli* ER2799 with pMM-1 and pMM-2 were indistinguishable and our results are in concordance with [31]. This indicates that both full-length and truncated forms of *aroA* CP4 confer similar glyphosate tolerance. Growth curves of cells harbouring expression plasmids are shown in Fig. 6.

CP4 EPSPS enzyme activity assay

Analysis of CP4 EPSPS enzyme activity revealed that both full-length and truncated forms showed similar specific activities (Tab. 2). These results indicate that the truncated CP4 EPSPS protein retained the full-length CP4 EPSPS activity. From these results, a conclusion was drawn that no significant alteration in EPSPS specific activity results from the truncation of the *aroA* CP4 gene.

CONCLUSIONS

The main objective of this work was to assess the food safety and the potential risk of HGT from the GM crops. We successfully generated chromosomal DNA from GM Roundup Ready maize and truncated *aroA* CP4 gene, which has been noticed during cloning experiments in *E. coli* within the sequence of cDNA which started at position 105 from 5' end. Another finding of this study was that despite the larger changes in DNA sequence of the *aroA* CP4 (82% similarity), under a strong selection pressure in certain point of HGT, the selected functional amino acid sequence was almost identical (99% identity) with the original one. Protein expression and purification was optimized for both full-length and truncated CP4 EPSPS proteins. The enzyme activities of both forms of proteins were found indistiguishable. Properties of both forms of CP4 EPSPS were tested in vivo by complementation test on a minimal medium into the *aroA* auxotrophic *E. coli* strain ER2799. In a minimal medium with the absence of aromatic amino acids, the transformants showed uniform growth. The liquid culture growth assay was performed to determine whether the *aroA* CP4 gene of the different form is able to confer glyphosate resistance in the intact cells. Both full-length and truncated forms of *aroA* CP4 genes showed equal performance and this helped to conclude that the truncated *aroA* CP4 gene confers the same resistance to glyphosate as the full-length gene. These comparative studies of the truncated form of CP4 EPSPS with full-length CP4 EPSPS showed that even the truncated form is functional. This should also be considered when assessing the risk of the gene transfer from GM food. Thus, the possibility of HGT of the truncated *aroA* CP4 to other organisms cannot be neglected and may potentially cause negative effects on human and animal health as well as on the ecosystem. However, this risk is considered to be quite low in reality because of the small length of the transgene. The truncation of the transgene and respective changes in the protein may cause the formation of new toxins and

allergens in the plant, thus it is also essential to consider the truncated forms of transgenic DNA at the risk analysis for possible negative effects. The results of this study may be important for biosafety issues relating to the consumption of food prepared using GMO.

Acknowledgements

This publication is the result of implementation of the project Centre of excellence for utilization of information on bio-macromolecules in disease prevention and in improvement of quality of life (ITMS 26240120003) supported by the Research and Development Operational Program funded by ERDF.

REFERENCES

1. Pirondini, A. – Marmiroli, N.: Environmental risk assessment in GMO analysis. Rivista di Biologia, *101*, 2008, pp. 215–246.
2. Phillips, T.: Genetically modified organisms (GMOs): Transgenic crops and recombinant DNA technology. Nature Education, *1,* 2008, pp. 1.
3. Pusztai, A. – Bardocz, S.: Potential health effects of foods derived from genetically modified (GM) plants – What are the issues? In: Traavik, T. – Lim, L. C. (Ed.): Biosafety First: Holistic approaches to risk and uncertainty in genetic engineering and genetically modified organisms. Trondheim : Tapir Academic Press, 2007, pp. 239–252.
4. Montserrat, C. F – Gil, J. M. – Traill, W. B.: Consumer acceptance, valuation of and attitudes towards genetically modified food: Review and implications for food policy. Food Policy, *33,* 2008, pp. 99–111.
5. Traavik, T. – Nielsen, K. M. – Quist, D.: Genetically engineered cells and organisms: substantially equivalent or different? In: Traavik, T. – Lim, L. C. (Ed.): Biosafety First: Holistic approaches to risk and uncertainty in genetic engineering and genetically modified organisms. Trondheim : Tapir Academic Press, 2007, pp. 137–152.
6. Mower, J. P. – Stefanović, S. – Hao, W. – Gummow, J. S. – Jain K. – Ahmed, D. – Palmer, J. D.: Horizontal acquisition of multiple mitochondrial genes from a parasitic plant followed by gene conversion with host mitochondrial genes. BMC Biology, *8*, 2010, pp. 150.
7. Thomason, B. – Read, T. D.: Shuffling bacterial metabolomes. Genome Biology, *7*, 2006, pp. 204.
8. Marri, P. R. – Hao, W. – Golding, G. B.: The role of laterally transferred genes in adaptive evolution. BMC Evolutionary Biology, *7*, 2007, pp. 21.
9. Marvier, M. – Van Acker, R. C.: Can crop transgenes be kept on a leash? Frontiers in Ecology and the Environment, *3*, 2005, pp. 93–100.
10. Andow, D. A. – Zwahlen, C.: Assessing environmental risks of transgenic plants. Ecology Letters, *9*, 2006, pp. 196–214.
11. Loureiro, I. – Escorial, C. – Garcia-Baudin, J. M. – Chueca, C.: Hybridization, fertility and herbicide resistance of hybrids between wheat and *Aegilops biuncialis.* Agronomy and Sustainable Environment, *29*, 2009, pp. 237–245.
12. Heinemann, J. A. – Traavik, T.: Problems in monitoring horizontal gene transfer in field trials of transgenic plants. Nature Biotechnology, *22*, 2004, pp. 1105–1109.
13. de Vries, J. – Wackernagel, W.: Microbial horizontal gene transfer and the DNA release from transgenic crop plants, Plant Soil, *226*, 2004, pp. 91–104.
14. Nielsen, K. M. – Bones, A. M. – Smalla, K. – van Elsas, J. D.: Horizontal gene transfer from transgenic plants to terrestrial bacteria – a rare event? FEMS Microbiology Reviews, *22*, 1998, pp. 79–103.
15. Scientific Opinion of the Panel on Genetically Modified Organisms on applications (EFSA-GMO-RX-MON810) for the renewal of authorisation for the continued marketing of (1) existing food and food ingredients produced from genetically modified insect resistant maize MON810; (2) feed consisting of and/or containing maize MON810, including the use of seed for cultivation; and of (3) food and feed additives, and feed materials produced from maize MON810, all under Regulation (EC) No 1829/2003 from Monsanto. EFSA Journal, *1149*, 2009, pp. 1–85.
16. Hug, K.: Genetically modified organisms: do the benefits outweigh the risks? Medicina, *44*, 2008, pp. 87–99.
17. van den Eede, G. – Aarts, H. – Buhk, H. J. – Corthier, G. – Flint, H. J. – Hammes, W. – Jacobsen, B. – Midtvedt, T. – van der Vossen, J. – von Wright, A. – Wackernagel, W. – Wilcks, A.: The relevance of gene transfer to the safety of food and feed derived from genetically modified (GM) plants. Food and Chemical Toxicology, *42*, 2004, pp. 1127–1156.
18. Brigulla, M. – Wackernagel, W.: Molecular aspects of gene transfer and foreign DNA acquisition in prokaryotes with regard to safety issues. Applied Microbiology and Biotechnology, *86*, 2010, pp. 1027–1041.
19. Alibhai, M. F. – Stallings, W. C.: Closing down on glyphosate inhibition – with a new structure for drug discovery. Proceedings of the National Academy of Sciences of USA, *98*, 2001, pp. 2944–2946.
20. Franz, J. E. – Mao, M. K. – Sikorski, J. A.: Glyphosate: A unique global herbicide. ACS Monograph 189. Washington D. C. : American Chemical Society, 1997, pp. 1–653.
21. Chen, L. – Pradhan, S. – Evans Jr., T. C.: Herbicide resistance from a divided EPSPS protein: the split Synechocystis DnaE intein as an *in vivo* affinity domain. Gene, *263*, 2001, pp. 39–48.
22. Monsanto Company (St. Louis, Missouri). Glyphosate-tolerant 5-enolpyruvylshikimate-3-phosphate synthases. Inventors: Barry, G. – Kishore, G. – Padgette, B. – Stephen, R. – Stallings, W. – William, C. IPC C12N 15/82 (20060101). United States Patent Application No. US 5633435. 1997-09-13.

23. Opitz, M. – Huang, C.Y. – Ayliffe, M. A. – Timmis, J. N.: Direct measurement of the transfer rate of chloroplast DNA into the nucleus. Nature, *422*, 1997, pp. 72–76.
24. Laemmli, U. K.: Cleavage of structural proteins during the assembly of the head of bacteriophage T7. Nature, *227*, 1970, pp. 680–685.
25. Forlani, G.: Properties of the 5-*enol*-pyruvyl-shikimate-3-phosphate synthase isoforms isolated from maize cultured cells. Journal of Plant Physiology, *150*, 1997, pp. 369–375.
26. Karplus, K.: SAM-T08, HMM-based protein structure prediction. In: Nucleic Acids Research [online], *37*, 2009, Suppl. 2 (Web Server Issue), pp. W492–W497. doi: 10.1093/nar/gkp403.
27. Thompson, J. D. – Higgins, D. G. – Gibson, T. J.: CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research, *22*, 1994, pp. 4673–4680.
28. Smith, P. K. – Krohn, R. I. – Hermanson, G. T. – Mallia, A. K. – Gartner, F. H. – Provenzano, M. D. – Fujimoto, E. K. – Goeke, N. M. – Olson, B. J. – Klenk, D. C.: Measurement of protein using bicinchoninic acid. Analytical Biochemistry, *150*, 1985, pp. 76–85.
29. Stallings, W. C. – Abdel-Meguid, S. S. – Lim, L. W. – Shieh, H. S. – Dayringer, H. E. – Leimgruber, N. K. – Stegeman, R. A. – Anderson, K. S. – Sikorski, J. A. – Padgette, S. R. – Kishore, G. M.: Structure and topological symmetry of the glyphosate target 5-enolpyruvylshikimate-3-phosphate synthase: a distinctive protein fold. Proceedings of the National Academy of Sciences USA, *88*, 1991, pp. 5046–5050.
30. Steinrucken, H. C. – Amrhein, N.: The herbicide glyphosate is a potent inhibitor of 5-enolpyruvylshikimic aid-3-phosphate synthase. Biochemical and Biophysical Research Communications, *94*, 1980, pp. 1207–1212.
31. Natarajan, S. – Stuchlík, S. – Kukučková, M. – Renczésová, V. – Vávrová, S. – Bargárová, Z. – Pálffy, R. – Celec, P. – Mačor, M. – Turňa, J.: Comparative study of two forms of *aroA* CP4 gene in *Escherichia coli.* Biologia*, 62*, 2007, pp. 265–269.

Received 19 September 2011; revised 22 November 2011; accepted 30 November 2011.