Regulatory Toxicology and Pharmacology 54 (2009) 143–153

Contents lists available at ScienceDirect

Regulatory Toxicology and Pharmacology

journal homepage: www.elsevier.com/locate/yrtph

Safety evaluation of the double mutant 5-enol pyruvylshikimate-3-phosphate synthase (2mEPSPS) from maize that confers tolerance to glyphosate herbicide in transgenic plants

Corinne Herouet-Guicheney [a,*], David Rouquié [a], Martine Freyssinet [b], Thomas Currier [c], Aris Martone [c], Junguo Zhou [c,1], Elizabeth E.M. Bates [d], Jean-Marc Ferullo [d], Koen Hendrickx [d], Dominique Rouan [d]

[a] Bayer CropScience SA, BP153, 355 rue Dostoïevski, 06130 Sophia Antipolis, France
[b] Bayer CropScience SA, Lyon, France
[c] Bayer CropScience, Research Triangle Park, USA
[d] Bayer BioScience N.V., Gent, Belgium

ARTICLE INFO

Article history:
Received 15 December 2008
Available online 20 March 2009

Keywords:
2mEPSPS protein
Safety
Transgenic
Genetically modified
Glyphosate tolerance

ABSTRACT

Glyphosate tolerance can be conferred by decreasing the herbicide's ability to inhibit the enzyme 5-enol pyruvylshikimate-3-phosphate synthase, which is essential for the biosynthesis of aromatic amino acids in all plants, fungi, and bacteria. Glyphosate tolerance is based upon the expression of the double mutant 5-enol pyruvylshikimate-3-phosphate synthase (2mEPSPS) protein. The 2mEPSPS protein, with a lower binding affinity for glyphosate, is highly resistant to the inhibition by glyphosate and thus allows sufficient enzyme activity for the plants to grow in the presence of herbicides that contain glyphosate. Based on both a review of published literature and experimental studies, the potential safety concerns related to the transgenic 2mEPSPS protein were assessed. The safety evaluation supports that the expressed protein is innocuous. The 2mEPSPS enzyme does not possess any of the properties associated with known toxins or allergens, including a lack of amino acid sequence similarity to known toxins and allergens, a rapid degradation in simulated gastric and intestinal fluids, and no adverse effects in mice after intravenous or oral administration (at 10 or 2000 mg/kg body weight, respectively). In conclusion, there is a reasonable certainty of no harm resulting from the inclusion of the 2mEPSPS protein in human food or in animal feed.

1. Introduction

Glyphosate is an active ingredient in certain herbicides that are used as non-selective agents for weed control in various crops. The mode of action of glyphosate is to specifically bind to and block the activity of the 5-enol pyruvylshikimate-3-phosphate synthase (EPSPS), an essential enzyme involved in the biosynthesis of aromatic amino acids in all plants, bacteria and fungi. The double mutant 5-enol pyruvylshikimate-3-phosphate synthase (2mEPSPS) protein, which is encoded by the *double mutant 5-enol pyruvylshikimate-3-phosphate synthase* (*2mepsps*) gene, has been widely used to confer glyphosate tolerant properties, to various crops such as corn, cotton, canola, and soybean, by using genetic engineering. The *2mepsps* coding sequence was produced by introducing two point mutations to the wild-type *epsps* gene cloned from corn (*Zea mays*) through *in vitro* DNA technologies. The resultant 2mEPSPS protein has a lower binding affinity for glyphosate, thus allowing sufficient enzyme activity for the plants to grow in the presence of glyphosate herbicide (data not shown).

In higher plants, EPSPS is synthesized from a nuclear gene in the form of a cytoplasmic precursor and imported into the plastids where it accumulates in its mature form (Forlani et al., 1994; Lebrun et al., 1997a). Transit peptides are typically cleaved from the pre-protein upon translocation to the plastids (Della-Cioppa et al., 1986). The *2mepsps* gene was introduced into the nuclear plant genome with a plastid transit peptide sequence fused to the 5′-end of the coding sequence. In most instances, the transit peptide was the optimized plastid transit peptide sequence (OTP) containing sequences from the ribulose-1,5-disphosphate carboxylase oxygenase (RuBisCO) small subunit genes of maize and of sunflower (*Helianthus annuus*) (Lebrun et al., 1997b). Additionally, appropriate promoters and terminators were inserted at the 5′ and 3′ ends of the *2mepsps* gene to promote its expression in various crops that express glyphosate tolerance.

* Corresponding author. Fax: +33 (0) 4 93 65 39 24.
E-mail address: corinne.herouet-guicheney@bayercropscience.com.
[1] Present Address: Johnson & Johnson Pharmaceutical, Raritan, USA.

0273-2300/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.yrtph.2009.03.005

The objective of this manuscript is to present safety evaluation of the 2mEPSPS protein, using an approach that is consistent with recommendations provided by the *Codex Alimentarius* (CAC, 2003a), International Organizations including the OECD (2001) and Regulatory Authorities (Paoletti et al., 2008). It starts with an analysis of the published information with a focus on the origin of the nucleotide sequences, the natural occurrence of similar EPSPS proteins and the potential risk of consuming them. More importantly, experimental studies were conducted with highly purified proteins to further address safety questions. Specifically, the potential for adverse effects in mammals caused by the 2mEPSPS protein is evaluated by considering the physico-chemical characteristics of the protein, by assessing its similarity to common allergens and toxins through bioinformatics studies, by determining the stability of the 2mEPSPS protein through digestibility studies as well as by evaluating the acute toxicity through *in vivo* tests in mice.

2. Materials and methods

2.1. Engineering of the 2mEPSPS protein

The wild-type *epsps* coding sequence was isolated from the maize genome and mutated by site-directed mutagenesis (Lebrun et al., 1997a). Specifically, the wild-type gene was mutated at position 102 (substituting threonine by isoleucine) and at position 106 (substituting proline by serine) giving rise to the double mutant gene (*2mepsps*) (Lebrun et al., 1997a). A methionine codon was added between the N-terminal of the 2mEPSPS protein sequence and the optimized plastid transit peptide. The mutations finally reside at positions 103 and 107 of the mature 2mEPSPS protein. The *2mepsps* gene encodes a 47 kDa mature protein consisting of 445 amino acids (Fig. 1).

2.2. Amino acid sequence homology and N-glycosylation site searches

2.2.1. Overall toxin and allergen homology searches

On March 2, 2009, an *in silico* search was conducted to identify any known allergens presenting an amino acid identity >35% with the 2mEPSPS protein, on a window of 80 amino acids. The complete amino acid sequence of the 2mEPSPS protein was compared to AllergenOnline (version 8.0, 1313 sequences). AllergenOnline (http://www.allergenonline.com) is a resource list that is focused on both known and putative allergens. This allergen database was specifically developed for predicting allergenicity of novel proteins by the Food Allergy Research and Resource Program's (FARRP). The database is updated annually by searching NCBI and International Union of Immunological Societies (IUIS) as well as by evaluating the candidate entries for evidence of allergenicity. Additions or subtractions to the database are done by a peer-review panel of international allergy experts to judge whether proteins are allergens or putative allergens based on predefined criteria. The criteria and references for evidence of allergenicity of the groups are provided on the website.

On March 2, 2009, another *in silico* search was conducted in an attempt to determine if the 2mEPSPS protein shares amino acid identities with known toxins. The complete amino acid sequence of the 2mEPSPS protein was then compared to all protein sequences (including potential toxins) present in six large reference public databases: Uniprot-Swissprot (release 56.7, 2009; 408,099 sequences), Uniprot-trEMBL (release 39.7, 2009; 7,001,017 sequences), PIR (Protein Identification Resources, release 80, 2004, 283,416 sequences), Nrl-3D (2007, 56,020 sequences), DAD (DDBJ Aminoacid sequence Database, 44.0, 2008, 2,561,319 sequences), and GenPept (release 169, 2009, 6,185,794 sequences). Visual inspection of the specific details of alignments was employed (*e.g.*, E-score, shared identity, length of the alignment, and gap frequency) due to the lack of a specific definition for significant criteria with toxins (International Life Science Institute, ILSI, 2009).

For all searches, the comparison algorithm used was BLASTP (standard protein–protein basic local alignment search tool program), which is maintained by the National Center for Biotechnology Information (NCBI, Altschul et al., 1997). The BLOSUM62 scoring matrix enabled the comparison of sequences with no less than 62% divergence (Henikoff and Henikoff, 1992).

2.2.2. Epitope homology search

On March 2, 2009, another type of bioinformatics study was carried out by comparing the 2mEPSPS protein sequence, split up into 8 amino acid blocks, with all known allergens present in the AllergenOnline database by using the FindPatterns program from the Genetic Computer Group (GCG) suite.

For this epitope search, matches of at least 8 contiguous and identical amino acids were used for identifying potential allergens. This size assumption was based on an extensive literature review (Hileman et al., 2002; Kleter and Peijnenburg, 2002; Stadler and Stadler, 2003; Thomas et al., 2005; Goodman, 2006) showing that high numbers of non-allergens have matching sequences of 7 amino acids or less with known allergens, hence any such match is probably random and cannot be interpreted as an indication of an allergenic potential.

2.2.3. N-Glycosylation search

The 2mEPSPS protein was also examined for any potential N-glycosylation sites based on its known consensus sequence. The sequences were Asparagine-Xaa-Serine/Threonine (where Xaa

```
1   MAGAEEIVLQ PIKEISGTVK LPGSKSLSNR ILLLAALSEG TTVVDNLLNS EDVHYMLGAL
61  RTLGLSVEAD KAAKRAVVVG CGGKFPVEDA KEEVQLFLGN AGIAMRSLTA AVTAAGGNAT
121 YVLDGVPRMR ERPIGDLVVG LKQLGADVDC FLGTDCPPVR VNGIGGLPGG KVKLSGSISS
181 QYLSALLMAA PLALGDVEIE IIDKLISIPY VEMTLRLMER FGVKAEHSDS WDRFYIKGGQ
241 KYKSPKNAYV EGDASSASYF LAGAAITGGT VTVEGCGTTS LQGDVKFAEV LEMMGAKVTW
301 TETSVTVTGP PREPFGRKHL KAIDVNMNKM PDVAMTLAVV ALFADGPTAI RDVASWRVKE
361 TERMVAIRTE LTKLGASVEE GPDYCIITPP EKLNVTAIDT YDDHRMAMAF SLAACAEVPV
421 TIRDPGCTRK TFPDYFDVLS TFVKN
```

Fig. 1. The *2mepsps* coding sequence codes for a protein of 445 amino acids almost identical to the wild-type EPSPS protein. The 2mEPSPS protein differs from the wild-type maize EPSPS at two positions (in blue and bold). A threonine was replaced by an isoleucine at position 103 and a proline was replaced by a serine at position 107. Two putative N-glycosylation sites are indicated at positions 118 and 394 (in green, bold, and underscored).

represents any amino acid except Proline) and Asparagine-Xaa-Cysteine. The algorithm used for carrying out the analyses was FindPatterns program from GCG suite.

2.3. 2mEPSPS protein production and characterization

2.3.1. 2mEPSPS protein production and purification from Escherichia coli (E. coli)

The *2mepsps* coding sequence was inserted into an *E. coli* cloning vector. The expression vector was introduced into an *E. coli* NovaBlue DE3 strain. The recombinant bacteria were grown in medium containing 25 μg/ml kanamycin for selection. The expression of 2mEPSPS protein was induced with 0.5 mM isopropyl β-D-thiogalactopyranoside (IPTG, United States Biochemical). Cells were resuspended in buffer A (10 mM NaH2PO4, 1 mM DTT, 1 mM EDTA, pH 7.8) and lysozyme was added. After homogenization, benzonase was also added to the suspension. After centrifugation, the supernatant was dialyzed several times against buffer A and the dialysates were adjusted to pH 7.8. Then the solution was then loaded onto an equilibrated anion exchange Q-Sepharose column (GE HealthCare). The 2mEPSPS protein was eluted by increasing the salt concentration from 0% to 100% buffer B (buffer A + 400 mM KCl). The 2mEPSPS protein fractions were pooled and precipitated with ammonium sulphate and loaded onto an equilibrated Phenyl Sepharose column (GE HealthCare). The 2mEPSPS protein was eluted by decreasing the salt concentration from 100% buffer C (buffer A + 1.1 M $(NH_4)2SO_4$, pH 7.8) to 100% buffer A. The 2mEPSPS containing fractions were pooled and dialyzed against buffer A. They were then loaded onto a Resource Q column (GE HealthCare). The protein was eluted by a stepwise increase of the salt concentration. The 2mEPSPS containing fractions were pooled and dialyzed against buffer F (0.1 M Tris, 2.7 mM KCl, 137 mM NaCl, 1 mM DTT, pH 7.5).

The *E.coli*-produced 2mEPSPS protein by was used in the digestibility and acute tests. Importantly, Detoxi-Gel™ (Pierce) was used to remove endotoxins from the protein batch used in the acute studies. Endotoxin concentrations were measured and calculated by using the Chromogenic Limulus Amebocyte Lysate (LAL) QCL-1000 kit (Cambrex).

2.3.2. 2mEPSPS protein production and purification from transgenic plants

The *in planta* 2mEPSPS protein was purified from GHB614 transgenic cotton. Fresh leaves were harvested, inserted into plastic bags with a zip closure and placed directly on dry ice. The frozen leaves were then stored at −10 °C or lower until grinding. Grinding was performed on the day of harvest. The sample for analysis was ground in a Waring Laboratory Blender prechilled with dry ice. Small amounts of dry ice were added to the blender periodically to ensure the sample remained frozen during preparation. The ground sample was stored in a freezer at approximately −20 °C overnight to allow the dry ice to dissipate before extraction.

The 2mEPSPS protein was extracted by mixing ground plant leaves at a ratio of 1 g of ground leaves to 5 mL of extraction buffer in a 50 mL polypropylene screw cap centrifuge tube. The extraction buffer contained 120 mM NaCl, 2.7 mM KCl, 100 mM sodium phosphate (pH 7.4), 0.1 mM DDT, 10 g/L polyvinylpyrrolidone (PVP-40, average molecular weight 40,000) and 0.4% octyl-β-D-glucopyranoside. Protease inhibitors (1 μg leupeptin, 1 mM phenyl-methanesulfonyl-fluoride (PMSF), 1 mM benzamidine HCl and 1 μg/ml antipain) were added to the extraction buffer immediately before use. The tube containing extraction buffer and ground leaf material was continuously inverted at a rate of about 20 rpm at room temperature for 30 min. The extract was clarified by centrifugation in a Beckman JA-17 rotor at 16,000 rpm for 10 min at 20 °C and further clarified by filtration through a 0.22 μ polyethersulfone filter.

The antibody affinity column used for this purification was prepared using approximately 2.5 mg of a monoclonal antibody to 2mEPSPS protein, which was covalently attached to a column (Pierce, Rockford, IL, product number 44894) with a bed volume of about 2 ml. The antibodies were covalently attached to the column according to the manufacturer's instructions.

The clarified extract was then added to an antibody affinity column equilibrated with extraction buffer without the protease inhibitors. Approximately 50 ml of extract was prepared. The extract was added to a column packed with a matrix containing the covalently attached anti-2mEPSPS antibodies. The column was then unplugged, and the entire amount of extract was allowed to flow through the antibody affinity column.

The antibody column was washed with about 15 ml of extraction buffer not containing protease inhibitors. The 2mEPSPS protein was then eluted by washing the column with 30 ml of acid Pierce elution buffer (pH 2.5, product number 21004). The sample was collected into a tube containing 1 ml of 1 M sodium phosphate buffer (pH 7.6) in order neutralize the acidic Pierce elution buffer as rapidly as possible. The entire collected volume was concentrated by centrifugation for 10 min at 4100*g* in two Amicon Ultra centrifugal filter devices (Millipore, Bedford, MA product number UFC901024). After the first centrifugation, the volume was increased to 15 mL with 0.1 M sodium phosphate, 0.15 M NaCl buffer (pH 7.2). This step was repeated until the final volume reached about 0.4 ml. All extraction and purification steps were performed at 20–25 °C. The purified protein solution was then stored at −10 °C or lower until further analyses were performed.

Isolation of 2mEPSPS protein from soybean and canola was done with the same antibodies and in essentially the same way as for cotton except for the amount of starting leaf material that was different and the ratio of leaf to extraction buffer that was also different in some cases. The reason for these differences is that canola and soybean express different amounts of 2mEPSPS protein. Moreover, different amounts of starting material were necessary to prepare sufficient amounts of extract to saturate the antibody affinity column.

2.3.3. Demonstration of structural and functional equivalence

The identity and integrity of the 2mEPSPS proteins produced in *E. coli* and in cotton, canola or soybean events were characterized by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS–PAGE), high performance liquid chromatography (HPLC)–electrospray mass spectroscopy (LC/MS) and N-terminal sequencing. The immunoreactivity of the 2mEPSPS proteins was confirmed by western blotting. The enzymatic activity was also evaluated.

2.3.3.1. SDS–PAGE. For analysis of the 2mEPSPS proteins produced in *E. coli* and cotton event GHB614, SDS–PAGE was performed using a Novex Bis–Tris 12% polyacrylamide gel (InVitrogen, product number NP0341BOX) and a MOPS SDS running buffer according to the manufacturer's instructions. Approximately 300 and 1200 ng of 2mEPSPS protein from *E. coli* and cotton GHB614, respectively, were loaded per lane. After electrophoresis, the gel was stained with Coomassie brilliant blue and destained with methanol, acetic acid, water until the background was clear.

2.3.3.2. Western blotting. For analysis of the 2mEPSPS proteins produced in *E. coli* and various transgenic plants, SDS–PAGE was performed using the Novex Bis–Tris system described above. The gel was blotted to a PVDF membrane (New England Nuclear, product number NEF1001) for western blotting according to the instructions provided by InVitrogen. A rabbit antibody to the 2mEPSPS protein (Bayer CropScience), which was linked to horse radish peroxidase was incubated with the membrane. A substrate that becomes luminescent when reacting with horse radish peroxidase

was used to reveal the position of the protein of interest (ECL Plus luminescent detection kit). Approximately 5 ng of protein from *E. coli*, GHB614 cotton, canola or soybean events were loaded in their respective lanes. All reagents used for western blotting were obtained from Amersham Pharmacia Biotech. MagicMark™ XP molecular weight (MW) markers came from InVitrogen Life Technologies.

2.3.3.3. High performance liquid chromatography (HPLC)–electrospray mass spectrometry (LC/MS). An HPLC-electrospray selected ion monitoring mass spectrometry (LC/MS) method was developed to analyze the peptides generated from a trypsin digest of 2mEPSPS protein. The method monitored the mass to charge ratio of the most abundant charge state of each peptide. Peptides with only 1–3 amino acids were not analyzed because they were statistically not unique to one particular protein. The peptides generated from a trypsin digest of the standard 2mEPSPS protein from *E. coli* were compared to the peptides isolated from the trypsin digest of the 2mEPSPS protein isolated from a GHB614 cotton leaf sample by an antibody column.

The 2mEPSPS protein from *E. coli* was denatured with 6 μl of Sodium 3-[(2-methyl-2-undecyl-1,3-dioxolan-4-yl)methoxy]-1-propanesulfonate (1 mg/ml, Waters, product number 186,001,861) and 10 μl of dithiothreitol (200 mM, Sigma–Aldrich) at 37 °C for 1 h and digested with 1 μg of Trypsin (Promega). The resulting peptides were injected onto a Perceptive Biosystem PepMap C18 2.1 × 250 mm column and a BioSuite C18 PA-B 2.1 × 250 mm column connected to an Ionics EP10+MS/MS instrument for analyses. The HPLC solvents were (A) acetonitrile 99% and 1% formic acid and (B) formic acid 1% at a flow rate of 100 μl/min. A gradient was used for both columns.

Each peptide obtained from the standard 2mEPSPs protein trypsin digest was identified by its retention time and mass to charge ratio. If a peptide was present in the trypsin digest of the plant 2mEPSPS protein, a peak was observed at the expected retention time and mass to charge ratio. The actual molecular weight (MW) of each peptide was determined by multiplying the number of protons indicative of the charge state by the mass to charge ratio of the detected ion and then subtracting the number of protons from this value.

2.3.3.4. N-terminal sequencing. The N-terminal sequences of the 2mEPSPS proteins expressed in *E.* coli, transgenic cotton event GHB614, canola or soybean were also analyzed in order to confirm the identity of 2mEPSPS protein. Sequencing was performed at Eurosequence B.V. (Groningen, The Netherlands) with a Perkin Elmer Applied Biosystems Model 494 Procise protein/peptide sequencer equipped with an on-line Perkin Elmer Applied Biosystems Model 140C PTH Amino Acid Analyzer (Hewick et al., 1981). The chemical process employed by the protein sequencer to determine the amino acid sequence was derived from the degradation method developed by Edman (1964).

2.3.3.5. Determination of the enzymatic activity. The enzymatic activities in *E. coli* and in leaf material of transgenic cotton event GHB614, canola and soybean were monitored according to the method described by Forlani et al. (1994) with shikimate-3-phosphate (S3P) and phosphoenolpyruvate (PEP) as substrates. The release of inorganic phosphate was measured based on the malachite green dye method (Lanzetta et al., 1979) with minor modifications.

2.4. Digestibility tests

2.4.1. Human simulated gastric fluid (SGF) assay

The pH in the lumen of the human stomach is typically between pH 1 and 2 under fasting conditions, although there is considerable intra- and inter-individual variation (Evans et al., 1988; Russell et al., 1994; Lindahl et al., 1997). In 2001, in an attempt to account for this pH range in the *in vitro* pepsin digestion assay, the scientific advisory panel of the FAO/WHO (2001) recommended performing the pepsin digestion assay at pH 2.0. In 2004, the results from the ILSI ring trial study indicated that there was no appreciable difference in the time to disappearance of the full-length protein or protein fragments at pH 1.2 or 2.0 (Thomas et al., 2004). However, there was greater consistency among laboratories at pH 1.2 than pH 2.0, perhaps because of differential buffering capacities of the test proteins, or solutions that might raise the reaction conditions beyond the optimal activity for pepsin. This result, together with the existence of an historical database of the *in vitro* digestive fate of proteins tested at pH 1.2 (Astwood, 1996; Fu et al., 2002; Kenna and Evans, 2000), including a study with EPSPS (Okunuki et al., 2002), suggested that pH 1.2 be utilized for analyses of protein digestibility by pepsin.

The SGF assay was performed in accordance with ILSI recommendations (Thomas et al., 2004) at pH 1.2. The 2mEPSPS protein was digested in a solution of pepsin (Sigma) at 37 °C for incubation times ranging from 0.5 to 60 min. Two reference proteins, horseradish peroxidase (Sigma, France) and ovalbumin (Sigma) were used as controls. The resulting solutions were analysed by SDS-PAGE as described by Laemmli (1970) using a Bio-Rad Mini-Protean III cell (10–20% gradient polyacrylamide Tris/Tricine, Bio-Rad, France). Molecular weight protein markers (Mark 12, InVitrogen – 2.5–200 kDa) were run on each gel. The gels were stained with the Coomassie brilliant blue method (InVitrogen) as described by Neuhoff et al. (1988). After rinsing, the gels were scanned by using a GS-800 calibrated densitometer (Bio-Rad).

2.4.2. Human simulated intestinal fluid (SIF) assay

The protocol was based on the procedure previously described (Thomas et al., 2004) with an adaptation to use SIF instead of SGF, *i.e.* presence of pancreatin rather than pepsin and pH 7.5 rather than pH 1.2. After separation by SDS–PAGE electrophoresis, the 2mEPSPS protein was transferred onto nitrocellulose membranes (New England Nuclear, MA) by electroblotting. For immunodetection, the membranes were incubated with specific polyclonal rabbit anti-2mEPSPS protein antibodies, at a dilution of 1:40,000 in PBS/Tween 20 buffer, for 1 h. After rinsing, the immunoreactivity was revealed by secondary goat anti-rabbit antibodies coupled with an alkaline phosphatase (Sigma) for 1 h, at a dilution of 1:2500 in a similar buffer. The detection of the alkaline phosphatase was done by placing the membrane in a solution of bromo-chloro-indoyl-phosphate and nitro-blue-tetrazolium (Sigma) for approximately 20 min in a dark room, then washed and dried. The protein molecular weight markers, run with each gel, were revealed with Coomassie brilliant blue. The membranes were scanned by using a GS-800 calibrated densitometer (Bio-Rad).

2.5. Acute studies in mice

To test the 2mEPSPS protein for any likelihood of direct toxicity, a single dose was administered to mice either by the oral route of exposure, which is considered to be the most important portal of entry via the digestive tract (Delaney et al., 2008), at 2000 mg/kg body weight (OECD, 2001) or by the intravenous route of exposure at 10 mg/kg body weight (Delaney et al., 2008).

2.5.1. Acute oral study

As recommended by the OECD TG 425 limit test guidance document (OECD, 2001), groups of 5 female OF1 mice were administered either 2mEPSPS protein or bovine serum albumin (negative

control) by oral gavage. The dose level was at 2000 mg/kg body weight. All animals were observed for clinical signs daily for 15 days whilst their body weights were measured weekly. At termination of the study period, animals were subjected to a necropsy including macroscopic examination.

2.5.2. Acute intravenous study

The parenteral route of exposure was used to ensure maximum systemic exposure as reported in the scientific literature (Delaney et al., 2008) with intravenous injections of toxic proteins (*e.g.*, bacterial toxins). Moreover, by excluding the confounding effects of an unknown amount of digestion and absorption at different stages of the digestion process, this route provided the advantage to study the direct potential intrinsic toxic properties of food proteins in cases when the physiological degradation mechanisms may not work or only partially work (Delaney et al., 2008; Dirks et al., 2005; Untersmayr et al., 2005). More importantly, with regards to surgical or hygienic applications, this study design would permit a direct assessment of the toxic potential of 2mEPSPS containing-absorbent cotton, which might have a direct contact with blood.

Groups of 5 female OF1 mice were given intravenous tail injections of the 2mEPSPS protein, aprotinin or melittin in physiological saline at the relatively high dose levels of 1 and 10 mg/kg body weight, at a constant volume of 10 ml/kg body weight. As for the acute oral study, all animals were observed daily for clinical signs for 15 days, with special attention given during the first 4 h, while their body weights were measured weekly. At study termination, animals were subjected to a necropsy including macroscopic examination.

The selection of negative and positive protein controls (melittin and aprotinin) was based on the need for well characterized and readily available pure proteins from commercial sources. The lethality of melittin (Sigma, negative and positive controls) is 3.1 mg protein/kg body weight (Schmidt, 1994). By selecting the doses of 1 and 10 mg/kg in the acute study design, no mortality was expected when we used the dose of 1 mg protein/kg and 100% mortality (within 10 min) was expected when we used the dose of 10 mg protein/kg. Aprotinin (Sigma, negative control) is known to be safe in humans and mice at very high doses. The dose level of 10 mg/kg body weight was higher than dose levels used in high dose efficacy studies with aprotinin, at 4.2 mg/kg body weight (Mossinger and Dietrich, 1998).

3. Results

3.1. History of safe use of the source of the gene

As defined in the publication by Constable et al. (2007), important factors in establishing a history of safe use for a plant include (1) the period over which the food has been consumed, (2) the way in which it has been prepared and used and at what intake levels, (3) the potential hazard associated to its consumption and, (4) the observations from animal and human exposure.

The wild-type *epsps* gene used to generate the *epsps* gene was isolated from maize (*Zea mays* L.). Therefore, an extensive literature review on maize was performed and is summarized in Table 1 (Jones et al., 1995; Halim et al., 1973; Maga, 1982; OECD, 2003). Of major importance for the toxicological evaluation, maize naturally contains insignificant levels of four anti-nutrient or toxic compounds: enzyme inhibitors, raffinose, phytic acid and DIMBOA (2,4-dihydroxy-7-methoxy-2H-1,4-benzoxazin-3 (4h)-one). In spite of the presence of these toxins, maize has been safely consumed by humans, domesticated animals, or wildlife for centuries and represents a large part of the dietary intake. It is also important to identify potential allergenic risk. The literature showed that maize is considered to be a very uncommon allergenic food (Lee et al., 2005; Moneret-Vautrin et al., 1998; Pasini et al., 2002; Pastorello et al., 2003) and known maize allergens are rare. Especially, the 2mEPSPS protein has no relation to the identified allergens.

More generally, maize has been regarded as beneficial and safe to eat by the animal and human populations for several generations, after different processing modifications or not. It can be easily concluded that maize has a long history of safe use. Then, it can be deemed safe as a source of the gene.

3.2. Characterization of the 2mEPSPS amino acid sequence

The 2mEPSPS protein does not exist in nature. In Fig. 1, the amino acid sequences of the wild-type maize EPSPS and 2mEPSPS proteins are reported. The 2mEPSPS protein differs from the wild-type maize EPSPS by only 2 amino acid substitutions that have been introduced into the modified enzyme, threonine by isoleucine at position 103 and proline by serine at position 107. The two enzymes are homodimers of 445 amino acids (molecular weight of 47 kDa by calculation) and are highly similar (>99% identity). The 2mEPSPS protein also shows high amino acid sequence similarities

Table 1
History of safe use of the source of the gene (*Zea mays*).

Well-known biology	*Zea* ssp. *mays* is a member of the *Maydeae* tribe of the grass family, *Poaceae*. Maize (corn) is the world's most widely grown cereal, reflecting its ability to adapt to a wide range of production environments
Well-known origin	Maize is indigenous to the Western Hemisphere and originated from the Mexican plateau or in the highlands of Guatemala. It is domesticated from teosinte by human selection
Large production	Worldwide production of maize is about 700 million tons a year. The United States, China, Brazil, and Mexico account for 70% of the global production. The US accounts for 81% of the production of sweet maize
Long length of use	Field maize has been grown for 8000 years in Mexico and Central America and for 500 years in Europe
High human exposure with safe records	In Latin America and Africa, maize is mainly used to feed humans. Globally, only 21% of total grain production is consumed as food. It is also a very suitable raw material for the manufacture of starch, and maize oil accounts for approximately 9% of domestic vegetable oil production
High animal exposure with safe records	In industrialised countries where maize is a major crop (grain and forage), it is also the principal component of livestock (ruminants) feeds. In Asia, it is also mainly used to feed animals
Very limited impact of endogenous anti-nutrients / toxins	Only four anti-nutrients exist in maize: phytic acid, enzyme inhibitors, raffinose, and 2,4-dihydroxy-7-methoxy-2H-1,4-benzoxazin-3 (4 h)-one (DIMBOA). However kernels have not been described to be normally toxic to humans, domesticated animals, or wildlife as the current varieties do not contain significant levels of endogenous toxins. Toxicity is only recorded, at certain levels of exposure to DIMBOA, for insects
No pathogenicity	No pathological records exist
Very few endogenous allergens and rare allergies	Maize does not belong to the major allergenic food list established in the US and in Europe. Only a few allergens among tens of thousand proteins are reported to cause rare clinical cases of allergies to maize, mainly respiratory. These are linked to common proteins present in grain cereals like rice, barley or wheat (*e.g.*, 9 kDa lipid transfer protein, 16 kDa trypsin inhibitor)

to other EPSPS from different crops. Sequence identities of the deduced amino acid sequence of 2mEPSPS to that of other EPSPS proteins found in nature are as follows:

Crops	Maize	Rice	Grape	Lettuce	Tomato	Canola
% Sequence identities	>99.5	86	79	77	75	75

3.3. Familiarity of the EPSPS enzyme family

It is expected that homologous proteins derived from a common ancestor have similar amino acid sequences, are structurally similar and share common function. Therefore, it is desirable to introduce a gene that encodes for a protein that is homologous to only proteins with safe records.

The family of EPSPS enzyme proteins is ubiquitous in plant, yeast and microbial food sources. It is widespread in nature. In addition, it is not present in animals, which do not synthesize their own aromatic amino acids. Based on literature review, this EPSPS enzyme family has never been recorded as pathogenic, toxic or allergenic. As a component of food and feed, the protein has never been associated with any health concerns. Thus, this family of wild-type EPSPS proteins can be deemed safe and forms the benchmark for the comparative safety assessment of the 2mEPSPS protein derived from transgenic crops.

3.4. Homology searches with known proteins

The 2009 overall homology BLAST results showed no evidence for any similarity between the 2mEPSPS protein and any known allergenic proteins. The epitope homology FindPatterns analysis confirmed that there was no evidence for any similarity to known allergens, i.e. no continuous 8 amino acids identity with known allergens, by using the AllergenOnline database.

The findings also showed that there was no similarity with known toxins based on the very high E-score value, which can identify functional homology, between the 2mEPSPS protein and known toxins.

As expected, the BLAST results showed that the 2mEPSPS protein only has high structural similarity with other wild-type EPSPS proteins, including the one from maize (>99%), for which no adverse effects have been reported following consumption. The high similarity of the wild-type EPSPS and the 2mEPSPS proteins, as described above, indicates a similar safe profile.

Furthermore, Fig. 1 shows two putative N-glycosylation sites, which are the most associated with allergenicity (Jenkins et al., 1996), at positions 118 and 394 on the 2mEPSPS sequence. However, these two putative sites were identical to those in the wild-type EPSPS protein (*i.e.*, not affected by the substitutions) (not shown). As the wild-type protein is safe to consume, it can be concluded that the two putative sites present no safety concerns.

3.5. Structural and functional equivalence of 2mEPSPS protein derived from E. coli and crops

The exact level of the 2mEPSPS protein in crops was evaluated. In GHB614 cotton, it was expressed at approximately 102 ± 2 mg 2mEPSPS protein/kg in delinted cottonseed. Because of the low expression level of the 2mEPSPS protein in the cotton, the following studies were conducted with 2mEPSPS protein produced in *E. coli*. In order to validate this approach, the structural and functional similarity between the bacterially produced protein and plant expressed protein was determined. The characteristics of the 2mEPSPS protein produced in *E. coli* were compared to those of the 2mEPSPS from cotton event GHB614 as described below. Incomplete data with soybean and canola were also generated and added into this manuscript as they corroborated the transformed cotton results.

3.5.1. Structural similarity

The theoretical N-terminal sequence of the 2mEPSPS protein deduced from the DNA gene sequence is: methionine, alanine, glycine, alanine, glutamic acid, glutamic acid, and isoleucine. For GHB614 cotton, the 2mEPSPS protein was isolated from leaves and the N-terminal sequence was determined by Edman degradation. The following primary sequence was obtained from the N-terminus: alanine, glycine, alanine, glutamic acid, and glutamic acid. This sequence is an exact match to the sequence deduced from the DNA sequence of the *2mepsps* gene for residues 2–6, confirming the identity of the protein isolated from GHB614 cotton.

The N-terminal methionine was not identified from the 2mEPSPS protein isolated from GHB614 cotton (data not shown). This could be due to post-translational modifications, as removal of a methionine is often found in proteins from both prokaryotic and eukaryotic organisms (Bradshaw et al., 1998). Therefore, this was not considered to be detrimental to the similarity approach used.

This was confirmed by Fig. 2A and B, which showed the same electrophoretic mobilities and same molecular weight of approximately 42 kDa for both proteins. This value is in line with the theoretical molecular weight of 47 kDa calculated from the amino acid sequence deduced from the DNA sequence. In addition, Fig. 2B also showed that the immunoreactivities of the two 2mEPSPS proteins were equivalent by western blotting.

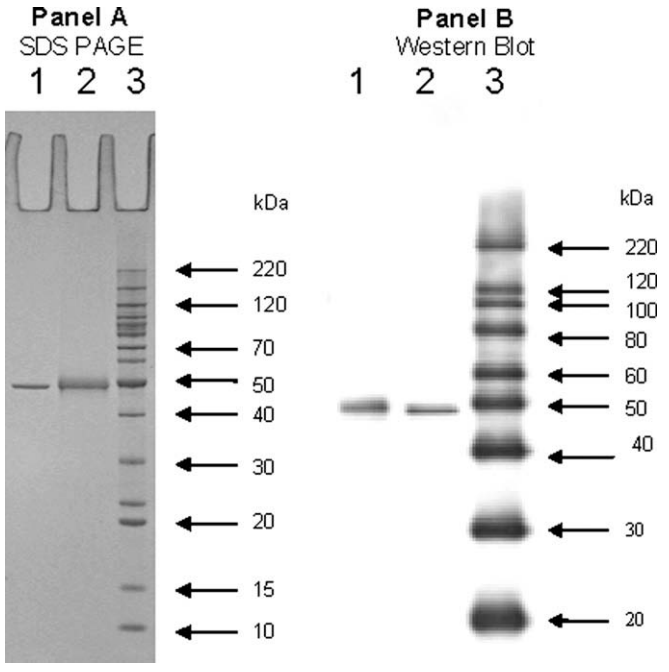

Fig. 2. Comparison of the 2mEPSPS protein from *E. coli* with the 2mEPSPS protein isolated from leaves of transgenic GHB614 cotton. (A) Shows the SDS–PAGE gel stained with Coomassie brilliant blue. Lane 1 contains approximately 300 ng of 2mEPSPS protein produced in *E. coli*. Lane 2 contains approximately 1200 ng 2mEPSPS protein from GHB614 cotton leaves. Lane 3 contains molecular weight markers of 220, 160, 120, 100, 90, 80, 70, 60, 50, 40, 30, 25, 20, 15, and 10°kDa. Only the underlined molecular weights are marked by arrows. (B) Shows a western blot. Lane 1 contains approximately 5 ng of 2mEPSPS protein from GHB614 cotton. Lane 2 contains approximately 5 ng of the 2mEPSPS protein from *E. coli*. Lane 3 contains molecular weight markers of 220, 120, 100, 80, 60, 50, 40, 30, and 20 kDa.

Among all these elements demonstrating the similarity of the two proteins from *E. coli* or from cotton event GHB614, the results of the high performance liquid chromatography (HPLC)–electrospray selected ion monitoring mass spectrometry peptide mapping data (Table 2) were the most compelling. Peptides with 1–3 amino acids, which are not indicative of a unique protein, were not analyzed (indicated as "NA*"). When the corresponding peptides from both proteins were detected, their calculated masses were identical. The absence of peptide 1–13 in the 2mEPSPS protein from GHB614 cotton was consistent with the Edman degradation data. Peptide 14–20 in the 2mEPSPS protein from GHB614 cotton was also not detected. Some possible explanations for non-detection of this peptide are: (1) partial digestion of the protein resulting in low levels of this peptide, (2) a matrix effect that prevents its detection, (3) a chemical modification occurring during protein purification or (4) a post-translational modification of the protein. This last hypothesis is unlikely for N-glycosylation since the protein is effectively targeted to the chloroplast where no post-translational modifications occur and this peptide does not contain any of the amino acid sequences necessary for N-glycosylation (see Table 2). By using the electrospray liquid chromatography/mass spectroscopy (LC/MS) method, the sequence coverage obtained for the 2mEPSPs protein produced from *E. coli* was 96% and the sequence coverage for the 2mEPSPs protein extracted from GHB614 cotton was 91.5%. This experiment demonstrates the structural identity of the two proteins over 91.5% of the amino acid sequence. Taken together with the N-terminal sequencing and the molecular weight (SDS–PAGE) data, it can be concluded that the two proteins are structurally undistinguishable.

Table 2
Masses of the tryptic peptides calculated from the selected ion monitored for the 2mEPSPS protein from *E. coli* and GHB614 cotton.

2mEPSPS residue number	Theoretical mass + charge [M+H]	Mass + charge [M+*n*H] of the ion selected for monitoring[a]	Calculated masses[b] of 2mEPSPS peptides from *E. coli*	Calculated masses[b] of 2mEPSPS peptides from GHB614 cotton leaf
1–13	1399.7	1399 [M+H]	1398	ND[d]
14–20	733.8	734 [M+H]	733	ND[d]
21–25	501.6	502 [M+H]	501	501
26–30	576.6	577 [M+H]	576	576
31–61	3342.9	1672 [M+2H]	3342	3342
62–71	1033	1033 [M+H]	1032	1032
72–74	289	289 [M+H]	288	288
75–75	175	NA[c]	NA[c]	NA[c]
76–84	790	790 [M+H]	789	789
85–91	805.9	807 [M+H]	806	806
92–106	1648.9	1648 [M+H]	1647	1647
107–128	2105	1053 [M+2H]	2104	2104
129–130	306	NA[c]	NA[c]	NA[c]
131–142	1296	1296 [M+H]	1295	1295
143–160	1907	954 [M+2H]	1906	1906
161–171	969	969 [M+H]	968	968
172–173	246	NA[c]	NA[c]	NA[c]
174–204	3219.8	1610[M+2H]	3218	3218
205–216	1435.8	1436 [M+H]	1435	1435
217–220	548.7	548 [M+H]	547	547
221–224	450.6	450 [M+H]	449	449
225–233	1103	1103 [M+H]	1102	1102
234–237	570.7	571 [M+H]	570	570
238–241	389	389 [M+H]	388	388
242–243	310	NA[c]	NA[c]	NA[c]
244–246	331	NA[c]	NA[c]	NA[c]
247–286	3870	1291 [M+3H]	3870	3870
287–297	1226	1226 [M+H]	1225	1225
298–312	1631.8	1631 [M+H]	1630	1630
313–317	605.6	606 [M+H]	605	605
318–318	147	NA[c]	NA[c]	NA[c]
319–321	397	NA[c]	NA[c]	NA[c]
322–329	905	905 [M+H]	904	904
330–351	2260.7	1131[M+2H]	2260	2260
352–357	733.8	734 [M+H]	733	733
358–359	246	NA[c]	NA[c]	NA[c]
360–363	534	534 [M+H]	533	533
364–368	589.8	589 [M+H]	588	588
369–373	591.7	592 [M+H]	591	591
374–392	2019	1010 [M+2H]	2018	2018
393–405	1533.6	1534 [M+H]	1533	1533
406–423	1882	1882 [M+H]	1881	1881
424–429	648.7	649 [M+H]	648	648
430–430	147	NA[c]	NA[c]	NA[c]
431–444	1679.9	1680 [M+H]	1679	1679
445–445	133	NA[c]	NA[c]	NA[c]

[a] The method for selected ion monitoring was developed using 2mEPSPS protein from *E. coli*. The method involved identification of the most abundant ion for each peptide and a determination of the retention time for each ion by HPLC.

[b] The mass spectrometer measures mass (m) divided by charge (z) with unit resolution. Mass (m) includes the mass of the peptide + the number of positive charges. For example, the ion detected for peptide 31–61 containing 2 protons was 1672. Thus its calculated mass is 2(1672) – 2 or 3342. The uncertainty in the calculated mass is obtained from the average mass determination uncertainty for the peptide using all the charge states detected in the full scan spectrum.

[c] NA, not analyzed. No attempt was made to detect digestion products of 1 or 2 amino acids. Some tripeptides were not analyzed due to instrument limitations.

[d] ND, not detected.

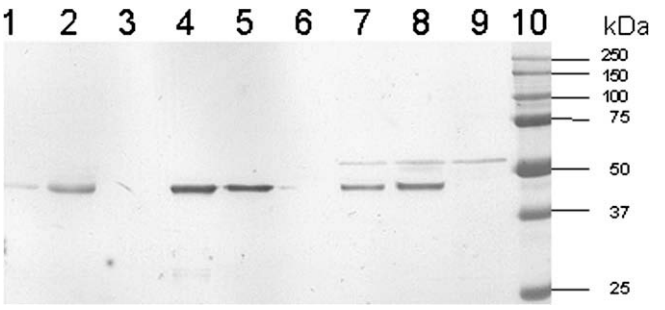

Fig. 3. Western blot of the 2mEPSPS protein in extracts of different plant species. Lanes 1 and 2 contain 2mEPSPS protein from extracts of GHB614 cotton. Lane 3 contains extract of non-transgenic cotton. Lanes 4–6 contain extracts of canola leaves from two transgenic events (lanes 4 and 5) and a non-transgenic plant (lane 6), respectively. Lanes 7–9 contain extracts of soybean leaves from two transgenic events and a non-transgenic plant, respectively. An additional band is visible in lanes 7–9. This band is resulting from a non-specific binding since the band appears in the wild type (lane 9). Lane 10 contains molecular weight markers of 250, 150, 100, 75, 50, 37, and 25 kDa.

The same approach was applied to other plant species transformed with the *2mepsps* gene and the same conclusion was raised. Fig. 3 shows that the electrophoretic mobilities and immunoreactivities of the 2mEPSPS proteins produced in GHB614 cotton, canola and soybean events are indistinguishable.

3.5.2. Functional similarity

The two 2mEPSPS proteins from *E. coli* and GHB614 cotton were both active enzymes (data not shown).

3.6. Digestibility assays

Fig. 4 shows that the 2mEPSPS protein was degraded very rapidly and completely in the SGF (pepsin, pH 1.2), within few seconds of incubation. The control proteins, horse radish peroxydase and ovalbumin, were rapidly and slowly digested, respectively, in SGF assay. As expected, control samples showed no significant protein degradation in the absence of pepsin (data not shown). These findings are in agreement with published data (Thomas et al., 2004).

Fig. 5 shows that, at time zero of incubation with SIF, the pancreatin bands at approximately 31 kDa were clearly visible. The azoalbumin control was digested by the pancreatin as expected (data not shown). At time zero of incubation, the protein degradation was so rapid (a few seconds) that the 2mEPSPS protein band was barely visible on the gel and it was not clearly visible after scanning the gel. The 2mEPSPS protein was completely degraded in the SIF (pancreatin, pH 7.5), within a few seconds.

3.7. Acute toxicology studies

3.7.1. Oral route of exposure

Control female OF1 mice treated with 2000 mg/kg of bovine serum albumin (negative control) showed no clinical signs of systemic toxicity as expected. Similarly, there were also no 2mEPSPS protein-related mortalities, no clinical signs or 2mEPSPS protein-related effects on body weight in female OF1 mice after an acute oral administration of the 2mEPSPS protein at 2000 mg/kg. Therefore, this study confirmed that the 2mEPSPS protein is not acutely toxic to mice when administered *per os* at 2000 mg/kg body weight.

3.7.2. Acute intravenous administration

The protein controls (melittin and aprotinin) demonstrated that the acute intravenous study was able to distinguish between toxic and non-toxic proteins. Control female OF1 mice treated with melittin at 1 mg/kg or with aprotinin at 1 and 10 mg/kg (negative controls) showed no visible signs of systemic toxicity. In contrast, melittin at 10 mg/kg (positive control) induced 100% mortality within 10 min as expected.

In this study, there were no mortalities or toxic effect in mice after acute intravenous administration of the 2mEPSPS protein at 1 or 10 mg/kg. Therefore, this second acute toxicology study confirmed that the 2mEPSPS protein is not acutely toxic when injected intravenously up to 10 mg/kg body weight, the highest dose that could be administered by the parenteral route of exposure.

4. Discussion

All the results of the analytical tests offer a multi-directional approach to demonstrate structural and functional equivalence of the 2mEPSPS proteins produced in *E. coli* and GHB614 transformed cotton. From these comparisons, it is reasonable to conclude that the 2mEPSPS protein produced in *E. coli* is representative of the 2mEPSPS protein produced in GHB614 cotton. Therefore, the safety data obtained for the 2mEPSPS protein produced in *E. coli* can be extrapolated to support the safety of the 2mEPSPS protein produced in cotton event GHB614.

The approach used for assessing the safety of 2mEPSPS protein largely complied with the ILSI and very recent regulatory recommendations (Delaney et al., 2008; Thomas et al., 2008; Paoletti et al., 2008). This is a weight of evidence approach, as outlined by the *Codex* recommendations Alinorm 03/34A (CAC, 2003b) and other regulatory agencies worldwide (*e.g.*, EFSA, 2006; FDA, 1992; Health Canada, 2001; etc.), recognizing that no single endpoint is sufficiently predictive of the allergenic potential of a novel protein. This is also a tiered approach as specific findings may trigger additional studies. In the TierI (hazard identification) step, this approach evaluates whether the gene is isolated from a safe source, the sequence similarity of the protein to known proteins, including those that are known to be allergenic or toxic, or those which have a familiar mode of action and a safe profile, the biological function of the protein and more generally its physico-biochemical properties and, the stability of the protein to human simulated digestive fluids. TierII (Hazard Characterization) studies are conducted when the results from TierI are not sufficient to allow a reasonable certainty of no harm and are then conducted on a case-by-case basis (Delaney et al., 2008).

Briefly, the *2mepsps* coding gene is obtained from a source with a history of safe use (maize). The 2mEPSPS is highly – structurally and functionally – similar to wild-type EPSPS found in nature, and possesses the same mode of action, allowing sufficient enzyme activity for the plants to grow in the presence of glyphosate herbicide.

Of major importance, the 2mEPSPS protein does not have amino acid sequence similarity to proteins that are known to be allergenic although the homology approach used is very conservative. Indeed sequence searches for matches of >50% identity (rather than >35%) over 80 or greater amino acids may be more realistic (Goodman et al., 2008; Thomas et al., 2008). In addition, the epitope search, an approach known to result in random false positive matches (Goodman et al., 2008), revealed no sequence identity with currently identified allergens, providing further evidence that the 2mEPSPS does not possess known allergenic potential.

Furthermore, although limited by the lack of toxin databases and the clear criteria in bioinformatics search of proteins for toxic potential (ILSI, 2009), the current analysis indicate that 2mEPSPS protein does not have amino acid sequence similarity to proteins that are known to be toxic. Instead, there is a high similarity between the 2mEPSPS proteins and other EPSPS proteins from different crops (maize, rice, lettuce, etc.), with a safe profile.

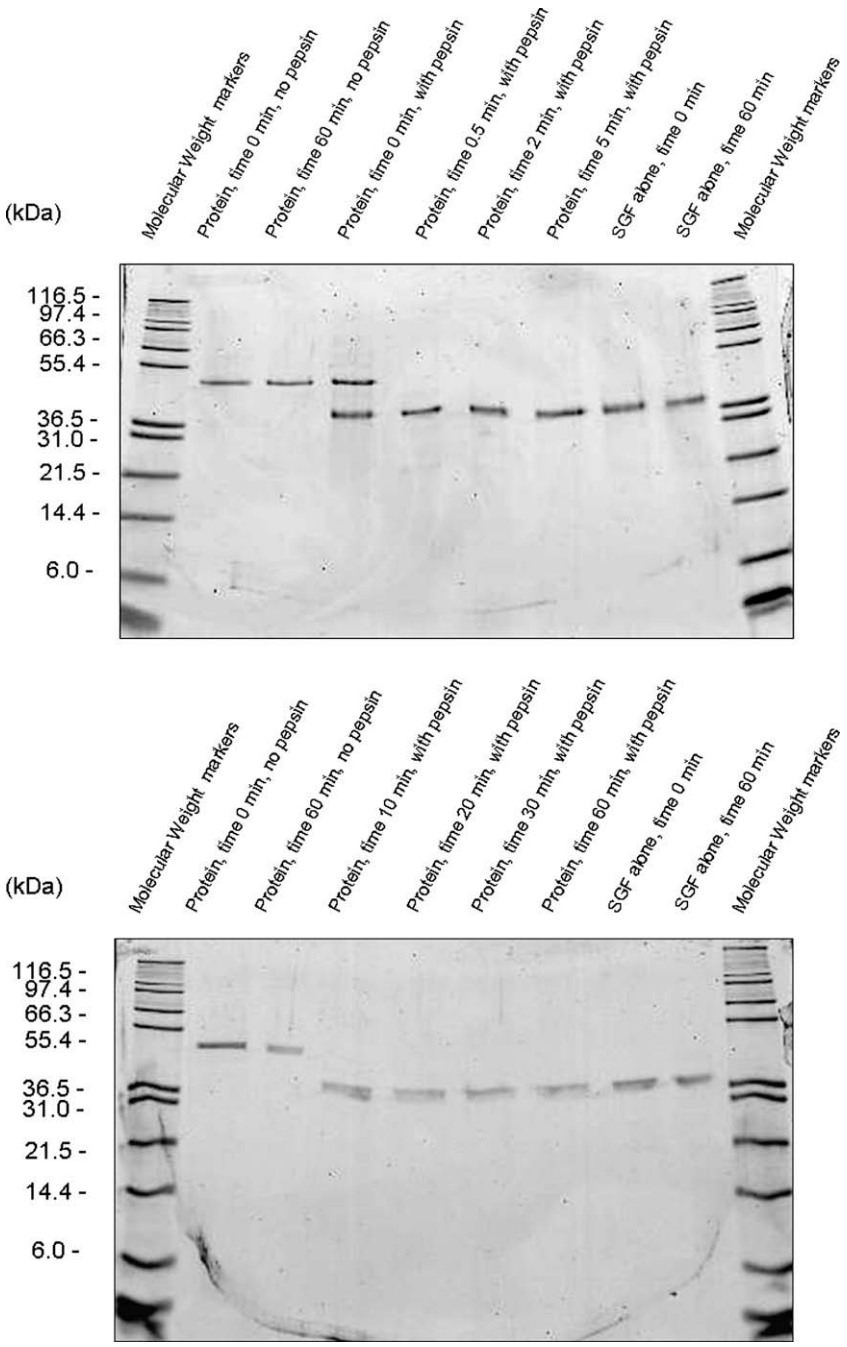

Fig. 4. Two SDS–PAGE analyses of the 2mEPSPS proteins after digestion in human simulated gastric fluid (SGF, pepsin, pH 1.2), at different time points. min = minutes; kDa = kDalton(s).

Furthermore, probable absence of such toxic or allergenic effects in humans and other mammals is supported by the rapid degradation of the protein in digestibility assays, implying that the protein will not be available for absorption in the lower digestive tract (Taylor and Lehrer, 1996; Astwood, 1996).

Of interest, many protein allergens are glycosylated, raising the possibility that the glycosyl groups may contribute to their allergenicity (Jenkins et al., 1996). As the two putative N-glycosylation sites are similar between the maize wild-type and the modified EPSPS, no impact on the safety profile of the EPSPS protein is expected. More generally, glycosylation of proteins occur in the endoplasmic reticulum (ER). Proteins that are not processed via the ER cannot be glycosylated, even if they contain N-glycosylation sites. It is the case for EPSPS and 2mEPSPS, as these proteins are transported into the chloroplast, where no glycosylation can occur (Lerouge et al., 1998). The mass spectrometry analysis of the 2mEPSPS protein confirms experimentally the absence of N-glycosylation in *E. coli* bacteria and transformed GHB614 cotton.

Although no concerns were raised in the TierI step, two acute mouse studies with two different routes of exposure were performed for strengthening the safety evaluation. Few proteins are toxic when ingested (Delaney et al., 2008). When a protein is toxic, it generally acts via acute mechanisms and at very low dose levels (Jones and Maryanski, 1991; Sjoblad et al., 1992). Moreover, whether a protein is toxic also depend on its property, the route of exposure and the dose received. The acute oral and intravenous

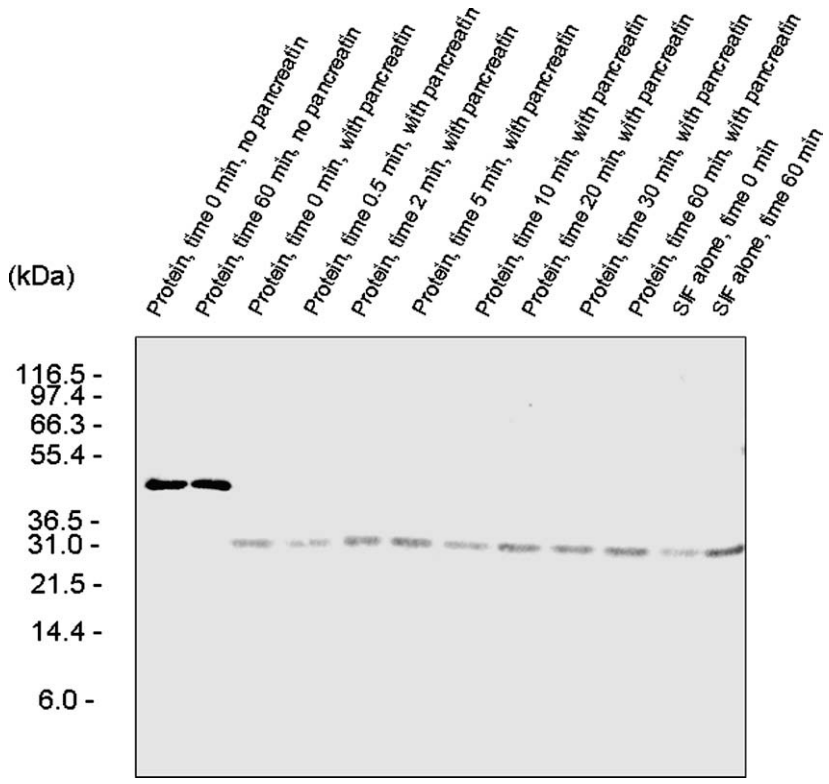

Fig. 5. Western blot analysis of the 2mEPSPS proteins after digestion in human simulated intestinal fluid (SIF, pancreatin, pH 7.5). Molecular weight markers are indicated on the left border of the gel.

toxicity studies did not identify evidence for potential toxicity of the 2mEPSPS protein. Accordingly, the oral No Observed Effect Level (NOEL) was determined to be greater than 2000 mg/kg/day.

Taking the transgenic GHB614 cotton as an example and using the existing "nuts and oilseeds" consumption databases (FAOSTAT, 2007), which list the raw agricultural commodity cottonseed, the maximum predicted daily intake of the 2mEPSPS protein is considered to be approximately 50 μg 2mEPSPS protein/day/kg body weight in humans. Even in the worst case scenario assuming all nuts and oilseeds were derived from GHB614 cotton, this would correspond to a minimal safety factor of 40,000. Given that the exposure to the 2mEPSPS protein will be nil in most food products as a result of plant processing (Thomas et al., 2007; *i.e.* limited presence of cotton proteins in food), this safety factor is largely underestimated for cotton.

5. Conclusion

Risk is defined as hazard (or inherent toxicity) of a product and exposure to the product. By using the step-wise, evidence-based and weight of evidence approach, the hazard is near zero in the case of the 2mEPSPS protein. *Zea mays*, source of the gene, has a good history of safe use. The 2mEPSPS protein is not toxic for mammals and do not possess any of the characteristics associated with food allergens. Therefore, it can be concluded that there is no expected risk associated with 2mEPSPS protein consumption.

Acknowledgments

The authors would like to thank Agnès Arnaud, Nadine Bautsoens, Chuck Haas, Bénédicte Labory-Carcenac, Stéphane Lebas, Dimitri Tack, Robert-Jan van der Klis, and Ann Vanhooren for their excellent technical assistance. The authors would like to extend their gratitude to Dr. Bernhard Weingaertner and his team for the assistance in the production of the 2mEPSPS protein. They also thank Dr. Regina Oberdörfer for her scientific support with regard to the nutritional data used for calculating the safety factor.

References

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.

Astwood, J.D., 1996. Stability of food allergens to digestion *in vitro*. Nat. Biotech. 14, 1269–1273.

Bradshaw, R.A., Brickey, W.W., Walker, K.W., 1998. N-terminal processing: the methionine aminopeptidase and N-acetyl transferase families. Trends Biochem. Sci. 23, 263–267.

Codex Alimentarius Commission (CAC), 2003a. Joint FAO/WHO Food Standards Programme, Food and Agriculture Organization. *Codex* Principles and Guidelines on Foods Derived from Biotechnology. 26th session, 30 June–7 July 2003, Rome, Italy.

Codex Alimentarius Commission (CAC), 2003b. Report on the 4th Session of the *Codex* Ad Hoc Intergovernmental Task Force on Foods Derived from Biotechnology. 11–14 March 2003, ALINORM 03/34A, Yokohama, Japan, pp. 1–56.

Constable, A., Jonas, D., Cockburn, A., Davi, A., Edwards, G., Hepburn, P., Herouet-Guicheney, C., Knowles, M., Moseley, B., Oberdörfer, R., 2007. History of safe use as applied to the safety assessment of novel foods and foods derived from genetically modified organisms. Food Chem. Toxicol. 45, 2513–2525.

Delaney, B., Astwood, J.D., Cunny, H., Eichen-Conn, R., Herouet-Guicheney, C., MacIntosh, S., Meyer, L.S., Privalle, L., Gao, Y., Mattsson, J., Levine, M., 2008. Evaluation of protein safety in the context of agricultural biotechnology. ILSI international food biotechnology committee task force on protein safety. Food Chem. Toxicol. 46, S71–S97.

Della-Cioppa, G., Bauer, S.C., Klein, B.K., Shah, D.M., Fraley, R.T., Kishore, G.M., 1986. Translocation of the precursor of 5-enol-pyruvylshikimate-3-phosphate synthase into chloroplasts of higher plants *in vitro*. Proc. Natl. Acad. Sci. USA 83, 6873–6877.

Dirks, C.G., Pedersen, M.H., Platzer, M.H., Bindslev-Jensen, C., Skov, P.S., Poulsen, L.K., 2005. Does absorption across the buccal mucosa explain early onset of food-induced allergic systemic reactions? J. Allergy Clin. Immunol. 115, 1321–1323.

European Food Safety Agency (EFSA), 2006. Guidance document of the Scientific Panel on Genetically Modified Organisms for the risk assessment of genetically modified plants and derived food and feed. EFSA J. 374, 1–115.

Edman, P., 1964. Determination of amino acid sequences in proteins. Thromb. Diath. Haemorrhag. Suppl. 13, 17–20.
Evans, D.F., Pye, G., Bramley, R., Clark, A.G., Dyson, T.J., Hardcastle, J.D., 1988. Measurement of gastrointestinal pH profiles in normal ambulant human subjects. Gut 29, 1035–1041.
FAOSTAT, 2007. FAOSTAT database. <http://faostat.fao.org/> (accessed 02.03.07.).
FAO/WHO (Food and Agriculture Organization of the United Nations/World Health Organization), 2001. Joint FAO/WHO Expert Consultation on Foods Derived from Biotechnology – Evaluation of Allergenicity of Genetically Modified Foods, 22–25 January 2001, Rome, Italy, pp. 1–27.
FDA, 1992. Statement of policy: foods derived from new plant varieties. Fed. Regist. 57, 22984.
Forlani, G., Parisi, B., Nielsen, E., 1994. 5-enol-pyruvyl-shikimate-3-phosphate synthase from *Zea mays* cultured cells. Plant Physiol. 105, 1107–1114.
Fu, T.J., Abbott, U., Hatzos, C., 2002. Digestibility of food allergens and non-allergenic proteins in simulated gastric and intestinal fluids – a comparative study. J. Agric. Food Chem. 50, 7154–7160.
Goodman, R.E., 2006. Practical and predictive bioinformatics methods for the identification of potentially cross-reactive protein matches. Mol. Nutr. Food Res. 50, 655–660.
Goodman, R.E., Vieths, S., Sampson, H.A., Hill, D., Ebisawa, M., Taylor, S.L., van Ree, R., 2008. Allergenicity assessment of genetically modified crops – what makes sense? Nat. Biotech. 26, 73–81.
Halim, A.H., Wassom, C.E., Mitchell, H.L., 1973. Trypsin inhibitor in corn (*Zea mays* L.) as influenced by genotype and moisture stress. Crop Sci. 13, 405–406.
Health Canada, 2001. Elements of precaution: recommendations for the regulation of food biotechnology in Canada. An Expert Panel Report on the Future of Food Biotechnology Prepared by the Royal Society of Canada at the Request of Health Canada, Canadian Food Inspection Agency and Environment Canada. January 2001, Ottawa, Canada, pp. 1–245.
Henikoff, S., Henikoff, J.G., 1992. Amino acid substitution matrices from protein blocks. Proc. Natl. Acad. Sci. USA 89, 10915–10919.
Hewick, R.M., Hunkapiller, M.W., Hood, L.E., Dreyer, W.J., 1981. A gas-liquid solid phase peptide and protein sequenator. J. Biol. Chem. 256, 7990–7997.
Hileman, R.E., Silvanovich, A., Goodman, R.E., Rice, E.A., Holleschak, G., Astwood, J.D., Hefle, S.L., 2002. Bioinformatic methods for allergenicity assessment using a comprehensive allergen database. Int. Arch. Allergy Immunol. 128, 280–291.
International Life Science Institute (ILSI), Health and Environmental Sciences Institute (HESI) Protein Allergenicity Technical Committee (PATC), 2009. Comments to the FIFRA Scientific Advisory Panel of the US Environmental Protection Agency for its February 25–27, 2009, meeting on plant-incorporated protectants (PIPs) test guidelines OPPTS 890.2150 Bioinformatics analysis of protein amino acid sequences (Docket ID, No EPA-HQ-OPP-2008-0835).
Jenkins, N., Parekh, R.B., James, D.C., 1996. Getting the glycosylation right: implications for the biotechnology industry. Nat. Biotechnol. 14, 975–981.
Jones, D.D., Maryanski, J.H., 1991. Safety considerations in the evaluation of transgenic plants for human foods. In: Levin, M.A., Strauss, H.S. (Eds.), Risk Assessment in Genetic Engineering. McGraw-Hill, New York, pp. 64–82.
Jones, S.M., Magnolfi, C.F., Cooke, S.K., Sampson, H.A., 1995. Immunologic cross-reactivity among cereal grains and grasses in children with food hypersensitivity. J. Allergy Clin. Immunol. 96, 341–351.
Kenna, J.G., Evans, R.M., 2000. Digestibility of proteins in simulated gastric fluid. The Toxicologist, 54, Abstract 666.
Kleter, G.A., Peijnenburg, A.A., 2002. Screening of transgenic proteins expressed in transgenic food crops for the presence of short amino acid sequences identical to potential, IgE – binding linear epitopes of allergens. BMC Struct. Biol. 2, 8. Available from: <http://www.biomedcentral.com/1472-6807/2/8>.
Laemmli, U.K., 1970. Cleavage of structural proteins during assembly of the head of the bacteriophage T4. Nature 227, 680–685.
Lebrun, M., Sailland, A., Freyssinet, G. 1997a. Mutated 5-enol pyruvylshikimate-3-phosphate synthase, gene encoding for said protein and transformed plants containing said gene. International patent publication W0 97/04103-A2. 06 February 97. 25p.
Lebrun, M., Leroux, B., Sailland, A., 1997b. Chimeric gene for the transformation of plants. US patent number 5,633,448.
Lerouge, P., Cabanes-Macheteau, M., Rayon, C., Fischette-Lainé, A.C., Gomord, V., Faye, L., 1998. N-Glycoprotein biosynthesis in plants: recent developments and future trends. Plant Mol. Biol. 38, 31–48.
Lindahl, A., Ungell, A.L., Knutson, L., Lennernas, H., 1997. Characterization of fluids from the stomach and proximal jejunum in men and women. Pharm. Res. 14, 497–502.
Lanzetta, P.A., Alvarez, L.J., Reinach, P.S., Candia, O.A., 1979. An improved assay for nanomole amounts of inorganic phosphate. Anal. Biochem. 100, 95–97.
Lee, S.H., Benmoussa, M., Sathe, S.K., Roux, K.H., Teuber, S.S., Hamaker, B.R., 2005. A 50 kDa maize gamma-zein has marked cross-reactivity with the almond major protein. J. Agric. Food Chem. 53, 7965–7970.
Maga, J.A., 1982. Phytate: its chemistry, occurrence, food interactions, nutritional significance, and methods of analysis. J. Agric. Food Chem. 30, 1–7.
Moneret-Vautrin, D.A., Kanny, G., Beaudouin, E., 1998. L'allergie alimentaire au maïs existe-t-elle? Allergie et immunologie 30, 230.
Mossinger, H., Dietrich, W., 1998. Activation of hemostasis during cardiopulmonary bypass and pediatric aprotinin dosage. Ann. Thorac. Surg. 65, S45–S50.
Neuhoff, V., Arold, N., Taube, D., Ehrhardt, W., 1988. Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using Coomassie Brilliant Blue G-250 and R-250. Electrophoresis 9, 255–262.
OECD, 2001. Test no. 425: acute oral toxicity – up and down procedure. In: OECD Guidelines for the Testing of Chemicals. 17 December 2001, pp. 1–14.
OECD, 2003. Environment, Health and Safety Publications. Series on Harmonisation of Regulatory Oversight in Biotechnology No. 27. Consensus Document on the Biology of *Zea mays subsp. mays* (Maize). ENV/JM/MONO(2003)11. July 23, 2003, pp 1–49.
Okunuki, H., Teshima, R., Shigeta, T., Sakushima, J., Akiyama, H., Goda, Y., Toyoda, M., Sawada, J., 2002. Increased digestibility of two products in genetically modified food (CP4-EPSPS and Cry1Ab) after preheating. J. Food Hyg. Soc. Jpn. 43, 68–73.
Paoletti, C., Flamm, E., Yan, W., Meek, S., Renckens, S., Fellous, M., Kuiper, H., 2008. GMO risk assessment around the world: some examples. Trends Food Technol. 19, S66–S74.
Pasini, G., Limonato, B., Curioni, A., Vincenti, S., Cristaudo, A., Cantucci, B., Dal BelinPeruffo, A., Giannattasio, M., 2002. IgE-mediated allergy to corn: a 50 kDa protein, belonging to the Reduced Soluble Proteins, is a major allergen. Allergy 57, 98–106.
Pastorello, E.A., Pompei, C., Pravettoni, V., Farioli, L., Calamari, A.M., Scibilia, J., Robino, A.M., Iametti, S., Fortunato, D., Bomnomi, S., Ortolani, C., 2003. Lipid-transfer protein is the major maize allergen maintaining IgE-binding activity after cooking at 100 degrees C, as demonstrated in anaphylactic patients and patients with positive double-blind, placebo-controlled food challenge results. J. Allergy Clin. Immunol. 112, 775–783.
Russell, T.L., Berardi, R.R., Barnett, J.L., Dermentzoglou, L.C., Jarvenpaa, K.M., Schmaltz, S.P., Schmidt, J.O., 1994. Toxinology of venoms from the honeybee genus *apis*. Toxicon 33, 917–927.
Schmidt, J.O., 1994. Toxinology of venoms from the honeybee genus *apis*. Toxicon. 33, 917–927.
Sjoblad, R., McClintock, J.T., Engler, R., 1992. Toxicological considerations for protein components of biological pesticide products. Regul. Toxicol. Pharmacol. 15, 3–9.
Stadler, M.B., Stadler, B.M., 2003. Allergenicity prediction by protein sequence. FASEB J. 17, 1141–1143.
Taylor, S., Lehrer, S.B., 1996. Principles and characteristics of food allergens. Crit. Rev. Food Sci. Nut. 36, S91–S118.
Thomas, K., Aalbers, M., Bannon, G.A., Bartels, M., Dearman, R.J., Esdaile, D.J., Fu, T.J., Glatt, C.M., Hadfield, N., Hatzos, C., Hefle, S.L., Heylings, J.R., Goodman, R.E., Henry, B., Herouet, C., Holsapple, M., Ladics, G.S., Landry, T.D., MacIntosh, S.C., Rice, E.A., Privalle, L.S., Steiner, H.Y., Teshima, R., Thomas, K., van Ree, R., Woolhiser, M., Zawodny, J., 2004. A multi-laboratory evaluation of a common *in vitro* pepsin digestion assay protocol used in assessing the safety of novel proteins. Regul. Toxicol. Pharmacol. 39, 87–98.
Thomas, K., Bannon, G., Hefle, S., Herouet, C., Holsapple, M., Ladics, G., MacIntosh, S., Privalle, L., 2005. *In silico* methods for evaluating human allergenicity to novel proteins: International Bioinformatics Workshop Meeting Report, 23–24 February 2005. Toxicol. Sci. 88, 307–310.
Thomas, K., Herouet-Guicheney, C., Ladics, G., Bannon, G., Cockburn, A., Crevel, R., Fitzpatrick, J., Mills, C., Privalle, L., Vieths, S., 2007. Evaluating the effect of food processing on the potential human allergenicity of novel proteins: international workshop report. Food Chem. Toxicol. 45, 1116–1122.
Thomas, K., Herouet-Guicheney, C., Ladics, G., McClain, S., MacIntosh, S., Privalle, L., Woolhiser, M., 2008. Current and future methods for evaluating the allergenic potential of proteins: international workshop report 23-25 October 2007. Food Chem. Toxicol. 46, 3219–3225.
Untersmayr, E., Bakos, N., Scholl, I., Kundi, M., Roth-Walter, F., Szalai, K., Riemer, A.B., Ankersmit, H.J., Scheiner, O., Boltz-Niculescu, G., Jensen-Jarolim, E., 2005. Anti-ulcer drugs promote IgE formation towards dietary antigens in adult patients. FASEB J. 19, 632–656.