MONSANTO

APPLICANT: **Monsanto Australia Limited**

A416
CP4 ESPS gene in Roundup Ready® Corn Line NK603

SUBMISSION: Application to Australia New Zealand Food Authority for the inclusion of corn containing the CP4 EPSPS gene by Monsanto in Standard A18 - Food Derived From Gene Technology

VOLUME: 1 of 1

SUPPORTING INFORMATION

DATE: 30 January 2002

PREPARED BY: Megan Shaw
Regulatory Product Manager

Title

Safety Assessment of Roundup Ready Corn Event NK603 Containing Genes Encoding CP4 EPSPS and CP4 EPSPS L214P

Authors

James D. Astwood, Cherian George, Murtaza Alibhai, Roderick McCoy, Linda Lahman, Bruce G. Hammond, John N. Leach & Andre Silvanovich

Final On

December 21, 2001

Performing Laboratories

Monsanto Company
Product Safety Center
800 North Lindbergh Blvd
St. Louis, MO 63197

Laboratory Project ID

MSL Number: 17600

[redacted]

Statement of No Data Confidentiality

No claim of confidentiality is made for any information contained in this study on the basis of its falling within the scope of FIFRA §10 (d),(1) (A), (B) or (C).

"We submit this material to the United States Environmental Protection Agency specifically under the requirements set forth in FIFRA as amended, and consent to the use and disclosure of this material by EPA strictly in accordance with FIFRA. By submitting this material to EPA in accordance with the method and format requirements contained in PR Notice 86-5, we reserve and do not waive any rights involving this material that are or can be claimed by the company notwithstanding this submission to EPA."

Company: Monsanto Company

Company Agent: _____

Title: _____

Signature: _____ Date: _____

Statement of Compliance

This report does not fall under the scope of the Good Laboratory Practice Standards as specified in 40 CFR Part 160 because it is not a study as defined in this part.

Submitter _____ Date: _____

Sponsor Lindell K. Lehman Date: 21 December 2001

Author _____ Date: Dec 21, 2001

MSL Number: **17600**

Title: **Safety Assessment of Roundup Ready Corn Event NK603 Containing Genes Encoding CP4 EPSPS and CP4 EPSPS L214P.**

Facility:
Monsanto Company
Product Safety Center
800 North Lindbergh Blvd
St. Louis, Missouri 63197

Authors:
James D. Astwood, *Ph.D.*
Cherian George, *Ph.D.*
Murtaza Alibhai, *Ph.D.*
Roderick McCoy, *M.D., Ph.D.*
Linda Lahman, *M.S.*
Bruce G. Hammond, *Ph.D.*
John N. Leach, *B.A.*
Andre Silvanovich, *Ph.D.*

Sponsor: Monsanto Company

Completion Date: December 21, 2001

Records Retention: Report related supporting documentation and final report will be retained at Monsanto-St. Louis.

Signatures of Final Report Approval:

James D. Astwood, *Ph.D.* Date: Dec. 21, 2001

Cherian George, *Ph.D.* Date: Dec. 21, 2001

Sponsor Representative Date: 21 December 2001

Summary of Quality Control Review

This report was reviewed to ensure that it accurately reflects the supporting documentation. The supporting documentation was audited for compliance to the Monsanto Company Guidelines for Keeping Research Records (GRR 10/1/99), and where applicable, to Monsanto SOP's.

Joan Rejda-Heath

Date: Dec. 21, 2001

Joan Rejda-Heath, *Ph.D.*
Quality Assurance Specialist
Monsanto Company

Table of Contents

List of Tables

List of Figures

Abbreviations

aa	amino acid
CFR	Code of Federal Regulations
CP4	A strain of *Agrobacterium tumefaciens*
CP4 EPSPS	EPSPS from *Agrobacterium tumefaciens ssp.* strain CP4
DTT	Dithiothreitol
EPSPS	5-*enol*pyruvylshikimate-3-phosphate synthase
E. coli	*Eschericia coli*
FIFRA	Federal Insecticide, Fungicide and Rodenticide Act
GLP	Good Laboratory Practice
kDa	kilodalton
mA	milliampere
mEPSPS	modified EPSPS from *Zea mays*
mM	millimolar
N0	Digestibility control without protein and zero incubation time
N9	Digestibility control without protein and incubated for as long as T=9
NOEL	No observable effect level
P0	Digestibility control without pepsin and zero incubation time
P9	Digestibility control without pepsin and incubated for as long as T=9
PAE	EPSPS protein from *Pseudomonas aeruginosa*
PAGE	Polyacrylamide gel electrophoresis
P/N	Product number, same as catalog number
RR	Roundup Ready
SDS	Sodium dodecyl sulfate
SGF	Simulated gastric fluid
SOP	Standard operating procedure
T	Time point
TCA	Trichloroacetic acid
Tricine	N-[tris(hydroxymethyl)methyl]glycine
Tris	Tris(hydroxymethyl)aminomethane
v/v	solute volume to solution volume
w/o	without
w/v	solute weight to solution volume

① natural mutation

② X-ray crystal structure modelling shows no 2° or 3° changes structural

③ same enzymatic activity

④ kinetic [illegible] region ?

⑤ HPLC protein heterologous in Hist region

1.0 Summary

Monsanto Company has developed Roundup Ready® corn event NK603 which is tolerant to glyphosate (the active ingredient in Roundup® herbicide) at the whole plant level. The inserted DNA in corn event NK603 has recently been sequenced. Roundup Ready corn event NK603 contains two *cp4 epsps* genes which differ by two nucleotides, one of which results in an amino acid substitution of proline for leucine at amino acid position 214 (L214P) in the encoded protein. Both *cp4 epsps* genes were present in the first (F1) and the second (BC1F1) generation after transformation.

The purpose of this document is to describe the structural and functional relationship between the CP4 EPSPS proteins encoded by these two genes in the context of food, feed and environmental safety of this product. It was shown that CP4 EPSPS and CP4 EPSPS L214P proteins are structurally and functionally equivalent. This conclusion was based on the demonstration that (1) proline substitutions naturally occur near position 214 in extant EPSPS proteins, (2) modeling using the known X-ray crystal structure of CP4 EPSPS showed that the L214P substitution does not alter the predicted secondary and tertiary structure of CP4 EPSPS, (3) equivalent EPSPS activity was observed for CP4 EPSPS and CP4 EPSPS L214P, (4) the variable loop region containing the proline substitution is not relevant to the enzymatic activity of EPSPSs generally, and (5) the CP4 EPSPS protein domain containing the proline substitution is highly heterogeneous in all known EPSPS proteins.

The EPSPS family of proteins is ubiquitous in plants, fungi and certain microbes; and is essential for the production of aromatic amino acids. On the basis of the history of use of EPSPS family of proteins in the food supply, the range of amino acid sequences present in this family of proteins and confirmatory safety data on the CP4 EPSPS protein, it was previously concluded that CP4 EPSPS protein poses no safety concern from a food, feed or environmental perspective. This conclusion has been confirmed by regulatory authorities around the world and by the commercial production of crop products containing the CP4 EPSPS protein on over 300 million acres over the past five years. CP4 EPSPS L214P was demonstrated to have equivalent functional activity to CP4 EPSPS, lack amino acid sequence similarity to toxins and allergens, and to be rapidly digested *in vitro*, as expected. Based on these data, CP4 EPSPS L214P was concluded to be structurally and functionally equivalent to the CP4 EPSPS protein and is safe for human and animal consumption.

Monsanto has conducted extensive agronomic and environmental safety studies with the Roundup Ready® corn event NK603. Comparisons of disease and pest susceptibilities, yield, morphology, weediness, outcrossing and other relevant characteristics confirmed that this event does not pose any increased environmental risks compared to traditional corn.

® Roundup and Roundup Ready are registered trademarks of Monsanto Technology LLC; St. Louis, Missouri

Finally, evaluations of equivalence in composition, animal nutrition, toxicity, and agronomic and environmental parameters between Roundup Ready corn event NK603 and traditional corn hybrids demonstrate that (1) the presence of the CP4 EPSPS L214P does not alter the composition of corn, (2) there are no biologically relevant pleiotropic effects due to the CP4 EPSPS L214P, (3) corn grain containing the NK603 insert is as safe and nutritious as traditional corn hybrids and (4) Roundup Ready corn event NK603 containing CP4 EPSPS and CP4 EPSPS L214P proteins does not present an increased risk to the environment compared to traditional corn.

It was concluded that CP4 EPSPS L214P is indistinguishable from CP4 EPSPS with respect to function, food, feed and environmental safety.

2.0 Introduction

Among the herbicidal active ingredients used in agriculture, glyphosate (*N*-phosphonomethyl glycine), commercialized under the trade name Roundup® herbicide, has been used globally as a safe and effective means of weed control (Baird *et al.,* 1971). Studies have shown that glyphosate inhibits the enzyme 5-*enol*pyruvylshikimate-3-phosphate synthase (EPSPS), a key enzyme in the shikimate pathway (Franz *et al.,* 1997). The EPSPS reaction is the penultimate step in the shikimic acid pathway for the biosynthesis of aromatic amino acids (phenylalanine, tyrosine, tryptophan). The shikimic acid pathway is only present in plants and microorganisms and is absent in mammals, fish, birds, reptiles, and insects and thus is a good target for novel herbicides (Alibhai and Stallings, 2001).

The *cp4 epsps* gene from *Agrobacterium ssp.* strain CP4 has been sequenced and encodes a 47.6 kDa CP4 EPSPS protein consisting of a single polypeptide of 455 amino acids (Padgette *et al.,* 1993). The CP4 EPSPS protein is functionally similar to plant EPSPS enzymes but has a much reduced affinity for glyphosate (Padgette *et al.,* 1993). In conventionally bred plants, glyphosate binds to the plant EPSPS enzyme and blocks the biosynthesis of aromatic amino acids, thereby starving plants of these essential nutrients (Steinrucken and Amrhein, 1980; Haslam, 1993). In Roundup Ready plants, which have been genetically modified to be tolerant to Roundup herbicide, the nutritional requirements for normal growth and development are met by the continued action of the glyphosate-tolerant CP4 EPSPS enzyme in the presence of glyphosate (Padgette *et al.,* 1996). A comprehensive safety assessment of the CP4 EPSPS protein has been described in the literature (Harrison *et al.,* 1996).

Monsanto Company has developed Roundup Ready corn event NK603 which is tolerant to glyphosate (the active ingredient in Roundup herbicide) at the whole plant level. The inserted DNA in NK603 has recently been sequenced (Kesterson *et al.,* 2001). Roundup Ready corn event NK603 contains two *cp4 epsps* genes which differ by two nucleotides, one of which results in an amino acid substitution, L214P, in the encoded protein. Both *cp4 epsps* genes

were present in the first (F1) and the second (BC1F1) generation after transformation; and are expected to be present in all subsequent breeding lines containing the NK603 insert.

The purpose of this document is to describe the structural and functional relationship between CP4 EPSPS proteins encoded by these two genes in the context of food, feed and environmental safety.

3.0 Diversity and Biological Activity of EPSPS Proteins

In order to accurately assess the potential functional consequences of the L214P substitution in one of the CP4 EPSPS proteins produced by the NK603 event, a taxonomical survey of known EPSPS proteins was performed to determine the extent in sequence diversity among EPSPS proteins. Bioinformatic analyses were performed using the *Agrobacterium ssp.* CP4 protein sequence as a query sequence to search a combined database composed of publicly available protein sequence data. The analysis identified 206 sequences that were subsequently classified as EPSPS proteins and the alignments of 21 of the EPSPS proteins are shown in Appendix A. There are numerous regions within the alignments that are highly variable including the sequence that flanks L214 in CP4 EPSPS. This flanking sequence highlighted in Table 1 shows that 14 EPSPS proteins have a proline within one amino acid of L214 of CP4 EPSPS in the aligned sequences.

Table 1. Tolerance of Proline Near Position 214 in Extant EPSPS Proteins

EPSPS Protein	%Identity to CP4 EPSPS	Amino Acid Alignment in Region Near L214P
CP4 EPSPS(Q9R4E4)	100	FGANLTVETD
CP4 L214P	99.8	FGANPTVETD
Tomato(M21071_1)	28.0	FGVFVEHSS.
Rice(AP002542_28)	26.7	FGVKAEHSD.
Maize(O24566)	28.0	FGVKAEHSD.
Maize GA21(A69537)	27.8	FGVKAEHSD.
Canola(X51475_1)	24.8	FGVSAEHSD.
C. jejuni(X89371_2)	35.5	MKAPIRVSN.
C. jejuni(AL139076_205)	35.7	MKAPIRVSN.
C. jejuni(AROA_CAMJE)	35.7	MKAPIRVSN.
P. aeruginosa(Q9HZ69)	50.0	FGYPVVVEG.
P. aeruginosa(AE004740_5)	50.0	FGYPVVVEG.
P. aeruginosa(G5712093)	54.1	FGYPVSVNG.
P. aeruginosa(X78413)	50.4	FGYPVVVEG.
S. aureus(AP003362_152)	35.4	FNIPIEAER.
S. aureus(Q99U25)	35.4	FNIPIEAER.
S. aureus(AP003134_5)	35.4	FNIPIEAER.
S. aureus(AROA_STAAU)	35.4	FNIPIEAEG.
S. aureus(BAB57626)	35.4	FNIPIEAER.
S. aureus(L05004_2)	35.4	FNIPIEAEG.
M. catarrhalis(MCA101591)	46.2	FGYPVQVDG.

This finding indicates that a proline in this region of EPSPS is a naturally occurring residue. Moreover, the extent to which this region flanking position 214 has sustained sequence heterogeneity while maintaining EPSPS function indicates that L214P in CP4 EPSPS is a well-tolerated substitution.

4.0 Structural comparison of EPSPS Proteins

The presence of sequence heterogeneity including the above examples of proline substitutions flanking L214 suggested that this region was potentially localized to the surface of the EPSPS protein and not essential for EPSPS function. The X-ray crystal structure analysis of EPSPS from *E. coli* and *Agrobacterium ssp.* CP4 indicate that the structures are highly superimposable even though they only possess 27.6 percent identity at the amino acid level (Figure 1). It is likely that EPSPS enzymes from different species have diverged from a common ancestor but have retained a shared three-dimensional structure. In addition, enzyme activity data comparing these enzymes shows that these proteins have similar enzyme specific activities further demonstrating that the overall three dimensional structure (not the precise amino acid sequence) is the key element in the EPSPS function (Padgette *et al.*, 1996).

DNA sequence analysis has shown that one of the *cp4 epsps* genes in corn event NK603 contains an L214P substitution in the CP4 EPSPS protein coding sequence. Comparison of CP4 EPSPS and CP4 EPSPS L214P shows that the three-dimensional structures of the two proteins are superimposable (Figure 2). X-ray crystal structure analysis of CP4 EPSPS indicates that the L214P substitution is not in the active site of the CP4 EPSPS enzyme (Figure 2). Examination of the region flanking residue 214 demonstrates that the L214P substitution in CP4 EPSPS has produced no significant perturbation in the secondary and tertiary structure of the protein and thus is not anticipated to confer functional changes to the protein (Figure 3). Bioinformatics and primary amino acid sequence analysis of EPSPS have shown that there are other EPSPS proteins that contain a proline residue in the region flanking position 214 of CP4 EPSPS (Table 1). Homology modeling of one of the EPSPS proteins from *Pseudomonas aeruginosa* (PAE) indicates that PAE contains a proline in the region flanking residue 214 of CP4 EPSPS and that a proline residue is one of the naturally occurring amino acids in this highly variable region of EPSPS enzymes (Figures 4 and 5). From a structural perspective, the L214P substitution in CP4 EPSPS is not anticipated to confer functional changes to the CP4 EPSPS enzyme.

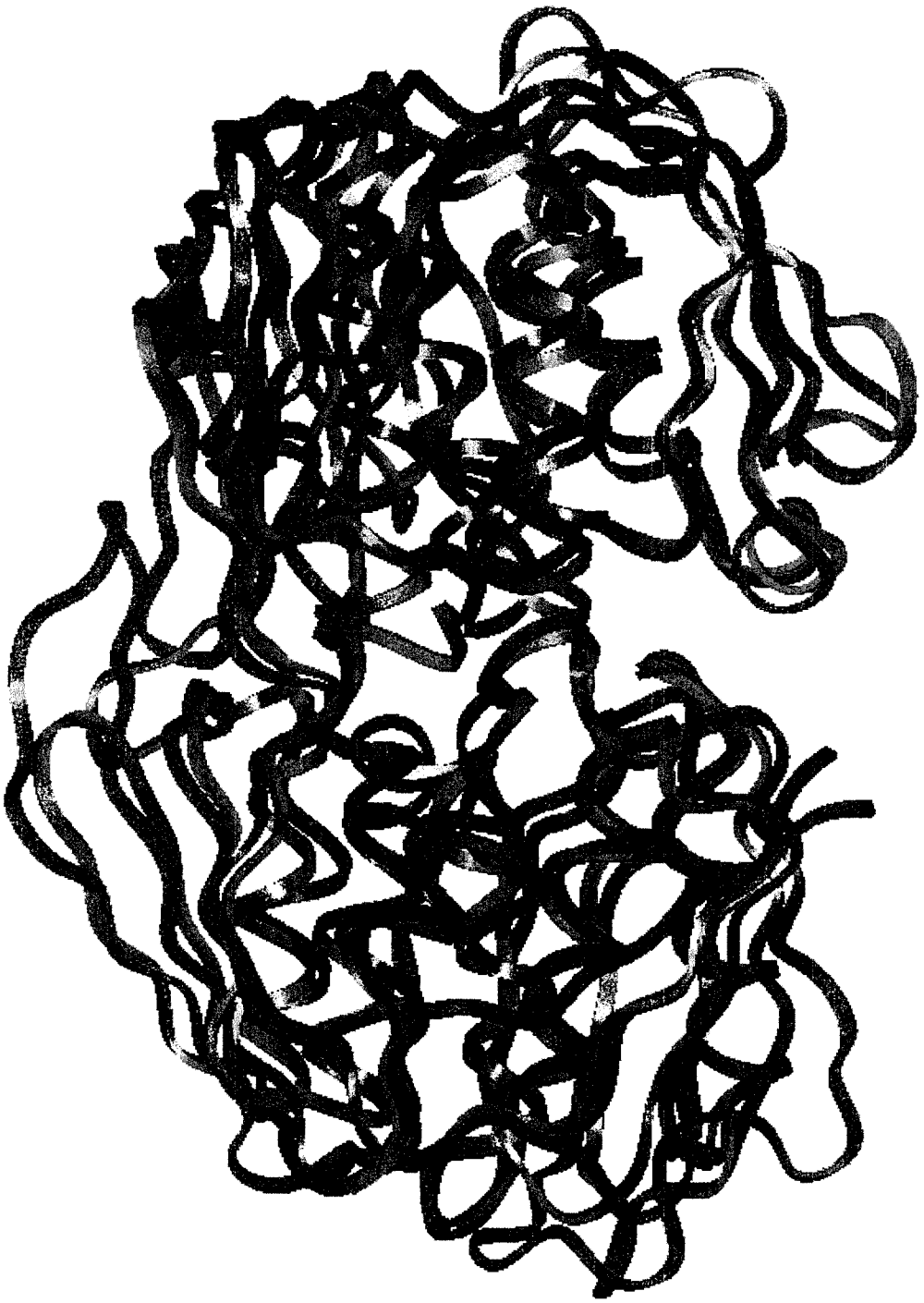


Figure 1. Structures of *E. coli* and CP4 EPSPS. The X-ray structures are represented as ribbon diagrams with the *E. coli* enzyme shown in blue and the CP4 EPSPS enzyme in red. Although *E. coli* and CP4 EPSPS enzymes share only 27.6% identity, X-ray crystal structure analyses indicate that both proteins are superimposable and share homologous structural characteristics that are important for function (Alibhai and Stallings, 2001).

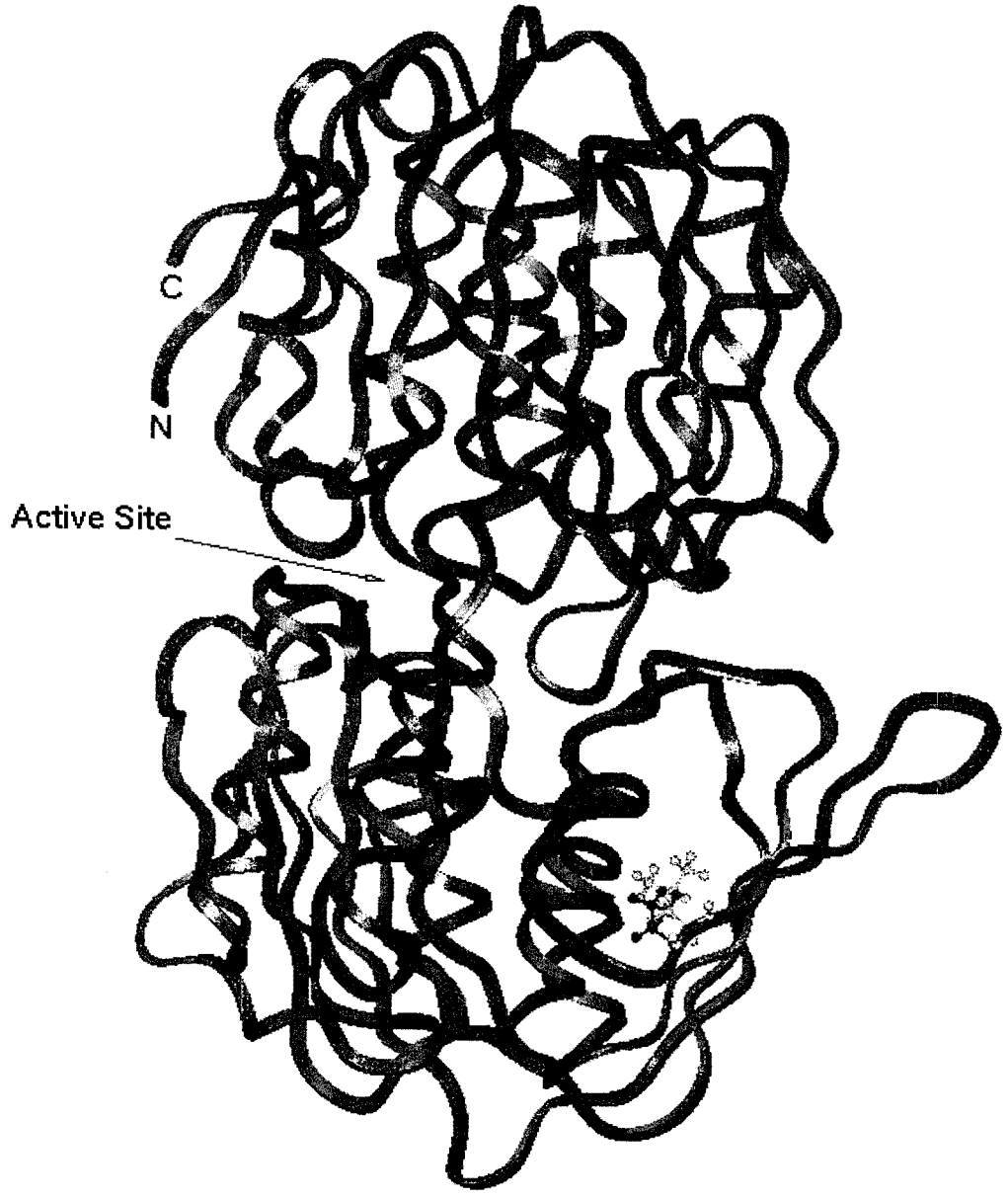


Figure 2. Structures of CP4 EPSPS and CP4 EPSPS L214P. The structures are shown as ribbon diagrams with the CP4 EPSPS in blue and CP4 EPSPS L214P in red. The Leu-214 residue in native CP4 EPSPS is displayed in yellow and the Pro-214 residue in CP4 EPSPS L214P is shown in green. The model of CP4 EPSPS L214P was generated by replacing the Leu residue with Pro using the InsightII package (Accelrys, CA). The model was then minimized using the Discover forcefield engine with the cvff forcefield using the steepest descent optimizer with 100 iterations. The Figure shows that Leu 214 (yellow) is not in the active site and the replacement of Leu with Pro does not perturb the structure of the CP4 EPSPS enzyme.

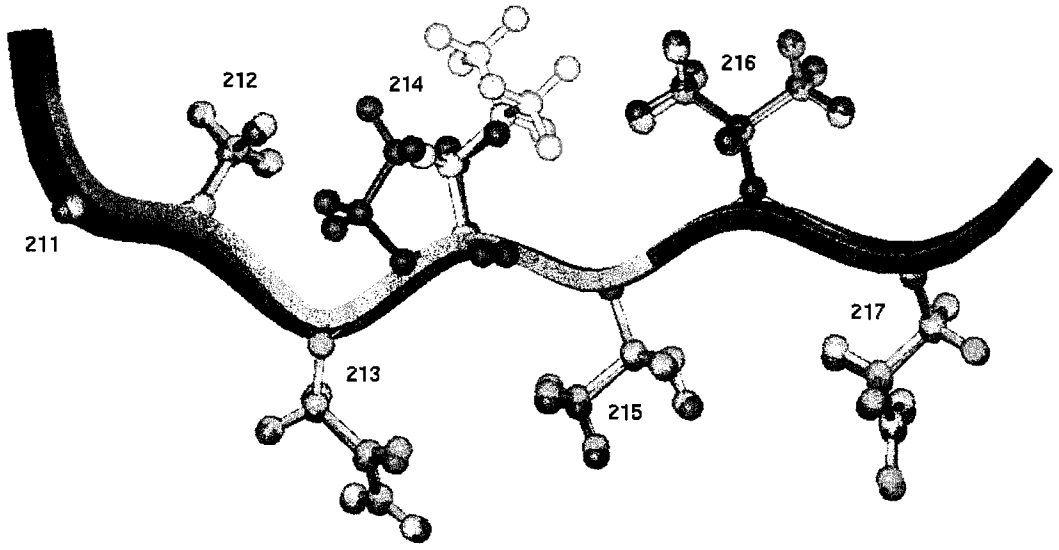


Figure 3. Ribbon diagram of CP4 EPSPS and CP4 EPSPS L214P near the residue 214 region of the enzymes. The CP4 EPSPS is shown in cyan and the CP4 EPSPS L214Pis shown in red. Side chains of Leu-214 in CP4 EPSPS and the Pro residue in CP4 EPSPS L214P are shown in yellow and green, respectively. A close-up view around the 214 regions of CP4 EPSPS and CP4 EPSPS L214P clearly shows that the mutation does not affect the structure in that region and hence is not anticipated to affect activity as verified by the results of the specific activity data.

Figure 4. Structures of CP4 EPSPS and the EPSPS protein from *Pseudomonas aeruginosa*. The proteins are shown as ribbon diagrams with the CP4 EPSPS enzyme shown in blue and the *P. aeruginosa* EPSPS enzyme in red. The side chains of Leu-214 in CP4 EPSPS and Pro-191 in *P. aeruginosa* enzyme are displayed in yellow and green, respectively. The structure of EPSPS from the *P. aeruginosa* enzyme was constructed using the Modeler module of the InsightII package (Accelrys, CA). Although *P. aeruginosa* and CP4 EPSPS enzymes share only 50.4% identity, X-ray crystal structure analysis indicate that both proteins are superimposable.

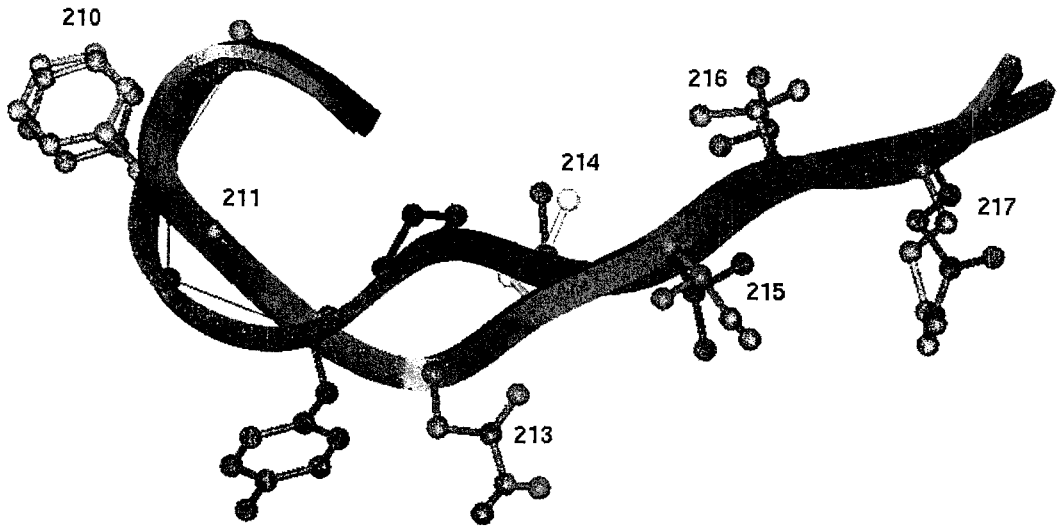


Figure 5. Ribbon diagram of CP4 EPSPS and EPSPS from *P. aeruginosa* near the residue 214 region of CP4 and the corresponding Pro-191 region of the *P. aeruginosa* enzyme. The CP4 EPSPS is shown in blue and the *P. aeruginosa* EPSPS is shown in red. Side chains of Leu-214 in CP4 EPSPS and the Pro residue in *P. aeruginosa* enzyme are shown in yellow and green, respectively. A close-up view around the 214 region indicates that this region is a highly variable region in the EPSPS enzymes. The Pro residue (green) in PAE lies very close to the Leu residue (yellow) in CP4 EPSPS indicating that there is natural variation in this region of the EPSPS enzyme and the mutation of L214P in CP4 EPSPS in this highly variable of EPSPS enzymes will not perturb the structure or affect activity.

5.0 Food and Feed Safety

5.1 CP4 EPSPS Protein Characterization

The CP4 EPSPS proteins produced in Roundup Ready corn event NK603 are functionally similar to a diverse family of EPSPS proteins typically present in food and feed derived from plant and microbial sources (Harrison, *et al.,* 1996). The EPSPS proteins are required for the production of aromatic amino acids in plants and microbes. The enzymology and known function of EPSPS proteins generally, and CP4 EPSPS specifically, indicate that this class of enzymes perform a well-described and understood biochemical role in plants. Since EPSPS catalyzes a non-rate limiting step in aromatic amino acid biosynthesis, the level or specific activity of EPSPS does not alter the flux though this metabolic pathway (Franz *et al.,* 1997), as is expected for a non-regulatory enzyme in a metabolic pathway.

Part of safety evaluation of these EPSPS proteins includes the known structural relationship between CP4 EPSPS and other EPSPS proteins found in food. This relationship is demonstrated by comparison of the amino acid sequences with conserved identity of the active site residues, and the expected conserved three-dimensional structure based on similarity of the amino acid sequence. From this perspective, CP4 EPSPS L214P and CP4 EPSPS are considered equivalent given extant sequence diversity of EPSPSs. Thus, the difference in the CP4 EPSPS amino acid sequences found in corn event NK603 is insignificant compared to the typical sequence diversity known in EPSPSs found in food (Padgette *et al.,* 1996).

The detailed enzymology of CP4 EPSPS reviewed by Padgette *et al.* (1996) and subsequent biochemical composition evaluations (Ridley *et al.,* 2000; McCann *et al.,* 2001; Pyla *et al.,* 2000), confirm that CP4 EPSPSs produced in corn event NK603 do not affect the levels of aromatic amino acids nor any of the other measured components in corn containing the NK603 event compared to traditional varieties of corn.

5.1.1 Purification and Characterization of CP4 EPSPS L214P from *E. coli*

CP4 EPSPS L214P was produced heterologously in *E. coli* by fermentation and purified to 98% homogeneity (Appendix C) by standard EPSPS purification techniques (Heeren, *et al.,* 1993). The specific enzymatic activity of the CP4 EPSPS L214P was determined to be equivalent to similarly purified CP4 EPSPS. The CP4 EPSPS L214P as a constituent of corn event NK603 confers the same enzymatic function as CP4 EPSPS.

5.2 History of Safe Use

EPSPS enzymes are found ubiquitously in all plant-derived foods. In addition to available data for CP4 EPSPS (Harrison *et al.,* 1996), evaluations of a modified *Zea mays* EPSPS (mEPSPS), which is ~28% or identical to CP4 EPSPS, have confirmed the rapid digestibility

and lack of toxicity of EPSPSs (Sidhu *et al.,* 1997; Astwood, 1997; Naylor, 1997). Both CP4 EPSPS and mEPSPS have been consumed widely in Roundup Ready soybeans and Roundup Ready corn, respectively (James, 2000).

5.3 Lack of Structural Homology of CP4 EPSPS Protein to Known Protein Toxins

Another aspect used for the assessment of potential toxic effects of proteins introduced into plants is to compare the amino acid sequence of the protein to known toxic proteins. Homologous proteins derived from a common ancestor have similar amino acid sequences, are structurally similar and often share common function. Therefore, it is undesirable to introduce DNA which encodes for a protein that is homologous to a protein which is toxic to animals and people. Homology is determined by comparing the degree of amino acid similarity between proteins using published criteria (Doolitte *et al.,* 1990). Consistent with CP4 EPSPS (Hileman and Astwood, 1999b) and mEPSPS (Astwood, 1997), CP4 EPSPS L214P protein showed no meaningful amino acid sequence similarity when compared to known protein toxins when evaluated using the FASTA algorithm (data not shown).

5.4 Digestibility *in vitro*

As part of the safety assessment of CP4 EPSPS L214P, *in vitro* digestion in simulated mammalian gastric fluid was used to assess the susceptibility of CP4 EPSPS L214P protein to proteolytic digestion. Rapid degradation of the protein correlates with limited exposure to the gastrointestinal tract and little likelihood that the protein can exert pharmacological, toxic or allergenic effects.

CP4 EPSPS protein has previously been shown to be rapidly digested *in vitro* (Harrison, *et al.*, 1996). Using enzyme activity assays and immunoblot analysis it was demonstrated that the half-life of CP4 EPSPS protein in simulated gastric fluid was less than 2 minutes when assayed for activity (the shortest time point assayed), and less than 15 seconds when assayed by immunoblot. Proteins which are rapidly degraded in the gastrointestinal tract are unlikely to confer toxicity, be allergenic or display untoward pharmacological activity (Astwood *et al.,* 1996; Astwood and Fuchs, 2000). The *in vitro* digestibility of CP4 EPSPS L214P in simulated gastric fluid was predicted to parallel CP4 EPSPS based upon structure and function data. This prediction was confirmed using *E. coli* produced, purified CP4 EPSPS L214P that was shown to have a half-life of less than 15 seconds in simulated gastric fluid (see Appendix D).

5.5 Lack of Acute Oral Toxicity of the CP4 EPSPS Protein in Mice

Few proteins are toxic when ingested and those that are toxic, typically act in an acute manner (Sjoblad *et al.*, 1992). To confirm the lack of acute toxicity, an oral toxicity study with CP4 EPSPS as the test material was performed on mice to directly assess any potential toxicity associated with the protein (Harrison *et al.,* 1996). Acute administration was

considered appropriate to assess the safety of CP4 EPSPS, since proteins that are toxic act via acute mechanisms (Sjoblad *et al.*, 1992). There were no treatment-related adverse effects in mice administered CP4 EPSPS protein by oral gavage at dosages up to 572 mg/kg body weight. Results from this study demonstrated that the CP4 EPSPS protein is not acutely toxic to mammals. This result was expected since CP4 EPSPS is from a ubiquitous family of proteins with a history of safe consumption, the CP4 EPSPS protein is readily digested in gastric and intestinal fluids *in vitro* and there is no biologically plausible mechanism of toxicity to animals.

5.6 Lack of Structural Homology or Physicochemical Similarity to Known Protein Allergens

Although there are no single predictive bioassay available to assess the allergenic potential of proteins in humans (FDA, 1992), the physicochemical and human exposure profile of the protein provides a basis for assessing potential allergenicity by comparing it to known protein allergens. Thus, important considerations contributing to the allergenicity of proteins ingested orally includes exposure and an assessment of the factors that contribute to exposure, such as stability to digestion, prevalence in the food, and consumption pattern (amount) of the specific food (Metcalfe *et al.*, 1996; Kimber *et al.*, 1999).

A key parameter contributing to the systemic allergenicity of certain food proteins appears to be stability to the peptic and acidic conditions of the digestive system (Astwood *et al.*, 1996; Fuchs and Astwood, 1996; FAO, 1995; Kimber *et al.*, 1999). Important protein allergens tend to be stable to peptic digestion and the acidic conditions of the stomach if they are to reach the intestinal mucosa where an adverse immune response can be initiated. As noted above, the *in vitro* assessment of the CP4 EPSPS L214P protein digestibility indicates that the protein, like other food-derived proteins, is very labile to digestion when compared to many clinically important food allergens.

Another significant factor contributing to the allergenicity of certain food proteins is their high concentrations in foods (Taylor, 1992; Taylor *et al.*, 1987; Fuchs and Astwood, 1996). Most food allergens are present as major protein components in the specific food, representing from 2-3% up to 80% of total protein (Fuchs and Astwood, 1996). In contrast, the CP4 EPSPS proteins are present at extremely low levels (generally less than 0.1% of the total protein content) in Roundup Ready corn plants (Sidhu and Ledesma, 1999).

It is also important to establish that the protein does not represent a previously described allergen and does not share potentially immunologically-relevant amino acid sequence segments or structure with a known allergen. An efficient way to determine whether the added protein is an allergen or is likely to contain cross-reactive structures is to compare the amino acid sequence of the introduced protein with those of all know allergens. A database of protein sequences associated with allergy and coeliac disease has been assembled from publicly available genetic databases (GenBank, EMBL, PIR and SwissProt). The amino acid

sequence of the CP4 EPSPS L214P was compared to these sequences and did not share any meaningful amino acid sequence similarity with any of the known allergens (data not shown) as had been demonstrated previously for both CP4 EPSPS (Hileman and Astwood, 1999a) and mEPSPS (Astwood, 1997).

5.7 Substantial Equivalence

Extensive studies have been conducted with Roundup Ready corn event NK603 which demonstrate that there are no unexpected or pleiotropic effects associated with the DNA insertions in this product or due to the expression of the two *cp4 epsps* genes.

Comparisons of the agronomic characteristics of event NK603 with conventional corn show no differences in phenotype, yield or other measured parameters (USDA, 2001). No differences were observed in composition and no differences in nutritional value were detected in the animal feeding studies described below.

5.7.1 Grain and Forage Composition of Corn Event NK603

The *cp4 epsps* and *cp4 epsps L214P* genes were both constituent in corn event NK603 during all evaluations of whole food and feed safety. To assess potential pleiotropic effects due to the insertion of the *cp4 epsps* genes into the maize genome or due to the enzyme activity associated with the introduced CP4 EPSPS and CP4 EPSPS L214P enzymes, a comprehensive comparison of 51 nutritional biochemical components of corn event NK603 to conventional corn was conducted for corn grown in 1998 and 2000 in the United States, and in 1999 for corn grown in the European Union. Grain was collected from replicated field trials and was measured for proximates, fibre, amino acids, fatty acids, vitamin E, nine minerals, phytic acid, trypsin inhibitor, and selected secondary metabolites. Forage was also collected and measured for proximates and fibre. All comparisons showed that corn event NK603 is compositionally equivalent to conventional corn hybrids (Ridley *et al.*, 2000, McCann *et al.*, 2001; Pyla *et al.*, 2000). These studies demonstrated that the insertion and expression of both *cp4 epsps* genes do not alter the biochemical flux of aromatic amino acids associated with the shikimic acid pathway (the site of EPSPS activity), as expected. In addition to establishing that corn event NK603 is as safe and nutritious as conventional corn hybrids, these studies confirmed that the enzymatic activity of the CP4 EPSPS L214P was not materially different than endogenous EPSPS. For example, the lack of changes in amino acid composition was consistent with the known enzymatic activity of CP4 EPSPS produced endogenously by *Agrobacterium ssp.* strain CP4, heterologously by *E. coli*, and heterologously by genetically modified crops such as Roundup Ready soybeans (Harrison *et al.*, 1996), Roundup Ready canola (Lee and Astwood, 2000), or Roundup Ready cotton (Nida, *et al.*,1996). Grain and forage composition of corn event NK603 were shown to be compositionally equivalent to traditional corn hybrids.

* This study not provided to ANZFA

5.7.2 Farm Animal Nutrition Studies

A series of animal feeding studies were performed previously with corn grain from event NK603 (which contains both *cp4 epsps* genes). These farm animal nutrition studies addressed the potential of any pleiotropic effect caused by the insertion process or site of insertion as well as the safety of any produced protein or peptide and any other constituent that could arise.

Broiler chicken performance was assessed while fed diets incorporating up to 64% grain from corn event NK603 over a 42 day period relative to diets containing grain from the parental and reference corn hybrids. No differences in a broad spectrum of parameters were observed, including weight gain, feed intake, feed efficiency, and carcass and quality measures. This study is highly sensitive to untoward pleiotropic effects due to the presence of unexpected enzyme activities in the grain of NK603 corn due to the rapid growth of the birds during the duration of the study – a greater than 50 fold increase in body weight was observed (Taylor *et al.,* 2001). Likewise, grower and finisher swine fed Roundup Ready corn event NK603 at diet incorporation rates of between 68% and 82 % respectively, showed equivalent performance to parental and reference corn diets (Stanisiewski *et al.,* 2001).

5.7.3 Lack of Subchronic Oral Toxicity in Rats

A 13 week subchronic oral toxicity study was conducted to compare the responses of rats fed diets containing grain from Roundup Ready corn event NK603 to rats fed (1) diets containing grain from a parental hybrid and (2) a series of diets containing grain from conventional corn hybrids. This study showed that at diet incorporation rates up to 33% corn, the highest dose tested, no differences in food consumption, organ weight, blood chemistry, hematology and urinalysis were observed between treatments (Dudek, 2001). This study reinforced both the safety and nutritional equivalence of grain produced from Roundup Ready corn event NK603 and conventional corn, including supporting the safety of the expressed proteins or peptides or other constituents of corn event NK603.

5.7.4 Conclusions from Animal Nutrition and Toxicological Evaluations

The lack of acute oral toxicity of CP4 EPSPS and mEPSPS with margins of exposure of ~183,000 fold and ~71,000 fold, respectively, calculated using average dietary consumption (Table 2), in combination with margins of safety established in whole grain animal feeding studies (Table 3), confirm the food and feed safety and nutritional equivalence of corn event NK603 relative to conventional corn hybrids including the constituent EPSPS proteins. Furthermore, the whole grain studies address the safety of both the corn grain and the two CP4 EPSPS proteins of corn event NK603; and any products derived thereof. The nutritional value or wholesomeness of corn event NK603, even when fed to animals at levels much higher than humans would encounter in the diet, was the same as conventional hybrids of corn.

Table 2. Margins of Exposure Computed for Humans Consuming Representative EPSPS Proteins

EPSPS Protein	Crop	NOEL (mg/kg)	Margin of Exposure[a,b,c]	
			50th percentile	90th percentile
CP4 EPSPS	Corn	572.0	183,000	58,000
mEPSPS	Corn	45.6	71,000	22,000

[a] Margin of Exposure = NOEL / upper bound estimate of dietary exposure rounded to nearest 1000; calculations assume that the genetically modified crop represents 100% of corn in food/feed consumed. This is a worst-case assumption since RR varieties represent only a fraction of the total varieties of corn that are used for food/feed.
[b] Dietary exposure was estimated by the Exposure 1® Chronic Dietary Exposure Analysis Program of Technical Assessment Systems, Inc.
[c] Corn consumption for the general US population averaged over the entire year is 0.2 g/kg/day at the 50th percentile and 0.63 g/kg/day at the 90th percentile (refer to Appendix B).

Table 3. Safety Margins in Corn Grain Diet Incorporation in Animal Feeding Studies

Feeding Study	Safety Margin*	
	50th percentile	90th percentile
42 day Broiler	285	91
103 day Swine	115	37
13 Week Rat	105	33

**Compared to Average US Population* . Corn consumption for the general US population averaged over the entire year is 0.2 g/kg/day at the 50th percentile and 0.63 g/kg/day at the 90th percentile (refer to Appendix B).

6.0 Environmental Safety

The conclusion of safety for food, feed or the environment has been confirmed by regulatory authorities around the world and by the commercial production of crop products containing the CP4 EPSPS protein on over 300 million acres over the past five years (James, 2000). Monsanto has conducted extensive agronomic and environmental safety studies with the Roundup Ready corn event NK603. Comparisons of disease and pest susceptibilities, yield, morphology, weediness, outrcrossing and other relevant characteristics confirmed that this event does not pose any increased environmental risks compared to traditional corn (USDA, 2001).

7.0 Conclusion

Taxonomic bioinformatics data show that significant sequence heterogeneity exists generally among EPSPSs (as low as 25% similarity over the entire molecule) and specific heterogeneity exists proximal to position 214 of CP4 EPSPS, including known examples of proline substitutions. The presence of known examples of prolines near this position suggests that substitutions in this domain of EPSPS are not critical to the structural arrangement or function of the protein. Modeling of a proline residue at position 214 in the known CP4 EPSPS X-ray crystallography structure shows that position 214 is located in a loop region on the adaxial surface plane, and is distal to the active site. Furthermore, modeling of other EPSPS proteins shows that significant structural elasticity exists among known EPSPS proteins, further supporting the conclusion that the L214P substitution would not be expected to effect EPSPS function.

The CP4 EPSPS L214P produced by corn event NK603 is well within the EPSPS protein family which has a demonstrated history of safe use and consumption. The enzyme activity of CP4 EPSPS has been well described, representing an endogenous activity common to all plant foods and some microbial species. CP4 EPSPS has been shown to lack sequence or functional similarity to toxins and allergens, is rapidly digested *in vitro*, lacks acute oral toxicity, and is produced at very low levels. Furthermore, the highly divergent mEPSPS has the same food, feed and environmental safety profile as CP4 EPSPS.

CP4 EPSPS L214P was produced heterologously in *E. coli* and was purified to 98% homogeneity. Characterization of the *E. coli* produced CP4 EPSPS L214P protein demonstrated equivalence in function to CP4 EPSPS. This confirmed predictions from structure/function modeling of the CP4 EPSPS L214P using the known CP4 EPSPS X-ray crystal structure as a reference. Bioinformatics evaluations confirmed that CP4 EPSPS L214P shares no sequence similarity to toxins and allergens. Finally, CP4 EPSPS L214P was found to be equivalently digested *in vitro* relative to CP4 EPSPS. These data confirm that CP4 EPSPS L214P may be considered equivalent to CP4 EPSPS with respect to structural and functional characteristics that are relevant to food and feed safety.

Finally, evaluations of equivalence in composition, animal nutrition, toxicity, and agronomic and environmental parameters between Roundup Ready corn event NK603 and traditional corn hybrids demonstrate that (1) the presence of the CP4 EPSPS L214P does not alter the composition of corn, (2) there are no biologically relevant pleiotropic effects due to the CP4 EPSPS L214P, (3) corn grain containing the NK603 insert is as safe and nutritious as traditional corn hybrids and (4) Roundup Ready corn event NK603 containing CP4 EPSPS and CP4 EPSPS L214P proteins does not present an increased risk to the environment compared to traditional corn.

It was concluded that CP4 EPSPS L214P is indistinguishable from CP4 EPSPS with respect to function, food, feed and environmental safety.

8.0 References

Alibhai, M.F and Stallings, W.C. (2001) Closing down on glyphosate inhibition – with a new structure for drug discovery. *Proc. Natl. Acad. Sci. USA* **98**: 2944-2946.

Astwood, J.D. (1997) Modified maize 5-enolpyruvylshikimate-3-phoshpate synthase (mEPSPS) has no significant sequence similarity to known allergens and toxins. Monsanto Technical Report MSL-15168, St. Louis, Missouri.

Astwood, J.D. and Fuchs, R.L. (2000) Status and safety of biotech crops. Pp 152-164. *In* Agrochemical discovery insect, weed and fungal control. Baker D.R. and N.K. Umetsu (eds.). ACS Symposium Series 774.

Astwood, J.D., Leach, J.N. and Fuchs, R.L. (1996) Stability of food allergens to digestion in vitro. *Nature Biotechnology* **14**:1269-1273.

Baird, D.D., Upchurch, R.P, Homesley, W.B. and Franz, J.E. (1971) *Proc. North Cent. Weed Control Conf.* **26**: 64-68.

Burnette, B.L., and Holden, L.R. (1994) Validation of the phosphate release enzymatic assay for the quantitation of functionally active 5-enol-pyruvyl-shikimate-3-phosphate synthase (EPSPS). Monsanto Technical Report MSL-13553. St. Louis, Missouri.

Deutcher, M.P. (1990) Guide to Protein Purification, in: *Methods in Enzymology* 182. Academic Press, Inc., Harcourt Brace Jovanovich, Publishers, New York.

Doolittle, R.F. (1990) Searching through sequence databases. *Methods in Enzymology* **183**: 99-110.

Dudek, B.R. (2001) Amended report:13 week feeding study in rats with grain from Roundup Ready corn (NK603) preceded by a 1-week baseline food consumption determination with PMI certified rodent diet #5002. Monsanto Technical Report MSL-17555.

FAO (Food and Agriculture Organization). (1995) Report of the FAO Technical Consultation on Food Allergies, Rome, Italy, November 13-14, 1995. FAO, Rome.

Franz, J.E., Mao, M.K., and Sikorski, J.A. (1997) *Glyphosate: A Unique Global Herbicide.* ACS Monograph 189. American Chemical Society, Washington D.C., pp 1-653.

Fuchs, R.L. and Astwood, J.D. (1996) Allergenicity assessment of foods derived from genetically modified plants. *Food Technology* **50**: 83-88.

Harrison, L.A., Bailey, M.R., Naylor, M.W., Ream, J.E., Hammond, B.G., Nida, D.L., Burnette, B.L., Nickson, T.E., Mitsky, T.A., Taylor, M.L., Fuchs, R.L., and Padgette, S.R. (1996) The expressed protein in glyphosate-tolerant soybean, 5-enolpyruvylshikimate-3-phosphate synthase from *Agrobacterium* sp. strain CP4, is rapidly digested in vitro and is not toxic to acutely gavaged mice. *Journal of Nutrition* **126**: 728-740.

Haslam, E. (1993) Shikimic Acid: metabolism and metabolites. John Wiley and Sons, Chichester, England.

Heeren, R.A., Padgette, S.R., and Gustafson, M.E. (1993) Purification of recombinant *Escherichia coli* CP4 5-enol pyruvyl shikimate 3-phosphate synthase for equivalence studies. Monsanto Technical Report MSL-12574, St. Louis, Missouri.

Hileman, R.E. and Astwood, J.D. (1999a) Bioinformatics analysis of CP4 EPSPS protein sequence utilizing an allergen database. Monsanto Technical Report MSL-16267, St. Louis, Missouri.

Hileman, R.E. and Astwood, J.D. (1999b) Bioinformatics analysis of CP4 EPSPS protein Sequence Utilizing Toxin and Public Domain Genetic Databases. Monsanto Technical Report MSL-16268, St. Louis, Missouri.

James, C. (2000) Global overview of commercialized transgenic crops: 2000. ISAAA Brief no. 23. ISAAA: Ithaca, N.Y.

Kesterson, N.K., Reiser, S.E., Cavato, T.A., and Lirette, R.P. (2001 Draft) PCR and DNA sequence analysis of the insert in Roundup Ready maize event NK603. Monsanto Technical Report MSL-17588, St. Louis, Missouri.

Kimber, I., N.I. Kerkvliet, S.L. Taylor, J.D. Astwood, K. Sarlo, and R.J. Dearman. (1999) Toxicology of protein allergenicity: Prediction and characterization. *Toxicological Sciences* **48**:157-162.

Lee, T.C. and Astwood, J.D. (2000) Assessment of the physicochemical equivalence of CP4 EPSPS and GOX proteins produced in glyphosate tolerant canola event 17209-2001 (Svalöf Weibull line 00.77143) relative to Roundup Ready® canola events RT73 and RT200 and *Escherichia coli*. Monsanto Technical Report MSL-16946, St. Louis, Missouri.

McCann, M.C., Breeze, M., Sorbet, R., Riordan, S., and Astwood, J.D. (2001) Compositional analysis of forage and grain collected from Roundup Ready(®) corn hybrid NK603 grown in Kihei, Hawaii during 2000. Monsanto Technical Report MSL-17234, St. Louis, Missouri.

Metcalfe, D.D., Astwood, J.D., Townsend, R., Sampson, H.A., Taylor, S.L. and Fuchs, R.L. (1996) Assessment of the allergenic potential of foods derived from genetically engineered crop plants. *Critical Reviews in Food Science and Nutrition* **36**(S):S165-S186.

Naylor, M. (1997) Acute oral toxicity study with modified maize 5-enolpyruvyl shikimate-3-phosphate synthase (mEPSPS) protein in albino mice. Monsanto Technical Report MSL-15235, St. Louis, Missouri.

Neuhoff, V., Norbert, A., Taube, D. and Wolfgang, E. (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.

Nida, D.L., Kolacz, K.H., Buehler, R.E., Deaton, W.R., Schuler, W.R., Armstrong, T.A., Taylor, M.L., Ebert, C.C., Rogan, G.J., Padgette, S.R., and Fuchs, R.L. (1996) Glyphosate-Tolerant Cotton: Genetic characterization and protein expression. *Journal of Agricultural and Food Chemistry* **44**: 1960-1966.

Padgette, S.R., Barry, G.F., Re, D.B., Eichholtz, D.E., Weldon, M., Kolacz, K.H. and Kishore, G.M. (1993) Purification, cloning, and characterization of a highly glyphosate-tolerant EPSP synthase from *Agrobacterium sp.* strain CP4. Monsanto Technical Report MSL-12738, St. Louis, Missouri.

Padgette, S., Re, D., Barry, G.F., Eichholtz, D., Delannay, X., Fuchs, R., Kishore, G. and Fraley, R. (1996) New weed control opportunities: Development of soybeans with a Roundup Ready™ gene, pp. 53-84. In Herbicide Resistant Crops, S. O. Duke (ed.), CRC, Boca Raton, Florida.

Pyla, P.D., Nemeth, M.A., Sorbet, R., and Astwood, J.D. (2000) Amendment to MSL 16700: Determination of secondary metabolite levels in Roundup Ready maize line NK603 grain harvested from 1998 U.S. field trials. Monsanto Technical Report MSL-16902, St. Louis, Missouri.

Ridley, W.P., George, C., Nemeth, M.A., Astwood, J.D., Breeze M.L. and Sorbet, R. (2000) Compositional analysis of forage and grain collected from Roundup Ready® maize event NK603 grown in 1999 E.U. field trials. Monsanto Technical Report MSL-16897, St. Louis, Missouri.

Schägger, H. and von Jagow, G. (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Analytical Biochemistry*, **166**: 368-379.

Sidhu, R.S. and Ledesma, B.E. (1999) Introduced protein levels and compositional analyses of Roundup Ready® corn line NK603 tissues produced in 1998 U.S. field trials. Monsanto Technical Report MSL-16278, St. Louis, Missouri.

Sidhu, R.S., Leach, J.N., Lee, T.C. and Astwood, J.D. (1997) Assessment of the digestibility of modified maize 5-enolpyruvylshikimate-3-phosphate synthase (mEPSPS) protein *in vitro* using mammalian digestive fate models. Monsanto Technical Report MSL-15169, St. Louis, Missouri.

Sjoblad, R.D., McClintock, J.T., and Engler, R. (1992) Toxicological considerations for protein components of biological pesticide products. *Regulatory Toxicol. and Pharmacol.* **15**: 3-9.

Stanisiewski, E.P., Hartnell, G.F., and Lewis, A.J. (2001 Draft) Performance of pigs fed diets containing Roundup Ready (NK603), parental or conventional corn grown during 2000 in Nebraska. Monsanto Technical Report MSL-17500, St. Louis, Missouri.

Steinrücken, H.C. and Amrhein, N. (1980) The herbicide glyphosate is a potent inhibitor of 5-enolpyruvyl-shikimic acid -3-phosphate synthase. *Biochem. Biophys. Res. Comm.* **94**: 1207-1212.

Taylor, S.L. 1992. Chemistry and detection of food allergens. *Food Technology* **46**:146-152.

Taylor, S.L., Lemanske Jr., R.F., Bush, R.K. and Busse, W.W. (1987) Food allergens: structure and immunologic properties. *Ann. Allergy* **59**:93-99.

Taylor, M.L., Hartnell, G.F., Riordan, S.G., Nemeth, M.A., George, B., Carpenter, D.M., and Astwood, J.D. (2001) Amended report for MSL-17107: Sponsor summary of report for study #2000-01-39-02 (Comparison of broiler performance when fed diets containing event NK603, parental line, or commercial corn). Monsanto Technical Report MSL 17458, St. Louis, Missouri.

USDA. 2001. Decision on Monsanto request (00-011-01p): Extension of determination of nonregulated status for glyphosate herbicide tolerant corn lines NK603. Environmental Assessment. Federal Register. **65**: 52693-52694.

U.S. FDA. 1992. Statement of policy: Foods derived from new plant varieties. Federal Register **57**: 22984-23005.

9.0 Appendix A – Clustal Sequence Alignments

These alignments were generated using the ClustalW multiple sequence alignment program. Amino acid numbering is relative to the CP4 EPSPS sequence. Residues in red demarcate the region surrounding position 214. Proline substitutions at or near position 214 are highlighted in blue.

```
                                  1         10         20         30         40
CP4                       .....MLHGA SSRPATARKS SGLSGTVR.I PGDKSISHRS FMFGGLASGE
CP4 NK603 VARIANT         .....MLHGA SSRPATARKS SGLSGTVR.I PGDKSISHRS FMFGGLASGE
P. AERUGINOSA(Q9HZ69)     .......MHN NDLIYLAQPG GSLSGTIR.V PGDKSISHRS IMLGSLAEGT
P. AERUGINOSA(X78413)     .......... .......... .........V PGDKSISHRS IMLGSLAEGT
P. AERUGINOSA(AE004740_5) .......... NDLIYLAQPG GSLSGTIR.V PGDKSISHRS IMLGSLAEGT
P. AERUGINOSA(G5712093)   .......... ......AQPG GRLSGRIR.V PGDKSISHRS IMLGSLAEGV
M. CATARRHALIS(MCA101591) .......... .......... ..ISGVIS.V PGDKSISHRS IMFGSLADGV
S. AUREUS(AP003362_152)   .......... MVSEQIIDIS GPLKGEIE.V PGDKSMTHRA IMLASLAEGT
S. AUREUS(AP003134_5)     .......... MVSEQIIDIS GPLKGEIE.V PGDKSMTHRA IMLASLAEGT
S. AUREUS(BAB57626)       .......... MVSEQIIDIS GPLKGEIE.V PGDKSMTHRA IMLASLAEGT
S. AUREUS(Q99U25)         .......... MVSEQIIDIS GPLKGEIE.V PGDKSMTHRA IMLASLAEGT
S. AUREUS(L05004_2)       .......... MVNEQIIDIS GPLKGEIE.V PGDKSMTHRA IMLASLAEGV
S. AUREUS(AROA_STAAU)     .......... MVNEQIIDIS GPLKGEIE.V PGDKSMTHRA IMLASLAEGV
C. JEJUNI(AROA_CAMJE)     .......... ...MKIYKLQ TPVNAILENI AADKSISHRF AIFSLLTQEE
C. JEJUNI(AL139076_205)   .......... ...MKIYKLQ TPVNAILENI AADKSISHRF AIFSLLTQEE
C. JEJUNI(X89371_2)       .......... ...MKIYKLQ TPVNAILENI AADKSISHRF AIFSLLTQEE
MAIZE                     .......... GAEEIVLQPI KEISGTVK.L PGSKSLSNRI LLLAALSEGT
MAIZE GA21                .......... GAEEIVLQPI KEISGTVK.L PGSKSLSNRI LLLAALSEGT
RICE                      .......... .......... .......... .......... ..........
CANOLA                    .......... KASEIVLQPI REISGLIK.L PGSKSLSNRI LLLAALSEGT
TOMATO                    .......... KPHEIVLXPI KDISGTVK.L PGSKSLSNRI LLLAALSEGR

                              50         60         70         80
CP4                       TRITGLLEGE DVINTGKAMQ AMGARIRK.. .EGDTWIIDG VGNGGLL...
CP4 NK603 VARIANT         TRITGLLEGE DVINTGKAMQ AMGARIRK.. .EGDTWIIDG VGNGGLL...
P. AERUGINOSA(Q9HZ69)     TEVEGFLEGE DALATIQAFR DMGVVIEGP. .QNGRVTVHG VGLHGLK...
P. AERUGINOSA(X78413)     TEVEGFLEGE DALATIQAFR DMGVVIEGP. .QNGRVTVHG VGLHGLK...
P. AERUGINOSA(AE004740_5) TEVEGFLEGE DALATIQAFR DMGVVIEGP. .QNGRVTVHG VGLHGLK...
P. AERUGINOSA(G5712093)   TEVEGFLEGE DALATLQAFR DMGVVIEGP. .HHGRVTIHG VGLHGLK...
M. CATARRHALIS(MCA101591) THVTGFLQGE DALATLQAFA DMGVKIER.. .NGDKVTIHG VGIDGLK...
S. AUREUS(AP003362_152)   SNIYKPLLGE DCRRTMDIFR LLGVDIKE.. .DEDKLVVNS PGYKAFK...
S. AUREUS(AP003134_5)     SNIYKPLLGE DCRRTMDIFR LLGVDIKE.. .DEDKLVVNS PGYKAFK...
S. AUREUS(BAB57626)       SNIYKPLLGE DCRRTMDIFR LLGVDIKE.. .DEDKLVVNS PGYKAFK...
S. AUREUS(Q99U25)         SNIYKPLLGE DCRRTMDIFR LLGVDIKE.. .DEDKLVVNS PGYKAFK...
S. AUREUS(L05004_2)       STIYKPLLGE DCRRTMDIFR HLGVEIKE.. .DDEKLVVTS PGYQVN....
S. AUREUS(AROA_STAAU)     STIYKPLLGE DCRRTMDIFR HLGVEIKE.. .DDEKLVVTS PGYQVN....
C. JEJUNI(AROA_CAMJE)     NKAQNYLLAQ DTLNTLEIIK NLGAKIEQ.. ...KDSCVKI IPPKEIL...
C. JEJUNI(AL139076_205)   NKAQNYLLAQ DTLNTLEIIK NLGAKIEQ.. ...KDSCVKI IPPKEIL...
C. JEJUNI(X89371_2)       NKAQNYLLAQ DTLNTLEIIK NLGAKIEQ.. ...KDSCVKI IPPKEIL...
MAIZE                     TVVDNLLNSE DVHYMLGALR TLGLSVEADK AAKRAVVVGC GGKFPVE.DA
GA21 RR MEPSPS            TVVDNLLNSE DVHYMLGALR TLGLSVEADK AAKRAVVVGC GGKFPVE.DA
RICE                      .......... ....MLEALK ALGLSVEADK VAKRAVVVGC GGKFPVEKDA
CANOLA                    TVVDNLLNSD DINYMLDALK KLGLNVERDS VNNRAVVEGC GGIFPASLDS
TOMATO                    TVVDNLLSSD DIHYMLGALK TLGLHVEDDN ENQRAIVEGC GGQFPVGKKS
```

```
                                90        100        110        120        130
CP4                             APEAPLDFGN AATGCRLTMG LVGVYDFDST FIGD..ASLT KRPMGRVLNP
CP4 NK603 VARIANT               APEAPLDFGN AATGCRLTMG LVGVYDFDST FIGD..ASLT KRPMGRVLNP
P. AERUGINOSA(Q9HZ69)           APPGPIYLGN SGTSMRLLSG LLAAQPFDST LTGD..ASLS KRPMNRVAKP
P. AERUGINOSA(X78413)           APPGPIYLGN SGTSMRLLSG LLAAQPFDST LTGD..ASLS KRPMNRVAKP
P. AERUGINOSA(AE004740_5)       APPGPIYLGN SGTSMRLLSG LLAAQPFDST LTGD..ASLS KRPMNRVAKP
P. AERUGINOSA(G5712093)         PAPGPIYLGN SGTSMRLLSG LLAAQDFDST LTGD..ASLS KRPMNRVANP
M. CATARRHALIS(MCA101591)       APKTPLYMGN SGTSMRLLAG ILSAQAFDSV MTGD..VSLS QRPMERVAVP
S. AUREUS(AP003362_152)         TPHQVLYTGN SGTTTRLLAG LLSGLGIESV LSGD..VSIG KRPMDRVLRP
S. AUREUS(AP003134_5)           TPHQVLYTGN SGTTTRLLAG LLSGLGIESV LSGD..VSIG KRPMDRVLRP
S. AUREUS(BAB57626)             TPHQVLYTGN SGTTTRLLAG LLSGLGIESV LSGD..VSIG KRPMDRVLRP
S. AUREUS(Q99U25)               TPHQVLYTGN SGTTTRLLAG LLSGLGIESV LSGD..VSIG KRPMDRVLRP
S. AUREUS(L05004_2)             TPHQVLYTGN SGTTTRLLAG LLSGLGNESV LSGD..VSIG KRPMDRVLRP
S. AUREUS(AROA_STAAU)           TPHQVLYTGN SGTTTRLLAG LLSGLGNESV LSGD..VSIG KRPMDRVLRP
C. JEJUNI(AROA_CAMJE)           SPNCILDCGN SGTAMRLMIG FLAGISGFFV LSGD..KYLN NRPMRRISKP
C. JEJUNI(AL139076_205)         SPNCILDCGN SGTAMRLMIG FLAGISGFFV LSGD..KYLN NRPMRRISKP
C. JEJUNI(X89371_2)             SPNCILDCGN SGTAMRLMIG FLAGISGFFV LSGD..KYLN NRPMRRISKP
MAIZE                           KEEVQLFLGN AGTAMRPLTA AVTAAGGNAT YVLDGVPRMR ERPIGDLVVG
GA21                            KEEVQLFLGN AGIAMRSLTA AVTAAGGNAT YVLDGVPRMR ERPIGDLVVG
RICE                            KEEVQLFLGN AGTAMRPLTA AVTAAGGNAT YVLDGVPRMR ERPIGDLVVG
CANOLA                          KSDIELYLGN AGTAMRPLTA AVTAAGGNAS YVLDGVPRMR ERPIGDLVVG
TOMATO                          EEEIQLFLGN AGTAMRPLTA AVTVAGGHSR YVLDGVPRMR ERPIGDLVDG

                                 140        150        160        170        180
CP4                             LREMGVQVKS .EDGDRLPVT LRG.PKTPTP ITYRVPMASA QVKSAVLLAG
CP4 NK603 VARIANT               LREMGVQVKS .EDGDRLPVT LRG.PKTPTP ITYRVPMASA QVKSAVLLAG
P. AERUGINOSA(Q9HZ69)           LREMGAVIET .GPEGRPPMT IRG.GQRLTG MHYDMPMASA QVKSCLLLAG
P. AERUGINOSA(X78413)           LREMGAVIET .GPEGRPPMT IRG.GQRLTG MHYDMPMASA QVKSCLLLAG
P. AERUGINOSA(AE004740_5)       LREMGAVIET .GPEGRPPMT IRG.GQRLTG MHYDMPMASA QVKSCLLLAG
P. AERUGINOSA(G5712093)         LREMGAVIET .AAEGRPPMT IRG.GHKLKG LTYTMPMASA QVKSCLLLAG
M. CATARRHALIS(MCA101591)       LRNMGAKIQS TGKKGTAPLS ITG.SQTLNA IEYQLPVASA QIKSCLILAS
S. AUREUS(AP003362_152)         LKLMDANIEG .IEDNYTPLI IK..PSVIKG INYQMEVASA QVKSAILFAS
S. AUREUS(AP003134_5)           LKLMDANIEG .IEDNYTPLI IK..PSVIKG INYQMEVASA QVKSAILFAS
S. AUREUS(BAB57626)             LKLMDANIEG .IEDNYTPLI IK..PSVIKG INYQMEVASA QVKSAILFAS
S. AUREUS(Q99U25)               LKLMDANIEG .IEDNYTPLI IK..PSVIKG INYQMEVASA QVKSAILFAS
S. AUREUS(L05004_2)             LKLMDANIEG .IEDNYTPLI IK..PSVIKG INYQMEVASA QVKSAILFAS
S. AUREUS(AROA_STAAU)           LKLMDANIEG .IEDNYTPLI IK..PSVIKG INYQMEVASA QVKSAILFAS
C. JEJUNI(AROA_CAMJE)           LTQIGARIYG RNEANLAPLC IE..GQKLKA FNFKSEISSA QVKTAMILSA
C. JEJUNI(AL139076_205)         LTQIGARIYG RNEANLAPLC IE..GQKLKA FNFKSEISSA QVKTAMILSA
C. JEJUNI(X89371_2)             LTQIGARIYG RNEANLAPLC IE..GQNLKA FNYKSEISSA QVKTAMILSA
MAIZE                           LKQLGADVDC FLGTDCPPVR VNGIGGLPGG KVKLSGSISS QYLSALLMAA
GA21                            LKQLGADVDC FLGTDCPPVR VNGIGGLPGG KVKLSGSISS QYLSALLMAA
RICE                            LKQLGADVDC FLGTECPPVR VKGIGGLPGG KVKLSGSISS QYLSALLMAA
CANOLA                          LKQLGADVEC TLGTNCPPVR VNANGGLPGG KVKLSGSISS QYLTALLMAA
TOMATO                          LKQLGAEVDC SLGTNCPPVR IVSKGGLPGG KVKLSGSISS QYLTALLMAA
```

Campylobacter

```
                                190        200                210        220
CP4                             LNTPGITTVI EPIMTR.... .DHTEKMLQG FGANLTVETD ADGVRTIRLE
CP4 NK603 VARIANT               LNTPGITTVI EPIMTR.... .DHTEKMLQG FGANPTVETD ADGVRTIRLE
P. AERUGINOSA(Q9HZ69)           LYAAGETSVT EPAPTR.... .DHTERMLRG FGYPVVVEG. ....STAKVE
P. AERUGINOSA(X78413)           LYAAGETSVT EPAPTR.... .DHTERMLRG FGYPVVVEG. ....STAKVE
P. AERUGINOSA(AE004740_5)       LYAAGETSVT EPAPTR.... .DHTERMLRG FGYPVVVEG. ....STAKVE
P. AERUGINOSA(G5712093)         LYAEGKTTVT EPAPTR.... .DHTERMLRG FGYPVSVNG. ....ATASVE
M. CATARRHALIS(MCA101591)       LWAKGTTTII EPEVSR.... .DHTERMLNA FGYPVQVDG. ....CKISVT
S. AUREUS(AP003362_152)         LFSNDTTVIK ELDVSR.... .NHTETMFRH FNIPIEAER. ..LSITTTPD
S. AUREUS(AP003134_5)           LFSNDTTVIK ELDVSR.... .NHTETMFRH FNIPIEAER. ..LSITTTPD
S. AUREUS(BAB57626)             LFSNDTTVIK ELDVSR.... .NHTETMFRH FNIPIEAER. ..LSITTTPD
S. AUREUS(Q99U25)               LFSNDTTVIK ELDVSR.... .NHTETMFRH FNIPIEAER. ..LSITTTPD
S. AUREUS(L05004_2)             LFSKEPTIIK ELDVSR.... .NHTETMFKH FNIPIEAEG. ..LSINTTPE
S. AUREUS(AROA_STAAU)           LFSKEPTIIK ELDVSR.... .NHTETMFKH FNIPIEAEG. ..LSINTTPE
C. JEJUNI(AROA_CAMJE)           FRADNVCTFS EISLSR.... .NHSENMLKA MKAPIRVSN. ..DGLSLEIN
C. JEJUNI(AL139076_205)         FRADNVCTFS EISLSR.... .NHSENMLKA MKAPIRVSN. ..DGLSLEIN
C. JEJUNI(X89371_2)             FRANNVCAFS EISLSR.... .NHSENMLKA MKAPIRVSN. ..DGLSLEIS
MAIZE                           PLALGDVEIE IIDKLISIPY VEMTLRLMER FGVKAEHSD. ..SWDRFYIK
GA21                            PLALGDVEIE IIDKLISIPY VEMTLRLMER FGVKAEHSD. ..SWDRFYIK
RICE                            PLALGDVEIE IIDKLISIPY VEMTLRLMER FGVKAEHSD. ..SWDRFYIK
CANOLA                          PLALGDVEIE IIDKLISVPY VEMTLKLMER FGVSAEHSD. ..SWDRFFVK
TOMATO                          PLALGDVEIE IIDKLISVPY VEMTLKLMER FGVFVEHSS. ..GWDRFLVK

                                  230        240        250        260        270
CP4                             GRGK.LTGQV IDVPGDPSST AFPLVAALLV PGSDVTILN. VLMNPTRTGL
CP4 NK603 VARIANT               GRGK.LTGQV IDVPGDPSST AFPLVAALLV PGSDVTILN. VLMNPTRTGL
P. AERUGINOSA(Q9HZ69)           SGHK.LSATH IEVPADISSA AFFLVAASIA EGSELVLQH. VGINPTRVGV
P. AERUGINOSA(X78413)           SGHK.LSATH IEVPADISSA AFFLVAASIA EGSELVLQH. VGINPTRVGV
P. AERUGINOSA(AE004740_5)       SGHK.LSATH IEVPADISSA AFFLVAASIA EGSELVLQH. VGINPTRVGV
P. AERUGINOSA(G5712093)         SGGK.LTATH IEVPGDISSS AFFLVAASIA EGSELVLEH. VGINPTRTGV
M. CATARRHALIS(MCA101591)       GGGR.LTATD IIVPADISSA AFFMVLGAIG GGEGLTIEK. VGMNPTRTGV
S. AUREUS(AP003362_152)         AIQH.IKPAD FHVPGDISSA AFFIVAALIT PESDVTIHN. VGINPTRSGI
S. AUREUS(AP003134_5)           AIQH.IKPAD FHVPGDISSA AFFIVAALIT PESDVTIHN. VGINPTRSGI
S. AUREUS(BAB57626)             AIQH.IKPAD FHVPGDISSA AFFIVAALIT PESDVTIHN. VGINPTRSGI
S. AUREUS(Q99U25)               AIQH.IKPAD FHVPGDISSA AFFIVAALIT PESDVTIHN. VGINPTRSGI
S. AUREUS(L05004_2)             AIRY.IKPAD FHVPGDISSA AFFIVAALIT PGSDVTIHN. VGINQTRSGI
S. AUREUS(AROA_STAAU)           AIRY.IKPAD FHVPGDISSA AFFIVAALIT PGSDVTIHN. VGINQTRSGI
C. JEJUNI(AROA_CAMJE)           PLKKPLKAQN IIIPNDPSSA FYFVLAAIIL PKSQIILKN. ILLNPTRIEA
C. JEJUNI(AL139076_205)         PLKKPLKAQN IIIPNDPSSA FYFVLAAIIL PKSQIILKN. ILLNPTRIEA
C. JEJUNI(X89371_2)             PLKKPLKAQN IIIPNDPSSA FYFALAAIIL PKSQIILKN. ILLNPTRIEA
MAIZE                           GGQKYKSPKN AYVEGDASSA SYFLAGAAIT GGTVTVEGCG TTSLQGDVKF
GA21                            GGQKYKSPKN AYVEGDASSA SYFLAGAAIT GGTVTVEGCG TTSLQGDVKF
RICE                            GGQKYKSPGN AYVEGDASSA SYFLAGAAIT GGTVTVQGCG TTSLQGDVKF
CANOLA                          GGQKYKSPGN AYVEGDASSA SYFLAGAAIT GETVTVEGCG TTSLQGDVKF
TOMATO                          GGQKYKSPGK AFVEGDASSA SYFLAGAAVT GGTVTVEGCG TSSLQGDVKF

                                280        290        300        310        320
CP4                             ILTLQEMGAD IEVINPRLAG GEDVADLRVR SS.TLKGVTV PEDRAPSMID
CP4 NK603 VARIANT               ILTLQEMGAD IEVINPRLAG GEDVADLRVR SS.TLKGVTV PEDRAPSMID
P. AERUGINOSA(Q9HZ69)           IEILRLMGGD LSLENQREVG GEPVADIRVR SA.RLKGIDI PEDLVPLAID
P. AERUGINOSA(X78413)           IEILRLMGGD LSLENQREVG GEPVADIRVR SA.RLKGIDI PEDLVPLAID
P. AERUGINOSA(AE004740_5)       IEILRLMGGD LSLENQREVG GEPVADIRVR SA.RLKGIDI PEDLVPLAID
```

```
P. AERUGINOSA(G5712093)     IDILRLMGAD ITLENQREVG GEPVADLRVR AA.KLKGIEI PEALVPLAID
M. CATARRHALIS(MCA101591)   IDILTLMGAD ITVMNEAVVG GEPIADITVR PS.DLHGIDI PEHLVPLAID
S. AUREUS(AP003362_152)     IDIVEKMGGN IQLFNQT.TG AEPTASIRIQ YTPMLQPITI EGELVPKAID
S. AUREUS(AP003134_5)       IDIVEKMGGN IQLFNQT.TG AEPTASIRIQ YTPMLQPITI EGELVPKAID
S. AUREUS(BAB57626)         IDIVEKMGGN IQLFNQT.TG AEPTASIRIQ YTPMLQPITI EGELVPKAID
S. AUREUS(Q99U25)           IDIVEKMGGN IQLFNQT.TG AEPTASIRIQ YTPMLQPITI EGELVPKAID
S. AUREUS(L05004_2)         IDIVEKMGGN IQLFNQT.TG AEPTASIRIQ YTPMLQPITI EGELVPKAID
S. AUREUS(AROA_STAAU)       IDIVEKMGGN IQLFNQT.TG AEPTASIRIQ YTPMLQPITI EGELVPKAID
C. JEJUNI(AROA_CAMJE)       YKILQKMGAK LEMTITQ.ND FETIGEIRVE SS.KLNGIEV KDN.IAWLID
C. JEJUNI(AL139076_205)     YKILQKMGAK LEMTITQ.ND FETIGEIRVE SS.KLNGIEV KDN.IAWLID
C. JEJUNI(X89371_2)         YKILQKMGAK LEMTITQ.ND FETIGEIRVE SS.KLNGIEV KDN.IAWLID
MAIZE                       AEVLEMMGAK VTWTETSVTV TGPPREPFGR KH..LKAIDV NMN...KMPD
GA21                        AEVLEMMGAK VTWTETSVTV TGPPREPFGR KH..LKAIDV NMN...KMPD
RICE                        AEVLEMMGAK VTWTDTSVTV TGPPREPYGK KH..LKAVDV NMN...KMPD
CANOLA                      AEVLEKMGCK VSWTENSVTV TGPSRDAFGM RH..LRAVDV NMN...KMPD
TOMATO                      AEVLEKMGAE VTWTENSVTV KGPPRNSSGM KH..LRAIDV NMN...KMPD

                             330        340        350        360        370
CP4                         EYPILAVAAA FAEGATVMNG LEELRVKESD RLSAVANGLK LNGVDCDEGE
CP4 NK603 VARIANT           EYPILAVAAA FAEGATVMNG LEELRVKESD RLSAVANGLK LNGVDCDEGE
P. AERUGINOSA(Q9HZ69)       EFPVLFVAAA CAEGRTVLRG AEELRVKESD RIQVMADGLK ALGVKAEPTP
P. AERUGINOSA(X78413)       EFPVLFVAAA CAEGRTVLRG AEELRVKESD RIQVMADGLK ALGVKAEPTP
P. AERUGINOSA(AE004740_5)   EFPVLFVAAA CAEGRTVLRG AEELRVKESD RIQVMADGLK ALGVKAEPTP
P. AERUGINOSA(G5712093)     EFPVLFVAAA CAEGRTVLTG AEELRVKESD RIQVMADGLL ALGVKCE...
M. CATARRHALIS(MCA101591)   EFPILFIAAS CAYGVTKLTG AKELRVKESD RIQVMADGLA TLGIDSKVLE
S. AUREUS(AP003362_152)     ELPVIALLCT QAVGTSTIKD AEELKVKETN RIDTTADMLN LLGFELQPTN
S. AUREUS(AP003134_5)       ELPVIALLCT QAVGTSTIKD AEELKVKETN RIDTTADMLN LLGFELQPTN
S. AUREUS(BAB57626)         ELPVIALLCT QAVGTSTIKD AEELKVKETN RIDTTADMLN LLGFELQPTN
S. AUREUS(Q99U25)           ELPVIALLCT QAVGTSTIKD AEELKVKETN RIDTTADMLN LLGFELQPTN
S. AUREUS(L05004_2)         ELPVIALLCT QAVGTSTIKD AEELKVKETN RIDTTADMLN LLGFELQPTN
S. AUREUS(AROA_STAAU)       ELPVIALLCT QAVGTSTIKD AEELKVKETN RIDTTADMLN LLGFELQPTN
C. JEJUNI(AROA_CAMJE)       EAPALAIAFA LAKGKSSLIN AKELRVKESD RIAVMVENLK LCGVEARELD
C. JEJUNI(AL139076_205)     EAPALAIAFA LAKGKSSLIN AKELRVKESD RIAVMVENLK LCGVEARELD
C. JEJUNI(X89371_2)         EAPALAIAFA LAKGKSSLIN AKELRVKESD RIAVMVENLK LCGVEARELD
MAIZE                       VAMTLAVVAL FADGPTAIRD VASWRVKETE RMVAIRTELT KLGASVEEGP
GA21                        VAMTLAVVAL FADGPTAIRD VASWRVKETE RMVAIRTELT KLGASVEEGP
RICE                        VAMTLAVVAL FADGPTAIRD VASWRVKETE RMVAIRTELT KLGASVEEGP
CANOLA                      VAMTLAVVAL FADGPTTIRD VASWRVKETE RMIAICTELR KLGATVEEGS
TOMATO                      VAMTLAVVAL FADGPTTIRD VASWRVKETE RMIAICTELR KLGATVVEGS

                             380        390        400        410        420
CP4                         TSLVVRGRP. DGKGLGNASG AAVATHLDHR IAMSFLVMGL VSENPVTVDD
CP4 NK603 VARIANT           TSLVVRGRP. DGKGLGNASG AAVATHLDHR IAMSFLVMGL VSENPVTVDD
P. AERUGINOSA(Q9HZ69)       DGIVIEG... .....GAFGG GEVWAHGDHR IAMSFSVASL RASGPIRIHD
P. AERUGINOSA(X78413)       DGIVIEG... .....GAFGG GEVWAHGDHR IAMSFSVASL RASGPIRIHD
P. AERUGINOSA(AE004740_5)   DGIVIEG... .....GAFGG GEVWAHGDHR IAMSFSVASL RASGPIRIHD
P. AERUGINOSA(G5712093)     .......... .......... .......... .......... ..........
M. CATARRHALIS(MCA101591)   DGIIIQGKGI QGEKNAIFGG GVIESHHDHR IAMSFAVASS RATDDIIIQG
S. AUREUS(AP003362_152)     DGLIIHPS.. .....EFKTN ATVDSLTDHR IGMMLAVASL LSSEPVKIKQ
S. AUREUS(AP003134_5)       DGLIIHPS.. .....EFKTN ATVDSLTDHR IGMMLAVASL LSSEPVKIKQ
S. AUREUS(BAB57626)         DGLIIHPS.. .....EFKTN ATVDSLTDHR IGMMLAVASL LSSEPVKIKQ
S. AUREUS(Q99U25)           DGLIIHPS.. .....EFKTN ATVDSLTDHR IGMMLAVASL LSSEPVKIKQ
```

```
S. AUREUS(L05004_2)          DGLIIHPS.. .....EFKTN AT.DILTDHR IGMMLAVACV LSSEPVKIKQ
S. AUREUS(AROA_STAAU)        DGLIIHPS.. .....EFKTN AT.DILTDHR IGMMLAVACV LSSEPVKIKQ
C. JEJUNI(AROA_CAMJE)        DGFEIEGG.. .....CELKS SKIKSYGDHR IAMSFAILGL LCG..IEIDD
C. JEJUNI(AL139076_205)      DGFEIEGG.. .....CELKS SKIKSYGDHR IAMSFAILGL LCG..IEIDD
C. JEJUNI(X89371_2)          DGFEIEGG.. .....CELKS SKIKSYGDHR IAMSFAILGL LCG..IEIDD
MAIZE                        DYCIITPP.. .....EKLNV TAIDTYDDHR MAMAFSLAAC AEVP.VTIRD
GA21                         DYCIITPP.. .....EKLNV TAIDTYDDHR MAMAFSLAAC AEVP.VTIRD
RICE                         DYCIITPP.. .....EKLNI TAIDTYDDHR MAMAFSLAAC ADVP.VTIRD
CANOLA                       DYCVITPP.. .....AKVKP AEIDTYDDHR MAMAFSLAAC ADVP.VTIKD
TOMATO                       DYCIITPP.. .....EKLNV TEIDTYDDHR MAMAFSLAAC ADVP.VTIKN

                               430        440
CP4                          ATMIATSFPE FMDLMAGLGA KILSD.....
CP4 NK603 VARIANT            ATMIATSFPE FMDLMAGLGA KIELSDTKAA
P. AERUGINOSA(Q9HZ69)        CANVATSFPN FLALCAQTGI RVAVENN...
P. AERUGINOSA(X78413)        CANVATSFPN FLALCAQTGI RV........
P. AERUGINOSA(AE004740_5)    CANVATSFPN FL..... ..........
P. AERUGINOSA(G5712093)      .......... .......... ..........
M. CATARRHALIS(MCA101591)    TETVNTSFPN FAEL..... ..........
S. AUREUS(AP003362_152)      FDAVNVSFPG FLPKLKLLEN EG........
S. AUREUS(AP003134_5)        FDAVNVSFPG FLPKLKLLEN EG........
S. AUREUS(BAB57626)          FDAVNVSFPG FLPKLKLLEN EG........
S. AUREUS(Q99U25)            FDAVNVSFPG FLPKLKLLEN EG........
S. AUREUS(L05004_2)          FDAVNVSFPG FLPKLKLLQN EG........
S. AUREUS(AROA_STAAU)        FDAVNVSFPG FLPKLKLLQN EG........
C. JEJUNI(AROA_CAMJE)        SDCIKTSFPN FIEILSNLGA RIDY......
C. JEJUNI(AL139076_205)      SDCIKTSFPN FIEILSNLGA RIDY......
C. JEJUNI(X89371_2)          SDCIKTSFPN FIEILSNLGA RIDY......
MAIZE                        PGCTRKTFPD YFDVLSTFVK N.........
GA21                         PGCTRKTFPD YFDVLSTFVK N.........
RICE                         PGCTRKTFPN YFDVLSTFVR N.........
CANOLA                       PGCTRKTFPD YFQVLESITK H.........
TOMATO                       PGCTRKTFPD YFEVLQKYSK H.........
```

10.0 Appendix B - Calculations for Safety Margins

Calculations for Table 2 - Acute Dose Comparisons

(1) CP4 EPSPS

Levels of CP4 EPSPS protein in corn grain (1st year US field trials, MSL 16278) is 15.6 ug/gram (highest production level in grain, mean production across sites is 10.9 ug/gram)

Human consumption of grain at the 50th percentile is 0.2 grams/kg body weight (average adult consumption for US adults[a])

Human consumption of CP4 EPSPS at 50th percentile = 0.2 grams/kg x 15.6 µg/gram = 3.12 µg/kg or 0.00312 mg/kg

Safety margin at 50th percentile consumption of grain = 572[b] mg/kg/0.00312mg/kg = 183,333

Human consumption of grain at the 90th percentile is 0.63 grams/kg body weight (average adult consumption for US adults[a])

Human consumption of CP4 EPSPS at 90th percentile = 0.63 grams/kg x 15.6 µg/gram = 9.8 µg/kg or 0.0098 mg/kg

Safety margin at 90th percentile consumption of grain = 572 mg/kg/0.0098 mg/kg = 58,367

(2) mEPSPS

Levels of mEPSPS protein in grain is, on average, 3.2 µg/gram.

Human consumption of grain at the 50th percentile is 0.2 grams/kg body weight (average adult consumption for US adults[a])

Human consumption of mEPSPS at the 50th percentile = 0.2 grams/kg x 3.2 µg/gram = 0.64 µg/kg or 0.00064 mg/kg

Safety margin at the 50th percentile consumption of grain = 45.6[c] mg/kg/0.00064 mg/kg = 71,250.

Human consumption of grain at the 90th percentile is 0.63 grams/kg body weight (average adult consumption for US adults[a])

[a] Analysis based on NFCS 1989-92 USDA food intake survey.
[b] Harrison *et al.*, 1996.
[c] Naylor, 1997.

Human consumption of mEPSPS at the 90th percentile = 0.63 grams/kg x 3.2 µg/gram = 2.0 µg/kg or 0.002 mg/kg

Safety margin at the 90th percentile consumption of grain = 45.6 mg/kg/0.002 mg/kg = 22,850

Calculations for Table 3 - Feeding Study Exposure Comparisons

(3) Poultry feeding study (data from Monsanto feeding studies)

Corn consumption = 57 grams/kg body weight/day

Levels of CP4 EPSPS protein in corn grain (1st year US field trials, MSL 16278) is 15.6 ug/gram (highest production level in grain, mean production across sites is 10.9 ug/gram)

Chicken consumption of CP4 EPSPS is 57 grams/kg x 15.6 ug/gram = 889 ug/kg/day

Human consumption of grain at the 50th percentile is 0.2 grams/kg body weight (average adult consumption for US adults[a])

Human consumption of CP4 EPSPS at 50th percentile = 0.2 grams/kg x 15.6 µg/gram = 3.12 µg/kg

Safety margin at the 50th percentile consumption of grain = 889 ug/kg/3.12 µg/kg = 285

Human consumption of grain at the 90th percentile is 0.63 grams/kg body weight (average adult consumption for US adults[a])

Human consumption of CP4 EPSPS at 90th percentile = 0.63 grams/kg x 15.6 µg/gram = 9.8 µg/kg.

Safety margin at the 50th percentile consumption of grain = 889 ug/kg/9.8µg/kg = 91

(4) Swine feeding study

Corn consumption = 23 grams/kg body weight/day

Levels of CP4 EPSPS protein in corn grain (1st year US field trials, MSL 16278) is 15.6 ug/gram (highest production level in grain, mean production across sites is 10.9 ug/gram)

Swine consumption of CP4 EPSPS is 23 grams/kg x 15.6 ug/gram = 359 ug/kg/day

Human consumption of grain at the 50th percentile is 0.2 grams/kg body weight (average adult consumption for US adults[a])

Human consumption of CP4 EPSPS at the 50th percentile = 0.2 grams/kg x 15.6 µg/gram = 3.12 µg/kg

Safety margin at the 50th percentile consumption of grain = 359 ug/kg/3.12 µg/kg = 115

Human consumption of grain at the 90th percentile is 0.63 grams/kg body weight (average adult consumption for US adults[a])

Human consumption of CP4 EPSPS at the 90th percentile = 0.63 grams/kg x 15.6 µg/gram = 9.8 µg/kg

Safety margin at the 90th percentile consumption of grain = 359 ug/kg/9.8 µg/kg = 37

(5) rat feeding study

Corn consumption = 21 grams/kg body weight/day

Levels of CP4 EPSPS protein in corn grain (1st year US field trials, MSL 16278) is 15.6 ug/gram (highest production level in grain, mean production across sites is 10.9 ug/gram)

Rat consumption of CP4 EPSPS is 21 grams/kg x 15.6 ug/gram = 328 ug/kg/day

Human consumption of grain at the 50th percentile is 0.2 grams/kg body weight (average adult consumption for US adults[a])

Human consumption of CP4 EPSPS at the 50th percentile = 0.2 grams/kg x 15.6 µg/gram = 3.12 µg/kg

Safety margin at the 50th consumption of grain = 328 ug/kg/3.12 µg/kg = 105

Human consumption of grain at the 90th percentile is 0.63 grams/kg body weight (average adult consumption for US adults[a])

Human consumption of CP4 EPSPS at the 90th percentile = 0.63 grams/kg x 15.6 µg/gram = 9.8 µg/kg

Safety margin at the 90th consumption of grain = 328 ug/kg/9.8 µg/kg = 33

11.0 Appendix C - Preparation and Characterization of Purified CP4 EPSPS L214P Protein

An *E. coli* heterlogous expression plasmid containing the *cp4 epsps L214P* gene was prepared. *E. coli* bacteria were transformed with the *cp4 epsps L214P* vector. The DNA sequence of the *cp4 epsps L214P* gene insert was confirmed prior to bacterial transformation. A small-scale fermentation of the newly transformed *E. coli* was conducted. The bacterial culture was grown in a shake flask at 37 °C in the presence of chloramphenicol and ampicillin for 5 hours. The temperature was dropped to 20 °C and was induced by IPTG. The culture was allowed to grow for an additional 5 hours. The cell paste from the fermentation was lysed using a French press apparatus. The protein, CP4 EPSPS L214P, was isolated from the cell lysate using a combination of hydrophobic interaction and anion exchange chromatography using Phenyl Sepharose 6 Fast Flow and Mono Q HR10/10 resins, respectively (Heeren *et al.,* 1993). The CP4 EPSPS L214P fraction was pooled, concentrated to 2.3 mg/ml protein concentration and assigned the lot number 7070519. The protein was found to be 98% pure based on densitometric analysis of a Colloidal-blue stained SDS-PAGE gel. The apparent molecular weight of the CP4 EPSPS L214P protein was estimated at ~46 kD based on relative electrophoretic mobility on SDS-PAGE. This compares very favorably with a theoretical MW of 47.6 kD. The identity of the CP4 EPSPS L214P protein was confirmed by N-terminal sequence data, western blot analysis and by MALDI-TOF mass spectroscopy. The specific activity of the CP4 EPSPS L214P protein was 4.24 Units/mg protein based on analysis using the standard phosphate release assay (Burnette and Holden, 1994). This compares very closely with a specific activity of 4.20 U/mg protein for the *E. coli*-produced CP4 EPSPS standard (lot # 5192245) when analyzed concurrently.

All experimental data was archived along with the data file for report, MSL-17600.

12.0 Appendix D – Digestibility of CP4 EPSPS L214P *in vitro*

Summary

The purpose of this experiment was to assess the *in vitro* digestibility of purified CP4 EPSPS L214P in simulated gastric fluid. Digestibility was assessed by SDS-PAGE gel staining. The results of this experiment show that CP4 EPSPS L214P protein was rapidly digested after incubation in SGF.

Test Substance

The test substance for this experiment was CP4 EPSPS L214P protein (Lot 7070519). This material was characterized to have a total protein concentration of 2.3 mg/ml in storage buffer [50 mM Tris, pH 7.5, 50 mM KCl, 2 mM DTT, and 25% (v/v) glycerol] with a purity of 98%. Prior to use in this experiment, the protein was diluted to 1.68 mg/ml with deionized water.

Test System

The test system for this experiment was simulated gastric fluid (SGF). SGF preparation was based on the recommendation of SOP BR-ME-0460-01 with the exception that a highly purified form of pepsin (Sigma P/N P-6887, Lot 99H7665) was used, so the amount of pepsin used to formulate SGF was reduced to provide a digestion environment of 10 pepsin activity units per 1 µg of test substance total protein. The pepsin concentration determined for SGF preparation was calculated from the specific activity on the product label (3,460 units/mg protein at 100% purity). The final formulation of SGF was 0.20% (w/v) sodium chloride, 0.26 mg/ml pepsin, pH to 1.2 with HCl.

Experimental Design

A schematic of the digestion experiment is shown below in Figure 2.

Digestibility of the Test Substance in SGF

The test substance was used in this experiment to generate an incubation time course of CP4 EPSPS L214P in SGF.

Digestions were prepared by adding CP4 EPSPS L214P protein to tubes containing SGF. Digestions were conducted so that 10 pepsin activity units were used per 1 µg of test substance total protein. Digestions were incubated at 37 °C in separate tubes for each of the incubation times. SGF digestions were quenched by addition of 0.2 M sodium carbonate to the test system. This has been shown in previous experiments to be an appropriate method of quenching SGF activity (Astwood *et al.*, 1996).

Zero incubation time points (T = 0) were quenched by addition of 0.2 M sodium carbonate to SGF prior to addition of the test substance.

The incubation time points generated in this experiment were 0, 15, 30 sec, and 1, 2, 4, 8, 15, 30, and 60 min.

Experimental Controls

Experimental controls were prepared to characterize the stability of the test substance in the test system (SGF) lacking pepsin. Volumes were the same as those used to prepare digestions of CP4 EPSPS L214P protein in SGF, but only 0 and 60 minute incubation time points were generated.

Experimental controls were prepared to characterize the test system (SGF) lacking the test substance, by adding CP4 EPSPS L214P protein to SGF. These experimental controls were prepared in a similar manner as described above. Volumes were the same as those used to prepare digestions of CP4 EPSPS L214P protein in SGF, but only 0 and 60 minute incubation time points were generated.

Analytical Methods

The digestibility of purified CP4 EPSPS L214P protein in SGF was assessed using SDS-PAGE gel staining. Gel staining methods are commonly used and extensively referenced in the scientific literature relevant to the purposes for which they are being employed (Deutcher, 1990; Schägger and von Jagow, 1987; Neuhoff *et al.*, 1988).

Digestive Fluid Activity Assay

Following SOP BR-ME-0460-01 with the exceptions noted in file, the activity of SGF was assessed before generating digestions of CP4 EPSPS L214P protein. This assessment of SGF activity demonstrated that the test system was active and appropriate for use in this experiment.

SDS-PAGE

Samples from the SGF *in vitro* digestion of purified CP4 EPSPS L214P protein were analyzed by SDS-PAGE using pre-cast 10-20% tricine mini-gels (NOVEX). This procedure is described in SOP BtC-PRO-026-01 with the following modifications. All SDS-PAGE runs conducted during this experiment used NOVEX brand tricine gels run with tricine buffers. For tricine gels, the upper buffer reservoir was filled with 1X Novex® Tricine SDS Running Buffer (InVitrogen, Carlsbad, CA) and the lower buffer reservoir was filled with 200 mM Tris, pH ≈8.9. Tricine SDS-PAGE gels were used because they have been shown to provide optimum resolution of low molecular weight proteins (Schägger and von Jagow, 1987).

Colloidal Blue Gel Staining

After separation of proteins by SDS-PAGE, gels were stained using a colloidal blue dye. Prior to staining, the gels were incubated in fix solution [40% (v/v) methanol, 7% (v/v) acetic acid] for approximately 30 minutes at room temperature. Gels were then stained overnight with Brilliant Blue G Colloidal dye (Sigma P/N B-2025) diluted 4:1 with methanol. Excess background was removed by washing the gels in destaining solution [25% (v/v) methanol, 10% (v/v) acetic acid] for approximately 1 minute followed by several washes in 25% (v/v) methanol.

This method was selected because it is sensitive to low amounts of protein (Neuhoff *et al.*, 1988).

Assessment of Digestibility by Colloidal Blue Gel Staining

The digestibility of CP4 EPSPS L214P protein was assessed by staining SDS-PAGE gels with colloidal blue. Based on the purity and predigestion concentration of the test substance, each lane was loaded with 544 ng of CP4 EPSPS L214P protein. Experimental controls without CP4 EPSPS L214P protein were loaded with the same volume as those containing CP4 EPSPS L214P protein, so that all other components of the digestion sample would be loaded at the same level. CP4 EPSPS L214P protein was digested below the limit of detection within 15 seconds in SGF (See below - Figure 1, lane 5).

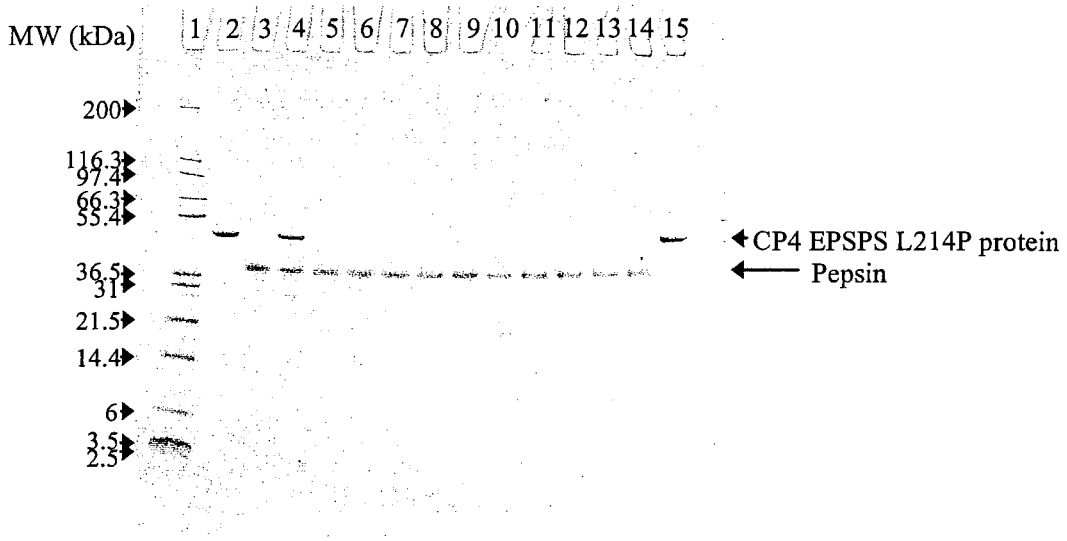
Experimental controls lacking pepsin (See below - Figure 1, lanes 2 and 15) demonstrated that degradation of CP4 EPSPS L214P protein was attributed to digestion by SGF and not instability of the test substance in pH 1.2 while incubating at 37 °C. Experimental controls without CP4 EPSPS L214P protein (See below - Figure 1, lanes 3 and 14) demonstrate a consistently low background for digestion samples incubated from 0 to 60 minutes.

A limit of detection was not determined for this analysis. However, the manufacturer (Sigma, St. Louis) reports a sensitivity comparable to silver staining, and this method has demonstrated a sensitivity of 10 ng or lower in various studies at Monsanto.

Conclusions

The results of this experiment showed that CP4 EPSPS L214P protein was rapidly digested after incubation in SGF. The sensitivity of the analytical methods employed demonstrated that approximately 98% or greater of the CP4 EPSPS L214P protein was digested within 15 seconds in SGF.

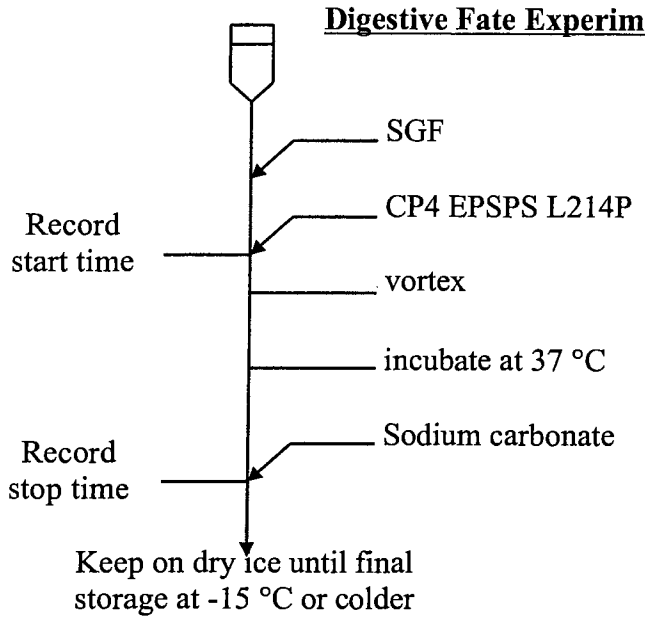
Appendix D - Figure 1.



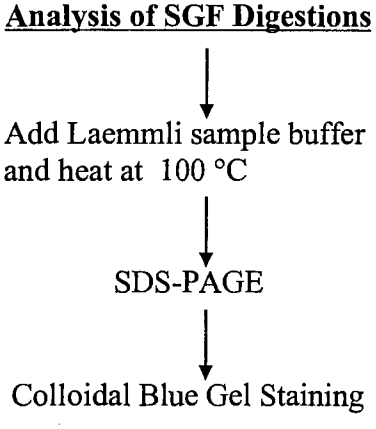
Colloidal Blue Stained Gel Showing the Digestion of Purified CP4 EPSPS L214P Protein in Simulated Gastric Fluid. Proteins were separated by SDS-PAGE using a 10→20% polyacrylamide gradient in a tricine buffered gel. Proteins were detected by staining with Brilliant Blue G stain. CP4 EPSPS L214P protein was loaded at ≈544 ng per lane based on purity corrected and pre-digestion estimates.

Lane	Description	Incubation time
1	Molecular weight markers	
2	Experimental control w/o pepsin (P0)	0 s
3	Experimental control w/o CP4 EPSPS L214P protein (N0)	0 s
4	CP4 EPSPS L214P protein in SGF, T = 0	0 s
5	CP4 EPSPS L214P protein in SGF, T = 1	15 s
6	CP4 EPSPS L214P protein in SGF, T = 2	30 s
7	CP4 EPSPS L214P protein in SGF, T = 3	1 min
8	CP4 EPSPS L214P protein in SGF, T = 4	2 min
9	CP4 EPSPS L214P protein in SGF, T = 5	4 min
10	CP4 EPSPS L214P protein in SGF, T = 6	8 min
11	CP4 EPSPS L214P protein in SGF, T = 7	15 min
12	CP4 EPSPS L214P protein in SGF, T = 8	30 min
13	CP4 EPSPS L214P protein in SGF, T = 9	60 min
14	Experimental control w/o CP4 EPSPS L214P protein (N9)	60 min
15	Experimental control w/o pepsin (P9)	60 min

Appendix D - Figure 2



Experimental controls are prepared by adding deionized water to SGF in place of the test substance. Additional experimental controls were prepared by adding the test substance to SGF lacking pepsin.



Study Title

PCR and DNA Sequence Analysis of the Insert in Roundup Ready® Maize Event NK603

Authors

Niki K. Kesterson
Steven E. Reiser
Tracey A. Cavato
Ronald P. Lirette

Study Completed On

January 10, 2002

Performing Laboratory

Monsanto Company
Product Characterization Center
Biotechnology Regulatory Sciences
700 Chesterfield Parkway North
St. Louis, MO 63198

Laboratory Project ID

Study 01-01-46-25
MSL-17588

[redacted]).

Information claimed confidential on the basis of its falling within the scope of FIFRA section 10 (d)(1)(A), (B), or (C) has been removed to a confidential appendix, and is cited by cross reference number in the body of the study.

"We submit this material to the United States Environmental Protection Agency specifically under requirements set forth in FIFRA as amended, and consent to the use and disclosure of this material by EPA strictly in accordance with FIFRA. By submitting this material to the EPA in accordance with the method and format requirements contained in PR Notice 86-5, we reserve and do not waive any rights involving this material that are or can be claimed by the company notwithstanding this submission to EPA."

Company: Monsanto Company

Company Agent: _____

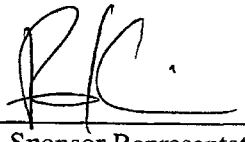
Title: _____

Signature: _____ Date: _____

Statement of Compliance

This study meets the requirements under GLP as specified in 40 CFR Part 160 with the following exception:

Sequence information, generated by the Monsanto Genomic Sequencing Center, was not generated in compliance with the GLP regulations, however all experiments conducted to confirm sequence data within this report were performed in compliance with the GLP regulations.

Submitter	Date
	1/10/02
Sponsor Representative	Date
	10-Jan-02
Study Director	Date

Quality Assurance Statement

Study Title: PCR and DNA Sequence Analysis of the Insert in Roundup Ready® Maize Event NK603

Study Number: 01-01-46-25

Reviews conducted by the Quality Assurance Unit confirm that the final report accurately describes the methods and standard operating procedures followed and accurately reflects the raw data of the study.

Following is a list of reviews conducted by the Monsanto AG Regulatory QAU on the study reported herein.

Dates Of Inspection / Audit	Phase	Date Reported To: Study Director	Date Reported To: Management
November 12, 2001	PCR/Sequence	November 21, 2001	November 21, 2001
January 8, 2002	Raw Data Audit	January 10, 2002	January 10, 2002
January 8, 2002	Draft Report Audit	January 10, 2002	January 10, 2002

C. Marie Braton

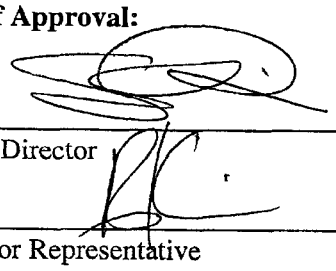
C. Marie Braton
Quality Assurance
Monsanto Regulatory, Monsanto Company

January 10, 2002
Date

Signatures of Approval

Study Number:	01-01-46-25
Title:	PCR and DNA Sequence Analysis of the Insert in Roundup Ready® Maize Event NK603
Primary Facility:	Monsanto Company 700 Chesterfield Parkway North St. Louis, MO 63198
Monsanto Genome Sequencing Center:	Monsanto Company Genomics Sequencing Center 800 North Lindbergh Boulevard St. Louis, MO 63167
Sponsor Representative:	Regulatory Affairs, Maize Team
Study Director:	Steven E. Reiser
Contributors:	Niki K. Kesterson and Tracey A. Cavato
Study Initiation Date:	November 9, 2001
Study Completion Date:	January 10, 2002
Records Retention:	All study specific raw data, protocols, final report and facility records will be retained at Monsanto, St. Louis.
Sample Storage:	Any study samples that are to be retained will be stored at Monsanto, St. Louis.

Signatures of Approval:



10-Jan-02
Study Director — Date

11/10/02
Sponsor Representative — Date

Table of Contents

Table of Contents (Cont'd)

Abbreviations

~	approximately
CTP2	chloroplast transit peptide
CP4 EPSPS	enzyme 5-enolpyruvylshikimate-3-phosphate synthase isolated from *Agrobacterium sp.* strain CP4
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
e35S	cauliflower mosaic virus (CaMV) promoter with the duplicated enhancer
EPSPS	5-enolpyruvylshikimate-3-phosphate synthase
NOS 3'	nopaline synthase 3' polyadenylation sequence
P-ract1/ract1 intron	rice actin promoter and intron
PCR	polymerase chain reaction
V	volts
*Zm*HSP70 intron	maize (*Zea mays*) *hsp70* gene (heat-shock protein)

I. SUMMARY

The molecular characterization of Roundup Ready® maize event NK603 has been previously reported (Deng *et al.*, 1999). This characterization, largely based on Southern blot analysis, demonstrated that one copy of the DNA restriction fragment used for transformation was present in maize event NK603, along with a 217-bp segment containing a portion of the enhancer region of the rice actin promoter inversely linked to the 3' end of the inserted transformation cassette. In the current study, PCR analyses and subsequent DNA sequencing were performed on the insert in Roundup Ready maize event NK603. These analyses confirmed the results of the previous characterization by demonstrating the expected linkage of the elements contained within the insert in maize event NK603. Furthermore, the DNA sequence of the entire insert in Roundup Ready maize event NK603 was obtained and is reported.

II. INTRODUCTION

A. Background. Roundup Ready maize plants were produced by insertion of a sequence encoding the CP4 EPSPS protein into the maize genome. The protein encoded by this sequence confers tolerance to glyphosate, the active ingredient in Roundup® herbicide. Roundup Ready maize event NK603 was produced by particle acceleration technology using a linear DNA restriction fragment from the plasmid vector PV-ZMGT32 (Figure 1). The DNA restriction fragment used for transformation contained two CP4 EPSPS gene expression cassettes. The first cassette contains the CP4 EPSPS coding sequence under the regulation of the rice actin promoter, a rice actin intron, a chloroplast transit peptide (CTP2) sequence and a nopaline synthase 3' polyadenylation sequence (NOS 3'). The second EPSPS gene cassette contains the CP4 EPSPS coding sequence under the regulation of the enhanced 35S cauliflower mosaic virus promoter (e35S), a maize heat-shock protein 70 intron (*Zm*HSP70), CTP2, and the NOS 3' polyadenylation sequence. Previous molecular characterization of the insert in event NK603 (Deng *et al.*, 1999) demonstrated that one copy of the DNA restriction fragment used for transformation is present in event NK603, along with a 217-bp segment containing a portion of the enhancer region of the rice actin promoter inversely linked to the 3' end of the inserted transformation cassette (Figure 2). In addition, the genomic DNA sequences flanking the insert were previously identified and confirmed (Cavato *et al.*, 2001).

B. Purpose. The purpose of this study was to generate overlapping PCR products spanning the length of the DNA insert containing the CP4 EPSPS coding regions in

® Roundup and Roundup Ready are registered trademarks of Monsanto Technology, LLC., St. Louis, Missouri.

maize event NK603, and to sequence those PCR products in order to obtain the sequence of the entire insert within Roundup Ready maize event NK603.

III. MATERIALS AND METHODS

A. Test Substance. The test substance was maize event NK603 (Lot # RDR-0111-11849-S). DNA from the test substance was isolated from grain representing the F1 generation prior to the start of this study. The raw data detailing the extraction is archived with this study.

B. Control Substance. The control substance was the non-transgenic maize line B73 (Lot # TCP-0005-10313-I). The control substance DNA used in this study was previously isolated from grain under Study # 01-01-50-15.

C. Reference Substances. The reference substances included the plasmid PV-ZMGT32 (used to generate maize event NK603) which was used as a positive control in some of the PCR analyses. The PCR products generated from the plasmid were used as size indicators for the appropriate PCR analyses. The 500 bp DNA Ladder from Invitrogen was used to estimate band sizes in the PCR analyses.

D. Test System. There was no test system. This study used analytical methods to analyze the maize event.

E. Test, Control, and Reference Substance Characterization. The identity of the test substance was confirmed as part of this study and the data are archived with this study. The identity of the control substance and plasmid PV-ZMGT32 was confirmed prior to the start of the study and the data are archived with Study # 01-01-50-15.

F. DNA Isolation. The DNA from the test substance was isolated from grain using Qiagen's DNeasy Plant Miniprep Kit according to the manufacturer's instructions. DNA from the control substance was isolated under Study # 01-01-50-15. An exact copy of the raw data detailing the extraction and quantitation was archived with this study.

G. DNA Quantitation. Quantitation of the DNA samples was performed during the study using a Hoefer DyNA Quant 200 Fluorometer (San Francisco, CA)(SOP BR-EQ-0065-01) using Roche molecular size marker IX as a calibration standard when quantitating genomic DNA.

H. PCR Analysis and Sequence Confirmation of the Organization of the Insert in Roundup Ready Maize Event NK603. The linkage of the elements contained within the insert was confirmed by generating four overlapping PCR products spanning the

length of the insert (Products A-D, Figure 3). The PCR analyses were conducted multiple times using 25 ng of genomic DNA or 10 ng of PV-ZMGT32 plasmid DNA, when appropriate, as a template in a 50 µl reaction volume. PCR products A and C contained a final concentration of 1.5 mM $MgCl_2$, 0.4 µM of each primer, 100-200 µM each dNTP, and 2.5 units of Red*Taq* DNA polymerase (Sigma Chemical Co.). PCR products B and D contained a final concentration of 2.5 mM $MgCl_2$, 0.2 µM of each primer, 400 µM each dNTP, and 2.5 units of TaKaRa LA Taq DNA polymerase (TAKARA SHUZO CO., LTD.). The reactions for product A were performed under the following cycling conditions: 94°C for 3 minutes; 35 cycles at 94°C for 30 seconds, 66°C for 30 seconds, 72°C for 1.5 minutes; 1 cycle at 72°C for 10 minutes. The reactions for product B were performed under the following cycling conditions: 94°C for 3 minutes; 35 cycles at 94°C for 30 seconds, 64°C for 30 seconds, 72°C for 3 or 3.5 minutes; 1 cycle at 72°C for 10 minutes. The reactions for product C were performed under the following cycling conditions: 94°C for 3 minutes; 35 cycles at 94°C for 30 seconds, 60°C for 30 seconds, 72°C for 1.5 minutes; 1 cycle at 72°C for 10 minutes. The reactions for product D were performed under the following cycling conditions: 94°C for 3 minutes; 35 cycles at 94°C for 30 seconds, 60°C for 30 seconds, 72°C for 3 minutes; 1 cycle at 72°C for 10 minutes. The PCR products were separated according to SOP # GEN-PRO-003-01 or BR-ME-315-01 using 1 % agarose gel electrophoresis at 130-200 V for ~0.3-1.5 hours. Following electrophoresis, all PCR products derived from Roundup Ready maize event NK603 were excised from the gel and purified using the QIAquick Gel Extraction following the procedure supplied by the manufacturer, with minor modifications. The purified PCR products were then sequenced with the initial PCR primers as well as primers designed internal to the amplified sequences. All sequencing was performed by the Monsanto Genomics Sequencing Center using dye-terminator chemistry. A consensus sequence was created from acceptable sequence runs using DNAstar/SeqManII version 5.01.

In addition to the four overlapping PCR products spanning the length of the insert, a fifth PCR product was generated with 196 ng of genomic DNA as a template using the same reaction conditions described above for products B and D (except 0.4 µM of each primer was used) and the following cycling conditions: 94°C for 1 minute; 14 cycles at 98°C for 10 seconds, 68°C for 5 minutes; 16 cycles at 98°C for 10 seconds, 68°C for 5 minutes plus 15 seconds per cylce; 1 cycle at 72°C for 10 minutes (data not shown). A portion of this PCR product was sequenced with primers internal to the product to confirm a very small section of the rice actin promoter. This sequence is included in the consensus sequence of the insert in Roundup Ready corn event NK603 (Figure 4).

I. Data Not Reported or Rejected. Sequence data which were not interpretable or only contained small stretches of high quality data were rejected. In addition, PCR analyses

using RedTaq DNA polymerase in an attempt to generate products B and D were rejected since no products were generated.

IV. RESULTS AND DISCUSSION

Organization and Sequence of the Inserted DNA in Roundup Ready Maize Event NK603. The organization of the elements within the insert was confirmed using PCR analysis by amplifying four overlapping regions of DNA which span the entire length of the insert. The location of the PCR products generated in relation to the insert, as well as the results of the PCR analyses, are shown in Figure 3. The control reactions containing no template DNA (lanes 4, 9, 14 and 18) and the non-transgenic B73 control reactions (lanes 2, 6, 11 and 16) did not generate PCR products with any of the primer sets, as expected. The plasmid PV-ZMGT32 was used as a positive control in two PCR analyses, products B and C, since these amplicons were located entirely within the inserted DNA in event NK603. In the analyses generating products B and C, Roundup Ready maize event NK603 and plasmid PV-ZMGT32 each generated the expected size PCR products of 3212 bp for Product B (lanes 7 and 8) and 1306 bp for Product C (lanes 12 and 13). Roundup Ready maize event NK603 also produced the expected size PCR products of ~982 bp for Product A (lane 3) and ~3134 bp for Product D (lane 17). The generation of the predicted size PCR products from Roundup Ready maize event NK603 establishes that the arrangement or linkage of the elements in the insert is the same as those in plasmid PV-ZMGT32, and that the elements within the insert are arranged as depicted in the schematic of the insert in Figure 2.

All of the PCR products generated in Roundup Ready maize event NK603 were subjected to DNA sequencing to further confirm the organization of the elements within the insert. The consensus sequence representing the insert in Roundup Ready maize event NK603, generated by compiling numerous sequencing reactions performed on the four PCR products that spanned the length of the insert, as well as the short stretch of sequence obtained from the fifth PCR product, is shown in Figure 4. The DNA sequence of the insert contains 6923 bases with base one equal to base 150 of plasmid PV-ZMGT32 and base 6706 equal to base 6855 in plasmid PV-ZMGT32 (Figure 1). In addition, there are 217 bp containing a portion of the enhancer region of the rice actin promoter positioned in the inverse orientation represented by bases 6707-6923 in Figure 4. While the sequence of two of the PCR products (A and D) also contained sequence from the maize genomic DNA flanking the insert, this is not reported here. The sequence of the maize DNA flanking the 5' and 3' ends of the insert has been previously reported (Cavato *et al.*, 2001).

The deduced amino acid sequences of the two CP4 EPSPS proteins present in Roundup Ready maize event NK603 were compared with the sequence of plasmid vector PV-ZMGT32. The sequence of the CP4 EPSPS protein from the first gene cassette is

identical between the insert and the plasmid. The sequence of the CP4 EPSPS protein from the second gene cassette differs by one amino acid from that of the plasmid. The amino acid at position 214 is a leucine in the plasmid and a proline in the plant insert. In addition to the single amino acid change, there is one silent nucleotide change in the second CP4 EPSPS coding region in the plant compared to the plasmid. These changes were confirmed in F1 generation material. The F1 generation is the progenitor for all NK603 generations, including the material used in all safety assessments, as well as, the material used to generate all commercial Roundup Ready corn varieties containing event NK603.

DNA sequencing of plasmid vector PV-ZMGT32 demonstrated that it contains 9308 bp. This is a net gain of one base compared to the 9307 bp that have been previously reported (Deng *et al.*, 1999). The difference in the number of base pairs has been reflected in the position of the *Mlu* I restriction sites and total number of bases illustrated in Figure 1. None of the differences observed from sequencing the plasmid vector were contained in the protein coding regions and these differences were present at the time of transformation. Therefore, there is no impact on the characterization of Roundup Ready maize event NK603 due to the minor modifications made in the depiction of the map for plasmid vector PV-ZMGT32.

V. CONCLUSIONS

Previous molecular analyses of Roundup Ready maize event NK603 demonstrated that there is one copy of the DNA restriction fragment used for transformation in event NK603, along with a 217-bp segment containing a portion of the enhancer region of the rice actin promoter inversely linked to the 3' end of the inserted transformation cassette (Deng *et al.*, 1999). As part of this study, PCR and DNA sequence analyses were performed to confirm the organization of the elements within the insert and describe the complete DNA sequence of the insert in Roundup Ready maize event NK603.

VI. REFERENCES

Cavato, T.A., Deng, M.Y., and Lirette, R.P. 2001. Amended Report for MSL-16857: Confirmation of the Genomic DNA Sequences Flanking the 5' and 3' Ends of the Insert in Roundup Ready® Corn Event NK603. MSL-17617, an unpublished study by Monsanto.

Deng, M.Y., Lirette, R.P., Cavato, T.A., and Sidhu, R.S. 1999. Molecular Characterization of Roundup Ready® (CP4 EPSPS) Corn Line NK603. MSL-16214, an unpublished study by Monsanto.

Fraley, R.T., Rogers, S.G., Horsch, R.B., Sanders, P.R., Flick, J.S., Adams, S.P., Bittner, M.L., Brand, L.A., Fink, C.L., Fry, J.S., Galluppi, G.R., Goldberg, S.B., Hoffmann, N.L., Woo, S.C. 1983. Expression of bacterial genes in plant cells. Proc Natl Acad Sci U.S.A., 80(15), 4803-07.

Harrison, L.A., Bailey, M.R., Leimgruber, R.M., Smith, C.E., Nida, D.L., Taylor, M.L., and Padgette, S.R. 1993. Equivalence of plant-and microbially-expressed proteins: CP4 EPSPS from glyphosate-tolerant soybeans and *E. coli*. MSL-12899, an unpublished study conducted by Monsanto.

McElroy, D., Zhang, W., Cao, J. and Wu, R. 1990. Isolation of an efficient actin promoter for use in rice transformation. Plant Cell. 2, 163-171.

Odell, J. T., Mag, F., and Chua, N.-H. 1985. Identification of DNA sequences required for activity of the cauliflower mosaic virus 35S promoter. Nature 313, 810-12.

Padgette, S., RE, D., Barry, G., Eichholtz, D., Delannay, X., Fuchs, R., Kishore, G., and Fraley, R. 1996. New weed control opportunities: development of soybeans with a Roundup Ready™ gene, p. 53-84. *In* S. O. Duke (ed.), Herbicide Resistant Crops. CRC Press, Boca Raton, Fl.

Rochester, D. E., Winer, J.A. and Shah, D. M. 1986. The Structure and Expression of Maize Genes Encoding the Major Heat Shock Protein, HSP70. EMBO J. 5, 451-458.

Van den Broeck, G., Timko, M.P., Kausch, A.P., Cashmore, A.R., Van Montagu, M., Herrera-Estrella, L. 1985. Targeting of a foreign protein to chloroplasts by fusion to the transit peptide from the small subunit of ribulose 1,5-bisphosphate carboxylase. Nature 313, 358-63.

Table 1. Summary of Genetic Elements in the Linear DNA Restriction Fragment from PV-ZMGT32 Used in the Transformation of Roundup Ready Maize Event NK603

Genetic Element	Size Kb	Bp Location in Insert (Figure 4)	Function (reference)
The first EPSPS gene cassette:			
P-ract/ract1 intron	1.4	51 – 1443	5' region of rice (*Oryzae sativa*) actin 1 gene containing the promoter, transcription start site and first intron (McElroy *et al.*, 1990).
CTP2	0.2	1460 – 1687	DNA sequence for chloroplast transit peptide, isolated from *Arabidopsis thaliana* EPSPS; transit peptide directs the CP4-EPSPS protein to the chloroplast, the site of aromatic amino acid synthesis (Van den Broeck *et al.*, 1985).
CP4 EPSPS	1.4	1688 – 3055	The 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) sequence isolated from *Agrobacterium* sp. Strain CP4 which imparts tolerance to glyphosate (Harrison *et al.*, 1993; Padgette *et al.*, 1996).
NOS 3'	0.3	3068 – 3323	The 3' nontranslated region of the nopaline synthase (NOS) gene from *Agrobacterium tumifaciens* T-DNA which terminates transcription and directs polyadenylation (Fraley *et al.*, 1983).
The second EPSPS gene cassette:			
e35S	0.6	3340 – 3952	The cauliflower mosaic virus (CaMV) promoter (Odell *et al.*, 1985) with the duplicated enhancer region used to drive expression of the *CP4 EPSPS* gene.
*Zm*HSP70	0.8	3982 – 4785	Intron from the maize (*Zea mays*) *hsp*70 gene (heat-shock protein) present to stabilize the level of gene transcription (Rochester *et al.*, 1986).
CTP2	0.2	4810 – 5037	DNA sequence for chloroplast transit peptide, isolated from *Arabidopsis thaliana* EPSPS; transit peptide directs the CP4-EPSPS protein to the chloroplast, the site of aromatic amino acid synthesis (Van den Broeck *et al.*, 1985).
CP4 EPSPS	1.4	5038 – 6405	The 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) sequence isolated from *Agrobacterium* sp. Strain CP4 which imparts tolerance to glyphosate (Harrison *et al.*, 1993; Padgette *et al.*, 1996).
NOS 3'	0.3	6418 – 6673	The 3' nontranslated region of the nopaline synthase (NOS) gene from *Agrobacterium tumifaciens* T-DNA which terminates transcription and directs polyadenylation (Fraley *et al.*, 1983).
Portion of enhancer from p-ract	0.2	6707-6923	The 217-bp fragment includes polylinker sequence (50 bp) and the first 167 bp of the enhancer region of the rice actin promoter. Neither the TATA box nor transcriptional initiation site is present within the fragment, which suggests that this fragment should not function as a promoter (Deng *et al.*, 1999).

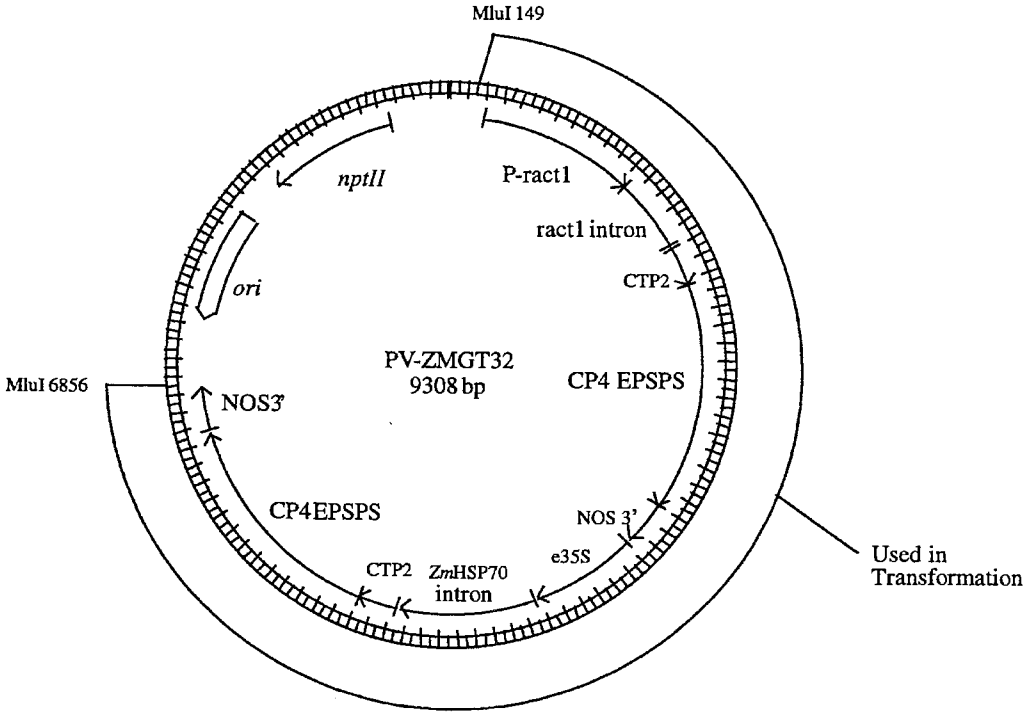


Figure 1. Plasmid Map of PV-ZMGT32. The *Mlu* I restriction fragment from plasmid PV-ZMGT32 was used to generate Roundup Ready maize event NK603. The vector PV-ZMGT32 contains 9308 bp, which represents a net gain of one base compared to the 9307 bp that have been previously reported. The net change in the number of base pairs has been reflected in the position of the *Mlu* I restriction sites and total number of bases.

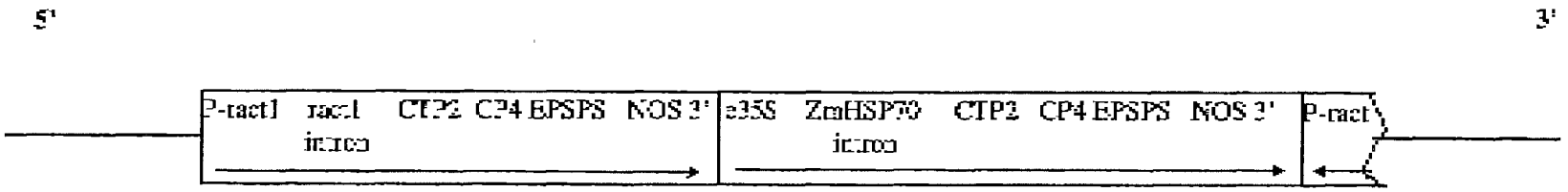


Figure 2. Schematic Representation of the Insert in Roundup Ready Maize Event NK603. This figure depicts the predicted insert for event NK603 as presented in Deng *et al.* (1999). There is one copy of the *Mlu* I restriction fragment from plasmid PV-ZMGT32 used for transformation, which contains two CP4 EPSPS gene cassettes. Immediately 3' of the second CP4 EPSPS cassette there is a 217-bp segment of the transformation cassette that contains a portion of the enhancer region of the rice actin promoter positioned in the inverse orientation.

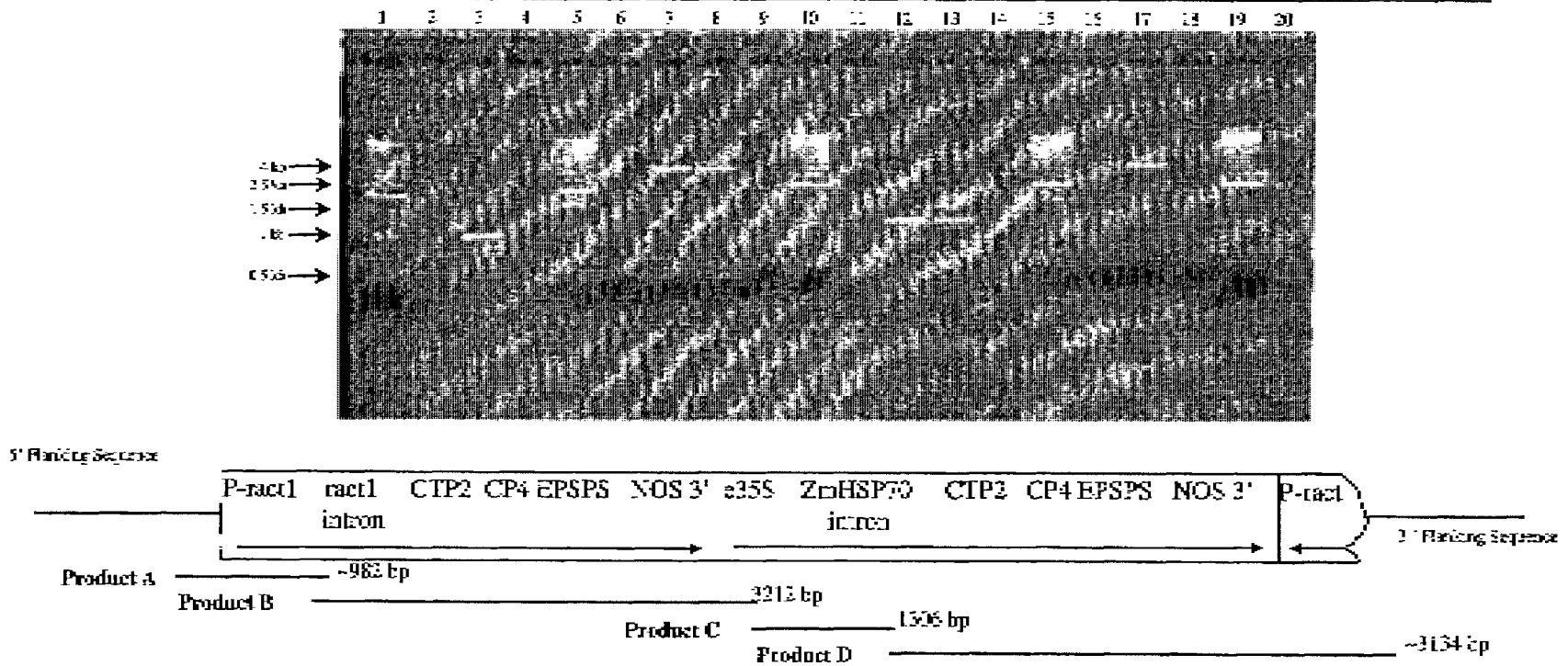


Figure 3. Overlapping PCR Analysis Demonstrating the Organization of the Inserted DNA in Roundup Ready Maize Event NK603. PCR analyses, which generated four overlapping products (A-D) demonstrating the linkage of the individual elements within the insert, were performed on PV-ZMGT32 plasmid DNA (lanes 8 and 13) and maize event NK603 genomic DNA extracted from grain tissue (lanes 3, 7, 12 and 17). Lanes 2, 6, 11 and 16 contain amplification mixture using 373 non-transgenic maize control DNA extracted from grain tissue as a template, while lanes 4, 9, 14 and 18 are control reactions containing no template DNA. Lanes 1, 5, 10, 15 and 19 contain Invitrogen 500 bp DNA ladder. Ten microliters of each of the PCR products were loaded on the gel. Lane 20 does not contain any sample.

→ Symbol denotes sizes obtained from MW markers on ethidium stained gel.

[CBI CROSS REFERENCE 1]

[CBI CROSS REFERENCE 1]

[CBI CROSS REFERENCE 1]

Appendix 1

Study Protocol

Monsanto Study #:	**01-01-46-25**
Study Title:	PCR and DNA Sequence Analysis of the Insert in Roundup Ready® Maize Event NK603
Sponsor:	Monsanto Company Biotechnology Regulatory Sciences 700 Chesterfield Parkway North St. Louis, MO 63198 Sponsor Representative: Regulatory Affairs, Maize Team Primary Contact: Linda K. Lahman Phone: 636-737-7653
Study Director:	Steven E. Reiser Monsanto Company - BB5K Biotechnology Regulatory Sciences 700 Chesterfield Parkway North St. Louis, MO 63198 Phone: 636-737-7696 FAX: 636-737-7662 e-mail: steven.e.reiser@monsanto.com
Primary Testing Facility:	Monsanto Company Biotechnology Regulatory Sciences 700 Chesterfield Parkway North St. Louis, MO 63198
Additional Testing Facility:	Monsanto Company Genomics Sequencing Center 800 North Lindbergh St. Louis, MO 63167 Primary Contact: Mary Blanchard Phone: 314-694-1184
Testing Facility Management Representative:	Patrick T. Weston Monsanto Company - BB5B Biotechnology Regulatory Sciences 700 Chesterfield Parkway North St. Louis, MO 63198 Phone: 636-737-5407 FAX: 636-737-7662 e-mail: patrick.t.weston@monsanto.com

[redacted]

Approved By:

Patrick T. Weston
Testing Facility Management Representative
Monsanto Company

Nov 8, 2001
Date

Linda K. Fahman
Regulatory Affairs, Maize Team
Sponsor Representative
Monsanto Company

8 Nov 2001
Date

Steven E. Reiser
Study Director
Monsanto Company

Nov. 9, 2001
Date

Reviewed By:

C Mari Bratton
Quality Assurance Specialist
Monsanto Company

Nov. 8, 2001
Date

[redacted]

[redacted]

1.0 Regulatory Compliance

1.1 GLP Compliance

This is a product characterization study as defined by section §160.135(b) of the Federal Insecticide, Fungicide and Rodenticide Act (FIFRA); Good Laboratory Practice Standards (40 CFR Part 160) intended to characterize the physical and/or chemical properties of a potential commercial product. This study will be conducted in compliance with all requirements of section §160.135(b), except for DNA sequencing by the Monsanto Genomics Sequencing Center, a non-GLP facility. Monsanto Regulatory QAU will provide oversight and distribute QA reports according to the Monsanto Regulatory QAU SOPs.

2.0 Purpose

The purpose of this study is to perform PCR analyses across the insert in Roundup Ready® maize event NK603. The DNA sequence of these PCR products will be determined.

3.0 Timelines

3.1	Proposed Experimental Start Date:	November, 2001
3.2	Proposed Experimental Termination Date:	December, 2001

4.0 Test, Control and Reference Substances

4.1 Test Substance

The test substance is maize event NK603 (Lot # RDR-0111-11849-S). DNA from the test substance was isolated from grain prior to the start of this study. The unique sample identification of the DNA used will be documented in the raw data.

4.2 Control Substance

The control substance is the non-transgenic control line B73 (Lot # TCP-0005-10313-I). The control substance DNA to be used in this study was previously isolated from grain under Study # 01-01-50-15. The unique

® Roundup Ready is a registered trademark of Monsanto Company, St. Louis, Missouri.

[redacted]

sample identification of the control DNA used will be documented in the raw data.

4.3 Reference Substances

The reference substances will include the plasmid PV-ZMGT32 (used to generate maize event NK603) which will be used as a positive control in some of the PCR analyses. The PCR products generated from the plasmid will be used as size indicators for the appropriate PCR analyses. The 500 bp DNA Ladder from Gibco BRL will be used to estimate band sizes in the PCR analyses. The unique identities of the reference standards will be documented in the raw data.

4.4 Characterization of Test, Control and Reference Substances

The identity of the test substance will be confirmed as part of the analyses of this study. The identity of the control substance and plasmid PV-ZMGT32 was confirmed prior to the start of the study and is archived with Study # 01-01-50-15. The chain of custody for the test and control samples will be confirmed by the Study Director. Given the qualitative nature of the analyses being performed, if the test and reference substances produce PCR products, they will be considered to be stable during storage at 2-8°C.

5.0 Description of Experimental Design

Overlapping PCR products will be generated across the entire length of the insert. These PCR products will then be purified and sent for sequencing to the Monsanto Genomics Sequencing Center (a non-GLP facility).

5.1 Analytical Methods

All methods will be conducted as described below or by other appropriate methods approved by the Study Director and documented in the raw data.

5.1.1 DNA Extraction

The DNA from the test substance was isolated prior to the start of the study using the Dneasy Plant Mini Kit from Qiagen. The raw data detailing the extraction will be archived with this study. DNA from the control substance was previously isolated using the Rogers and Bendich (1985) method as a part of study # 01-01-50-15. An exact copy of the raw data, detailing the extractions and quantitation will be archived with this study. All previously extracted DNAs have been stored at 2-8°C. If necessary, additional DNA will be extracted under this protocol using methods approved by the Study Director. The source of any additional material will be added by amendment, if necessary.

[redacted]

5.1.2 DNA Quantitation

Any needed DNA quantitation will be conducted using Hoefer's DyNA Quant 200 Fluorometer according to SOP # BR-EQ-0065-01 using Roche's Molecular Weight Marker IX as a calibration standard. Two

readings will be averaged to obtain a final reading if they are within 10% of each other. If they are not, a third reading will be taken. If there are two readings within 10% of each other, the outlier will be rejected and the remaining two will be averaged. If none of the readings are within 10% of each other, the three readings will be averaged together.

5.1.3 Polymerase Chain Reaction

PCR amplification will be performed using genomic DNA template from both the test and control substances following standard PCR methodologies which will be documented in the raw data associated with this study. Accepted PCR reactions will have visible products in the appropriate lanes on the agarose gel (i.e. no bands of the expected size amplicon in the no template DNA control and control substance lanes, and a visible band in the test substance lane). Plasmid PV-ZMGT32 will be used as a positive control in PCR reactions when appropriate. (The exception will be for those reactions in which one or both primers anneal to the flanking genomic DNA. In such cases, plasmid PV-ZMGT32 is not a valid control). Various primers homologous to the inserted DNA and the flanking genomic DNA will be used in the PCR analyses.

5.1.3.1 Agarose Gel Electrophoresis of PCR Products

PCR products will be separated on an agarose gel according to SOP # GEN-PRO-003-01. After electrophoresis, the DNA from the test substance amplifications will be purified from the agarose matrix using an extraction kit following the procedure supplied by the manufacturer. Acceptable gels will have a dye front moving in a uniform path down the gel, and molecular weight marker bands and DNA will be visible on the agarose gel under UV illumination. Reasons why gels are rejected will be documented in the raw data.

5.1.3.2 Sequencing of Purified PCR Products

Purified PCR products will be mixed with appropriate primers and submitted to the Monsanto Genomics Sequencing Center for sequencing. Acceptable sequencing runs will produce raw data in the form of chromatograms which are interpretable. Reasons for rejecting sequencing runs will be documented in the raw data. A consensus sequence will be created from the acceptable runs in the primary contig using DNAstar/SeqManII version 4.03.

[redacted]

6.0 Control of Bias

A PCR containing no template DNA will be prepared with each primer set to serve as a negative control. When applicable, the plasmid PV-ZMGT32 will serve as a positive control in some PCR analyses.

7.0 Proposed Statistical Methods

No statistical analyses will be performed in this study

8.0 Records to be Maintained

Records will be maintained of all sample transfers, analyses, the protocol and all deviations and amendments thereto and copies of all letters memoranda and other correspondence related to this study. These documents may include: photocopies, computer generated hard copies or hand-written notes that describe the procedures used to generate data for this study. Upon completion of the study, all study records, protocol and final report will be archived at Monsanto.

9.0 Changes to the Protocol

Planned changes to the protocol will be documented in the form of written protocol amendments and signed by the Study Director. Amendments become part of the protocol and will be archived with the protocol. All other changes will be in the form of written protocol deviations and will be filed with the raw data. All changes to the protocol will be addressed in the final report.

10.0 References

Rogers, S.O. and Bendich, A.J., 1985. Extraction of DNA from milligram amounts of fresh, herbarium and mummified plant tissues. *Plant Mol. Biol.* 5:69-76.

[redacted]