



# UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

---

FACULTAD DE QUÍMICA

## PROGRAMA DE MAESTRÍA Y DOCTORADO EN CIENCIAS BIOQUÍMICAS

Identificación y caracterización de la proteína heteróloga  
RR en fracciones de soya transgénica.

**T E S I S**

Que para obtener el grado de:

MAESTRA EN CIENCIAS (BIOQUÍMICAS)

P r e s e n t a :

**MARÍA DEL CARMEN FONG CÁRDENAS**

**Asesora: Dra. Amanda Gálvez Mariscal**



México, D.F

2010



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Este trabajo fue realizado bajo la dirección de la Dra. Amanda Gálvez Mariscal en el laboratorio 312 del Departamento de Alimentos y Biotecnología de la Facultad de Química, UNAM.

Durante el desarrollo de este trabajo se contó con apoyo de una beca del CONACYT.

## AGRADECIMIENTOS

A la Universidad Nacional Autónoma de México por haberme dado la oportunidad de ser parte de ésta.

A la Dra. Amanda Gálvez por la oportunidad que me dio de realizar mi tesis de Maestría, por su apoyo y por su paciencia.

Al Dr. Roberto Arreguín por su apoyo técnico y sobre todo por sus enseñanzas.

A la Dra. Amelia Farrés por el préstamo del equipo de PCR para la realización de esta parte de la tesis y por sus consejos.

Al Dr. Javier Plasencia de la Parra por el préstamo del equipo para cuantificar ADN y apoyos diversos de su laboratorio.

A la Dra. Maricarmen Quirasco por su apoyo para la realización de este trabajo.

Al Dr. Rogelio Rodríguez Sotres por el préstamo del lector de placas de ELISA.

## **GRACIAS:**

A Dios, por todo.

A mis papás, por su apoyo que me brindaron desde el inicio de mi carrera, ya que son una parte importante de mí y que sin ustedes no hubiera logrado esta meta tan importante.

A mis hermanitos Toño y Francisco por todo su apoyo.

A mi esposo Carlos, por su comprensión, apoyo y paciencia.

A Alejandra por tu amistad, consejos, por tu paciencia y comprensión que me has brindado.

A Diana, por tu amistad, apoyo en la carrera y sobre todo por los consejos que me has dado.

A mis compañeros de laboratorio y del PUAL Vanesa, Denise, Carolina, Gaby, Marcela y Berenice por su apoyo en la realización de esta tesis.

A Alejandra Barrios, por el apoyo, consejos y amistad que me brindas.

## DEDICATORIA

A mis papás.

Los quiero mucho.

---

**Abreviaturas**

<b>ADN</b>	Ácido Desoxirribonucleico
<b>B.M.</b>	Biología molecular
<b>CIAD</b>	Centro de Investigación en Alimentación y Desarrollo
<b>DEAE</b>	Di – Etil - Amino - Etil
<b>ELISA</b>	“Enzyme-linked Immunosorbent assay”
<b>EPSPS</b>	5-enolpyruvilshikimato-3-fosfato sintasa
<b>FF</b>	Fast Flow (Flujo rápido)
<b>GM</b>	Genéticamente Modificado(a)
<b>OGM</b>	Organismo Genéticamente Modificado
<b>PCR</b>	Reacción en Cadena de la Polimerasa
<b>Q</b>	Sal cuaternaria de amonio
<b>R.A</b>	Grado reactivo
<b>SE</b>	Soya Esperanza
<b>SH</b>	Soya Héctor
<b>SC</b>	Soya Cajeme
<b>STA</b>	Soya comprada en Tonalá, Chiapas
<b>STO</b>	Soya comprada en Tlatelolco
<b>SFL</b>	Soya comprada en una tienda en el sur de la ciudad
<b>SNA</b>	Soya comprada en NUTRISA, Satélite
<b>T.A.</b>	Temperatura ambiente

---

---

## ÍNDICE

- I. RESUMEN**
- II. INTRODUCCIÓN**
- III. ANTECEDENTES**
- IV. 1. Generalidades**
  - 2. Área global de los cultivos transgénicos**
    - 2.1 Países que cultivan transgénicos**
    - 2.2 Principales cultivos**
    - 2.3 Expectativas**
  - 3. Soya (*Glicine max*)**
    - 3.1 Soya tolerante a herbicidas**
  - 4. Herbicida Glifosato**
    - 4.1 Evaluación de la inocuidad de los cultivos GM**
  - 5. Enzima 5-enolpiruvilshikimato-3-fosfato sintasa (EPSPS)**
  - 6. Métodos de purificación de la EPSPS a partir de diferentes organismos**
- V. HIPÓTESIS**
- VI. OBJETIVOS**
- VII. MATERIALES Y MÉTODOS**
  - 1. Estrategia experimental**
  - 2. Obtención y pre-tratamiento de la materia prima**
    - 2.1 Homogenización de la materia prima**
    - 2.2 Extracción de grasa de las muestras**
    - 2.3 Extracción de proteínas**
  - 3. Cuantificación de proteína**
  - 4. ELISA**
    - 4.1 Preparación de la placa**

- 4.2 Protocolo de ELISA
- 5. Detección de secuencias GM
  - 5.1 Extracción de ADN
  - 5.2 PCR punto final
  - 5.3 PCR para transgén
  - 5.4 Purificación del producto de PCR
- 6. Secuenciación
- 7. PCR tiempo real para las muestras de soya
  - 7.1 Preparación de la mezcla de reacción
- 8. Separación de proteínas por electroforesis
- 9. Isoelectroenfoque preparativo
- 10. Electroelución
- 11. Cromatografía de intercambio iónico

## **VIII. RESULTADOS Y DISCUSIÓN**

- 1. Verificación de las muestras
- 2. Cuantificación de ADN por PCR tiempo real
  - 2.1 Cuantificación absoluta del gen endógeno y del transgén
  - 2.2 Cuantificación del material transgénico en las muestras analizadas
- 3. Secuenciación del producto de PCR
  - 3.1 Alineamiento de la secuencia
- 4. Estudios preliminares para la purificación de la CP4EPSPS
- 5. Comparación del perfil electroforético de las muestras en geles de poliacrilamida
- 6. Isoelectroenfoque preparativo
  - 6.1 ELISA
- 7. Cromatografía de intercambio aniónico
  - 7.1 Columna de intercambio aniónico DEAE-FF

**7.2 Columna de intercambio aniónico Q-FF**

**8. Espectroscopía de masas**

**9. Secuenciación**

**IX. CONCLUSIONES**

**X. REFERENCIAS**

**XI. ANEXOS**

## I. RESUMEN

La comercialización de productos de OGMs es un hecho a nivel mundial. Esto indica que los productos biotecnológicos deberán ser identificados y etiquetados, bajo la normatividad correspondiente. Los aspectos regulatorios, ya sean, nacionales e internacionales hacen evidente la necesidad no sólo de técnicas de detección, sino también de una metodología para la evaluación de riesgos. Para ello se debe evaluar el riesgo para cada población en particular con base en el patrón de consumo del alimento. En el caso particular de la soya, nuestro objeto de estudio, la ingestión es alta y va en aumento. La demanda de soya se incrementa un 3% al año (Norsoy, 2002). El país importa el 97% de la soya que consume y solo el 3% de la soya importada no es transgénica. Esto significa que el 95% del consumo de soya es de la transgénica (ISAAA, 2006). Se han diseñado lineamientos generales sobre la evaluación de riesgos potenciales. La mayoría de los estudios están enfocados en la alergenicidad (Harrison, 1996). En animales de laboratorio alimentados con soya GM y no GM, se han encontrado algunas diferencias significativas en hepatocitos y células acinares del páncreas (Hammond, 1996; Malatesta, et al., 2002).

Por otro lado, en el Centro de Investigación en Alimentación y Desarrollo (CIAD), la Dra. Ana Ma. Calderón estudió el tejido pancreático de ratas alimentadas con aislados de soya genéticamente modificada (GM), encontrando una disminución de los niveles de ARNm de tripsinógenos, así como de la actividad de amilasa pancreática, y del área de gránulos de zimógenos en las células acinares del páncreas (Magaña-Gómez, et al., 2007).

Por lo anterior, se decidió localizar e identificar la proteína heteróloga de la soya GM entre las demás fracciones de una muestra de soya transgénica, para que una vez purificada se pueda incluir en estudios posteriores y evaluar su ingestión subcrónica en ratas.

El objetivo principal de este proyecto fue la purificación y caracterización de la proteína 5-enolpiruvilshikimato-3-fosfato sintasa (EPSPS) de la soya Roundup Ready. Ante la imposibilidad de trabajar con variedades isogénicas, se obtuvieron mezclas de variedades de diferentes regiones del país a las cuales se les identificó como GM realizando reacciones de PCR.

Primero se determinó cualitativamente (PCR punto final), cuáles de ellas eran GM, posteriormente se les realizó PCR tiempo real (Rt-PCR) para cuantificar el porcentaje del transgén y determinar cuál era la muestra con mayor contenido de material transgénico y trabajar con ésta. Después se probaron diferentes metodologías para la extracción y purificación de la proteína transgénica, quedando como metodología final la cromatografía de intercambio aniónico, con columnas Q-FF y DEAE-FF.

Se separó una proteína de aproximadamente de 43 kDa que al corroborar su identidad, permitió ver que no se trataba solamente del péptido de tránsito de *Petunia hybrida* sino que también contenía el de *Arabidopsis thaliana*. También se determinó que no solamente se trataba de una proteína heteróloga, ya que en otros estudios se han identificado más de una, dando pauta a la posible existencia de al menos cuatro isoformas de la misma. También se podría tratar de cuatro proteínas de fusión (Rang, 2005) y que en trabajos posteriores se tendrían que purificar y caracterizar cada una de ellas.

## II. INTRODUCCIÓN

Las técnicas de la ingeniería genética han permitido crear organismos con características nuevas comparados con su línea isogénica no transformada genéticamente. Estos son los llamados organismos genéticamente modificados (OGM) o transgénicos. Aun cuando estos términos son empleados casi indistintamente existe algo que los diferencia. El término OGM se aplica de manera general a organismos cuyo genoma ha sido modificado con información genética propia o de otro organismo. Ejemplo del primer caso es el tomate FlavSavr®, al cual se le insertó una copia (del mismo tomate) de la secuencia nucleotídica codificante de la poligalacturonasa, enzima relacionada con la maduración del fruto, pero en sentido inverso para evitar la transcripción del gen y así retardar el proceso de maduración (Smith et al., 1988). Un segundo caso se trata de organismos a los que se les ha insertado una secuencia nucleotídica proveniente de otra especie, para conferirle la nueva característica. Tal es el caso del maíz Bt, al cual se le insertó la secuencia del gen de una endotoxina producida originalmente en *Bacillus thuringiensis*. El término “transgénico” se aplica exclusivamente a casos como el segundo. De esta manera, un OGM presenta una nueva secuencia de ácido desoxirribonucleico (ADN), lo que lo diferencia del organismo original. Esta nueva pieza de ADN típicamente contiene uno o varios genes que han sido modificados de tal manera que son capaces de manifestarse en el nuevo organismo. Usualmente, el o los genes insertados determinan la presencia específica de nuevos productos génicos (Nelson y Cox, 2000). Como siempre sucede, el surgimiento de nuevas tecnologías despierta preocupaciones, tanto entre la comunidad científica, como entre el público en general y el uso de los productos de la ingeniería genética no fue la excepción. Pero existe un claro ejemplo del compromiso mundial de atender la problemática de la bioseguridad de plantas GM, que está plasmado en el protocolo sobre bioseguridad para el movimiento trans-fronterizo de OGM (Bolívar, 2004). El análisis de los posibles riesgos y beneficios debe hacerse considerando el cultivo de plantas transgénicas y plantas no transgénicas para valorar y ponderar si los posibles riesgos ecológicos son mayores, menores o iguales que con las plantas no-GM. Hails

(2000) hace un análisis del debate que existe actualmente entre los defensores y detractores de las plantas GM y sugiere que el mejor árbitro en la toma de decisiones debe ser el análisis científico, a través de ensayos de campo, donde se analicen diversas variables ecológicas que pueden ser alteradas, tratándose cada uno de los casos de diferente manera. Más extrema ha sido la preocupación que han exteriorizado algunos grupos sobre la posibilidad de que el consumo de transgenes sea dañino a la salud o que los genes de plantas transgénicas pudieran incorporarse a las células del ser humano y causar cáncer. No hay evidencia de que esto pudiera suceder y, si así lo fuese, debe aclararse que las plantas normales contienen miles de genes y que si el ADN de las plantas causara daños a la salud o se pudiese incorporar a los cromosomas del ser humano, esto vendría sucediendo desde que el hombre apareció en la Tierra. La posibilidad de que la proteína codificada por un gen introducido en plantas transgénicas tenga una función adicional a la esperada es ciertamente real. Existe evidencia de que muchas proteínas tienen funciones totalmente distintas a las que se les conocía, por lo que es necesario realizar un análisis profundo de las plantas GM que se pretenden liberar al ambiente y/o ser consumidas por el humano. Muchos de los alimentos que se consumen actualmente en algunos países son GM, o contienen ingredientes derivados de OGMs, por ejemplo, enzimas que se utilizan durante la producción de queso, cereales como el maíz o leguminosas como la soya resistente a ataques de plagas, a enfermedades y a herbicidas. México tiene un particular patrón de alimentación, diferente al de otros países. En las evaluaciones de inocuidad es importante evaluar a cada población en particular, con base en el patrón de consumo del alimento. La evaluación debe validarse para cada evento de modificación genética que se ha introducido o se planea introducir puesto que (aunque posea la misma construcción), es genética, proteínica y metabólicamente diferente (COFEPRIS).

Un aspecto crucial en el análisis de organismos genéticamente modificados en los alimentos es la cuantificación, ya que existe un límite máximo permitido (0.9%) en alimentos que si es superado requiere de etiquetado, de acuerdo a lo estipulado

por la Unión Europea (Reglamento 1829/2003), lo que afecta el comercio exterior. Así que, lo que pretenden los países desarrollados es mantener por debajo del nivel permitido el contenido indeseable de material GM en la producción de alimentos no-GM, así como el mantener y comercializar por separado los OGM debido a su alto valor agregado (como el caso de los de segunda generación). Existen tres sistemas para diferenciar en varios grados y con varias metodologías entre productos con características parecidas. 1) Rastreabilidad, 2) Segregación y 3) Preservación de la Identidad. La rastreabilidad es la capacidad para rastrear OGM y sus productos derivados en cualquier punto del mercado a través de la cadena de producción y distribución. Se podría distinguir entre productos de diferente calidad, rastrear con fines de control tanto de calidad como para la salud, proveer datos para estudios epidemiológicos, lograr la transparencia para el consumidor ya que se podría saber en dónde está el producto deseado y en que concentración. El término rastreabilidad puede ser usado en diferentes contextos, como productos, datos, para calibraciones o bien en sistemas de programación como sistemas de identificación por radio frecuencia (Moe, 1998; Regattieri, 2007). Un sistema de segregación es la separación física de cultivos o lotes de ingredientes en todos los pasos del sistema de abasto (producción, manipulación y distribución), para evitar el mezclado, lo que implica, la limpieza de las instalaciones, del transporte y de los almacenes y el monitoreo de la presencia de material GM. No se requiere de la aplicación simultánea de un sistema de rastreabilidad. Un sistema de Preservación de Identidad (IP) es un sistema de manejo y comercialización de cultivos que permite la identificación de la naturaleza y fuente de los materiales empleados. Esto garantiza que ciertas características del cultivo o alimento se mantengan del campo a la mesa. Este sistema responde al cómo, qué, cuándo, dónde, quién y por qué de un producto en particular. Todavía no hay una distinción clara entre sistemas de segregación IP, lo que tienen en común es que el control inicia desde el campo, pero en países centro de origen con posibilidades de flujo genético, estos sistemas deberían apoyar la vigilancia y monitoreo.

El área de mayor auge en la biotecnología agrícola es la producción de cereales y leguminosas más baratos al protegerlos del ataque de insectos y del efecto de herbicidas. Estas modificaciones se han hecho en ciertos cultivos básicos como el maíz, algodón, canola y soya. En el caso particular de la soya, se estima que el 60% de la producción mundial de este cultivo (54.4 millones de Has) es transgénica siendo la principal modificación genética la tolerancia al herbicida glifosato. Este cultivo es producido principalmente en EUA y Argentina (James, 2006). La demanda de soya transgénica se incrementa un 3% al año y se canaliza principalmente para consumo animal, formulación de alimentos balanceados, así como para la obtención de derivados texturizados, aceite y fibra dietética para consumo humano (AGBIOS). A su vez, la proteína de soya es generalmente utilizada para la formulación de emulsificantes y aditivos alimentarios que se consumen en menores cantidades en los alimentos procesados (Hurts et al., 1999). En México, el 97% de la soya que se consume es importada lo que equivale a 4.3 millones de toneladas anuales (USDA, 2005).

En México, los ensayos de OGMs permitidos en el 2006 fueron en algodón, plátano, calabacita, alfalfa, maíz y soya Solución Faena, que es el nombre comercial en México de la soya Roundup Ready (RR) de la compañía Monsanto. Los ensayos de campo se realizaron en el sureste del país (Chiapas, Yucatán y Quintana Roo) y en la Planicie Huasteca (SENASICA / SAGARPA).

Las proteínas heterólogas de las que se tiene más conocimiento son las de cultivos transgénicos comercializados actualmente y que tienen un propósito agronómico como el de proporcionar resistencia a insectos o tolerancia a herbicidas (Watson et al., 1992). Existe una cantidad importante de datos que evalúan los riesgos de alergenicidad y toxicidad de estas proteínas que consisten en alimentar a los animales con el grano entero o bien, realizar evaluaciones nutricias con las proteínas heterólogas purificadas. Para esto último, generalmente se requiere sobreexpresarlas en vectores bacterianos para aislarlas posteriormente y darlas como alimento a los animales de laboratorio (Hammond et

al., 1996; Cromwell et al., 2002). Una desventaja de este tipo de estudios es que las proteínas sobreexpresadas en sistemas bacterianos no experimentan las modificaciones post-traduccionales, como la glicosilación que debería existir si la fuente fuera la propia planta transgénica. Por lo tanto, existe poca información desde el punto de vista bioquímico acerca del comportamiento y las funciones de estas proteínas, así como de sus efectos en los animales alimentados con ellas.

La mayoría de los estudios sobre la evaluación de riesgos a la salud están enfocados en la evaluación de la toxicidad y la alergenicidad (Harrison et al., 1996). También se realizan estudios para evaluar riesgos derivados del consumo de transgénicos en los que se miden variables como la composición proximal del producto (Padgett, 1996), cambios de peso en los animales de estudio, eficiencia alimentaria y se evalúan aspectos macroscópicos de algunos órganos (Brake, 2004). Cuando se concluye que el alimento es tan seguro como su contraparte convencional, por no encontrar diferencias significativas entre los indicadores estudiados (contenido de proteínas, lípidos, carbohidratos, cenizas y fibra) (Padgett, 1996), el producto se autoriza para consumo humano y/o animal. En un caso particular, se encontraron diferencias en el peso de órganos y aspecto del hígado de ratas alimentadas con soya Roundup Ready, que expresa la proteína 5-enolpiruvil-shikimato-3-fosfato sintasa cepa CP4 (CP4-EPSPS), tolerante al herbicida glifosato. Sin embargo, los autores no atribuyen dichas diferencias a la dieta recibida por los animales, sino a su porcentaje de proteína (Hammond et al., 1996).

Malatesta y colaboradores (2002a, 2002b) evaluaron el consumo crónico de soya transgénica en ratones, a través de variables microscópicas. El grupo alimentado con soya transgénica mostró una disminución en la actividad de la amilasa pancreática, así como en el área de los gránulos de zimógenos en las células acinares del páncreas (Malatesta et al., 2002a). En hígado, el núcleo de los hepatocitos mostró una morfología irregular en los ratones alimentados con la

soya transgénica (Malatesta et al., 2002b), respecto a los ratones alimentados con la soya convencional.

En el Centro de Investigación en Alimentación y Desarrollo (CIAD), la Dra. Ana Ma. Calderón de la Barca estudió el tejido pancreático en ratas alimentadas con aislados de soya GM, encontrando una disminución de los niveles del ARNm de tripsinógenos, así como una disminución en la actividad de amilasa pancreática y en el área de gránulos de zimógenos en las células acinares del páncreas (Magaña-Gómez et al., 2007). El presente proyecto, en colaboración con el estudio del CIAD, pretende caracterizar las fracciones de aislados de soya RR en las que se encuentre la proteína heteróloga. Para ello, es importante separar la proteína nueva de las otras proteínas, para que una vez purificada se identifique plenamente como la heteróloga. Una vez conocida la (o las) fracción(es) de la soya en las que se expresa, se podría generar un aislado proteínico especial de soya, enriquecido con la proteína heteróloga. Con éste podría garantizarse que en estudios del efecto de su consumo en animales de laboratorio, los posibles cambios observados puedan atribuirse al consumo de la fracción enriquecida con la proteína heteróloga, en comparación con un aislado de soya similar, pero obtenido a partir de soya no GM. También, una vez separada la proteína heteróloga del resto, se podría calcular de manera directa el porcentaje de ésta con respecto a la proteína total mediante geles 2D y métodos de proteómica.

### **III. ANTECEDENTES**

#### **1. Generalidades**

Los esfuerzos biotecnológicos de la industria y de algunas instancias que realizan investigación en el área de la agricultura han dado como resultado la comercialización de cultivos como el maíz y la soya con nuevos rasgos agronómicos. Algunas de las modificaciones más importantes incluyen la resistencia a insectos y tolerancia a herbicidas (James, 2006). Esto ha sido el resultado de insertar nuevas secuencias de ADN en plantas para la expresión de proteínas que les confieran nuevas características (Gachet, 1999).

#### **2. Área global de los OGM's**

La importancia de los cultivos transgénicos ha trascendido a nivel mundial en los últimos diez años. En un periodo de siete años (1996-2002), la superficie sembrada con plantas transgénicas aumentó aproximadamente 35 veces, de 1.7 millones de hectáreas hasta 58.7 millones de hectáreas. En el 2003, los cultivos modificados genéticamente ocupaban del 21 al 22% de la superficie sembrada mundialmente y los principales fueron soya (62%), maíz (21%), algodón (12%) y canola (5%; James, 2003).

En el 2004, el área de cultivos genéticamente modificados siguió en aumento: 20% en comparación del 15% que hubo en el 2003, es decir, en el 2004 aumentaron 13.3 millones de hectáreas, siendo el segundo incremento más alto que se ha dado en cultivos genéticamente modificados.

El año 2005 marcó la primera década de comercialización de cultivos GM. En esta década (1996-2005), el área global de los cultivos biotecnológicos se incrementó 50 veces, de 1.7 millones de hectáreas en 1996 a 90 millones de hectáreas en el 2005 (James, 2005).

En el 2006, el primer año de la segunda década (2006-2015), el área global de estos cultivos incrementó un 13%, es decir, hubo un aumento de 12 millones de

hectáreas para este año, llegando a un total de 102 millones de hectáreas (James, 2006).

En la Tabla 1 se presenta un resumen de la superficie sembrada con cultivos genéticamente modificados:

**Tabla 1. Superficie sembrada con cultivos genéticamente modificados a nivel mundial**

<b>Año</b>	<b>Millones de Has</b>
1996	1.7
1997	11.0
1998	27.8
1999	39.9
2000	44.2
2001	52.6
2002	58.7
2003	67.7
2004	81
2005	90
2006	102
<b>Total</b>	<b>384.6</b>

Fuente: James, 2006

## **2.1 Países que cultivan OGMs**

El número de países que siembran cultivos genéticamente modificados ha incrementado de seis en 1996 a nueve en 1998, 12 en 1999, 21 en el 2005 hasta 22 en el 2006.

Algunos países llegaron a plantar hasta 50,000 millones de hectáreas o más en el 2005. México siembra por primera vez cultivos GM en 2004, junto con otros países como: EUA, Argentina, Brasil, Canadá, China, Paraguay, India, Sudáfrica, Uruguay, Australia, Rumania, Filipinas, España, Colombia, Irán, Honduras, Portugal, Alemania, Francia, la República Checa y Slovakia (Tabla 2; James, 2006).

**Tabla 2. Países que cultivan OGM's**

País	Millones de Has (2004)	Millones de Has (2005)	Millones de Has (2006)
EUA	47.6	49.8	54.6
Argentina	16.2	17.1	18
Brasil	5.0	9.4	11.5
Canadá	5.4	5.8	6.1
China	3.7	3.3	3.5
Paraguay	1.2	1.8	2.0
India	0.5	1.3	3.8
Sudáfrica	0.5	0.5	1.4
Uruguay	0.3	0.3	0.4
Australia	0.2	0.3	0.2
México	0.1	0.1	0.1
Rumania	0.1	0.1	0.1
Filipinas	<0.1	0.1	0.1
España	<0.1	0.1	0.1
Colombia	<0.1	<0.1	<0.1
Irán	--	<0.1	<0.1
Honduras	<0.1	<0.1	<0.1
Portugal	--	<0.1	<0.1
Alemania	<0.1	<0.1	<0.1
Francia	--	<0.1	<0.1
República Checa	--	--	<0.1
Slovakia	--	--	<0.1
<b>Total</b>	<b>81.0</b>	<b>90.0</b>	<b>102</b>

Fuente: Clive James, 2006

En México, en el 2005, se cultivaron los siguientes productos GM: algodón Solución Faena, algodón Bollgard y algodón LL25, sembrados en Chihuahua, la Planicie Huasteca, Baja California y la Comarca Lagunera. Por otro lado se sembró plátano con el gen LT en Colima, calabacita en Morelia, Michoacán, papa en el Edo. de México, Soya Solución Faena en Sinaloa, Tamaulipas, San Luis Potosí, Chiapas, Campeche, Yucatán y Quintana Roo (SENASICA/SAGARPA).

## 2.2 Principales Cultivos

La soya tolerante al herbicida glifosato continuó como el principal cultivo en el mundo en el 2006, ocupando 54.4 millones de hectáreas, que representan un 60% del área total sembrada con cultivos GM (90 millones de hectáreas). En países como EUA, Argentina, Brasil, Paraguay, Canadá, Rumania, Sudáfrica y México se ha ido incrementando su comercialización. En la Tabla 3 se mencionan los principales cultivos GM (James, 2006).

**Tabla 3. Cultivos genéticamente modificados dominantes**

<b>Cultivo</b>	<b>Millones de hectáreas</b>
Soya Tolerante a Herbicida (HT)	54.4
Maíz Bt	11.3
Maíz Bt / HT	6.5
Algodón Bt	4.9
Canola HT	4.6
Algodón Bt / HT	3.6
Maíz HT	3.4
Algodón HT	1.3
Total	90.0

Fuente: Clive James, 2006

Nota: HT : Tolerante a herbicida Bt : Con información genética de *Bacillus thuringiensis* (resistente a insectos)

### 2.3 Expectativas

La industria biotecnológica espera que el número y proporción de pequeñas industrias del ramo así como el de agricultores se incremente en la próxima década (2006-2015). También, se espera que en el continente Asiático acepte esta clase de cultivos GM y se incremente el consumo de sus productos tanto para consumo animal como humano. Mientras, en Europa se espera que se autoricen 17 variedades de maíz transgénico, así como que el área sembrada incremente a 150 millones de hectáreas en el mundo a finales del 2015 (ISAAA).

Sin embargo, no todos los países han aceptado los cultivos GM y algunos han impuesto regulaciones muy estrictas. Tal es el caso de la Unión Europea, cuyas leyes estipulan que productos que contengan más de un 0.9% de ingredientes proveniente de organismos genéticamente modificados deberán ser etiquetados (Rott, 2004).

Los aspectos regulatorios nacionales e internacionales hacen evidente la necesidad no sólo de técnicas de detección y cuantificación de transgénicos sino también de metodologías para la evaluación de riesgos alérgicos y toxicológicos, principalmente. En la práctica comercial el monitoreo y la vigilancia

de los cultivos transgénicos se basan en detecciones moleculares de ADN y de proteínas heterólogas, mediante las técnicas de PCR y ELISA, respectivamente (Vollenhofer y col., 1999; Takagi y col., 2006).

### **3. Soya (*Glycine max*)**

La soya, *Glycine max* (L.), es un miembro de la familia de las Leguminosas, subfamilia *Papilionoideae* (Hermann, 1962), que está presente en la cadena alimentaria desde hace más de 5,000 años como un producto básico de la dieta asiática. En China, por ejemplo, ha sido cultivada desde el año 3000 a.C. Desde su introducción en los Estados Unidos en 1800 hasta la actualidad, este mismo producto ha sido modernizado de diversas formas, para atraer la atención de consumidores (SOLAE).

Los granos de soya están compuestos por un 30% de hidratos de carbono (de los cuales un 15% es fibra), 18% de aceite (85% no saturado), 14% de humedad y un 35 a 40% de proteína. Esto la hace un alimento accesible para consumo humano, ya que puede ser derivado hacia diferentes productos (Del Valle, 1981). Además, su proteína contiene una gran cantidad de aminoácidos esenciales siendo una de las razones de su importancia para la industria alimentaria. La soya está considerada como una fuente completa de proteínas y de alta calidad. Además, es una buena fuente de fósforo, potasio, vitaminas del grupo B, zinc, hierro y vitamina E, que también actúa como antioxidante. Hoy en día, la soya ha sido modificada genéticamente para conferirle nuevas características como tolerancia a herbicidas y alto contenido de ácido oleico, entre otras (AGBIOS).

### 3.1 Soya tolerante a herbicidas

Padgett y colaboradores (1995) desarrollaron, identificaron y caracterizaron esta línea celular creando una línea de soya transgénica con un sistema novedoso para controlar las malezas, ya que es tolerante al herbicida glifosato y equivalente en su valor alimenticio con su contraparte tradicional.

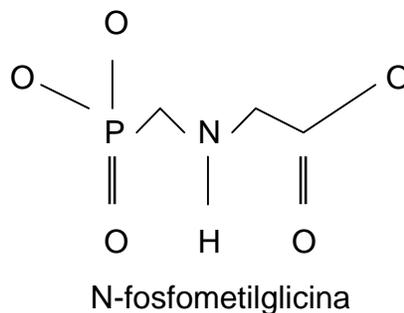
La variedad de soya utilizada en este estudio fue la Roundup® Ready (marca registrada en Estados Unidos) evento de transformación GTS 40-3-2. En México su nombre comercial es Solución Faena. El fenotipo que esta variedad de soya presenta es tolerancia al glifosato, ingrediente activo del herbicida Roundup® (Malik y col., 1989). Esta línea fue desarrollada como una opción para el control de malezas para el cultivo de soya.

El método de transformación utilizado para este caso fue biobalística o aceleración de partículas, el cual se basa en el “bombardeo” de células vegetales con microproyectiles de oro o tungsteno recubiertos con ADN, acelerados mediante una descarga eléctrica o helio presurizado (Luis, 2000). De esta manera, el gen *epsps* aislado de la bacteria *Agrobacterium tumefaciens* cepa CP4 fue introducido (Padgett, y col., 2005) creándose así, la variedad GM que expresa la enzima 5-enolpiruvilshikimato-3-fosfato sintasa (EPSPS) de origen bacteriano (CP4 EPSPS). Esto le confiere tolerancia al glifosato, ya que el herbicida se une a esta enzima en una conformación diferente a la que inhibe la catálisis de otras EPSPS. Esto se debe a que en la posición 100 de la secuencia de aminoácidos tiene una alanina en sustitución de una glicina, lo que impide que el glifosato se una covalentemente al sitio activo (Funke y col, 2006).

#### 4. Herbicida Glifosato (Roundup®)

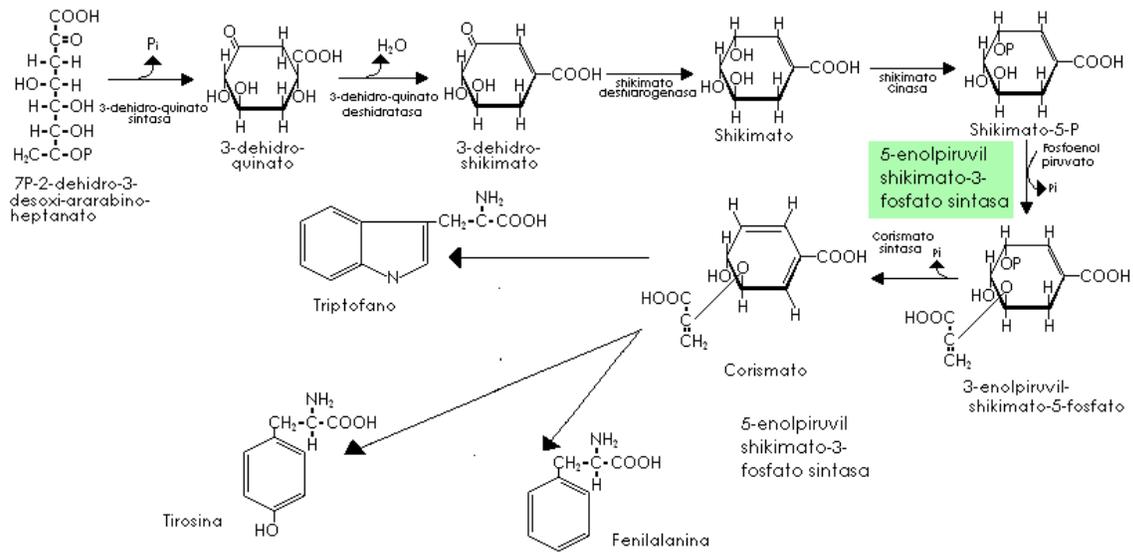
Uno de los mayores retos que afronta el agricultor al producir sus cultivos es controlar las malezas, ya que reducen drásticamente el rendimiento y la calidad de los cultivos. Para esto, se emplean herbicidas, como el glifosato, que pueden afectar no sólo a la maleza sino a la planta de interés. Por ello, el desarrollo de cultivos GM tolerantes a glifosato en los años 80's (Gasser, 1989) causó un gran impacto en las prácticas agrícolas de ese entonces. Este ofrecía al agricultor una nueva opción para el control de malezas, mediante el uso de un herbicida seguro para el ambiente, según la opinión de la empresa productora, compatible con los sistemas de conservación, capaz de incrementar la humedad del suelo, reducir la erosión y, en especial, proveer un control con un buen balance costo-beneficio (Holt y col, 1993).

El glifosato es un herbicida foliar activo, sistémico, que actúa en post-emergencia, de amplio espectro, usado para matar plantas no deseadas como pastos anuales y perenes, hierbas de hoja ancha y especies leñosas. Es un ácido usado en forma de sales, por lo general, la sal isopropilamina de glifosato o sal isopropilamina de N-fosfometil-glicina (Figura 1).



**Figura 1. Estructura química del glifosato (Voet, 1995).**

La acción del glifosato se debe a la inhibición de la EPSPS, lo cual impide a su vez la biosíntesis de aminoácidos aromáticos (fenilalanina, tirosina y triptofano), usados en la síntesis de proteínas esenciales para el crecimiento y sobrevivencia de las plantas (Figura 2). Como consecuencia, la planta expuesta al glifosato finalmente muere (Rubin, 1982).



**Figura 2. Reacción donde interviene la EPSPS**

Fuente: Forlani, 1994

Al ser un herbicida de amplio espectro, Roundup® mata a casi todos los tipos de plantas excepto a aquéllas que tienen el gen de la tolerancia. Por consiguiente, el agricultor puede aplicar un solo herbicida en sus campos sembrados con semillas tolerantes al herbicida y puede usar Roundup® eficazmente en la mayoría de las etapas de desarrollo de las plantas. El glifosato es absorbido por las hojas y tallos verdes y traslocado hacia las raíces y órganos vegetativos subterráneos, ocasionando la muerte total de las malezas emergidas. Los efectos son lentos, sobre todo en especies perennes, donde después de transcurridos 4-5 días desde la aplicación, comienza el amarillamiento de hojas y tallos que culminan con la muerte total de las malezas (Dupont®).

#### 4.1 Evaluación de la inocuidad de los cultivos GM

La evaluación de la inocuidad de un alimento producido con plantas GM se concentra en estudiar la posible analogía entre las secuencias de aminoácidos de la proteína heteróloga con las de toxinas y factores antinutrimientales proteínicos conocidos, así como su estabilidad térmica durante su procesamiento y su

degradación enzimática en modelos apropiados, representativos de los sistemas gástrico e intestinal.

Los estudios toxicológicos diseñados apropiadamente deben cumplir con los protocolos aprobados internacionalmente, con respecto al tiempo, dosis, número de animales, grupos control y análisis estadístico. Ello con el fin de evaluar el efecto que las nuevas proteínas expresadas en el OGM pueden tener en la salud de los consumidores potenciales, considerando su función y concentración en el alimento. Así mismo, estos estudios deben contemplar la presencia de la(s) nueva(s) proteína(s) expresadas como consecuencia de la transformación genética, la presencia potencial de otros nuevos constituyentes y los posibles cambios en los niveles de los constituyentes normales (COFEPRIS).

El bajo potencial de toxicidad de la línea transgénica GTS-40-30-2 se demostró haciendo un alineamiento de la secuencia de aminoácidos de la CP4EPSPS y realizando estudios de toxicidad aguda en ratones. Se determinó que la secuencia de aminoácidos de la CP4 EPSPS tiene una relación muy cercana a la secuencia de la EPSPS endógena de la soya. Además, no se encontró homología con alguna proteína toxigénica conocida, lo que llevó a considerar que no tiene algún efecto potencial adverso para el humano. Los estudios de la toxicidad oral aguda con la proteína CP4 EPSPS pura, sobreexpresada en bacterias, no revelaron efectos perjudiciales a los ratones, a una dosis de 572 mg/kg de peso, la cual es aproximadamente 1300 veces más que la del consumo de la CP4 EPSPS de la soya (Harrison, 1996).

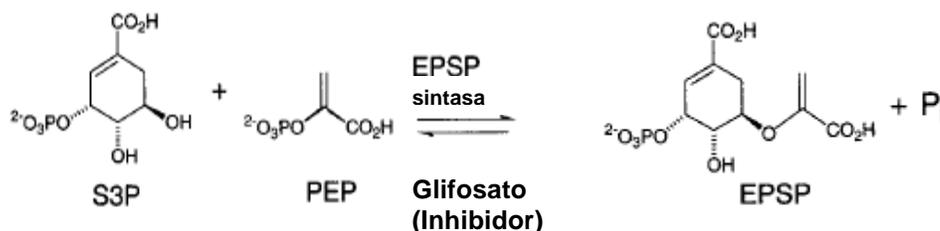
La absorción del glifosato en el tracto gastrointestinal en varias especies de mamíferos es aproximadamente del 30-36% de la dosis y, a través de la piel, es aproximadamente el 3%. El glifosato no es metabolizable en mamíferos pero su eliminación es rápida si se absorbe, siendo excretado aproximadamente el 80% ya sea por la orina, o por las heces antes de cumplir 24 horas de administración oral. Hasta ahora, no existe evidencia de la acumulación del glifosato en el cuerpo del

animal en estudio (Williams y col., 2000). Estudios de “irritación” en animales como conejos y cerdos han mostrado que el glifosato no irrita la piel de ninguno de estos dos tipos de animales, pero sí puede irritar los ojos de los conejos (AGBIOS). La velocidad de degradación del herbicida depende de la actividad microbiana que exista en el suelo, agua o algún sedimento. Cuando el glifosato se metaboliza, el metabolito formado de mayor proporción es el ácido aminometilfosfónico (AMPA), el cual se degrada a dióxido de carbono y fósforo inorgánico (Pi). Cabe mencionar que el glifosato se degrada al ponerse en contacto con el suelo (Carlisle y Trevors, 1988).

### 5. Enzima 5-enolpiruvilshikimato-3-fosfato sintasa (EPSPS)

La 5-enolpiruvilshikimato-3-fosfato sintasa (EPSPS) es una enzima que forma parte de la ruta de biosíntesis de aminoácidos aromáticos y otros compuestos aromáticos en plantas (Figura 2). Esta enzima cataliza la formación de 5-enolpiruvilshikimato-3-fosfato (EPSP), a partir del shikimato-3-fosfato (S-3-P) y el fosfoenolpiruvato (PEP; Stauffer, 2001).

Como ya se ha mencionado previamente, cuando las plantas tradicionales son tratadas con glifosato, éstas no pueden producir los aminoácidos aromáticos necesarios para sobrevivir y mueren, inhibiendo la catálisis de la EPSPS (Figura 3; Heldt, 2005).

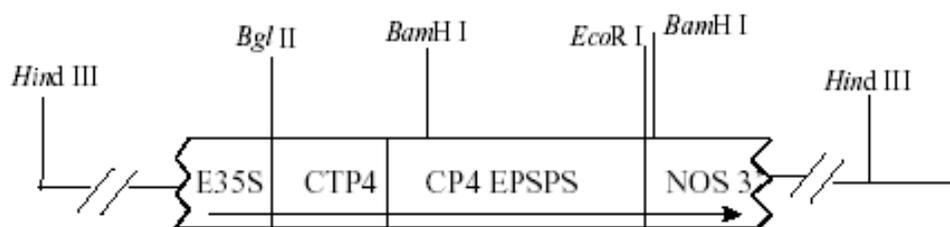


**Figura 3. Reacción catalizada por la 5-enolpiruvilshikimato-3-fosfato sintasa, blanco de acción del glifosato**

Fuente: Heldt, 2005

La secuencia de la proteína CP4EPSPS sufrió una mutación de un aminoácido: Ala-100 por Gly-96 en su sitio activo, este cambio provoca que no se pueda inhibir la enzima CP4EPSPS por acción del glifosato (Funke, 2006). La ruta de la biosíntesis de aminoácidos aromáticos está presente en hongos, plantas y microorganismos, pero no en animales (Jaworski, 1972). Así, el glifosato puede servir como herbicida, sin dañar a los animales y a los organismos que cuenten con la ruta (Lewis, 1999). Además, el herbicida glifosato (N-fosfometilglicina) es un inhibidor específico de la EPSPS, es decir, no afecta a otras enzimas de la ruta (Steinrücken, 1980).

La empresa Monsanto desarrolló varios cultivos comerciales resistentes a glifosato, aislando y clonando el gen de la enzima EPSPS, para ser insertado en varias especies (Monsanto, 2000). Para ello, utilizaron genes de varios microorganismos, así como varios péptidos de transferencia al cloroplasto de diferentes microorganismos y plantas (Barry y col., 1997). Finalmente, el gen utilizado fue el del microorganismo *Agrobacterium tumefaciens* cepa CP4 como se mencionó anteriormente. La construcción desarrollada incluye el gen de interés, más algunos elementos regulatorios accesorios para controlar la expresión de la enzima (Figura 4):



**Figura 4. Esquema de la construcción del evento GTS-40-30-2.** Se muestra el promotor proveniente del virus del mosaico de la coliflor (E35S), el péptido de tránsito al cloroplasto (CTP4), el transgen (CP4EPSPS) y el terminador NOS 3.

Fuente: AGBIOS

Los elementos presentes en dicha construcción son el promotor CaMV35S del virus del mosaico de la coliflor, una secuencia CTP4 (Cloroplast Transit Peptide) de *Petunia hybrida*, la cual codifica para un péptido de transferencia que

permite a la EPSPS translocarse al cloroplasto, el gen codificante de la proteína CP4EPSPS y el terminador NOS proveniente del microorganismo *A. tumefaciens* y que señala el término de la transcripción. El gen completo de la CP4 EPSPS (Figura 5) insertado en la soya tiene un tamaño de 2439 pb (Gen Bank No. AB209952).

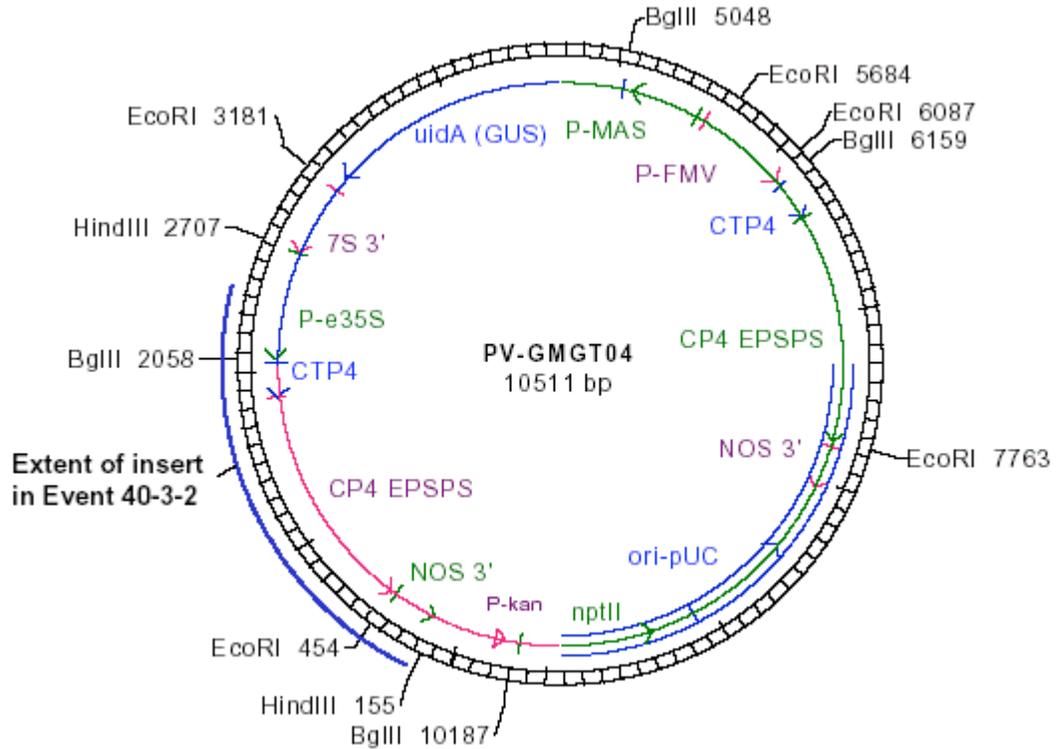
```

1  tggaaaagga aggtggctcc taaaaatgcc atcattgcca taaaggaaag gccatcgttg
61  aagatgcctc tgccgacagt ggtcccaaaag atggaccccc acccagcagg agcatcgtagg
121  aaaaagaaga cgttccaacc acgtcttcaa agcaagtgga ttgatgtgat atctccactg
181  acgtaagggg tgacgcacaa tcccactatc cttcgcaaga cccttctctc atataaggaa
241  gttcatttca tttggagagg acacgctgac aagctgactc tagcagatct ttcaagaatg
301  gcacaaatta acaacatggc acaagggata caaacctta atccaattc caatttccat
361  aaacccaag  ttctaaatc  ttcaagtttt  cttgtttttg  gatctaaaaa  actgaaaaat
421  tcagcaaat  ctatgttgg  tttgaaaaaa  gattcaattt  ttatgcaaaa  gttttgttcc
481  tttaggattt  cagcatcagt  ggctacagcc  tgcattgctc  acggtgcaag  cagccggccc
541  gcaaccgccc  gcaaatcctc  tggcctttcc  ggaaccgtcc  gcattcccgg  cgacaagtcc
601  atctcccacc  ggtccttcat  gttcggcggt  ctcgcgagcg  gtgaaacgcg  catcaccggc
661  cttctggaag  gcgaggacgt  catcaatacg  ggcaaggcca  tgcaggccat  gggcgccagg
721  atccgtaagg  aaggcgacac  ctggatcac  gatggcgctc  gcaatggcgg  cctcctggcg
781  cctgaggcgc  cgctcgattt  cggcaatgcc  gccacgggct  gccgcctgac  catgggctc
841  gtcggggtct  acgatttoga  cagcaccttc  atcggcgagc  cctcgctcac  aaagcgccc
901  atgggcccgc  tgttgaacc  gctgcgcgaa  atggcgctgc  agtgaaatc  ggaagacggt
961  gaccgtcttc  ccgttacctt  gcgcgggccc  aagacgccga  cgcgatcac  ctaccgctg
1021  ccgatggcct  ccgcacaggt  gaagtccgcc  gtgctgctcg  ccggcctcaa  cacgcccggc
1081  atcacgacgg  tcatcgagcc  gatcatgac  tgcgatcata  cggaaaagat  gctgcagggg
1141  tttggcgcca  acctaccgt  cgagcggat  gcggacggcg  tgcgcaccat  ccgctggaa
1201  ggccgcgcca  agctcacgg  ccaagtcatc  gacgtgccgg  gcgaccctc  ctgcagggc
1261  ttcccgtggt  ttgcccctc  gcttgttcc  ggctccgacg  tcaccatcct  caacgtgctg
1321  atgaaccca  cccgcaccgg  cctcatctg  acgtgcagg  aaatgggccc  cgacatcgaa
1381  gtcatacaac  tgcccttgc  cggcggcgaa  gacgtggcgg  acctgcgctg  tcgctcctcc
1441  acgctgaagg  gcgtcacggt  gccggaagac  cgcgcgcctc  cgatgatcga  cgaatatccg
1501  attctcgctg  tcgcccgcg  cttcgcggaa  gggcgaccg  tgatgaacgg  tctgaaagaa
1561  ctcccgctca  aggaaagcga  ccgctctcg  gccgtcgcca  atggcctcaa  gctcaatggc
1621  gtggattgcg  atgaggcgga  gacgtcgctc  gtcgtgctg  gccgcctga  cggcaagggg
1681  ctgggcaacg  cctcgggccc  cgcctcgcc  acccatctg  atcacgcat  cgccatgagc
1741  ttctctgca  tgggctcgt  gtcgaaaaac  cctgtcacgg  tggacgatgc  cacgatgatc
1801  gccacgagct  tcccggagt  catggacctg  atggccgggc  tggcgcgcaa  gatcgaactc
1861  tccgatacga  aggtgctctg  atgagctoga  attcgagctc  ggtaccggat  ccaattcccg
1921  atcgttcaaa  catttgcaa  taaagttct  taagattgaa  tcctgttgcc  ggtcttgcca
1981  tgattatcat  ataattctg  ttgaattacg  ttaagcatgt  aataattaac  atgtaatgca
2041  tgacgttatt  tatgagatgg  gtttttatga  ttagagtccc  gcaattatac  atttaatacg
2101  cgatagaaaa  caaaatatag  cgcgcaaac  aggataaatt  atcgcgcgcg  gtgtcatcta
2161  tgttactaga  tcggggatcg  atccccacc  ggtccttcat  gttcggcggg  ctgcgagcgc
2221  gtgaaacgcg  catcacggc  cttctggaag  gcgaggacgt  catcaatacg  ggcaaggcca
2281  tgcaggccat  gggcgccagg  atccgtaagg  aaggcgacac  ctggatcac  gatggcgctg
2341  gcaatggcgg  cctcctggcg  cctgagcgc  cgctcgattt  cggcaatgcc  gccacggcct
2401  gccgcctgac  catgggctc  gtcggggtct  acgatttcaa  gcgcatcatg  ctgggaa

```

**Figura 5. Secuencia completa del gen CP4 EPSPS.** Las bases 1 a la 265 pertenecen al promotor 35S del Virus del Mosaico de la Coliflor (sin color), la base 266 a la 297, son de región no traducida (amarillo), de la base 298 a la 510 se encuentra la región codificante del péptido de tránsito hacia el cloroplasto de *Petunia híbrida* (verde), de la 511 a la 1878 pertenece al gen *epsps* de *Agrobacterium tumefaciens* (azul), de la base 1885 a la 2186 corresponden al terminador NOS de *Agrobacterium tumefaciens* (gris) y de la 2186 a la 2439 (rojo), corresponden a un fragmento repetido truncado de la CP4 EPSPS.

El mapa del plásmido, proveniente de *Agrobacterium tumefaciens*, utilizado para la transformación de la línea RR es el que se muestra en la Figura 6:



**Figura 6. Plásmido PV-GMGT04**

Fuente: Padgette, 1995

Mediante análisis de Southern blot, realizados al ADN genómico del evento GTS 40-3-2, se demostró que hay dos sitios de integración en tandem, uno de ellos contiene una copia funcional del gen CP4 EPSPS de 1375 pb y el segundo sitio contiene un segmento no funcional del gen CP4 EPSPS de 72 pb (Monsanto, 2000). La expresión de este gen da como resultado una proteína aproximadamente de 46 kDa (polipéptido de 456 aminoácidos) y los niveles de expresión fueron cuantificados por ELISA teniendo una concentración promedio de 239 µg/g de tejido fresco en semillas y 495 µg/g de tejido fresco en hojas (AGBIOS). Los niveles de CP4EPSPS en soja tolerante a herbicida glifosato fueron de aproximadamente 0.03% en tejido fresco y 0.08% de la proteína total (Padgette y col., 1995).

## 6. Métodos de purificación de la EPSPS de diferentes organismos

Para estudiar una proteína particular es necesario separarla del resto de las proteínas y disponer de técnicas que permitan determinar sus propiedades. En la mayoría de los casos, se deben aplicar métodos diferentes de modo secuencial para purificar completamente a una proteína (Nelson, 2000).

Algunos métodos de purificación utilizados para separar la enzima EPSPS, proveniente de diversas fuentes y organismos se basan en el uso de columnas de intercambio iónico y exclusión molecular principalmente, como se muestra en la Tabla 4:

**Tabla 4. Métodos de purificación reportados para la EPSPS, a partir de diversos organismos**

Organismo	Métodos de Purificación	Fuente	Referencias
Maíz ( <i>Zea mays</i> )	Sulfato de amonio DEAE-Sephacel™ Celulosa fosfato	Línea celular	Forlani, y col., 1994
<i>E. coli</i>	Sulfato de amonio DEAE-Sephacel™ Fenilsefarosa Fosfocelulosa	Línea celular	Lewendon y Coggins, 1983
<i>Petunia hybrida</i>	Sulfato de amonio Sephadex G-25 DEAE-celulosa Hidroxiapatita Fenilagarosa Sephacryl S-200™	Línea celular	Steinrücken, y col., 1986
Chícharo ( <i>Pisum sativum</i> )	Sulfato de amonio Centrifugación a altas velocidades DEAE- sefarosa Fenil-sefarosa fosfocelulosa	Hoja	Mousdale y Coggins, 1984

<b>Organismo</b>	<b>Métodos de purificación</b>	<b>Fuente</b>	<b>Referencias</b>
Sorgo ( <i>Sorghum vulgare</i> )	Sulfato de amonio DEAE-cellulose Hidroxiapatita Fenilagarosa	Hoja	Ream, y col., 1988
<i>Mycobacterium tuberculosis</i>	Q-Sepharose	Línea celular	Oliveira, y col., 2003
Soya ( <i>Glicine max</i> )	Sulfato de amonio Q-Sepharose FF Fenil-sefarosa Mono Q 10/10 Fenil-Superosa 5/5	Semillas	Harrison, y col., 1996

Es importante mencionar que Padgette y col., en 1995, caracterizaron la línea de soya GTS 40-30-2, transformada con el vector PV-GMGT04 y la línea 61-67-1 (también tolerante al herbicida) transformada con el vector PV-GMGT05 que contiene también el gen CP4EPSPS, enfocándose principalmente en la línea 40-30-2 y no en la 61-67-1, por no ser comercial esta última. En estudios realizados por Harrison y col. (1996) purificaron a la CP4EPSPS de semillas de soya de la línea 61-67-1, como se describió anteriormente y, con base en esta purificación, realizaron estudios toxicológicos donde concluyeron que la soya tolerante a herbicidas es tan segura y nutritiva como la soya tradicional.

A partir de estas metodologías de purificación, para este trabajo se decidió utilizar las columnas de intercambio iónico para la purificación de la variedad Solución Faena, (línea 40-3-2), para identificarla plenamente como la proteína heteróloga. Una vez conocida la(s) fracción(es) de la soya en las que se recupera la CP4EPSPS, se podría generar un aislado proteínico especial de soya, enriquecido con la proteína heteróloga. Ello, para garantizar que en los estudios de alimentación con animales de laboratorio los posibles cambios observados en su histología sean debidos al consumo de la fracción enriquecida con la proteína heteróloga, comparándolo con un aislado de soya similar, pero obtenido a partir de soya no GM.

#### **IV. Hipótesis.**

Si las plantas transgénicas de soya Roundup® Ready expresan únicamente la construcción publicada por Monsanto en las bases de datos, entonces solamente se producirá una proteína heteróloga a partir del gen insertado.

## **V. Objetivo General**

Separar las proteínas expresadas en granos de soya Roundup® Ready e identificar la(s) proteína(s) heteróloga(s).

## **Objetivos específicos**

- Determinar cuáles de las muestras de granos de soya con las que se cuenta son transgénicas.
- Cuantificar el porcentaje de soya transgénica en cada muestra.
- Implementar un sistema que permita detectar a la proteína heteróloga de las diversas fracciones proteínicas de la soya por medios inmunológicos.
- Desarrollar un protocolo para enriquecer a la proteína EPSPS por Cromatografía de intercambio iónico.



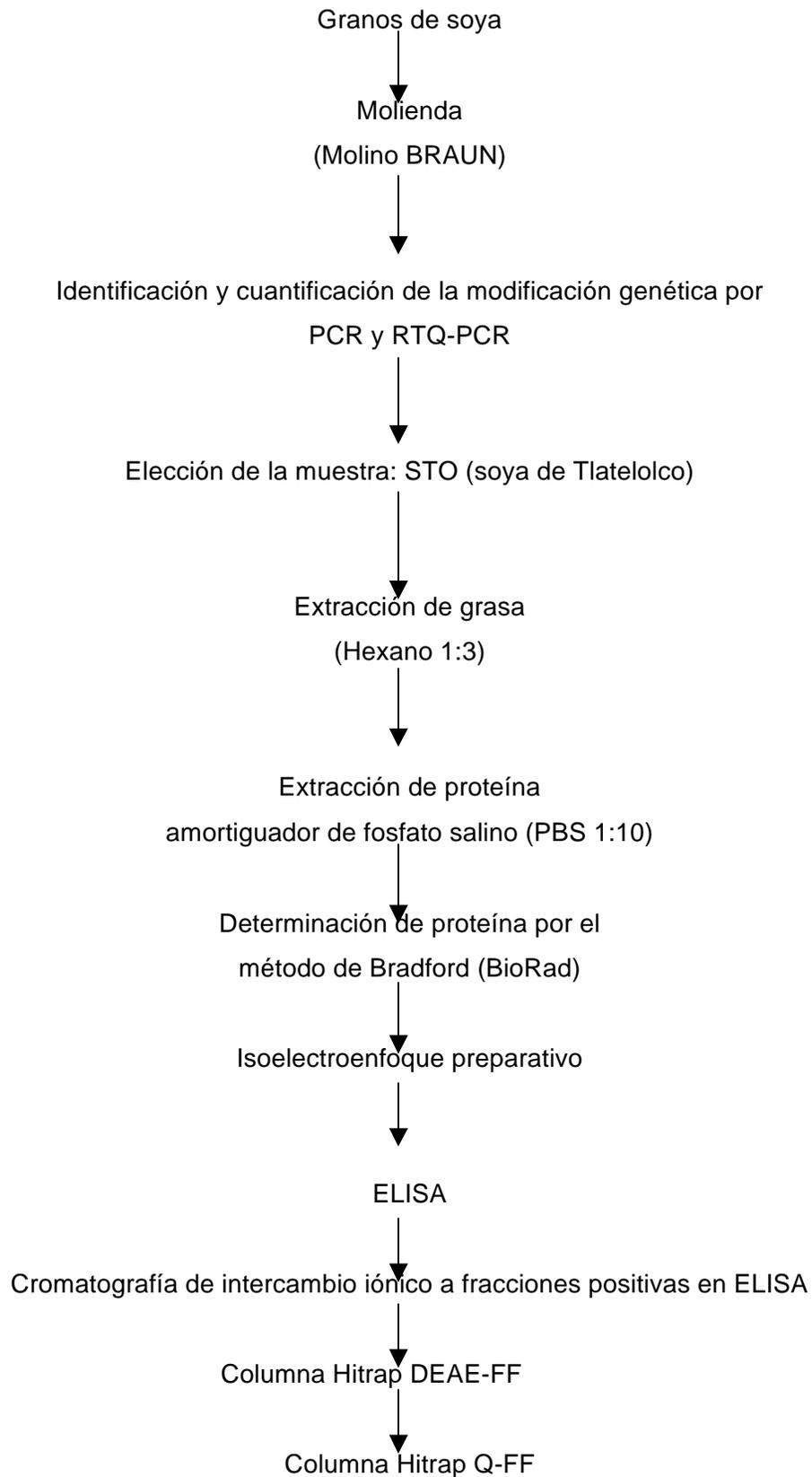
## VI. MATERIALES Y MÉTODOS

### **Descripción de la estrategia general.**

Para la separación de la proteína CP4 EPSPS de las demás proteínas de la soya se obtuvieron muestras de grano de soya de diferentes partes de la República Mexicana, con la finalidad de caracterizarlas y decidir cuál de ellas se utilizaría en la parte experimental de esta tesis.

Primero, se realizó PCR punto final y PCR tiempo real (RTQ-PCR), con el objeto de identificar en cuáles muestras había granos transgénicos. Una vez seleccionada la muestra con base en los resultados obtenidos de las PCRs, se extrajo la fracción lipídica con el objeto de quitar la mayor cantidad de elementos que pudieran afectar a la separación. La extracción de la fracción lipídica se realizó con hexano, como solvente, en una proporción 1:3 (peso/volumen). Posteriormente, se realizó la extracción de proteína con amortiguador de fosfato salino (PBS) en relación 1:10 ( 10 g de muestra de soya : 100 mL del amortiguador). Una vez obtenido el extracto proteínico de soya, se sometió a un isoelectroenfoque preparativo con la finalidad de comenzar a separar a las proteínas con base en su punto isoelectrónico. Finalmente, se seleccionó la fracción por ELISA y se sometió a cromatografía de intercambio aniónico.

## Estrategia experimental



## 2. Obtención y pre-tratamiento de la materia prima.

Las muestras SH, SE, SC1 y SC2 fueron proporcionadas por la Dra. Ana María Calderón del CIAD de Hermosillo, Sonora, provienen de diferentes regiones del norte del país. Además de éstas, se adquirieron otras muestras de frijoles de soya provenientes de diferentes regiones del Edo. de México, Distrito Federal y del estado de Chiapas (Tabla 5).

**Tabla 5. Muestras de granos de soya obtenidas en distintas regiones del país**

Clave	Descripción de la muestra	Cant. aprox. (Kg)
SH	Soya Héctor	2
SE	Soya Esperanza	2
SC1*	Soya Cajeme (1)	3
SC2*	Soya Cajeme (2)	2
STO	Soya comprada en Tlatelolco	1
STA	Soya comprada en Tonalá, Chis.	1
SFL	Soya comprada en una tienda en el sur de la ciudad	1
SNA	Soya comprada en NUTRISA, Satélite	1

\* La diferencia entre SC1 y SC2 radica en la fecha de recepción en el laboratorio, aunque son de la misma variedad.

El producto se almacenó a 4°C en bolsas de plástico con cierre hermético durante el transcurso del presente estudio.

### **2.1 Homogenización de la materia prima.**

La muestra se trituró en un molino marca BRAUN tres veces hasta obtener un polvo fino.

### **2.2 Extracción de la fracción lipídica de las muestras**

Para la extracción de grasa se utilizó hexano como solvente de extracción en relación 1:3 (peso/volumen). Las muestras fueron agitadas en una placa agitadora (Marca LEO Orbit Shaker, Modelo 3518, Lab-Line Instruments, Inc.) durante 30 minutos a 125 rpm a 4°C y posteriormente centrifugadas a 10000 rpm / 20 min / 4°C en una centrifuga BECKMAN, modelo J2-MC, rotor JA-10. Se eliminó el sobrenadante y se repitió la extracción tres veces más.

### **2.3 Extracción de proteína**

Las proteínas solubles fueron extraídas en amortiguador PBS, pH 7.4 de la siguiente manera:

1. Pesar 0.5 g de cada muestra
2. Adicionar amortiguador PBS relación 1:5 (peso/volumen)
3. Mezclar en placa agitadora hasta estabilizar el pH a 7.4
4. Centrifugar a 10000 rpm, 30 minutos, 4°C (Centrifuga BECKMAN, modelo J2-MC, rotor JA-10)
5. Pasar el sobrenadante a un tubo limpio

## **3. Cuantificación de proteína soluble por el método de Bradford (BIORAD)**

Esta determinación se basa en los cambios de color del azul brillante G-250 de Coomassie, como respuesta a diferentes concentraciones de proteína soluble presente en la muestra. La absorbancia del complejo colorante-proteína se da en el rango de 265-595 nm, incrementando la absorción a 595 nm en donde es determinada. El colorante se une principalmente a los residuos de aminoácidos básicos y aromáticos, especialmente arginina. Se pueden presentar interferencias

por alguna interacción de la proteína con otros compuestos como amortiguadores o detergentes (Bradford, 1976).

A un volumen de 800  $\mu$ L de solución proteínica se le agrega un volumen de 200  $\mu$ L de Azul de Coomasie, se mezclan en un vórtex, se espera cinco minutos y se toma la lectura, en celdas de volumen reducido, a una absorbancia de 595 nm.

#### **4. ELISA (Ensayo de Inmuno-adsorción ligado a enzimas).**

El método de ELISA se fundamenta en el uso de anticuerpos específicos para reconocer a la proteína de interés (Van Weemen, 1971). Este procedimiento es capaz de discriminar proteínas presentes en el producto analizado, de entre cientos y a veces miles de distintas proteínas presentes en la misma muestra (Brett et al., 1999). Este es un método de detección basado en el reconocimiento específico de la proteína heteróloga expresada por el organismo transgénico por un anticuerpo acoplado a una enzima. El anticuerpo reconoce la proteína específica previamente extraída a partir del OGM y se une a ella. El complejo es posteriormente detectado mediante una reacción enzimática cuyo producto es cromogénico (Jensen, 2001).

El kit utilizado para este estudio se adquirió por la compañía FUMAPLANT, S.C, principales distribuidores de la compañía AGDIA, compañía productora de este tipo de kits. El kit utilizado en este estudio es el CP4-EPSPS con número de catálogo PSP 74000, que cuenta con una placa de 96 pozos recubierta con anticuerpos específicos para la proteína CP4-EPSPS, los amortiguadores necesarios para la extracción y el lavado, el amortiguador para la enzima, el conjugado de la enzima peroxidasa, el sustrato para la enzima y los controles positivo y negativo.

La detección se hizo siguiendo las instrucciones del fabricante, descritas a continuación:

#### **4.1 Preparación de la placa de ELISA CP4-EPSPS**

La placa generalmente se organizó de la siguiente manera:

En las primeras tres líneas se colocó la curva patrón para lo cual el control positivo del kit (liofilizado) se reconstituyó con 2.5 mL de amortiguador MEB, el cual tuvo una concentración de 1%. A partir de esta concentración se realizaron las diluciones necesarias para obtener la curva siguiente: 1, 1/10, 1/100, 1/1000 y 1/10000. El blanco de reactivos se colocó en los primeros dos pozos: (A1 y B1) utilizando el amortiguador de extracción. En los demás pozos se colocaron los extractos de las muestras de soya o las fracciones a evaluar.

#### **4.2 Protocolo de ELISA**

1. Colocar 100  $\mu$ L de cada muestra en su respectivo pozo, según el esquema planeado.
2. Dejar incubando la placa en una cámara "húmeda" toda la noche a 4°C.
3. Una vez lista la incubación vaciar el contenido volteando la placa rápidamente.
4. Lavar los pozos con amortiguador PBST 1X, llenando los pozos con el amortiguador de lavado PBST 1X y vaciar enseguida. Repetir de 6 a 7 veces para asegurar la ausencia de la unión inespecífica a otras proteínas.
5. Realizar un último lavado con un tiempo de incubación de 3 minutos a temperatura ambiente en la cámara "húmeda"
6. Vaciar el contenido y dejar escurrir en papel absorbente, sacudir hasta que estén secos los pozos.
7. Añadir 100  $\mu$ L del conjugado de la enzima.
8. Dejar incubando 60 minutos en la cámara "húmeda".
9. Añadir 100  $\mu$ L de sustrato para la enzima.
10. Dejar incubando 20 minutos.
11. Leer a 630 nm.

La lectura se realizó en un lector de microplacas marca Benchmark, modelo 680 series microplate readers (BIORAD). Visualmente, los pozos que desarrollaron un color azul indicaron resultados positivos. Cuando el color es azul muy claro indica un resultado negativo, al igual que la ausencia de color. Este mismo procedimiento se aplicó a las muestras de las fracciones obtenidas en cada paso posterior a las columnas de intercambio iónico como método de identificación de la enzima en estudio.

## **5. Detección de secuencias GM**

Este experimento se realizó con el propósito de verificar cuáles de las muestras eran transgénicas y decidir con cuál se trabajaría posteriormente.

### **5.1 Extracción de ADN**

La extracción se realizó con el reactivo DNAzol™ (GIBCO-BRL), que permitió el aislamiento de ADN genómico de muestras sólidas o líquidas de plantas, animales y bacterias. Está compuesto por una mezcla de guanidina-detergente, lo que permitió la lisis de las células y la solubilización del ADN, posteriormente precipitado con isopropanol, permitiendo la extracción del 70-100% del ADN de las muestras (DNAzol™ Reagent, Life Technologies Form 3819) (INVITROGEN).

Para la extracción de ADN se realizan los siguientes pasos:

1. Pesar entre 40 y 60 mg de muestra.
3. Agregar a la muestra un mL de DNAzol y homogeneizar agitando el tubo por inversión aproximadamente 10 veces.
4. Centrifugar 10 minutos a 10,000 rpm a temperatura ambiente.
5. Pasar el sobrenadante (aprox. 0.9 mL) a un tubo.
6. Agregar 0.5 mL de isopropanol frío (-20°C) y mezclar por inversión.
7. Incubar a 4°C por 10 minutos.
8. Centrifugar 10 minutos a 10,000 rpm a temperatura ambiente y eliminar el sobrenadante por decantación.

9. Lavar con 1 mL de etanol al 70% y centrifugar 2 minutos a 10,000 rpm a temperatura ambiente y eliminar el sobrenadante por decantación.
10. Dejar secar el precipitado dejando el tubo abierto por unos 15 minutos.
11. Disolver el ADN en 50  $\mu$ L de NaOH 10 mM e incubar 15 minutos a 37°C para ayudar a disolverlo.
12. Centrifugar 5 minutos a 10,000 rpm a temperatura ambiente para eliminar el material insoluble y tomar el sobrenadante, pasarlo a un tubo limpio.
13. Neutralizar con 1/20 del volumen de Tris 1 M, pH 8.

## **5.2 Reacción en Cadena de la Polimerasa (PCR) Punto final**

El principio de esta técnica es la amplificación de secuencias específicas de ADN a partir de un par de secuencias cortas de ADN de cadena sencilla, conocidas como “cebadores”, “primers” u “oligos” que delimitan cierta región de ADN molde que será amplificada por la enzima polimerasa Taq (Gachet, 1999).

La metodología de PCR consta de tres pasos. *Desnaturalización*: La doble hélice del ADN molde es separada mediante un calentamiento a una temperatura de 94°C, suficiente para romper los puentes de hidrógeno que unen a las cadenas.

*Alineamiento*: Se disminuye la temperatura hasta aproximadamente 55°C, lo que permite que los cebadores, que delimitan la región de interés sean unidos (hibridación) a su cadena complementaria. El molde de ADN, los cebadores, la ADN polimerasa y los nucleótidos son adicionadas desde el inicio de la reacción.

*Extensión*: Finalmente, la ADN polimerasa sintetiza la nueva cadena mediante la adición del nucleótido complementario de cada base, a 72°C que es la temperatura óptima de esta enzima. Una vez que el ciclo es completado, puede ser repetido, y por cada ciclo el número de copias se duplicará, teniendo como resultado una amplificación exponencial (Vollenhofer, 1999).

### 5.3 Prueba de PCR para el transgén

Las secuencias de los cebadores fueron obtenidas en la página web de la base de datos de detección de OGMs del International Center for Genetic Engineering and Biotechnology (ICGEB) y fueron sintetizados por la compañía INVITROGEN. Estos cebadores fueron aplicados con éxito a diferentes materias primas, ingredientes y alimentos que contienen soya, por ejemplo: piensos, harinas, concentrados, harinas y soya texturizada (ICGEB). Estos oligos (ICGEB), fueron creados para detectar la región específica de la secuencia donde el gen y el elemento regulador se fusionan, en este caso, amplificando parte de la región promotora del CaMV 35S y parte del péptido de tránsito (CTP) del gen EPSPS de *Petunia híbrida*, lo que identifica en forma inequívoca al evento de transformación específico de que se trate.

Cebadores para detectar el transgén: RR1-F y RR1-R

5´-catttggagaggacacgctga-3´ (sentido) Tm=64°C Cebador RR1-F

5´-gagccatggttgaatttggcc-3´ (antisentido) Tm=66°C Cebador

RR1-R

Tamaño del amplicón : 74pb

Los cebadores descritos se visualizan en diferentes colores sobre la secuencia completa, descrita anteriormente (figura 5).

```

1 tggaaaagga aggtggctcc tacaatgcc atcattgcca taaaggaaag gccatcgttg
61 aagatgcctc tgccgacagt ggtcccaaag atggaccccc acccagcagg agcatcgttg
121 aaaaagaaga cgttccaacc acgtcttcaa agcaagtgga ttgatgtgat atctccactg
181 acgtaaggga tgacgcacaa tcccactatc cttcgcaaga cccttctctc atataaggaa
241 gttcatttca tttggagagg acacgctgac aagctgactc tagcagatct tcaagaatg
301 gcacaaatta acaacatggc acaagggata caaacctta atccaattc caatttccat
361 aaacccaag ttctaaatc ttcaagttt cttgtttttg gatctaaaaa actgaaaaat
421 tcagcaaat ctatgttgg tttgaaaaa gattcaattt ttatgcaaaa gttttgttcc
481 tttaggattt cagcatcagt ggctacagcc tgcatgcttc acggtgcaag cagccggccc
541 gcaaccgccc gcaaatcctc tggcctttcc ggaaccgctc gcatcccg cgacaagtgc
601 atctcccacc ggtccttcat gttcggcggc ctcgagagcg gtgaaacgcg catcaccggc
661 cttctggaag gcgaggacgt catcaatacg ggcaaggcca tgcaaggccat gggcgccagg
721 atccgtaagg aaggcgacac ctggatcatc gatggcgctc gcaatggcgg cctcctggcg
781 cctgaggcgc cgctcgattt cggcaatgcc gccacgggct gccgcctgac catgggcctc
841 gtcggggtct acgatttoga cagcaccttc atcggcgacg cctcgctcac aaagcgccc
901 atgggcccgc tgttgaacc cgtgcgcgaa atgggctgac aggtgaaatc ggaagcggg
961 gaccgtcttc ccgttacctt gcgcgggccc aagacgccga cgccgatcac ctaccgctg

```

```

1021 ccgatggcct ccgcacaggt gaagtccgcc gtgctgctcg ccggcctcaa cagccccggc
1081 atcacgacgg tcatcgagcc gatcatgacg tgcgatcata cggaaaagat gctgcagggc
1141 ttggcgcca accttaccgt cgagacggat gcggacggcg tgcgcacccat ccgcctggaa
1201 ggccgcggca agctcaccgg ccaagtcatc gacgtgccgg gcgaccctgc ctgcagggcc
1261 ttcccgtgg ttgcggccct gcttgttccg ggctccgacg tcaccatcct caacgtgctg
1321 atgaacccca cccgcaccgg cctcatcctg acgctgcagg aaatggggcg cgacatcgaa
1381 gtcacaaacc tgcgccttgc cggcggcgaa gacgtggcgg acctgcgcgt tcgctcctcc
1441 acgctgaagg gcgtcacggt gccggaagac cgcgcgcctc cgatgatcga cgaatatccg
1501 attctcgctg tcgccccgc cttcgcggaa gggcgaccg tgatgaacgg tctggaagaa
1561 ctccgcgtca aggaaagcga ccgcctctcg gccgtcgcca atggcctcaa gctcaatggc
1621 gtggattgcy atgagggcga gacgtcgctc gtcgtgcytg gccgccctga cggcaagggg
1681 ctggcaacg cctcggggcg cgcgcgtccc accatctcg atcaccgcat cgccatgagc
1741 ttctctgca tgggcctcgt gtcggaaaac cctgtcacgg tggacgatgc cacgatgatc
1801 gccacgagct tcccggagtt catggacctg atggccgggc tggcgcgaa gatcgaactc
1861 tccgatacga aggtcgctg atgagctcga attcgagctc ggtaccggat ccaattcccg
1921 atcgttcaaa catttgcaa taaagtttct taagattgaa tcctggtgcc ggtcttgcca
1981 tgattatcat ataattctg ttgaattacg ttaagcatgt aataattaac atgtaatgca
2041 tgacgttatt tatgagatgg gttttatga ttagagtccc gcaattatac atttaatac
2101 cgatagaaaa caaaatatag cgcgcaaac aggataaatt atcgcgcgcy gtgtcatcta
2161 tgttactaga tcggggatcg atccccacc ggtccttcat gttcggcggt ctgcgagcgy
2221 gtgaaacgcy catcaccggc cttctggaag gcgaggacgt catcaatacgy ggcaagggca
2281 tgcaggccat gggcgccagg atccgtaagg aaggcgacac ctggatcatc gatggcgctg
2341 gcaatggcgy cctcctggcy cctgagggcg cgcctcgatt cggcaatgcc gccacgggct
2401 gccgcctgac catgggcctc gtcggggtct acgatttcaa gcgcatcatg ctgggaa

```

Todas las Reacciones en Cadena de la Polimerasa se llevaron a cabo en un termociclador de la marca Techgene ( Modelo FTGENE 2D, Inglaterra).

Se realizó con el siguiente programa de tiempos y temperaturas:

<i>Desnaturalización</i>	3 min a 94°C	1 ciclo
	40 s a 94°C	
<i>Alineamiento</i>	40 s a 60°C	35 ciclos
	40 s a 72°C	
<i>Extensión</i>	7 min a 72°C	1 ciclo

Se realizaron varias corridas para las diferentes muestras con que se contaba. Se utilizó como control positivo la misma muestra de soya STO. Para la visualización de los productos se hicieron electroforesis en geles de agarosa al 2%. En la Tabla 6 se presentan los reactivos empleados para la PCR.

**Tabla 6. Reactivos empleados en la PCR para detectar el gen RR**

Reactivo	[inicial]	[Final]	Marca
Amortiguador PCR	10 X	2 X	Invitrogen
MgCl <sub>2</sub>	50 mM	1.5 mM	Invitrogen
dNTP´s	2.5 mM	0.2 mM	Invitrogen
Primer RR1-F	10 µM	0.5 µM	Invitrogen
Primer RR1-R	10 µM	0.5 µM	Invitrogen
Taq polimerasa	2.5 U	1 µL/reacción	Invitrogen
ADN molde	1 µL es el volumen determinado experimentalmente para que se lleve a cabo la reacción		
H <sub>2</sub> O	c.b.p. 20 µL		

#### **5.4 Purificación del producto de PCR. Método Gen Elute Agarose Spin Column (SIGMA).**

La purificación del producto de PCR se realizó con el propósito de eliminar los residuos de los reactivos de la reacción (dNTP´s, sales, cebadores, enzima, etc.) lo que permite evitar interferencias para una futura secuenciación de los productos de la PCR. Ésta se realizó en las minicolumnas Gen Elute Agarose Spin Columns (SIGMA) diseñadas para purificar fragmentos de ADN lineal de 100 pb hasta 10 kb de geles de agarosa. En este método se recupera de un 40 a 45% de los fragmentos cortados del gel. Se basa en el uso de una columna que contiene una serie de membranas y filtros, que retienen a la agarosa y las impurezas dejando pasar al ADN durante diez minutos de centrifugación. Una ventaja de este método es que no se requiere fundir la agarosa, ni electroeluir, ni de una digestión enzimática para la purificación del producto (Sigma-Aldrich).

El procedimiento seguido fue el recomendado por el fabricante, descrito a continuación:

### Procedimiento para la purificación de un producto de PCR

Todos los pasos se llevaron a cabo a temperatura ambiente.

- 1.- Correr un gel con las muestras que se desee purificar.
- 2.- Seleccionar la(s) banda(s) de interés y cortarla(s).
- 3.- Cada banda se coloca en una columna que contiene la membrana y debajo se coloca el tubo colector.
- 4.- Se realiza un prelavado con agua o con amortiguador TE (Tris 10 mM, pH 8.0, EDTA 1 mM).
- 5.- Centrifugar por 10 minutos a 13,000 rpm (eliminar el filtrado).
- 6.- Centrifugar ahora la banda de interés por 10 minutos a 13,000 rpm.
- 7.- El eluido es el ADN, al cual se le adiciona 20  $\mu$ L de acetato de sodio 3 M, pH 5.2. Se adiciona 2 mL de etanol absoluto frío y se deja incubando a temperatura ambiente durante dos horas.
- 8.- Centrifugar a 13,000 rpm por 15 minutos.
- 9.- Eliminar el sobrenadante.
- 10.- Lavar con 2 mL de etanol al 70%.
- 11.- Resuspender el pellet de ADN en agua (aproximadamente 30  $\mu$ L).
- 12.- Almacenar a  $-20^{\circ}\text{C}$ .

En la Tabla 7 se presentan los reactivos empleados para la purificación de los productos de PCR.

**Tabla 7. Reactivos empleados en la purificación de los productos de PCR**

REACTIVO	CONDICIONES DE USO	GRADO	DESCRIPCIÓN
Amortiguador TE	T.A.	B.M.	Tris-EDTA, pH 8.0
Acetato de sodio	T.A.	B.M.	3 M, pH 5.2
Etanol absoluto	100%	R.A.	
Etanol al 70%	$-20^{\circ}\text{C}$	R.A.	
Agua	Desionizada	Estéril	

T.A. = Temperatura ambiente

B.M.= Biología Molecular

## **6. Secuenciación**

Con la finalidad de corroborar la identidad del producto de PCR, la secuenciación se llevó a cabo en la Unidad de Biología Molecular en el Instituto de Fisiología Celular, UNAM. Los requisitos para la secuenciación de los productos de PCR fueron 15  $\mu\text{L}$  de ADN de soya los cuales tenían una concentración de 50  $\text{ng}/\mu\text{L}$  de ADN y 20  $\mu\text{L}$  del cebador RR1-F correspondiendo a una concentración de 20  $\text{pmol}/\mu\text{L}$  de ADN.

## **7. PCR tiempo real (RTQ-PCR) para las muestras de soya**

La cuantificación se realizó mediante un kit especialmente diseñado para la cuantificación del transgén Roundup Ready de soya: LightCycler GMO Soya Quantification Kit (ROCHE) en el equipo LightCycler Software Versión 4.0, modelo 2.0 (ROCHE, Switzerland).

Es un método altamente específico que permite cuantificar la cantidad de ácidos nucleicos mediante la PCR. La cuantificación del material genético se realiza durante la reacción y en comparación con la PCR punto final no requiere de análisis posterior a la reacción (electroforesis).

La detección de la reacción se llevó a cabo midiendo la fluorescencia generada en cada ciclo de amplificación y detectada por el sistema óptico del equipo. La cantidad de fluorescencia detectada es directamente proporcional a la cantidad de producto de amplificación generado (Heid et al., 1996).

Los resultados de la RTQ-PCR consisten en curvas de amplificación (Schmittgen et al., 2000). Esta cinética de amplificación se puede dividir en tres fases: la lineal, la exponencial y la estacionaria, siendo en la exponencial donde se lleva a cabo la cuantificación, en donde es necesario definir el punto de intensidad de fluorescencia o umbral de detección (en el cual todas las muestras pueden ser comparadas entre sí). Al ciclo al cual cada muestra consigue llegar a este umbral

de detección, se le llama Cp o Ct (ciclo de umbral de detección; Rodríguez, 2004). Mientras mayor cantidad de ADN de interés se encuentre en una muestra, menor número de ciclos requerirá para alcanzar el umbral de detección.

La cuantificación del contenido de OGMs requiere de una curva estándar tanto para el gen endógeno como para el transgén. Las diluciones para dichas curvas se muestran en la Tabla 8:

**Tabla 8. Curva patrón para gen endógeno y gen de referencia.**

Número de dilución	Dilución	Concentración gen de referencia endógeno (ng / $\mu$ L)	Concentración gen de referencia RR (ng / $\mu$ L)
1	Sin diluir	100	-
2	1:4	25	-
3	1:16	6.25	6.25
4	1:64	1.56	1.56
5	1:256	-	0.39
6	1:1024	-	0.098

El cálculo del porcentaje total de transgénicos se basa en los datos obtenidos en el gráfico de log concentración y Ct's de las curvas estándar. A partir de los datos de Ct's de las muestras, por interpolación se obtienen las concentraciones de las mismas y son éstas las que se sustituyen en la siguiente ecuación:

$$\frac{\text{Concentración de ADN GM}}{\text{Concentración del gen endógeno}} = \% \text{ OGM}$$

Nota: No es necesario multiplicar por 100, ya que el equipo ya está considerando que las concentraciones obtenidas se encuentran en porcentaje (Roche-Applied-Science).

### 7.1 Preparación de la mezcla de reacción para la RTQ-PCR

Se prepara en un tubo de 1.5 mL en hielo, para 20  $\mu$ L de reacción, se mezclan los reactivos que se mencionan en la Tabla 9, en el orden mencionado, después se mezcla pipeteando repetidamente, sin utilizar el vortex.

**Tabla 9. Mezcla de reacción para la detección de OGMs por RTQ-PCR**

Reactivo	Volumen
H <sub>2</sub> O (vial 7)	11 $\mu$ L
Mezcla de detección (vial 1) ó LightCycler GMO Soya Mezcla de detección del gen de referencia (vial 2)	2 $\mu$ L
Solución enzimática Mezcla de reacción (vial 3) Resulta de la mezcla del vial 3 y 4	2 $\mu$ L
Total	15 $\mu$ L

- Nota: Para preparar mezcla de reacción para más de una muestra, se multiplica el volumen de reacción por el número de reacciones a correr más una adicional.

Se colocan 15  $\mu$ L de la mezcla de reacción arriba mencionada en los capilares y 5  $\mu$ L de ADN de cada muestra.

Las condiciones de reacción se muestran en la Tabla 10.

**Tabla 10. Condiciones de reacción para la detección de OGMs por Rt-PCR**

Modo de análisis	Número de ciclos	Segmento	Temperatura	Tiempo	Modo de adquisición
<b>Pre-incubación</b>					
Ninguno	1		95°C	10 min	Ninguno
<b>Amplificación</b>					
Cuantificación	45	Desnaturalización	95°C	10 s	Ninguno
		Alineamiento	60°C	20 s	Único
		Extensión	72°C	72 s	Ninguno
<b>Enfriamiento</b>					
Ninguno	1		40°C	30 s	Ninguno

## **8. Separación de proteínas por electroforesis.**

Se realizaron geles de poliacrilamida formados por dos zonas: una zona de concentración (4.5% acrilamida, pH 6.8) y otra donde se llevó a cabo la separación de las proteínas (pH 8.8). Las tinciones se realizaron con azul de Coomasie, así como con tinción con plata. En este caso, el kit para teñir proteínas con plata fue el PlusOne Silver Staining kit, Protein (GE Healthcare, Amersham Biosciences AB) el cual tiñe proteínas en geles de poliacrilamida. Esta técnica de visualización detecta la mayoría de las proteínas en una concentración de nanogramos y es 100 veces más sensible que la tinción con azul de Coomassie. Su sensibilidad es de 0.2 a 0.6 ng de proteína por banda (GEHealthcare).

La separación de las distintas proteínas bajo condiciones desnaturalizantes se llevó a cabo por el método de SDS-PAGE (electroforesis con dodecil sulfato de sodio). El SDS es un detergente que confiere a las proteínas una carga negativa, en conjunto con la ruptura de los puentes disulfuro existentes por acción del agente reductor  $\beta$ -mercaptoetanol, logrando un desdoblamiento de las proteínas en una densidad igual de carga y una forma similar. El pH del sistema se regula con un amortiguador de TRIS a pH 6.8.

## **9. Isoelectroenfoque preparativo**

Esta técnica se realizó como un paso previo a la purificación, ya que el isoelectroenfoque preparativo se basa en la separación de las proteínas debido a un gradiente dado por las anfolinas (pH de 5 a 7), en el que las proteínas migran por su carga y punto isoeléctrico. Las fracciones proteínicas se obtienen en solución, lo que resulta con la ventaja de que pueden ser utilizadas para otras técnicas. El isoelectroenfoque preparativo se realizó en un Rotofor PREPARATIVE IEF CELL (BIORAD). El equipo consta de una cámara cilíndrica con un dedo de cerámica refrigerante dentro de dicha cámara. Generalmente se colocan 55 mL de muestra por cuatro horas a 12 W de potencia, con una rotación de 1 rpm para estabilizar los disturbios gravitacionales. La membrana divide la cámara en 20 compartimentos, correspondiendo a 20 fracciones con diferente pH cada una.

## **10. Electroelución**

La electroelución permite extraer a la proteína de un gel y tenerla en solución para posteriores experimentos. Se realizaron varios geles preparativos con la finalidad de cortar las bandas correspondientes a las proteínas de entre 40-50 kDa, ya que la proteína de interés oscila entre los pesos moleculares de 45 y 47 kDa, para después electroeluirlos. Los geles se corrieron en cámara BioRad® Mini Protean® II System vertical con vidrios de 8 x 6 cm.

La electroelución se llevó a cabo en equipo Hoeffer Gel Eluter (Hoeffer Scientific Instruments, San Francisco, CA, U.S.A). Se realizaron varias electroeluciones de las bandas cortadas de diferentes geles, se colectaron en el mismo amortiguador de corrida de electroforesis y se concentraron en Microcon®, de corte molecular de 3000 WL, el cual es un método que se basa en el uso de centrifugación para poder concentrar o dializar una muestra, dependiendo del tamaño de la membrana que se utilice como filtro. En este caso se utilizó un corte molecular de 3000 WL.

## **11. Cromatografía de intercambio iónico**

La separación y purificación de proteínas utilizando este método se basa principalmente en las diferencias de carga y la magnitud de las cargas eléctricas netas que presentan los aminoácidos de las proteínas a un pH determinado. La matriz de la columna es un polímero sintético que tiene unidos grupos cargados; cuando éstos son aniónicos se denominan intercambiadores catiónicos y si tiene unidos grupos catiónicos se denominan intercambiadores aniónicos. La afinidad de cada proteína por los grupos cargados de la columna está afectada por el pH y la concentración de iones salinos libres en competencia en la solución que la rodea. La separación puede optimizarse cambiando el pH y/o la concentración de sal de la fase móvil de forma que se cree un gradiente de pH o sal (Nelson, 2000).

Las columnas utilizadas fueron las siguientes:

- **Columna Hitrap DEAE-FF (Fast-Flow).** Columna de carácter aniónico (Dietilaminoetil) capaz de retener la proteína de interés, eluyendo las demás que no lo son. Las condiciones de la corrida fueron:

Un gradiente lineal de 0-100 %. Flujo de 1 mL/min. Se inyectaron 3.6 mg de proteína.

Se utilizaron los siguientes reguladores:

- ❖ Amortiguador de equilibrio: Tris-HCl 50 mM, pH 7.1
- ❖ Amortiguador de elución: Tris-HCl 50 mM, NaCl 100 mM, pH 7.1

- **Columna Hitrap Q-FF (Fast-Flow).** Columna de carácter aniónico (es una sal de amonio cuaternario) que retiene la proteína de interés. Las condiciones de la corrida fueron:

Un gradiente lineal de 0-100 %. Flujo de 1 mL/min. Se inyectaron 198 µg de proteína.

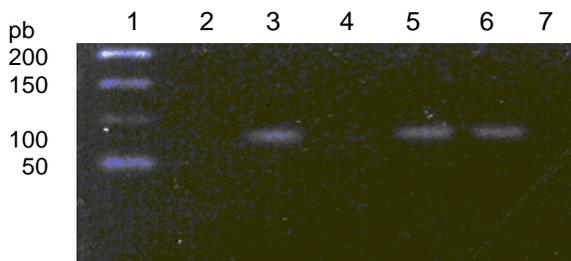
Se utilizaron los siguientes reguladores:

- ❖ Amortiguador de equilibrio: Tris-HCl 50 mM, pH 7.5
- ❖ Amortiguador de elución: Tris-HCl 50 mM, NaCl 1.0M, pH 7.5

## VI. RESULTADOS Y DISCUSIÓN

### 1. Verificación de muestras transgénicas por PCR punto final

Las muestras de soya utilizadas para el desarrollo de esta tesis se sometieron primeramente a PCR punto final para determinar cuáles eran las GM, por medio de la detección de un fragmento que amplifica parte del promotor y parte del CTP característico de la soya Roundup® Ready. Éste fue detectado utilizando los cebadores RR1-F y RR1-R, los cuales amplifican un fragmento de 74 pb (ICGEB). Los fragmentos de ADN fueron separados por medio de una electroforesis en gel de agarosa y en la Figura 7 se presentan visualizados en un transiluminador de UV, los resultados de la detección donde puede observarse una banda que coincide con el tamaño esperado de 74 pb correspondiente al transgén. Ésta aparece en las muestras STO y SFL lo que indica que contienen granos genéticamente modificados. Cabe recordar que este evento de soya RR está aprobado por la Secretaría de Salud para consumo humano y animal. Así mismo, existen diferentes estados de la República Mexicana que cuentan con permisos para la siembra de esta soya GM.



**Figura 7. Productos de la amplificación por PCR del transgén RR a partir de diferentes muestras de harinas de soya.**

Carril 1: Marcador Molecular XIII

Carril 2: Harina de soya SNA

Carril 3: Harina de soya STO

Carril 4: Harina de soya STA

Carril 5: Harina de soya SFL

Carril 6: Control positivo

Carril 7: Control negativo (reactivos sin ADN molde)

Posteriormente, se realizó PCR a las muestras provenientes del norte del país, utilizando las mismas condiciones que en el gel anterior. En la Figura 8 se muestra ausencia de productos de amplificación del transgén, excepto en el control positivo, carril 6.



**Figura 8. Productos de la amplificación por PCR del transgén RR a partir de diferentes muestras de harina de soya.**

Carril 1: Marcador molecular XIII

Carril 2: Harina de soya SE

Carril 3: Harina de soya SH

Carril 4: Harina de soya SC1

Carril 5: Harina de soya SC2

Carril 6: Control positivo

Carril 7: Control negativo (reactivos sin ADN molde)

Lo anterior indica que ninguna de las muestras contienen secuencias que permitan el apareamiento con estos cebadores, por lo tanto, no se trata de productos de soya RR. También, demuestra que las muestras de soya cosechadas en el 2006, provenientes del norte del país, no contienen secuencias transgénicas.

La PCR mostró ser un método rápido, eficiente y sensible para propósitos de control y discriminación entre muestras GM y No-GM. Los cebadores utilizados en este estudio muestran una alta especificidad, facilitando así, la confirmación por secuenciación.

## 2. Cuantificación de ADN por PCR tiempo real (RTQ-PCR).

A diferencia de la PCR punto final que evidencia la presencia o ausencia de una secuencia nucleotídica específica, la PCR en tiempo real es un método que permite cuantificar el número de copias iniciales de dicha secuencia (Rodríguez, 2004). La cuantificación se lleva a cabo mediante la medición de la fluorescencia generada en cada ciclo de amplificación (Heid et al., 1996) ya que la cantidad de ésta es proporcional a la cantidad generada del producto de amplificación.

La cuantificación por PCR en tiempo real tanto del gen endógeno como del transgén se hizo con el equipo LightCycler, Software Versión 4.0 que cuenta con un carrusel con 32 aberturas, donde se pueden colocar igual número de capilares. Cada muestra se cuantificó por duplicado, con el siguiente arreglo en el equipo (Tabla 11):

**Tabla 11. Organización de las muestras en la corrida de la RTQ-PCR**

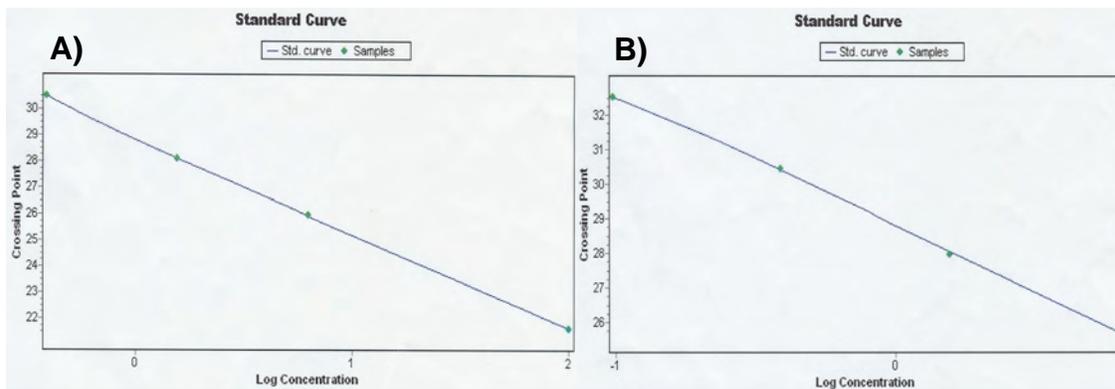
No. de pozo	Clave	Muestra
1	GMO 6.25	Curva estándar del transgén
2	GMO 1.56	
3	GMO 0.39	
4	GMO 0.098	
5	GMO 6.25	Calibrador
6	M1	STA
7	M1	STA
8	M2	STO
9	M2	STO
10	M3	SFL
11	M3	SFL
12	M4	SNA
13	M4	SNA
14	M5	SH
15	M5	SH

No. de pozo	Clave	Muestra
16	Lec 100	Curva estándar del gen endógeno
17	Lec 25	
18	Lec 6.25	
19	Lec 1.56	
20	Lec 6.25	Calibrador
21	M1	STA
22	M1	STA
23	M2	STO
24	M2	STO
25	M3	SFL
26	M3	SFL
27	M4	SNA
28	M4	SNA
29	M5	SH
30	M5	SH
31	NTC transgén	Controles negativos
32	NTC Lec	

NTC: Control negativo de reactivos de PCR sin ADN molde (siglas en inglés de No Template Control)

Los resultados de la PCR en tiempo real arrojan el valor del ciclo en el que la fluorescencia supera el umbral de detección de dicha fluorescencia. A este ciclo se le conoce como Ct y, en términos prácticos, significa cuántos ciclos le tomó a cada muestra generar la cantidad suficiente de fluorescencia para llegar y superar el umbral de detección. El valor de Ct es directamente proporcional a la cantidad inicial de molde y es fundamento para calcular la cantidad de ADN en este caso. Mientras mayor cantidad de ADN se encuentre en una muestra, menor número de ciclos requerirá para alcanzar el umbral de detección.

En todo ensayo de cuantificación es necesario evaluar el rango dinámico lineal en el cual pueden estar las muestras y determinar la eficiencia de amplificación y sensibilidad de la técnica, para cada uno de los genes que se evaluarán. Para ello, se utiliza una curva estándar, a partir de una serie de diluciones de ADN por duplicado. Las curvas estándar obtenidas en esta tesis se presentan en la Figura 9. A partir de los resultados de Ct se construye una gráfica: Ct vs. log de la concentración.



**Figura 9. Curva estándar del A) gen endógeno y B) transgén** realizada con el Software de cuantificación del LightCycler para OGM.

Entre los parámetros principales para considerar optimizado un ensayo de PCR en tiempo real están que la linealidad de la curva estándar tenga un valor de  $R^2 > 0.980$  ó  $r > 0.990$  y una eficiencia de amplificación del 90–110 % ó que el valor de la pendiente se encuentre en -3.1 y -3.6 (European Network, 2005). De acuerdo a los resultados presentados en la Tabla 12, tanto para el gen endógeno como para la cuantificación del ADN modificado, los parámetros del ensayo de PCR en tiempo real estuvieron dentro de los límites aceptables; por lo tanto, se consideraron válidos los datos de la cuantificación de cada gen de las muestras individuales y se procedió a analizarlos.

**Tabla 12. Valores de la curva de amplificación para la cuantificación del gen endógeno y del transgén**

	Gen endógeno	Transgén
Eficiencia de la reacción (%)	97.34	98.56
Linearidad de la curva estándar (r)	0.9865	0.9832
Rango de Ct's de la curva de mayor a menor concentración	20 – 31	25 – 32

### 2.1 Cuantificación absoluta del gen endógeno y del transgén

Una vez establecidos y verificados los parámetros de la PCR en tiempo real, se hizo la cuantificación absoluta del gen endógeno y del transgén para después determinar el porcentaje de material genéticamente modificado de las muestras. Los resultados se presentan en las Tablas 13 y 14, respectivamente.

**Tabla 13. Valores de Ct de la curva estándar y las muestras analizadas, para la cuantificación del gen endógeno**

Muestra	Valor de Ct
Lec 100	21.6
Lec 25	23.83
Lec 6.25	26.12
Lec 1.56	28.12
Lec 0.39	30.5
STA	32.15
STO	27.65
SFL	26.53
SNA	27.53
SH	26.48
SE	31.88
SC1	25.09
SC2	25.68

En este caso, se observa que las muestras STA y SE no cayeron dentro de la curva estándar del gen endógeno, por lo que, no se utilizarán para cuantificar la concentración de ADN modificado genéticamente.

**Tabla 14. Valores de Ct de la curva estándar y las muestras analizadas, para la cuantificación del transgén**

Muestra	Valor de Ct
GMO 6.25	25.75
GMO 1.56	27.97
GMO 0.39	30.44
GMO 0.098	31.51
STA	37.49
STO	28.17
SFL	26.55
SNA	ND
SH	39.24
SE	37.5
SC1	ND
SC2	ND

ND: muestra en la que no se detectó la secuencia del transgén.

En la Tabla 14 se observa que el ADN molde de la curva estándar comienza a amplificar a partir del ciclo 25 y las muestras a partir del ciclo 26.5. La primera muestra en amplificar fue SFL, seguida por la muestra STO. Las otras muestras se vuelven detectables, en ciclos posteriores al punto más diluido de la curva estándar (ciclo 31), lo que significa que no contienen secuencias modificadas genéticamente que puedan ser amplificadas. Estas muestras se consideran no transgénicas y la señal de fluorescencia que emiten puede deberse a que los cebadores se hayan apareado inespecíficamente.

## 2.2 Cuantificación del material transgénico en las muestras analizadas

Con los datos de la cuantificación absoluta del gen endógeno y del transgén se obtuvo el contenido de material transgénico (en porcentaje) en las muestras analizadas. Los resultados se presentan en la Tabla 15.

**Tabla 15. Resultados de la cuantificación por PCR tiempo real del material transgénico en muestras de soya**

Muestra	Ct transgén	[transgén] ng / $\mu$ L	Ct endógeno	[endógeno]	% OGM
STA	37.49	0.0049219	32.15	0.394076	LOQ
STO	28.17	1.434055	27.65	0.383694	3.74
SFL	26.55	3.845315	26.53	0.3811528	10.08
SNA	LOQ	ND	27.53	0.3834209	ND
SH	39.24	0.0016958	26.48	0.381039	LOQ
SE	37.5	0.004892	31.88	0.393445	LOQ
SC1	LOQ	ND	25.09	0.37791	ND
SC2	LOQ	ND	25.68	0.3792354	ND

ND: No detectado

LOQ: Por debajo del límite de cuantificación

En la Tabla 15 se observa que las muestras de soya de Tlatelolco (STO) y Flor de Liz (SFL) son las que presentan el mayor porcentaje de material transgénico. Estos valores incluso se hayan por arriba del límite establecido internacionalmente (0.9%). En las muestras STA, SH y SE, el Ct del transgén estuvo por arriba del estándar más diluido de la curva estándar, lo que significa que está por debajo del límite de cuantificación (Rott, 2004), por lo que se clasificaron como no transgénicas. Las muestras SNA, SC1 y SC2, fueron negativas a la presencia del transgén y también se clasificaron como no transgénicas.

---

### 3. Secuenciación del producto de PCR.

La secuenciación se llevó a cabo en el laboratorio de Biología Molecular en el Instituto de Fisiología de la UNAM a cargo de la Dra. Laura Ongay. Para ello, se eligió la muestra STO ya que fue una de las que produjo una banda definida y abundante. Primero se purificó el producto de PCR con columna (GenElute Agarose Spin Column). Esto se realizó con la finalidad de corroborar la identidad de los productos de la PCR (fragmentos de 74 pb). Se obtuvo la secuencia:

**cgtaggctctagcagatctttcaagaatggcacaattaacaacatgntcaagnnnn**

#### 3.1 Alineamiento de la secuencia.

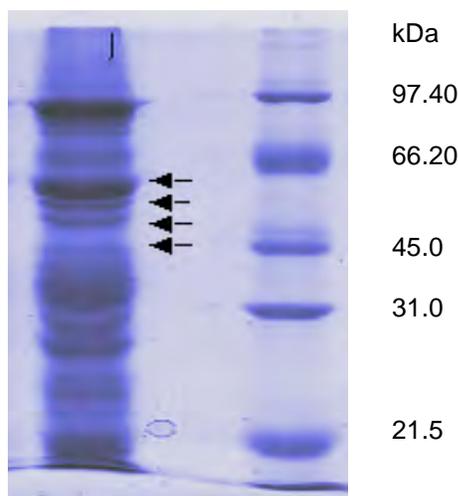
Para conocer las diversas secuencias que presentan alguna semejanza con el producto secuenciado se realizó un alineado con el programa Clustal W1.8 en la base de datos de GeneBank (NCBI), con el número de acceso ID:1158863884-26254-122923331555.BLASTQ4. Las secuencias con las que se encontró la más alta semejanza se presentan en la Figura 10. En ésta se observa que la secuencia de la muestra STO tuvo alta semejanza con tres secuencias depositadas previamente en el GenBank correspondientes a las secuencias codificantes de la proteína heteróloga de la soya RR y con la secuencia de *Petunia hybrida* correspondiente al CTP4, péptido de tránsito, presente en el transgén. Esto corroboró que la secuencia de la muestra STO que se amplificó, corresponde a la zona entre el promotor 35S y el péptido de tránsito del plásmido empleado en la transformación de la soya.



La segunda parte de esta tesis se enfocó en la detección y caracterización de la(s) proteína(s) heterólogas en la muestra STO ya que de las analizadas fue una de las que tuvo mayor porcentaje de ADN modificado genéticamente. Por ello, es de esperarse que sea una de las muestras con mayor contenido neto de la(s) proteína(s) heterólogas que se busca caracterizar.

#### 4. Estudios Preliminares para la purificación de la CP4EPSPS

En estudios preliminares se realizó una electroforesis en gel de poliacrilamida en condiciones desnaturalizantes del extracto proteínico de la muestra STO (Figura 11). De éste se cortaron las bandas de las fracciones de 45 a 60 kDa (indicadas por las flechas), nombrando Banda 1 a la de más arriba, Banda 2 a la inmediata inferior y así, sucesivamente, hasta la Banda 5. Estas fueron analizadas por ELISA para la inmunodetección de la proteína CP4EPSPS.



**Figura 11.** Electroforesis en gel de poliacrilamida en condiciones desnaturalizantes del extracto proteínico de la muestra STO, teñido con Azul de Coomassie. Las bandas señaladas con flechas, aledañas a la fracción de 46 kDa (correspondiente a la CP4-EPSPS) fueron cortadas y sometidas a ELISA.

Cabe mencionar que a la muestra SC1, previamente identificada por PCR como no transgénica, se le extrajo la proteína, diluyéndola y concentrándola con la

finalidad de verificar por ELISA si daba alguna lectura, pero no dió ninguna. Por lo tanto, fue incluida como control negativo para todas las pruebas de ELISA.

Los resultados de este estudio preliminar se muestran en la Tabla 16, en donde se observa que las lecturas no fueron tan altas, quizá porque la proteína de interés se encontraba en baja concentración o, posiblemente, porque al ser extraída del gel en condiciones desnaturalizantes, el epítotope no quedó expuesto a los anticuerpos correspondientes. Sin embargo, dado que fueron positivas, es un indicativo de que se tratan de isoformas de la CP4EPSPS, que conservan alta semejanza y probablemente patrones de plegamiento entre ellas, ya que son reconocidas por el anticuerpo diseñado específicamente contra la proteína heteróloga.

**Tabla 16.** Lecturas de la placa de ELISA para las bandas de la muestra STO y la muestra SC1

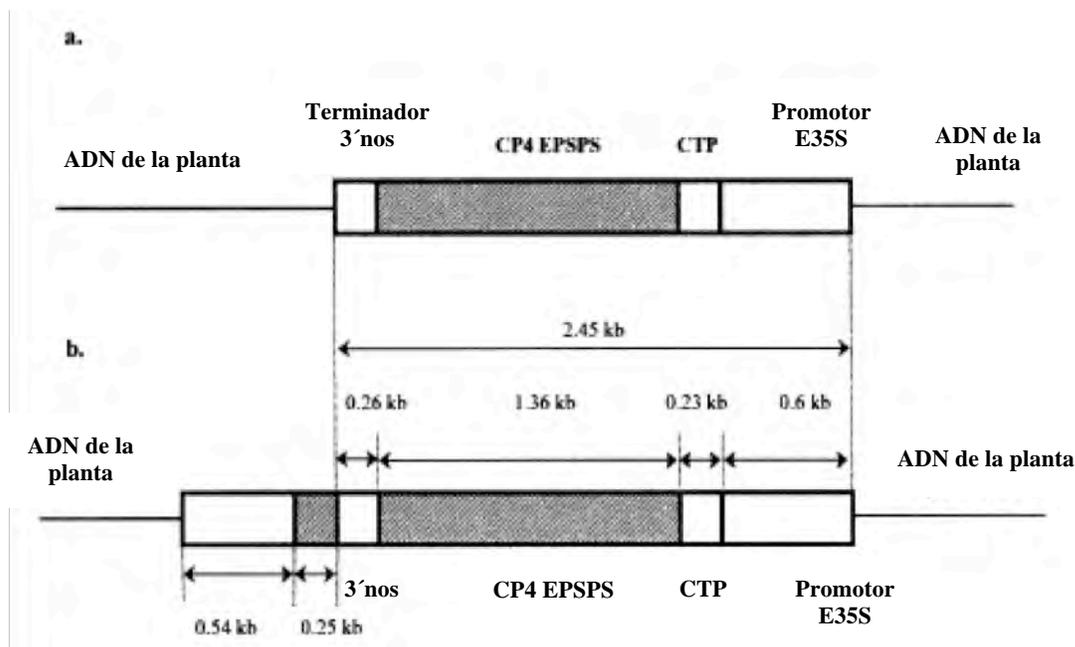
Muestra	Lectura Abs = 630 nm
PBS	0.042
Banda 1	0.045
Banda 2	0.112
Banda 3	0.098
Banda 4	0.101
Banda 5	0.094
Control +	1.134
Control +	1.452
SC1 (200 µg)	0.065
SC1 (1 mg)	0.064
SC1 (2 mg)	0.06
SC1 (2 µg)	0.55
SC1 (100 µg)	0.053

Estas isoformas pueden variar un poco en pesos moleculares y no tener las mismas actividades enzimáticas ya que no se expresan en cantidades iguales (Gottlieb, 1982).

La presencia de isoformas es posible ya que estudios previos hechos en sorgo y maíz (Ream, 1988 y Forlani, 1994) también han encontrado isoformas de la proteína EPSPS, por lo que es muy posible que lo mismo ocurra en la soya. De hecho, previamente, en un trabajo de Asia-Pacific Economic Cooperation (APEC, 2001), se evidenciaron por Western blot la proteína CP4EPSPS de 47 kDa y otras proteínas de menor peso molecular que inmunoreaccionaron igual que la heteróloga, a partir de proteínas extraídas de soya evento GTS-40-3-2.

Por otro lado, Windels y col. (2001), publicaron que ambos extremos de la construcción tenían secuencias no reportadas por la compañía Monsanto. Es decir, encontraron un fragmento de 181 pb en la región corriente arriba del promotor 35S.

Aparte de esta copia no funcional, en la región 3', después del terminador NOS se encontró un fragmento adicional de 254 pb de la secuencia de la CP4EPSPS y un segundo fragmento de 540 pb, el cual no corresponde a alguna secuencia conocida. En la Figura 12 se muestra el fragmento repetido de la CP4EPSPS de 0.25kb y el de 0.54kb que no corresponde a ninguna secuencia conocida.



**Figura 12.** Comparación de a) la estructura del inserto reportado por Padgett y col. (1995) y b) la estructura del inserto deducida por Windels (2001).

La caracterización y el aislamiento de la secuencia de la unión entre el ADN de la planta y el gen insertado permitió concluir que durante la integración del gen insertado ocurrieron algunos rearrreglos en el sitio de unión entre el terminador 3'NOS y el ADN genómico de la planta. Esta caracterización se realizó con el fin de establecer un método para la detección e identificación de productos derivados de la biotecnología. Por otro lado, Rang y col. (2005), encontraron cuatro diferentes ARN mensajeros (ARNm) obtenidos a partir de hojas de soya RR, derivados de la falta de reconocimiento por parte de la polimerasa a la señal del terminador. Como resultado, también se transcribió la secuencia reportada por Windels (2001) situada después de dicho terminador. Rang, trabajó con ADN complementario (ADNc), obtenido a partir de ARN de hojas de soya RR. El análisis de las secuencias se realizó por medio de PCR, utilizando cebadores específicos para el gen insertado, obteniéndose cuatro productos de diferente peso molecular. La secuenciación de cada una de estas mostró la existencia de cuatro moléculas de cDNA con diferentes secuencias. Las diferencias en las

secuencias se debieron a la ablación de fragmentos de ADN y no a rearrreglos del mismo, dando pie a la posible existencia de proteínas de fusión (Windels, 2001).

**Como parte de esta tesis se tradujeron, *in silico*, la secuencia a aminoácidos. En éstas se observó que el amino terminal de cada fragmento fue igual a la CP4EPSPS y que sólo existieron diferencias en la parte de la secuencia donde se une el fragmento repetido. De esta forma, encontraron que los cuatro ARNm descritos en dicho trabajo pueden codificar proteínas de fusión distintas la CP4EPSPS.**

**Con la secuencia de aminoácidos propuesta por Rang y col. (2005) y traducidas *in silico* (anexo II) se realizó su alinamiento y se les calculó el punto isoeléctrico (Expasy) y el peso molecular de cada proteína teórica, lo que se presenta a continuación:**

```

1
130
CP4EPSPS 3 MLHGASSRPA TARKSSGLSG TVRIPGDKSI SHRSFMFGGL ASGETRITGL LEGEDVINTG KAMQAMGARI RKEGDTWIID GVGNGGLLAP EAPLDFGNAA
TCRLTMGLV GYVDFDSTFI GDASLTKRPM
CP4EPSPS 5 MLHGASSRPA TARKSSGLSG TVRIPGDKSI SHRSFMFGGL ASGETRITGL LEGEDVINTG KAMQAMGARI RKEGDTWIID GVGNGGLLAP EAPLDFGNAA
TCRLTMGLV GYVDFDSTFI GDASLTKRPM
CP4EPSPS 6 MLHGASSRPA TARKSSGLSG TVRIPGDKSI SHRSFMFGGL ASGETRITGL LEGEDVINTG KAMQAMGARI RKEGDTWIID GVGNGGLLAP EAPLDFGNAA
TCRLTMGLV GYVDFDSTFI GDASLTKRPM
CP4EPSPS 4 MLHGASSRPA TARKSSGLSG TVRIPGDKSI SHRSFMFGGL ASGETRITGL LEGEDVINTG KAMQAMGARI RKEGDTWIID GVGNGGLLAP EAPLDFGNAA
TCRLTMGLV GYVDFDSTFI GDASLTKRPM
Consensus MLHGASSRPA TARKSSGLSG TVRIPGDKSI SHRSFMFGGL ASGETRITGL LEGEDVINTG KAMQAMGARI RKEGDTWIID GVGNGGLLAP EAPLDFGNAA
TCRLTMGLV GYVDFDSTFI GDASLTKRPM

260
201
CP4EPSPS 3 GRVLNPLREM GVQVKSEDDG RLPVTLRGP TPTPITYRVP MASAQVKS AV LLAGLNTPGI TTVIEPIMTR DHTEKMLQGF GANLTVETDA DGVRTIRLEG
RGKLTGQVID VPGDPSSTAF PLVAALLVPG
CP4EPSPS 5 GRVLNPLREM GVQVKSEDDG RLPVTLRGP TPTPITYRVP MASAQVKS AV LLAGLNTPGI TTVIEPIMTR DHTEKMLQGF GANLTVETDA DGVRTIRLEG
RGKLTGQVID VPGDPSSTAF PLVAALLVPG
CP4EPSPS 6 GRVLNPLREM GVQVKSEDDG RLPVTLRGP TPTPITYRVP MASAQVKS AV LLAGLNTPGI TTVIEPIMTR DHTEKMLQGF GANLTVETDA DGVRTIRLEG
RGKLTGQVID VPGDPSSTAF PLVAALLVPG
CP4EPSPS 4 GRVLNPLREM GVQVKSEDDG RLPVTLRGP TPTPITYRVP MASAQVKS AV LLAGLNTPGI TTVIEPIMTR DHTEKMLQGF GANLTVETDA DGVRTIRLEG
RGKLTGQVID VPGDPSSTAF PLVAALLVPG
Consensus GRVLNPLREM GVQVKSEDDG RLPVTLRGP TPTPITYRVP MASAQVKS AV LLAGLNTPGI TTVIEPIMTR DHTEKMLQGF GANLTVETDA DGVRTIRLEG
RGKLTGQVID VPGDPSSTAF PLVAALLVPG

261
390
CP4EPSPS 3 SDVTILNVLM NPTRTGLLIT LQEMGADIEV INPRLAGGED VADLRVRSST LKGVTVPEDR APSMIDEYPI LAVAAAFAG ATVMNGLEEL RVKESDRLSA
VANGLKLNKV DCDEGETSLV VRGRPDGKGL
CP4EPSPS 5 SDVTILNVLM NPTRTGLLIT LQEMGADIEV INPRLAGGED VADLRVRSST LKGVTVPEDR APSMIDEYPI LAVAAAFAG ATVMNGLEEL RVKESDRLSA
VANGLKLNKV DCDEGETSLV VRGRPDGKGL
CP4EPSPS 6 SDVTILNVLM NPTRTGLLIT LQEMGADIEV INPRLAGGED VADLRVRSST LKGVTVPEDR APSMIDEYPI LAVAAAFAG ATVMNGLEEL RVKESDRLSA
VANGLKLNKV DCDEGETSLV VRGRPDGKGL
CP4EPSPS 4 SDVTILNVLM NPTRTGLLIT LQEMGADIEV INPRLAGGED VADLRVRSST LKGVTVPEDR APSMIDEYPI LAVAAAFAG ATVMNGLEEL RVKESDRLSA
VANGLKLNKV DCDEGETSLV VRGRPDGKGL
Consensus SDVTILNVLM NPTRTGLLIT LQEMGADIEV INPRLAGGED VADLRVRSST LKGVTVPEDR APSMIDEYPI LAVAAAFAG ATVMNGLEEL RVKESDRLSA
VANGLKLNKV DCDEGETSLV VRGRPDGKGL

391
520
CP4EPSPS 3 GNASGAAVAT HLDHRIAMSF LVMGLVSENP VTVDDATMIA TSFPEFMDL FGLASGETR ITGLEGEDV INTGKAMQAM GARIRKEGDT WIIDGVGNNG
LLAPEAPLDF GNAATGCRILT MGLVGVYDFK
CP4EPSPS 5 GNASGAAVAT HLDHRIAMSF LVMGLVSENP VTVDDATMIA TSFPEFMDL FGLASGETR ITGLEGEDV INTGKAMQAM GARIRKEGDT WIIDGVGNNG
LLAPEAPLDF GNAATGCRILT MGLVGVYDFK
CP4EPSPS 6 GNASGAAVAT HLDHRIAMSF LVMGLVSENP VTVDDATMIA TSFPEFMDL ----- --LEGEDV INTGKAMQAM GARIRKEGDT WIIDGVGNNG
LLAPEAPLDF GNAATGCRILT MGLVGVYDFK
CP4EPSPS 4 GNASGAAVAT HLDHRIAMSF LVMGLVSENP VTVDDATMIA TSFPEFMDL ----- MAGLGAKELE SDTGKAMQAM GARIRKEGDT WIIDGVGNNG
LLAPEAPLDF GNAATGCRILT MGLVGVYDFK
Consensus GNASGAAVAT HLDHRIAMSF LVMGLVSENP VTVDDATMIA TSFPEFMDL. .... ..GLLEGE#V I#TGKAMQAM GARIRKEGDT WIIDGVGNNG
LLAPEAPLDF GNAATGCRILT MGLVGVYDFK

521 543
CP4EPSPS 3 RIMLGNFSEI ISIFLGISAV TGE
CP4EPSPS 5 RIMLGNFSEI ISIFLGISAV TGE
CP4EPSPS 6 RIMLGNFSEI ISIFLGISAV TGE
CP4EPSPS 4 RIMLGNFSEI ISIFLGISAV TGE
Consensus RIMLGNFSEI ISIFLGISAV TGE

```

**Tabla 17. Resultados de los cálculos obtenidos de las proteínas teóricas**

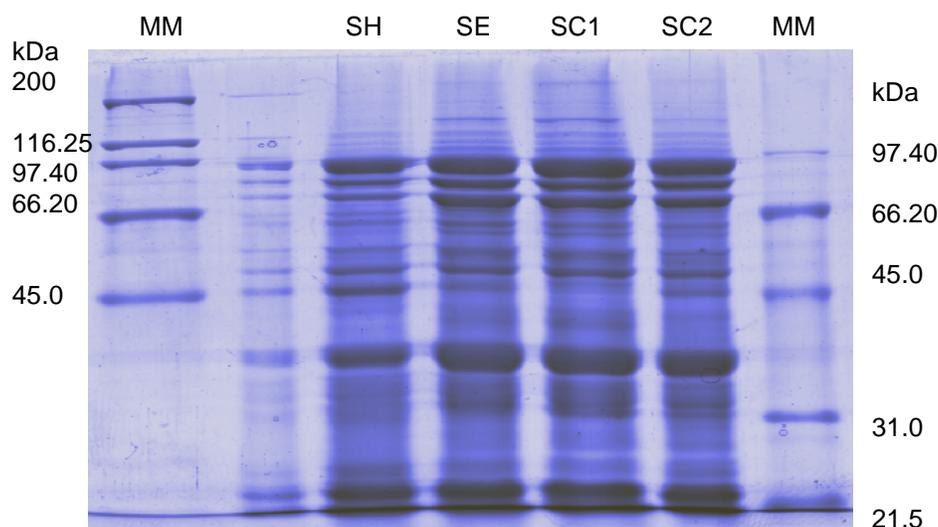
<b>Fragmento</b>	<b>Punto Isoeléctrico</b>	<b>Peso molecular (Da)</b>
CP4 EPSPS 3	5.05	56503.76
CP4 EPSPS 4	5.09	55687.94
CP4 EPSPS 5	5.01	56863.20
CP4 EPSPS 6	5.00	55371.48

De acuerdo a los resultados previos de este estudio y a los publicados, cabe la posibilidad de la existencia de más de una proteína CP4 EPSPS. Sin embargo, aun no se sabe si son el resultado de isoformas de la misma o de proteínas de fusión, como lo reporta Rang y col. (2005). Faltan más estudios para comprobar lo anterior puesto que los pesos moleculares teóricos obtenidos en este estudio (Tabla 17) son mayores a 47 kDa y no corresponden a los reportados por APEC mencionados en los antecedentes.

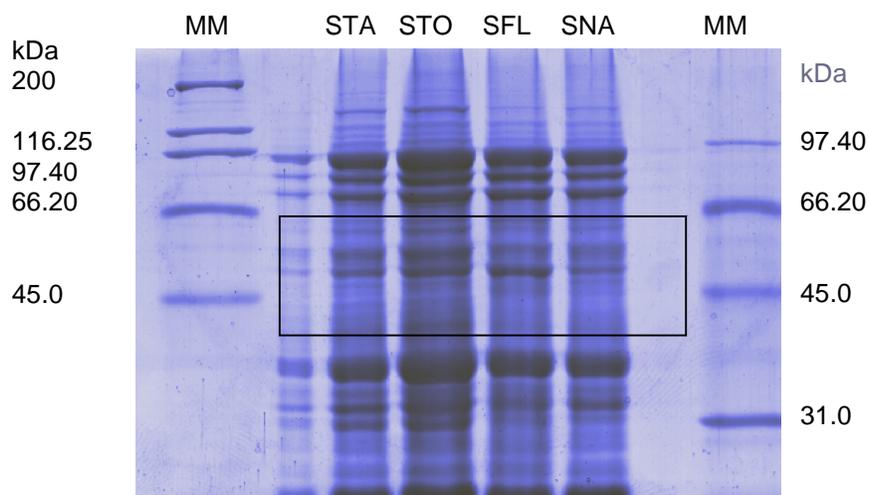
Es importante mencionar que los pesos moleculares obtenidos mediante densitometría de las bandas del gel de electroforesis mostrado en la Figura 11, fueron parecidos a los calculados teóricamente, corroborando así que existen al menos cuatro isoformas de esta proteína, que requerirán ser caracterizadas.

## 5. Comparación del perfil electroforético del extracto proteínico en geles de poliacrilamida.

Los extractos proteínicos de las muestras de soya GM y No-GM fueron analizados mediante electroforesis en geles de poliacrilamida en condiciones desnaturalizantes para comparar dichas muestras y observar si habría alguna diferencia entre éstas. En la Figura 13 se presenta el gel a partir de los extractos proteínicos de las muestras del norte del país, identificadas como No-GM.



**Figura 13.** Gel de electroforesis en condiciones desnaturalizantes al 10%, teñido con azul de Coomassie. Se muestra el patrón de bandeo de los extractos proteínicos de las muestras de soya provenientes del norte del país (SH, SE, SC1 y SC2). En los carriles de los extremos se localizan los marcadores de peso molecular (Bio-Rad). MM. Marcador Molecular



**Figura 14.** Gel de electroforesis en condiciones desnaturalizantes al 10%, teñido con azul de Coomasie, de los extractos proteínicos STA, STO, SFL y SNA. En los carriles de los extremos se localizan los marcadores de peso molecular (MM).

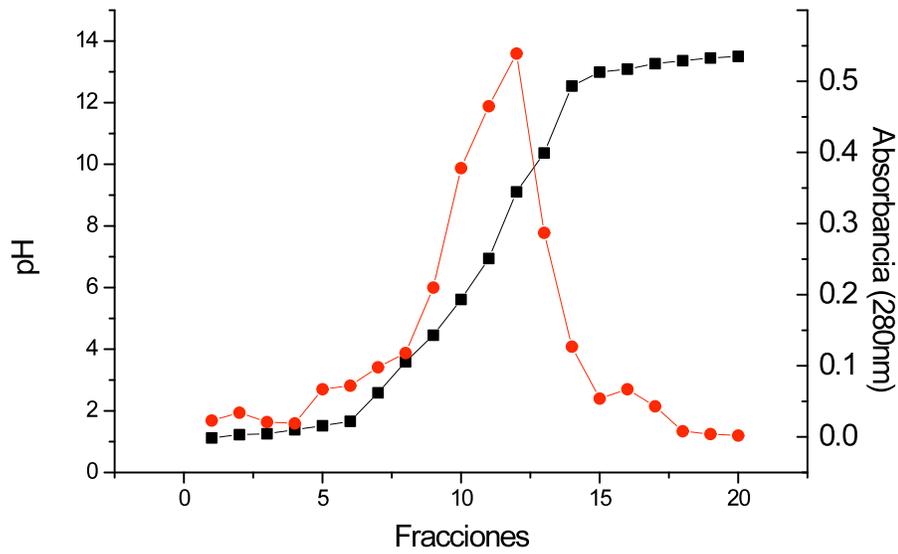
En la Figura 14 se muestra el perfil proteico de las muestras de la región central del país, consideradas GM, dos de las cuales contienen un porcentaje considerable de granos transgénicos (STO: 3.74 % y SFL:10.08 %) según la cuantificación por PCR en tiempo real. No se encontraron diferencias entre las muestras de granos transgénicos y las de no transgénicos. Sin embargo, sí las hubo entre los geles (correspondientes a diferentes zonas del país) y entre las mismas muestras que se presentan en cada gel, una mayor cantidad de proteínas en la región correspondiente a los 45 y 66.20 kDa. No se sabe si estas diferencias pudieran deberse a la expresión de la proteína transgénica respecto a las demás, es decir, que pueda incrementar o disminuir la concentración de las proteínas endógenas de la soya. También, podrían atribuirse dichas diferencias a la variedad de la soya, la zona y condiciones de cultivo, así como al estadio de desarrollo de la planta.

## 6. Isoelectroenfoque preparativo

**Esta técnica se basa en la separación de las proteínas según un gradiente de pH generado por las anfolinas, que provoca que las proteínas migren de acuerdo a su punto isoeléctrico, con la ventaja de que las fracciones proteínicas obtenidas se obtienen en solución, lo que facilita su procesamiento posterior por otras técnicas. Otra de las ventajas de este método es su capacidad de generar muestras homogéneas y que se puede trabajar con pequeñas cantidades. Sin embargo, tiene la desventaja de que algunas proteínas pueden precipitarse (insolubilizarse) por encontrarse en su punto isoeléctrico (Smith, 1980).**

A cada fracción obtenida se le midió el pH y la absorbancia a 280 nm para estimar la cantidad de proteína contenida en cada una de ellas y, posteriormente, determinar en cuál(es) fracción(es) se encuentra la proteína de interés, para continuar trabajando con ellas. Este experimento se realizó con 50 mL de extracto de soya STO y 1.2 mL de anfolinas (se utilizó una mezcla de 0.6 mL de anfolinas de 3.5 – 5 y 0.6 mL de anfolinas de 6 – 8), para generar el gradiente de pH.

En la Figura 15 se observa el gráfico del isoelectroenfoque, en donde se muestra el gradiente de pH formado (línea con cuadros) y las absorbancias a 280 nm (línea con círculos). De las 20 fracciones obtenidas, dos de ellas fueron mayoritarias, la fracción 11 y la 12, que presentaron una lectura de absorbancia de 0.465 y 0.539 respectivamente.



**Figura 15. Gráfico de las mediciones de pH y absorbancia del isoelectroenfoque preparativo.** La línea con cuadros muestra el gradiente de pH formado y la línea con círculos las absorbancias a 280 nm.

Cada una de las 20 fracciones colectadas fue sometida a la prueba de ELISA para identificar en qué fracción(es) está(n) presente(s) las proteínas capaces de reaccionar con los anticuerpos anti-CP4EPS de el kit antes mencionado.

### 6.1 ELISA

Este análisis permitió elucidar cuáles de las fracciones obtenidas por el isoelectroenfoque preparativo contenían a la proteína de interés. Los resultados se presentan en la Tabla 18. Enseguida se describe el acomodo de las fracciones y los controles. Se utilizaron 24 pozos para las 20 fracciones, dos pozos para el control negativo, el cual es el amortiguador PBS, y dos pozos más para el control positivo, que en este caso es extracto proteínico de la muestra STO.

**Tabla 18. Valores de lectura de ELISA de las 20 fracciones obtenidas por isoelectroenfoque**

<b>Fracción</b>	<b>Absorbancia (630nm)</b>
PBS	0.044
PBS	0.045
1	0.051
2	0.042
3	0.049
4	0.046
5	0.044
6	0.050
7	0.987
8	1.234
9	1.438
10	1.347
11	1.198
12	1.374
13	0.996
14	0.198
15	0.052
16	0.057
17	0.043
18	0.044
19	0.044
20	0.048
Ctrl +	1.567
Ctrl +	1.601

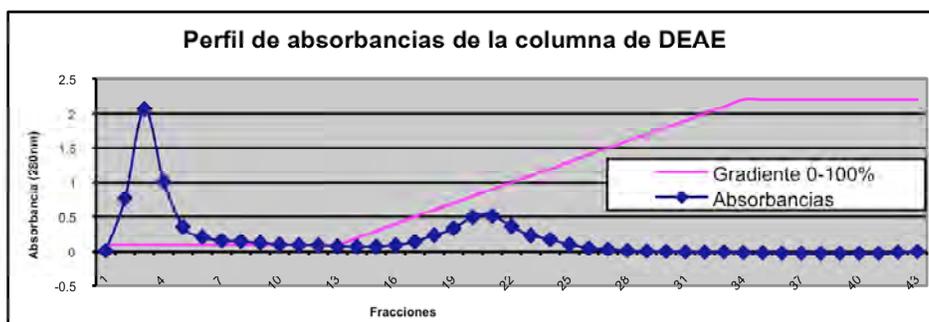
Las fracciones 7 a 14 son las que presentaron una lectura mayor a 560 nm, es decir que en éstas se encuentra la proteína de interés capaz de reaccionar inmunológicamente con los anticuerpos anti-CP4EPSPS y formar el compuesto colorido. La fracción nueve fue la que tuvo mayor cantidad de proteína CP4EPSPS, teniendo valores cercanos al control positivo de la reacción.

## **7. Cromatografía de intercambio iónico**

Con base en el punto isoeléctrico teórico de 5.18 de la proteína heteróloga buscada, obtenido de la base de datos expasy, y después de varios ensayos preliminares en los que se combinaron tipos de columnas y condiciones, se eligió la combinación de dos tipos de columnas de intercambio aniónico: DEAE-FF y Q-FF. La muestra analizada fue el extracto proteínico de la fracción nueve obtenida del isoelectroenfoque preparativo obteniéndose los siguientes resultados.

### **7.1 Columna de intercambio aniónico DEAE-FF**

Primero se trabajó con la columna de DEAE-FF, con las condiciones mencionadas en materiales y métodos. Después de la corrida, se obtuvieron 43 fracciones, a las cuales se les midió su absorbancia a 280 nm. Las últimas fracciones (28-43) fueron descartadas por su baja o nula concentración de proteína. En las fracciones obtenidas se observan dos picos donde cabía la posibilidad que se encontrara la proteína de interés, ya que según las lecturas de absorbancias a 280nm, es donde había mayor concentración de proteínas (Figura 16). El primer pico corresponden a las fracciones que no se retienen en la columna, mientras que el segundo pico muestra las fracciones que se retiene y que se van eluyendo conforme aumenta el gradiente del amortiguador salino.



**Figura 16. Perfil de absorbancias de la columna DEAE-FF.** Se utilizó un gradiente lineal de 0-100 %, flujo de 1ml/min. Buffer de equilibrio: Tris-HCl 50 mM, pH 7.1 y buffer de elución: Tris-HCl 50 mM, NaCl 100 mM, pH 7.1.

Se les realizó el ELISA a las fracciones 1 a la 27, de acuerdo con las condiciones mencionadas en materiales y métodos, para detectar la proteína CP4EPS. Se consideraron como muestras positivas las que presentaron una lectura mayor a 0.3 en el kit de ELISA. Si las lecturas dan entre 0.1 y 0.3, los resultados son indeterminados, requiriendo un mayor análisis y cuando las lecturas están por debajo de 0.1 se consideran negativas. Los resultados se presentan en la Tabla 19.

**Tabla 19. Lecturas de la placa de ELISA de las fracciones de la columna de DEAE-FF**

Fracción	Absorbancia (630 nm)
PBS	0.066
PBS	0.065
1	0.059
2	1.459
3	1.38
4	1.006
5	0.913
6	0.771
7	0.948
8	0.311

Fracción	Absorbancia a 630 nm
9	0.264
10	0.184
11	0.203
12	0.144
13	0.069
14	0.21
15	0.245
16	1.127
17	1.543
18	1.71
19	1.429
20	1.389
21	0.238
22	0.994
23	1.307
24	0.924
25	0.461
26	0.247
27	0.216
Ctrl + (200 µg)	1.729
Ctrl + (500 µg)	1.64
Ctrl + (1 mg)	1.827

Las muestras que están en rojo en la tabla 19 (fracciones 2-8, 16-20 y de la 22 a la 25), son las que presentaron una lectura mayor a 0.3, indicando que efectivamente se encuentra la enzima CP4EPSPS en dichas fracciones.

A continuación (Tabla 20), se presenta la eficiencia de recuperación de la columna DEAE-FF, respecto a la proteína total y la heteróloga.

**Tabla 20.** Eficiencia de recuperación de la columna DEAE-FF.

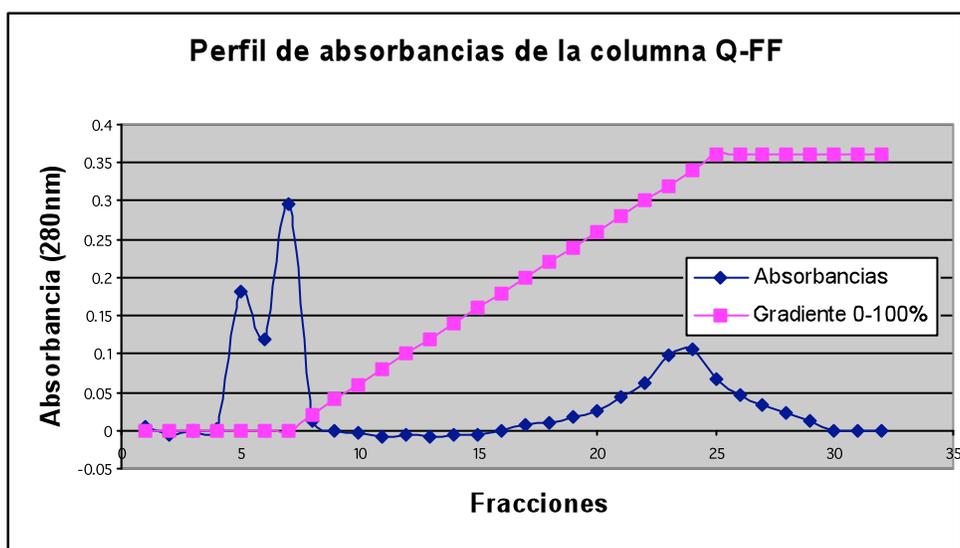
	<b>Al inicio</b>	<b>A la salida de la columna</b>	<b>% de recuperación</b>
<b>Proteína total</b>	<b>3.6 mg</b>	<b>2.264 mg</b>	<b>62.8 %</b>
<b>CP4EPSPS</b>	<b>26.64 µg</b>	<b>15.63 µg</b>	<b>58.6%</b>

Con base en su punto isoeléctrico teórico y las condiciones de corrida para esta columna, se hubiera esperado que la proteína se hubiera retenido hasta el final del gradiente. Como su punto isoeléctrico teórico es de 5.18, tendría más grupos cargados negativamente y en una columna de intercambio aniónico, caracterizada por tener unidos grupos catiónicos ( $-\text{CH}_2\text{-N}(\text{CH}_3)_3^+$ ), la afinidad de la proteína por la columna sería mayor y la separación se daría con base en el cambio de la concentración de sal de la fase móvil, de tal forma que se cree un gradiente salino (Nelson, 2001). Basándose en los resultados, cabe la posibilidad de que existan isoformas de la proteína, o bien, la concentración de proteína hubiera sido excesiva, provocando que no se retuviera y se eluyera desde el inicio de la corrida.

## **7.2 Columna de intercambio aniónico Q-FF**

Posteriormente, se utilizó la columna Q-FF, esperando que al igual que la columna anterior, si la columna tiene grupos catiónicos, al tener un punto isoeléctrico de 5.18, la proteína tendría mayor afinidad por la columna y la separación se daría con base al cambio de la concentración de sales de la fase móvil. Se inyectó la fracción de mayor lectura de la columna anterior (fracción 18 con una lectura de 1.71), en las condiciones que se mencionan en Materiales y Métodos y se obtuvieron 32 fracciones. A cada fracción obtenida se le midió la absorbancia a 280 nm para poder discriminar muestras y realizar el ELISA sólo a las que contenían proteína.

En la figura 17 se presenta el perfil de absorbancias de la columna Q-FF.



**Figura 17. Perfil de absorbancias de las fracciones obtenidas con la columna Q-FF.** Se utilizó un gradiente de 0-100 %, flujo de 1mLmin., buffer de inicio: Tris-HCl 50mM, pH 7.5 y buffer de elución: Tris-HCl 50 mM, NaCl 1.0 M, pH 7.5

De acuerdo a la lecturas de absorbancia, se les realizó ELISA a las fracciones 4-8 y 18-30, ya que las demás muestras presentaban una absorbancia muy pequeña y en la mayoría de los casos negativa. Los resultados del ELISA se muestran en la Tabla 21, presentándose en rojo las fracciones consideradas como positivas.

**Tabla 21. Lectura de la placa de las fracciones de la columna Q-FF**

Fracción	Absorbancia (630 nm)
PBS	0.044
PBS	0.045
4	0.059
5	0.083
6	0.292
7	0.277
8	0.258
18	0.159
19	0.608

Fracción	Absorbancia a 630 nm
20	1.006
21	1.188
22	1.235
23	1.211
24	1.118
25	1.066
26	0.981
27	0.727
28	0.477
29	0.251
30	0.174
Ctrl. +	1.47
Ctrl. +	1.403
SC1	0.061

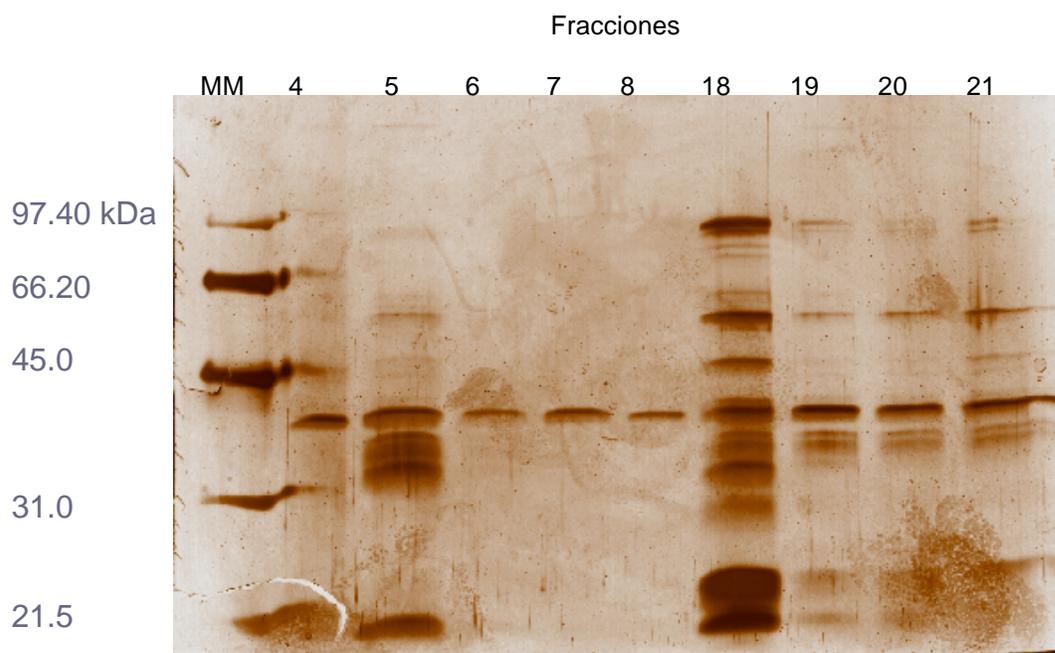
A continuación (Tabla 22) se presenta la eficiencia de recuperación de la columna Q-FF, respecto a la proteína total y la CP4EPSPS.

**Tabla 22.** Eficiencia de recuperación de la columna Q-FF

	Al inicio	A la salida de la columna	% de recuperación
<b>Proteína total</b>	<b>198 µg</b>	<b>38.6 µg</b>	<b>19.5 %</b>
<b>CP4EPSPS</b>	<b>23 µg</b>	<b>2.07 µg</b>	<b>9.13 %</b>

A las fracciones positivas se les realizó una electroforesis en gel de poliacrilamida en condiciones desnaturizantes para observar su patrón de bandeo, con la finalidad de corroborar si en alguna de las fracciones se contaba con bandas específicas. Se realizaron dos geles, uno para las fracciones 4 - 8 y otro para las fracciones 18 - 30. Se sabía que había proteína porque el ELISA fue positivo para estas fracciones; sin embargo, como no se pudo cuantificar por ningún método se sabía que se contaba con muy poca cantidad de proteína en

cada fracción. Por lo tanto, se realizaron geles de acrilamida de 1 mm de espesor, se cargaron 30  $\mu$ L en cada carril que fue lo que más se podía cargar y se tiñó con plata. El patrón de bandeo después de pasar las muestras por las columnas de DEAE-FF y Q-FF se puede observar en las figuras 20 y 21. También es importante mencionar que la cantidad de 2.07  $\mu$ g de proteína heteróloga es demasiado baja como para ocasionar efectos adversos a la salud humana o animal e incluso al ambiente.

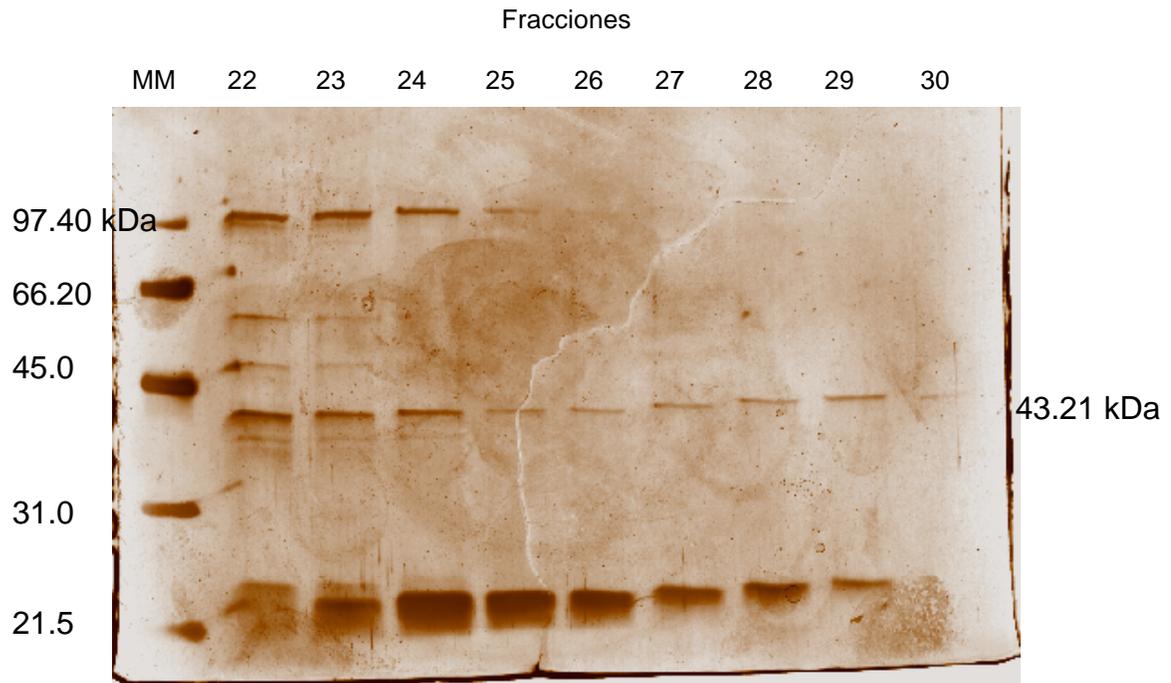


**Figura 18.** Gel de electroforesis en condiciones desnaturalizantes SDS PAGE al 10%, teñido con plata. Muestra el perfil electroforético de las proteínas después de las columnas DEAE y Q.

En la Figura 18 se observa que en los carriles 6, 7 y 8, aparece una banda aislada (43.21 kDa) que podría tratarse de la proteína buscada. Aunque la lectura en ELISA resultó baja (0.292, 0.277 y 0.258, respectivamente) podría ser que esté muy diluida ( 0.12 ng/mL) y no se tenga la cantidad suficiente para una lectura mayor. Para identificarla plenamente, se decidió secuenciar dicha banda.

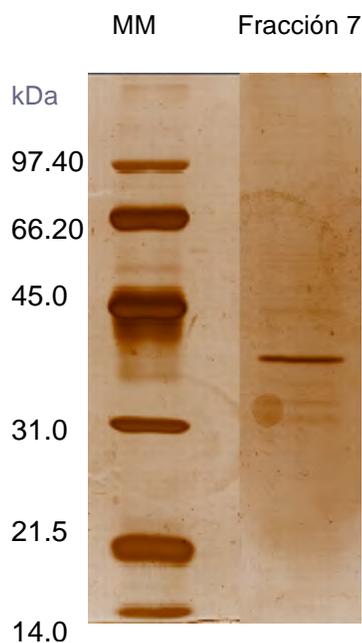
En la figura 19 se observa el perfil electroforético de las fracciones 22 a la 30. En éstas, las fracciones 22 a 28 tuvieron una señal positiva en ELISA pero se

observa más de una banda. Es decir, presuntamente pueden existir isoformas de la proteína EPSPS (Markert, 1975), ya que en sorgo se ha reportado tres (Ream, y col., 1988) y en maíz dos (Forlani, y col., 1994).



**Figura 19.** Gel de electroforesis en condiciones desnaturalizantes SDS PAGE al 10%, teñido con plata. Muestra el perfil electroforético de las proteínas después de las columnas DEAE y Q.

Con la finalidad de aislar la proteína de la fracción siete para su secuenciación y análisis por espectroscopía de masas, se realizó una electroforesis en gel de poliacrilamida en condiciones desnaturalizantes (Figura 20).



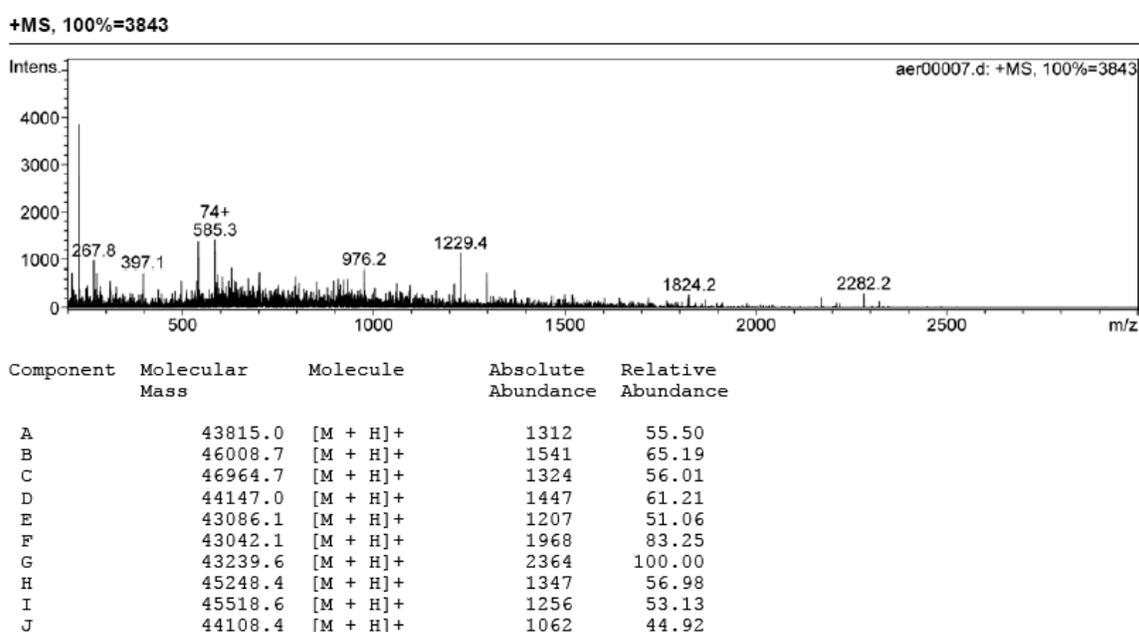
**Figura 20. Perfil electroforético de la fracción 7 obtenida a partir de columnas de intercambio aniónico DEAE y Q.** Gel de electroforesis en condiciones desnaturalizantes SDS PAGE al 10%, teñido con plata. Muestra la proteína que se logró separar mediante el uso de las columnas de intercambio aniónico DEAE y Q. La banda presenta un peso molecular aparente de 43 KDa.

Con referencia a los marcadores de peso molecular, se determinó el peso de la banda obtenida en la fracción siete y se obtuvo una masa molecular de 43.21 kDa. Esta banda se analizó por espectrometría de masas, para caracterizar el peso molecular que tiene.

## 8. Espectrometría de Masas

La espectrometría de masas se realizó a la proteína separada en la fracción siete con un peso molecular aproximado de 43 kDa (datos proporcionados por la densitometría). Para ello, se cortó la banda y se mandó al Laboratorio de Espectroscopia, en el Instituto de Química para analizarla en el espectroscopio Esquire 6000.

A continuación se muestra el reporte de desconvoluciones de los componentes del espectro de masas.



**Figura 21. Espectro de masas de la fracción siete.**

De acuerdo a la Figura 21, se observa que la fracción siete no se compone de una sola proteína, sino que dicho resultado muestra la presencia de diez componentes (A – J) de diferente pesos moleculares. La que se encuentra en mayor cantidad es una proteína de 43239.6 Da, basándonos en los resultados de la abundancia relativa arrojados en el estudio. Esto es una evidencia más que nos lleva a pensar que puede existir más de una proteína heteróloga, ya sean isoformas, proteínas de fusión o bien, podrían ser formas de degradación proteolítica.

## 9. Secuenciación

Con la finalidad de corroborar la identidad de los productos de la cromatografía de intercambio iónico, se mandó a secuenciar la banda correspondiente a la fracción 7. De ésta, se lograron secuenciar siete aminoácidos, los cuales fueron: M A Q V S I C. Cabe mencionar que se realizó por duplicado y se obtuvo el mismo resultado.

A ésta secuencia se le realizó un BLAST en el GeneBank, resultando con un 85% de homología con la EPSPS de *Arabidopsis thaliana* (Figura 22).

ID:1158870222-1160-41818478720.BLASTQ4

>sp|P05466|AROA\_ARATH 3-phosphoshikimate 1-carboxyvinyltransferase, chloroplast precursor (EC 2.5.1.19) (5-enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) - *Arabidopsis thaliana*

MAQVSRICNGVQNPSLISNLSKSSQRKSPLSVSLKTQQHPRAYPISSSWGLKKSGMTLIG  
SELRPLKVMSSVSTA EKASEIVLQPIREISGLIKLPGSKSLSNRILLLAALSEGTTVVDN  
LLNSDDINYMLDALKRLGLNVETDSENNRAVVEGCGGIFPASIDSKSDIELYLGNAGTAM  
RPLTAAVTAAGGNASYVLDGVPRMRERPIGDLVVGLKQLGADVECTLTGTNCPVVRVNANG  
GLPGGKVKLSGSISSQYLTALLMSAPLALGDVEIEIVDKLISVPYVEMTLKLMERFGVSV  
EHSDSWDRFFVKGGQKYKSPGNAYVEGDASSACYFLAGAAITGETVTVEGCGTTSLQGDV  
KFAEVLEKMGCKVSWTENSVTVTGPPRDAFGMRHLRAIDVNMNKMPDVAMTLAVVALFAD  
GPTTIRDVASWRVKETERMIAICTELRKL GATVEEGSDYCVITPPKKVKTAEIDTYDDHR  
MAMAFSLAACADVPIITINDSGCTRKTFPDYFQVLERITKH

**Figura 22.** Secuencia de aminoácidos obtenida de la fracción siete con el algoritmo BLAST con la base de datos del GenBank.

La construcción genética reportada por Monsanto especifica que sus elementos fueron el promotor CaMV35S del virus del mosaico de la coliflor, el péptido de tránsito de *Petunia hybrida*, el transgén de *Agrobacterium tumefaciens* cepa CP4 y el terminador NOS, también de *A. tumefaciens*. Sin embargo, el resultado obtenido del Blast anterior, no concuerda con dicha construcción. Esto podría deberse a la existencia de otro plásmido de transformación con el que se ensayó y probablemente transformó el cultivo en cuestión. Se realizó una búsqueda en la base de datos de las Patentes de los Estados Unidos

---

(USPTO.GOV), con la finalidad de encontrar una posible respuesta a la alineación encontrada, puesto que existía la posibilidad de que la información molecular proporcionada por Monsanto en los expedientes regulatorios no correspondiera con el resultado de esta secuenciación.

Se encontraron una serie de patentes correspondientes a la soya Roundup Ready®, entre ellas, las patentes 5,717,084, y la 5,728,925 (Herrera-Estrella y col., 1998) en las que se reivindica un gen quimérico que codifica para un péptido de tránsito y un péptido heterólogo.

En estas patentes aparecen las siguientes secuencias de los péptidos de tránsito:

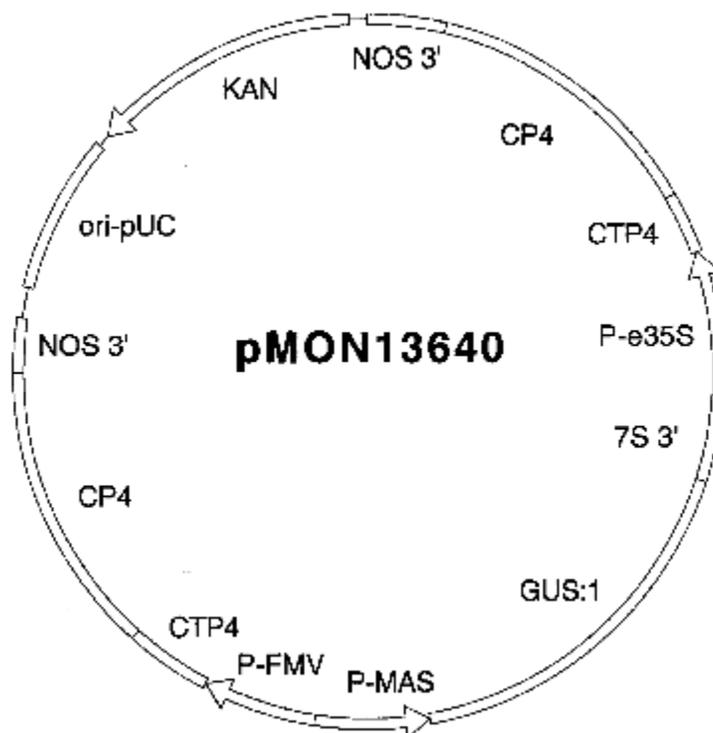
MASMISSSAVTTVSRASRGQSAAVAPFGGLKSMTGFPVKKVNTDITSITSNGGRVKC

MAASSSSSMALSSPTLAGKQLKLNPSSEIGAARPT

Correspondiendo al gen de la ribulosa-1,5-bisfosfato carboxilasa, y al gen de unión de la clorofila a/b respectivamente.

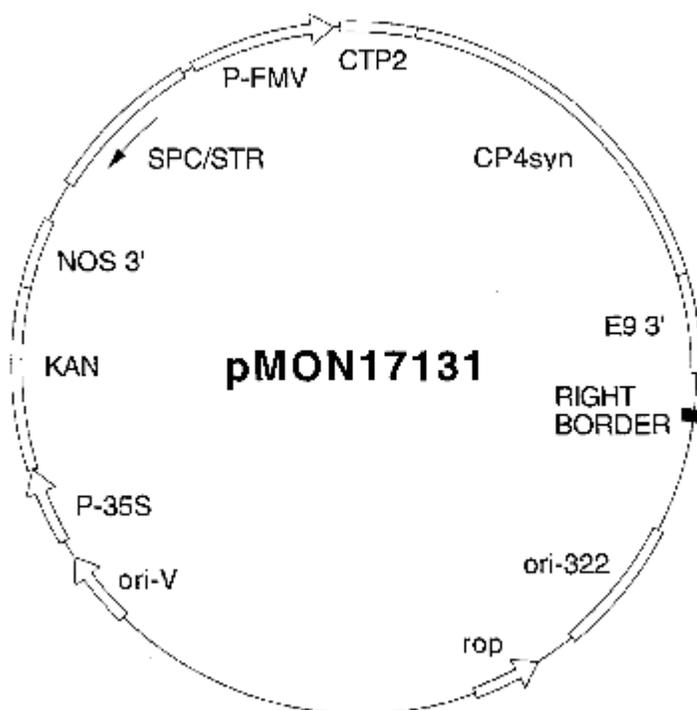
Además de estos péptidos de tránsito, en la patente 5,633,435 (Barry y col., 1997), se encontraron transformaciones con otros CTPs provenientes de diferentes organismos como la subunidad pequeña de la RUBISCO, Ferredoxina, Ferredoxina oxidoreductasa y la Tioredoxina. En esta misma patente también reportaron péptidos de tránsito sintéticos como el creado entre el CTP de la EPSPS de *Arabidopsis thaliana* y secuencias que codifican para la CP4EPSPS. Éste fue construido por mutagénesis dirigida y presenta un sitio de restricción Sph1 y se le denominó CTP2 (Anexo I). Posteriormente se construyó otro péptido de tránsito con un nuevo sitio de restricción a EcoRI, denominado CTP3, construido fusionando la secuencia del CTP de *Arabidopsis* con la CP4EPSPS. Ambos péptidos de tránsito importaron tan eficientemente la CP4EPSPS al cloroplasto como el control con el que hasta entonces se contaba que era el de *Petunia*. Esta patente menciona que la soya fue transformada al principio por el

método de biobalística, descrita (Christou y col., 1988) con el plásmido pMON13640 (Figura 23) en el que se observa que solamente contiene el péptido de tránsito de *Petunia*, el CTP4.



**Figura 23. Plásmido pMON13640.** Contiene el péptido de tránsito al cloroplasto de *Petunia hybrida* (CTP4).

La compañía Monsanto continuó realizando más construcciones genéticas y entre ellas utilizaron el plásmido pMON17131 (Figura 24) para transformar diferentes plantas como *Arabidopsis*, soya, maíz, trigo, papa, tomate, algodón, lechuga y betabel (Barry y col., 1997). En este plásmido se observa que utilizaron el CTP2 en la construcción, el cual proviene de *Arabidopsis*. Indicando que probablemente la compañía Monsanto estuvo o está mezclando los diferentes eventos de soya que obtuvo en las diferentes transformaciones y que por ese motivo se identificó el CTP2 de *Arabidopsis* y no de *Petunia*.



**Figura 24. Plásmido pMON17131.** Contiene el péptido de tránsito al cloroplasto de *Arabidopsis thaliana* (CTP2).

Por lo tanto, el producto obtenido a partir de la secuenciación para este estudio: **M A Q V S I C**, alinea con la secuencia de *Arabidopsis*, reportada en la patente número 5,633,435 de Barry y col., (1997) mencionado anteriormente. Es decir, la alineación es con el péptido CTP2 y no con el CTP4 del péptido de *Petunia* declarado por Monsanto para la construcción empleada en la transformación de la soya.

Así mismo, se puede llegar a presumir que no se está expresando una proteína heteróloga sino de al menos cuatro, las cuales podrían tratarse de isoformas o bien, de proteínas de fusión, así también podría ser que la ingeniería genética empleada por los desarrolladores de este OGM hayan utilizado una mezcla de diferentes plásmidos para la obtención de esta soya. O se trate de una mezcla de soyas GM con diferentes construcciones genéticas.

No se descarta totalmente la posibilidad de haber purificado la EPSPS propia de la soya, pero las probabilidades son mínimas, ya que, primero, el kit utilizado para la identificación de la heteróloga es específico para detectar a la transgénica. Segundo, en la secuenciación del fragmento de ADN hubiera alineado también con alguna secuencia de la soya silvestre, al igual que con la secuenciación del amino terminal, hubiera alineado con algún fragmento de la soya convencional, ya que la secuencia MAQVSIC es de *Arabidopsis thaliana*, utilizándose el CTP2 de ésta para obtener el plásmido pMON17131 utilizado para realizar las primeras transformaciones con soya, maíz y trigo entre otros.

Para confirmar la separación de la CP4EPSPS se podría concentrarse la fracción donde se encuentra esta proteína y posteriormente mandando a secuenciar la parte del sitio activo al glifosato para observar los aminoácidos que fueron sustituidos.

---

---

## VIII. CONCLUSIONES

La contribución hasta el momento de este trabajo fue la verificación e identificación de muestras de soya GM con base en técnicas moleculares como PCR punto final, así como la cuantificación por RTQ-PCR. La ventaja de esta última es que requiere un arreglo óptico que permite detectar fluorescencia, es más caro en comparación de la PCR convencional, pero es más rápido y, en ocasiones, más específico (Meyer, 1999). Una vez seleccionada la muestra de soya transgénica se logró la separación de uno de los péptidos de la CP4EPSPS, utilizando herramientas de purificación como el isoelectroenfoque preparativo y la cromatografía de intercambio aniónico (DEAE y Q). La cromatografía de intercambio iónico es el método que se utiliza con mayor frecuencia para purificación de proteínas. Ya que es un método en el que se ha observado una alta resolución en la separación de las proteínas, es relativamente fácil de realizar y de bajo costo. En primera instancia, se separó una banda equivalente a 43.21 kDa, que da señal positiva en ELISA. No es de extrañarse que esta proteína se visualice con un peso molecular menor, ya que Monsanto ha reportado que se ha llegado a visualizar esta proteína aproximadamente de 46 kDa en geles de SDS y Western blot, con un peso molecular específico de 47.6 kDa dado a las condiciones del gel, específicamente al tamaño del poro del mismo (Monsanto 2000). Es necesario corroborar que no haya proteínas heterólogas “enmascaradas” en esa banda. Al mandar a secuenciar la banda que se separó, se detectó que el péptido de tránsito es de un organismo diferente al reportado, es decir, proviene de *Arabidopsis thaliana* y no de *Petunia hybrida*. Esto no quiere decir que no esté presente el gen del péptido de *Petunia*, sino que al encontrarse al menos cuatro péptidos heterólogos, según los resultados del ELISA, es altamente probable que uno de los péptidos encontrados sí contenga el CTP de *Arabidopsis*. Además, existen diferentes trabajos donde sugieren la posibilidad de la existencia de más de un péptido heterólogo presente en la soya GTS 40-3-2. Uno de estos, es el trabajo reportado por Rang, 2005 donde demuestra la existencia de cuatro ARNm, obtenidos a partir del salto que hace la polimerasa del

codón de término. Esto lleva a la transcripción del fragmento adicional de menor tamaño que también se encuentra en la construcción de este evento, dando pie a la existencia de proteínas de fusión.

Tomando en cuenta los archivos regulatorios con información pública de libre acceso que la Compañía Monsanto ha presentado a las autoridades competentes, no se ha reportado la existencia de más de una proteína heteróloga en la soya.

De este trabajo experimental se concluye que pueden existir por lo menos la expresión de cuatro péptidos heterólogos, es decir, que podrían ser isoformas o bien, proteínas de fusión. Sin embargo las evidencias anteriores no permiten concluir que dicha expresión pueda repercutir en las proteínas endógenas de soya o entre ellas mismas aunque existe la posibilidad de que puedan provocar algún efecto en la planta.

### **Perspectivas**

Se recomienda la purificación de estos cuatro péptidos por cromatografía de exclusión molecular o electroforesis capilar y la caracterización enzimática y cinética de los mismos.

Se recomienda también que la secuenciación debería de ser entre el final del péptido de tránsito y el comienzo de la proteína EPSPS, ya que cuando se realiza algún BLAST de este tipo de proteínas, los algoritmos predictivos no toman en cuenta a la secuencia del péptido de tránsito, para llevar a cabo la predicción, y sólo se toma en cuenta el péptido maduro para la comparación entre proteínas. Así que, cuando se mandará a secuenciar la proteína obtenida, sí se lograría identificar y comprobar, principalmente, qué péptido de tránsito se ha utilizado para la transformación.

---

---

## IX. REFERENCIAS

- Ahmed, F. E. 2002. Detection on genetically modified organisms in foods. *Trends in Biotechnology*. 20 (5): 215-223
- Ausubel F.M., Brent R., Kingstor R. E., Mobre D.D., Seidman J.G., Smith J. A., Struhl K., 2000. Current protocols in molecular biology, Quantitation of DNA and RNA with absorption and fluorescence spectroscopy. A.3.D.1, Eds. John Wiley & Sons Inc.
- Barry; Gerard F. (St. Louis, MO), Kishore; Ganesh M. (Chesterfield, MO), Padgett; Stephen R. (Grover, MO), Stallings; William C. (Glencoe, MO). 1997. Glyphosate-tolerant 5-enolpyruvylshikimate-3-phosphate synthases. Monsanto Company (St. Louis, MO). United States Patent **5,627,061**
- Bolivar Zapata, F., 2004. Fundamentos y casos exitosos de la biotecnología moderna. El Colegio Nacional. Primera Edición.
- Bradford, M. M. 1976. A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding. *Analytical Biochemistry*. 72. 248-254.
- Brake, D. G. and Evenson, D. P. 2004. A generational study of glyphosate-tolerant soybeans on mouse fetal, postnatal, pubertal and adult testicular development. *Food and Chemical Toxicology*. 42: 29-36
- Brett G.M., Chambers S. J, Huang L., Morgan M. R. A. 1999. Design and development of immunoassays for detection of proteins, *Food control*. Vol.10, 401-406.
- Carlisle, S. M and Trevors, J. T. 1988. Glyphosate in the Environment. Water, Soil and Air Pollution 39, 409-420. Pesticides in Water: Report of the Working Party on the Incidence of Pesticides in Water, Department of the Environment, HSMO.
- Christou, P., McCabe, D. E., and Swain, W. F. 1988. Stable transformation of Soybean Callus by DNA-Coated Gold Particles. *Plant Physiol*. 87:671-674.
- Council Directive 2003/29/EC of the European Parliament and of the Council of 8 May 2003.
- Cromwell, G.L., Lindemann, M. D., Randolph, J. H., Parker, G. R., Coffey, R. D., Laurent, K. M., Armstrong, C. L., Mikel, W. B., Stanisiewski, E. P., and Hartnell, G. F. 2002. *J. Anim. Sci*. 80: 708-715
- Del Valle FR. 1981. Nutritional qualities of soya protein as affected by processing. *J Am Oil Chem Soc*. 419-29.

- 
- European Network of GMO laboratories. Definition of minimum performance requirement for Analytical methods of GMO destin. Version 25-01-2005
  - Forlani, G., Parisi, B. and Nielsen, E. 1994. 5-enol-Piruvyl-Shikimate-3-Phosphate Synthase from *Zea mays* Cultured Cells. *Plant Physiol.* 105:1107-1114
  - Funke, T., Han, M.L., Healy-Fried, Fischer, M. and Schönbrunn. 2006. Molecular basis for the herbicide resistance of Roundup Ready crops. *PNAS.* 103:13010-13015
  - Gachet, E., Martin, G. G., Vigneau, F., and Meyer, G. 1999. Detection of genetically modified organisms (GMO's) by PCR: a brief review of methodologies available. *Trends in Food Science and Technology.* 9: 380-388
  - Gasser, C. 1989. Genetically engineering plants for crop improvement. *Science.* 244:1293-1299
  - Gottlieb, L. D. 1982. Conservation and Duplication of Isoenzymes in Plants. *Science.* 216: 373-380
  - Hails, R.S. 2000. Genetically Modified Plants-The Debate Continues. *Trends in Environmental Ecology.* 15: 14-18.
  - Hammond, B. G., Vicini, J. L., Hartnell, G. F., Naylor, M., Knight, C. D., Robinson, E. H., Fuchs, R. L., and Padgett, S. R. 1996. The feeding value of soybeans fed to rats, chickens, catfish and dairy cattle is not altered by genetic incorporation of glyphosate tolerance. *Journal of Nutrition.* 126: 717-727
  - Hammond, B. G., Rogers, S. G. and Fuchs, R. L. 1999. Limitations of whole food feeding studies en food safety assessment. In: Food Safety Evaluation. OECD Documents, 85-97
  - 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 Padgett, S. R. 1996. The Expressed Protein in Glyphosate-Tolerant Soybean, 5-enolpyruvylshikimate-3-Phosphate Synthase from *Agrobacterium* sp. Strain CP4, Is Digested In Vitro and Is not Toxic to Acutely Gaviged Mice. *Journal of Nutrition.* 126: 728:740.
  - Heid, C. A., Stevens, J., Livak, K. J., and William, P. M. 1996. Real time quantitative PCR. *Genome Res.* 6 (10): 986-994
  - Hermann, F. J. 1962. A Revision of a Genus *Glicine* and its immediate allies. *US Dept. Agric. Tech Bull.* 1268: 1- 72

- 
- Heldt H.W. 2005. "Plant biochemistry". Elsevier Academic Press. Burlington, USA. 3ª Edición.
  - Herrera-Estrella; Luis (Gent, BE), Van Den Broeck; Guido (Gent, BE), Van Montagu; Marc (Brussel, BE), Schreier; Peter (Cologne, DE), Schell; Jeff (Cologne, DE), Bohnert; Hans J. (Tucson, AZ), Cashmore; Anthony R. (Woodside, NY), Timko; Michael P. (New York, NY), Kausch; Albert P. (Durham, NH). 1998. Chimaeric gene coding for a transit peptide and a heterologous polypeptide. United States Patent
  - Herrera-Estrella; Luis (Ghent, BE), Van Den Broeck; Guido (Ghent, BE), Van Montagu; Marc (Brussels, BE), Schreier; Peter (Cologne, DE), Schell; Jeff (Cologne, DE), Bohnert; Hans J. (Tucson, AZ), Cash more; Anthony R. (Woodside, NY), Timko; Michael P. (New York, NY), Kausch; Albert P. (Durham, NH). 1998. Chimaeric gene coding for a transit peptide and a heterologous peptide. United States Patent
  - Holt, J.S., Powles, S.B. and Holtum J. 1993. Mechanisms and agronomic aspects of herbicide resistance. *Annu. Rev. Plant. Physiol. Plant Mol. Biol.* 44:203-229
  - Hurts, C. D., Knight, A., and Bruce, I. J. 1999. PCR detection of genetically modified soya and maize in foodstuffs. *Molecular Breeding.* 5: 579-586
  - James, C. 2006. Global Status of Commercialized Biotech/GM Crops: 2006. ISAAA Briefs No.35. Ithaca, NY
  - James, C. 2005. Global Status of Commercialized Biotech/GM Crops: 2005. ISAAA Briefs No. 34. Ithaca, NY
  - James C. 2003. Global Review of Commercialized Transgenic Crops: 2002 Feature: Bt Maize. International Service for the Acquisition of Agro-biotech Applications. Briefs No. 29. ISAAA: Ithaca, NY, USA.
  - Jaworski, E.G. 1972. Mode of action of N-phosphonomethylglycine: Inhibition of aromatic amino acid biosynthesis. *J Agric Food Chem.* 20:1195-1198
  - Jensen A. 2001. A review, with notes on future needs and directions. National Veterinary Institute, Section of Food & Feed Microbiology, Oslo Norway, Version June 22<sup>nd</sup>.
  - Jonas DA, Antignac E, Antoine JM, Classen HG, Huggett A, Knudsen I, Mahler J, Ockhuizen T, Smith M, et al. 1986. The safety assessment of novel foods: Guidelines prepared by ILSI Europe Novel Food Task Force. *Food Chem Toxicol.* Oct;34:931-40.

- 
- Lewendon, A. and Coggins, J. R. 1983. Purification of 5-enolpiruvylshikimate-3-phosphate synthase from *Escherichia coli*. *Journal of Biochemistry*. 213: 187-191
  - Magaña-Gómez, J. A. 2007. Tesis: Evaluación histológica y de expresión de genes del páncreas de ratas alimentadas con soya transgénica. Hermosillo, Sonora: Centro de Investigación en Alimentación y Desarrollo. México.
  - Malatesta, M., Caporaloni, C., Gavaudan, S., Rocchi, M. B. L., Serafini, S., Tiberi, C., and Gazzanelli, G. 2002<sup>a</sup> Ultrastructural Morphometrical and Immunocytochemical Analyses of Hepatocyte Nuclei from Mice Fed on Genetically Modified Soybean. *Cell Structure and Function*. 27 (4): 173-180
  - Malatesta, M., Caporaloni, C., Rossi, L., Battistelli, S., Rocchi, M.B., Tonucci, F. and Gazzanelli G. 2002<sup>b</sup> Ultrastructural analysis of pancreatic acinar cells from mice fed on genetically modified soybean. *Journal of Anatomy*. 201(5): 409–415
  - Malik, J., Barry, G and Kishore, G. 1989. The herbicide glyphosate. *BioFactors* 2:17-25
  - Markert, C. L. 1975. Isozymes. Molecular Structure. Academic Press. USA.
  - Meyer, R. 1999. Development and application of DNA analytical methods for the detection of GMO's in food, *Food Control*.10:391-399
  - Moe, T. 1998. Perspectives on Traceability in Food Manufacture. *Food Science and Technology*: 9: 211-214
  - Monsanto Company. Update Molecular Characterization and Safety Assessment of Roundup Ready® Soybean Event 40-30-2. Confidential Report. Product Safety Center. St. Louis, MO, USA.
  - Mousdale, D. M. and Coggins, J. R. 1984. Purification and properties of 5-enolpiruvylshikimate 3-phosphate synthase from seedlings of *Pisum sativum* L. *Planta*. 160: 78-83
  - Nair RS, Fuchs RL, Schuette SA. 2002. Current methods for assessing safety of genetically modified crops as exemplified by data on Roundup Ready soybeans. *Toxicol Pathol*. Jan-Feb;30:117-25.
  - Nelson, D. L., y Cox, M.M. 2000. Lehninger Principios de Bioquímica. Ediciones Omega. Tercera Edición. 132
  - Oliveira, J. S., Mendes, M. A., Palma, M. S., Basso, L. A. and Santos D. S. 2003. One- Step purification of 5-enolpiruvylshikimate-3-phosphate synthase enzyme from *Mycobacterium tuberculosis*. *Protein Expresion and Purification*. 28. 287-292

- 
- Padgett, S. R., Kolacz, K. H., Re, D. B., LaVallee, B. J., Tinius, C. N., Rhodes, W. K., Otero, Y. I., Barry, G. F., Eichholtz, D. A., Peschke, V. M., Nida, D. L., Taylor, N. B. and Kishore, G. M. 1995. Developmet, Identification, and Caraceterization of a Glyphosate-Tolerant Soybean Line. *Crop Sci.* 35: 1451-1461
  - Padgett, S. R., Taylor, N. B., Nida, M. R., MacDonald, J., Holden, L. R. and Fuchs R. L. 1996. The Composition of Glyphosate-Tolerant Soybean Seeds Is Equivalent to That of Conventional Soybeans. *Journal of Nutrition.* 126: 702-716
  - Rang, A., Linke, B. and Jansen, B. 2005. Detection of RNA variants transcribed from the transgene in Round up Ready. *Eur Food Res Technol.* 220: 438 – 443
  - Ream, J. E., Steinrücken H. C., Porter, C. A. and Sikorski, J. A. 1988. Purification and properties of 5-enolpiruvylshikimate 3-phosphate synthase from Dark-Grown Seedlings of *Sorghum bicolor*. *Plant Physiol.* 87: 232-238
  - Regattieri, A., Gamberi, M. and Manzini, R. 2007. Traceability of Food Products: General Framework and experimental evidence. *Journal of Food Engineering:* 81: 347- 356.
  - Rodríguez, L.D., 2004. Development of Molecular-Based Techniques for the Detection, Identification and Quantification of Fodd-Borne Pathogens. Univertitat de Girona. ISBN: 84-688-8617-3. Diposit legal: GI-1127-2004
  - Rott, M. E., Lawrence, T. S., Wall, E. M., and Green, M. J. 2004. Detection and Quantification of Rounup Ready Soy in Foods by Conventional and Real-Time Polymerase Chain Reaction. *J. Agric. Food. Chem.* 52: 5223-5232
  - Rubin, J. L., Gaines, C. G. and Jensen, R. A. 1982. Enzymological Basis for Herbicidal Action of Glyphosate. *Plant Physiol.* 70: 833-839
  - Schmittgen, T. D., and Zakrajsek, B. A. 2000. Effect of experimental treatment on housekeeping gene expression:validation by real time, quantitative PCR. *J Biochem Biophys Methods.* 46: 69-81
  - Smith, A. K. 1980. Soybeans: Chemistry and Technology. Volume 1. Proteins. AVI Publishing Company, Inc. USA
  - Smith, CJS, Watson CF, Ray J, Bird CR, Morris PC, Schuch W, Grierson D. 1988. Antisense RNA inhibition of polygalacturonase gene expression in transgenic tomatoes. *Nature.* 334:724-6.
  - Stauffer, M.E., Young, J. K., and Evans, J. N. 2001. Shikimate-3-phosphate binds to the Isolated N-terminal domain of 5-enolpyruvylchikimate-3-phosphate synthase. *Biochemistry.* 40:3951-3957

- 
- Steinrücken, H.C. and Amrhein, N. 1980. The herbicide glyphosate is a potent inhibitor of 5-enolpyruvyl shikimic acid-3-phosphate synthase. *Biochem. Res. Comm.* 94:1207-1212
  - Steinrücken, h. C., Schulz, A., Amrhein, N., Porter, C. A. and Fraley, R. T. 1986. Overproduction of 5-enolpiruvylshikimate-3-phosphate Synthase in a Glyphosate-Tolerant *Petunia hybrida* Cell Line. *Archives of Biochemistry and Biophysics.* 244: 169-178
  - Takagi, K., Teshima, R., Nakajima, O., Okunuki, H., and Sawada, J. 2006. Improved ELISA method for screening human antigen-specific IgE and its application for monitoring specific IgE for novel proteins in genetically modified foods. *Regulatory Toxicology and Pharmacology.* 1-7
  - USDA, 2005. Market Improvement of Northern Soybeans Through Direct Marketing, Part II. Final Report.
  - Van Weemen, B. K. And Schuurs, A.H. W. M. 1971. Immunoassay Using Antigen-Enzyme Conjugates. *FEBS LETTERS.* 15: 232-236
  - Voet, D. 1995. Biochemistry. Second Edition. Jonh Wiley and Sons, Inc. USA.
  - Vollenhofer, S., Burg, K., Schmidt, J., and Kroath, H. 1999. Genetically Modified Organisms in Food-Screening and Specific Detection by Polimerase Chain Reaction. *J. Agric. Food. Chem.* 47: 5038-5043
  - Watson, J. D., Gilman, M., Witkowski, J., and Zoller, M. 1992. Recombinant DNA. Scientific American Books. 477
  - Williams, G. M., Kroes, R and Munro, I. C. 2000. Safety evaluation and risk assessment of the herbicide Roundup and its active ingredient, glyphosate, for humans. *Regulatory Toxicology and Pharmacology.* 31:117-165
  - Windels, P., Taveniers, T. Depicker, A., Van Bockstaele, E and De Loose, M. 2001.Characterization of the Roundup Ready soybean insert. *Eur. Food Res. Technol.* 213:107-112
  - Zimmermann, A., Hemmer, W., Liniger, M., Lüthy, J., and Pauli, U. 1998. A Sensitive Detection Method For Genetically Modified MaisGard™ Corn Using a Nested PCR-system. *Lebens-Wiss.u.-Technol.* 31: 664-667

---

**Páginas electrónicas**

- [www.agbios.com](http://www.agbios.com) AGBIOS
- [www.isaaa.org](http://www.isaaa.org) ISAAA
- [www.biotech.jrc.it](http://www.biotech.jrc.it) ICGEB
- [www.cofepris.gob.mx](http://www.cofepris.gob.mx) COFEPRIS
- [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov) NCBI
- [www.roche-applied-science.com](http://www.roche-applied-science.com) ROCHE
- [http://apec.biotec.or.th/pdf/part\\_3.pdf](http://apec.biotec.or.th/pdf/part_3.pdf) APEC
- [www.invitrogen.com](http://www.invitrogen.com) INVITROGEN
- [www.uspto.gov](http://www.uspto.gov) USPTO
- [www.expasy.org](http://www.expasy.org) EXPASY
- [www.sigma-aldrich.com](http://www.sigma-aldrich.com) SIGMA-ALDRICH
- [www.senasica.sagarpa.gob.mx](http://www.senasica.sagarpa.gob.mx) SENASICA/SAGARPA
- [www.solae.com](http://www.solae.com) SOLAE
- [www.gehealthcare.com](http://www.gehealthcare.com) GEHEALTHCARE

## Anexos

### Anexo I

A continuación se presentan las figuras que muestran las secuencias de ADN del péptido de tránsito al cloroplasto (CTP) y la secuencia de aminoácidos codificada derivada del CTP de la EPSPS de *Arabidopsis thaliana* (Anexo I, Fig. 9) conteniendo el sitio de restricción Sph1, a la cual se le denominó CTP2.

Las figuras 10A y 10B del anexo I muestran las secuencias de ADN y aminoácidos del péptido de tránsito derivada del gen EPSPS de *Arabidopsis thaliana* y que contiene el sitio de restricción EcoRI, la cual se refiere como CTP3.

En el anexo II se muestra los diferentes alineamientos del blast realizado al producto de la secuenciación del amino terminal que se realizó a la fracción siete obtenida después de las columnas de intercambio iónico.

## Anexo I

U.S. Patent

May 27, 1997

Sheet 25 of 70

5,633,435

```

60 AGATCTATCG ATAAGCTTGA TGTAATTGGA GGAAGATCAA AATTTTCAAT CCCCATTCCTT
113 CGATTGCTTC AATTGAAGTT TCTCCG ATG GCG CAA GTT AGC AGA ATC TGC AAT
    Met Ala Gln Val Ser Arg Ile Cys Asn
    1 5
161 GGT GTG CAG AAC CCA TCT CTT ATC TCC AAT CTC TCG AAA TCC AGT CAA
    Gly Val Gln Asn Pro Ser Leu Ile Ser Asn Leu Ser Lys Ser Ser Gln
    10 15 20 25
209 CGC AAA TCT CCC TTA TCG GTT TCT CTG AAG ACG CAG CAG CAT CCA CGA
    Arg Lys Ser Pro Leu Ser Val Ser Leu Lys Thr Gln Gln His Pro Arg
    30 35 40
257 GCT TAT CCG ATT TCG TCG TCG TGG GGA TTG AAG AAG AGT GGG ATG ACG
    Ala Tyr Pro Ile Ser Ser Ser Trp Gly Leu Lys Lys Ser Gly Met Thr
    45 50 55
305 TTA ATT GGC TCT GAG CTT CGT CCT CTT AAG GTC ATG TCT TCT GTT TCC
    Leu Ile Gly Ser Glu Leu Arg Pro Leu Lys Val Met Ser Ser Val Ser
    60 65 70
318 ACG GCG TGC ATG C
    Thr Ala Cys Met
    75

```

**Figura 9.** Secuencia del péptido de tránsito de *Arabidopsis thaliana* mostrando el sitio de restricción SphI (CTP2).

U.S. Patent

May 27, 1997

Sheet 26 of 70

5,633,435

AGATCTATCG	ATAAGCTTGA	TGTAATTGGA	GGAAGATCAA	AATTTTCAAT	CCCCATTCIT	60
CGATTGCTTC	AATTGAAAGTT	TCTCCG	ATG GCG CAA	GTT AGC AGA	ATC TGC AAT	113
		Met	Ala Gln Val	Ser Arg Ile	Cys Asn	
		1		5		
GGT GTG CAG AAC CCA	TCT CTT ATC TCC AAT	CTC TCG AAA	TCC AGT CAA			161
Gly Val Gln Asn Pro Ser	Leu Ile Ser Asn	Leu Ser Lys	Ser Ser Gln			
10	15	20	25			
CGC AAA TCT CCC TTA	TCG GTT TCT CTG AAG	ACG CAG CAG	CAT CCA CGA			209
Arg Lys Ser Pro Leu Ser	Val Ser Leu Lys	Thr Gln Gln	His Pro Arg			
	30	35	40			
GCT TAT CCG AIT TCG	TCG TCG TGG GGA	TTG AAG AAG	AGT GGG ATG ACG			257
Ala Tyr Pro Ile Ser Ser	Trp Gly Leu Lys	Lys Ser Gly	Met Thr			
	45	50	55			

Figura 10 A

U.S. Patent		May 27, 1997	Sheet 27 of 70	5,633,435												
TTA	ATT	GGC	TCT	GAG	CTT	GAG	CTT	AAG	GTC	ATG	TCT	TCT	GTT	TCC	305	
Leu	Ile	Gly	Ser	Glu	Leu	Arg	Pro	Lys	Val	Met	Ser	Ser	Val	Ser		
		60									70					
ACG	GCG	GAG	AAA	GCG	TCG	GAG	ATT	GTA	CTT	CAA	CCC	ATT	AGA	GAA	ATC	353
Thr	Ala	Glu	Lys	Ala	Ser	Glu	Ile	Val	Leu	Gln	Pro	Ile	Arg	Glu	Ile	
		75									85					
TCC	GGT	CTT	ATT	AAG	TTG	CCT	GGC	TCC	AAG	TCT	CTA	TCA	AAT	AGA	ATT	401
Ser	Gly	Leu	Ile	Lys	Leu	Pro	Gly	Ser	Lys	Ser	Leu	Ser	Asn	Arg	Ile	
		90													105	
																402
																C

**Figura 10B.** Secuencia del péptido de tránsito de *Arabidopsis thaliana* mostrando el sitio de restricción EcoRI (CTP3).

## Anexo II

Secuencias de aminoácidos, obtenidas de las traducciones in silico de las propuestas por Rang, 2005.

Primer fragmento, llamado:

### CP4 EPSPS 3

MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDG  
VNGGLLAPEAPLDFGNAATGCRLTMGLVGVDYDFDSTFIGDASLTKRPMGRVNLNPLREMGVQVKSEGDRLPVTLRGPKTP  
TPITYRVPMASAQVKSAVLLAGLNTPGITTVIEPIIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPG  
DPSSTAFPLVAALLVPGSDVTILNVLNPNTRTGLIILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLKGVTVPEDRAPSM  
IDEYPIILAVAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNQVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHR  
IAMSFLVMGLVSENPTVDDATMIATSFPEFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDGVGNG  
GLLAPEAPLDFGNAATGCRLTMGLVGVDYDFKRIMLGNFSEIISIFLGISAVTGE

pI/Mw teóricos: 5.05 / 56503.76

Segundo fragmento:

### CP4 EPSPS 4

MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDG  
VNGGLLAPEAPLDFGNAATGCRLTMGLVGVDYDFDSTFIGDASLTKRPMGRVNLNPLREMGVQVKSEGDRLPVTLRGPKTP  
TPITYRVPMASAQVKSAVLLAGLNTPGITTVIEPIIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPG  
DPSSTAFPLVAALLVPGSDVTILNVLNPNTRTGLIILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLKGVTVPEDRAPSM  
IDEYPIILAVAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNQVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHR  
IAMSFLVMGLVSENPTVDDATMIATSFPEFMGLMAGLAKIELSDTGKAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAP  
LDFGNAATGCRLTMGLVGVDYDFKRIMLGNFSEIISIFLGISAVTGE

pI/Mw teóricos: 5.09 / 55687.94

Tercer fragmento:

### CP4 EPSPS 5

MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDG  
VNGGLLAPEAPLDFGNAATGCRLTMGLVGVDYDFDSTFIGDASLTKRPMGRVNLNPLREMGVQVKSEGDRLPVTLRGPKTP  
TPITYRVPMASAQVKSAVLLAGLNTPGITTVIEPIIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPG  
DPSSTAFPLVAALLVPGSDVTILNVLNPNTRTGLIILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLKGVTVPEDRAPSM  
IDEYPIILAVAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNQVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHR  
IAMSFLVMGLVSENPTVDDATMIATSFPEFMGLMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDGV  
GNGGLLAPEAPLDFGNAATGCRLTMGLVGVDYDFKRIMLGNFSEIISIFLGISAVTGE

pI/Mw teóricos: 5.01 / 56863.20

Cuarto fragmento:

### CP4 EPSPS 6

MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDG  
VNGGLLAPEAPLDFGNAATGCRLTMGLVGVDYDFDSTFIGDASLTKRPMGRVNLNPLREMGVQVKSEGDRLPVTLRGPKTP  
TPITYRVPMASAQVKSAVLLAGLNTPGITTVIEPIIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPG  
DPSSTAFPLVAALLVPGSDVTILNVLNPNTRTGLIILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLKGVTVPEDRAPSM  
IDEYPIILAVAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNQVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHR  
IAMSFLVMGLVSENPTVDDATMIATSFPEFMGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFG  
NAATGCRLTMGLVGVDYDFKRIMLGNFSEIISIFLGISAVTGE

pI/Mw teórico: 5.00 / 55371.48

### Anexo III

A continuación se presentan los alineamientos obtenidos a partir del blast del producto de la secuenciación de la fracción siete obtenido de las columnas de intercambio iónico.

#### Blast *Arabidopsis thaliana*

Los aminoácidos obtenidos a partir de la proteína mandada a secuenciar se alinearon y esto fue lo que se obtuvo:

Blast RID : 1159380265-18197-138564513526.BLASTQ2

```
>X gi|15225450|ref|NP_182055.1| 3-phosphoshikimate 1-
carboxyvinyltransferase [Arabidopsis thaliana]
  gi|2583124|gb|AAB82633.1| 5-enolpyruvylshikimate-3-phosphate (EPSP)
synthase [Arabidopsis
thaliana]
  gi|63003818|gb|AAY25438.1| At2g45300 [Arabidopsis thaliana]
  gi|110742505|dbj|BAE99170.1| 5-enolpyruvylshikimate-3-phosphate (EPSP)
synthase [Arabidopsis
thaliana]
Length=520
```

```
Score = 22.7 bits (46), Expect = 1430
Identities = 7/8 (87%), Positives = 7/8 (87%), Gaps = 1/8 (12%)
```

```
Query 1 MAQVS-IC 7
      MAQVS IC
Sbjct 1 MAQVSRIC 8
```

```
gi|1351978|sp|P05466|AROA_ARATH 3-phosphoshikimate 1-
carboxyvinyltransferase, chloroplast precursor
(5-enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase)
  gi|295790|emb|CAA29828.1| EPSP [Arabidopsis thaliana]
Length=520
```

```
Score = 22.7 bits (46), Expect = 1430
Identities = 7/8 (87%), Positives = 7/8 (87%), Gaps = 1/8 (12%)
```

```
Query 1 MAQVS-IC 7
      MAQVS IC
Sbjct 1 MAQVSRIC 8
```

```
gi|66618|pir||XUMUVS 3-phosphoshikimate 1-carboxyvinyltransferase (EC
2.5.1.19) precursor
- Arabidopsis thaliana
Length=520
```

Score = 22.7 bits (46), Expect = 1430  
 Identities = 7/8 (87%), Positives = 7/8 (87%), Gaps = 1/8 (12%)

Query 1 MAQVS-IC 7  
 MAQVS IC  
 Sbjct 1 MAQVSRIC 8

Blast péptido 1

1159379853-776-5263836019.BLASTQ4

5-enol-pyruvylshikimate-3-phosphate synthase class 2 precursor  
 [Agrobacterium sp. CP4]  
 Length=527

Score = 970 bits (2508), Expect = 0.0, Method: Composition-based stats.  
 Identities = 527/527 (100%), Positives = 527/527 (100%), Gaps = 0/527 (0%)

Query 1 MAQINNMAQGIQTLNPNNSNFHKPQVPKSSSFLVFGSKKLNANSMLVLKKDSIFMQKFC 60  
 MAQINNMAQGIQTLNPNNSNFHKPQVPKSSSFLVFGSKKLNANSMLVLKKDSIFMQKFC  
 Sbjct 1 MAQINNMAQGIQTLNPNNSNFHKPQVPKSSSFLVFGSKKLNANSMLVLKKDSIFMQKFC 60

Query 61 SFRISASVATACMLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRIT 120  
 SFRISASVATACMLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRIT  
 Sbjct 61 SFRISASVATACMLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRIT 120

Query 121 GLLEGEDVINTGKAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLTMG 180  
 GLLEGEDVINTGKAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLTMG  
 Sbjct 121 GLLEGEDVINTGKAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLTMG 180

Query 181 LVGVYDFDSTFIGDASLTKRPMGRVLNPLREMGVQVKSSEDGDRLPVTLRGPKTPTPIYR 240  
 LVGVYDFDSTFIGDASLTKRPMGRVLNPLREMGVQVKSSEDGDRLPVTLRGPKTPTPIYR  
 Sbjct 181 LVGVYDFDSTFIGDASLTKRPMGRVLNPLREMGVQVKSSEDGDRLPVTLRGPKTPTPIYR 240

Query 241 VPMASAQVKSAVLLAGLNTPGITTVIEPIMTCDHTEKMLQGFGANLTVETDADGVRTIRL 300  
 VPMASAQVKSAVLLAGLNTPGITTVIEPIMTCDHTEKMLQGFGANLTVETDADGVRTIRL  
 Sbjct 241 VPMASAQVKSAVLLAGLNTPGITTVIEPIMTCDHTEKMLQGFGANLTVETDADGVRTIRL 300

Query 301 EGRGKLTGQVIDVPGDPSSTAFPLVAALLVPGSDVTILNVLNPNTRTGLLITLQEMGADI 360  
 EGRGKLTGQVIDVPGDPSSTAFPLVAALLVPGSDVTILNVLNPNTRTGLLITLQEMGADI  
 Sbjct 301 EGRGKLTGQVIDVPGDPSSTAFPLVAALLVPGSDVTILNVLNPNTRTGLLITLQEMGADI 360

Query 361 EVINLRLAGGEDVADLRVRSSTLKGVTVPEDRAPPIDEYPI LAVAAAFAEGATVMNGLE 420  
 EVINLRLAGGEDVADLRVRSSTLKGVTVPEDRAPPIDEYPI LAVAAAFAEGATVMNGLE  
 Sbjct 361 EVINLRLAGGEDVADLRVRSSTLKGVTVPEDRAPPIDEYPI LAVAAAFAEGATVMNGLE 420

Query 421 ELRVKESDRLSAVANGLKLNQVDCDEGETSLVVVRGPRPDGKGLGNASGA AVATHLDHRIAM 480  
 ELRVKESDRLSAVANGLKLNQVDCDEGETSLVVVRGPRPDGKGLGNASGA AVATHLDHRIAM  
 Sbjct 421 ELRVKESDRLSAVANGLKLNQVDCDEGETSLVVVRGPRPDGKGLGNASGA AVATHLDHRIAM 480

Query 481 SFLVMGLVSENPTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 527  
 SFLVMGLVSENPTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA  
 Sbjct 481 SFLVMGLVSENPTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 527

gi|18266432|gb|AAL67577.1|AF464188\_1 CP4EPS [Glycine max]  
 gi|27549260|gb|AAO17037.1| CP4EPS protein [synthetic construct]  
 Length=455

Score = 814 bits (2103), Expect = 0.0, Method: Composition-based stats.  
Identities = 452/455 (99%), Positives = 452/455 (99%), Gaps = 0/455 (0%)

Query	73	MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG	132
		MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG	
Sbjct	1	MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG	60
Query	133	KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI	192
		KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI	
Sbjct	61	KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI	120
Query	193	GDASLTKRPMGRVLNPLREMGVQVKSEGDRLPVTLRGPKTPTPI TYRVPMASAQVKS AV	252
		GDASLTKRPMGRVLNPLREMGVQVKSEGDRLPVTLRGPKTPTPI TYRVPMASAQVKS AV	
Sbjct	121	GDASLTKRPMGRVLNPLREMGVQVKSEGDRLPVTLRGPKTPTPI TYRVPMASAQVKS AV	180
Query	253	LLAGLNTPGITTVIEPIMTCDHTEKMLQGFANLTVETDADGVRTIRLEGRGKLTGQVID	312
		LLAGLNTPGITTVIEPIMT DHTEKMLQGFANLTVETDADGVRTIRLEGRGKLTGQVID	
Sbjct	181	LLAGLNTPGITTVIEPIMTRDHTKMLQGFANLTVETDADGVRTIRLEGRGKLTGQVID	240
Query	313	VPGDPSSTAFPLVAALLVPGSDVTILNLVLMNPTRTGLILTLQEMGADIEVINRLAGGED	372
		VPGDPSSTAFPLVAALLVPGSDVTILNLVLMNPTRTGLILTLQEMGADIEVIN RLAGGED	
Sbjct	241	VPGDPSSTAFPLVAALLVPGSDVTILNLVLMNPTRTGLILTLQEMGADIEVINRLAGGED	300
Query	373	VADLRVRSSTLKGVTVPEDRAPMIDEYPI LAVAAAAFAEGATVMNGLEELRVKESDRLSA	432
		VADLRVRSSTLKGVTVPEDRAP MIDEYPI LAVAAAAFAEGATVMNGLEELRVKESDRLSA	
Sbjct	301	VADLRVRSSTLKGVTVPEDRAPSMIDEYPI LAVAAAAFAEGATVMNGLEELRVKESDRLSA	360
Query	433	VANGLKLNQVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP	492
		VANGLKLNQVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP	
Sbjct	361	VANGLKLNQVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP	420
Query	493	VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA	527
		VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA	
Sbjct	421	VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA	455

gi|8469107|sp|Q9R4E4|ARO\_AGRSC 3-phosphoshikimate 1-carboxyvinyltransferase (5-enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS)  
gi|114794058|pdb|2GG4|A Chain A, Cp4 Epsp Synthase (Unliganded)  
gi|114794059|pdb|2GG6|A Chain A, Cp4 Epsp Synthase Liganded With S3p  
gi|114794060|pdb|2GGA|A Chain A, Cp4 Epsp Synthase Liganded With S3p And Glyphosate  
Length=455

Score = 810 bits (2093), Expect = 0.0, Method: Composition-based stats.  
Identities = 451/455 (99%), Positives = 451/455 (99%), Gaps = 0/455 (0%)

Query	73	MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG	132
		M HGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG	
Sbjct	1	MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG	60
Query	133	KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI	192
		KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI	
Sbjct	61	KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI	120
Query	193	GDASLTKRPMGRVLNPLREMGVQVKSEGDRLPVTLRGPKTPTPI TYRVPMASAQVKS AV	252
		GDASLTKRPMGRVLNPLREMGVQVKSEGDRLPVTLRGPKTPTPI TYRVPMASAQVKS AV	
Sbjct	121	GDASLTKRPMGRVLNPLREMGVQVKSEGDRLPVTLRGPKTPTPI TYRVPMASAQVKS AV	180
Query	253	LLAGLNTPGITTVIEPIMTCDHTEKMLQGFANLTVETDADGVRTIRLEGRGKLTGQVID	312
		LLAGLNTPGITTVIEPIMT DHTEKMLQGFANLTVETDADGVRTIRLEGRGKLTGQVID	

Sbjct	181	LLAGLNTPGITTVIEPIMTRDHTKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID	240
Query	313	VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINRLAGGED	372
Sbjct	241	VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED	300
Query	373	VADLRVRSSTLKGVTVPEDRAPPIDEYPI LAVAAAFAEGATVMNGLEELRVKESDRLSA	432
Sbjct	301	VADLRVRSSTLKGVTVPEDRAP MIDEYPI LAVAAAFAEGATVMNGLEELRVKESDRLSA	360
Query	433	VANGLKLNQVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSEN	492
Sbjct	361	VANGLKLNQVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSEN	420
Query	493	VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA	527
Sbjct	421	VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA	455

gi|114794061|pdb|2GGD|A Chain A, Cp4 Epsp Synthase Ala100gly Liganded With S3p  
And Glyphosate  
Length=455

Score = 809 bits (2089), Expect = 0.0, Method: Composition-based stats.  
Identities = 450/455 (98%), Positives = 450/455 (98%), Gaps = 0/455 (0%)

Query	73	MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG	132
Sbjct	1	MHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG	60
Query	133	KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI	192
Sbjct	61	KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNA TGCRLTMGLVGVYDFDSTFI	120
Query	193	GDASLTKRPMGRVLNPLREMGVQVKSEGDRLPVTLRGPKTPTPIYRVPMASAQVKS	252
Sbjct	121	GDASLTKRPMGRVLNPLREMGVQVKSEGDRLPVTLRGPKTPTPIYRVPMASAQVKS	180
Query	253	LLAGLNTPGITTVIEPIMTCDHTKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID	312
Sbjct	181	LLAGLNTPGITTVIEPIMTRDHTKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID	240
Query	313	VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINRLAGGED	372
Sbjct	241	VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED	300
Query	373	VADLRVRSSTLKGVTVPEDRAPPIDEYPI LAVAAAFAEGATVMNGLEELRVKESDRLSA	432
Sbjct	301	VADLRVRSSTLKGVTVPEDRAP MIDEYPI LAVAAAFAEGATVMNGLEELRVKESDRLSA	360
Query	433	VANGLKLNQVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSEN	492
Sbjct	361	VANGLKLNQVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSEN	420
Query	493	VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA	527
Sbjct	421	VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA	455

> [gi|15964006|ref|NP\\_384359.1](#)  3-phosphoshikimate 1-carboxyvinyltransferase  
[Sinorhizobium meliloti  
1021]

[gi|21362380|sp|Q92SV5|ARO\\_A\\_RHIME](#) 3-phosphoshikimate 1-carboxyvinyltransferase  
(5-enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS)

[gi|15073182|emb|CAC41690.1|G](#) PUTATIVE 3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE PROTEIN  
[Sinorhizobium meliloti]  
Length=455

Score = 770 bits (1988), Expect = 0.0, Method: Composition-based stats.  
Identities = 410/455 (90%), Positives = 428/455 (94%), Gaps = 0/455 (0%)

Query	73	MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG	132
		M HG++ RPATARKSS L GT+RIPGDKSISHRSFMFGGLA+GETRITGLLEGEDVINTG	
Sbjct	1	MSHGSNPRPATARKSSDLKGTLRIPGDKSISHRSFMFGGLAAGETRITGLLEGEDVINTG	60
Query	133	KAMQAMGARIRKEGDTWIIDGVGNGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI	192
		KAMQAMGARIRKEGDTWIIDGVGNGLLAPEAPLDFGNA TGCRMTMGLVGVYDFDSTFI	
Sbjct	61	KAMQAMGARIRKEGDTWIIDGVGNGLLAPEAPLDFGNAGTGCRLTMGLVGVYDFDSTFI	120
Query	193	GDASLTKRPMGRVNLNPLREMGVQVKSSEDGDRLPVTLRGPKTPITPYRVPMSAQVKSVA	252
		GDASLTKRPMGRVL+PLREMGVQVKS +GDRLPVTLRGPKTP ITYRVPMSAQVKSVA	
Sbjct	121	GDASLTKRPMGRVLDPLREMGVQVKSAGDRLPVTLRGPKTPNPITYRVPMSAQVKSVA	180
Query	253	LLAGLNTPGITTVIEPIMTCDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID	312
		LLAGLNTPGITTV+EP+MT DHTEKMLQGFGANLTVETDA+GVRTIRLEGRGKLTGQVID	
Sbjct	181	LLAGLNTPGITTVVEPVMTRDHTKMLQGFGANLTVETDAEGVRTIRLEGRGKLTGQVID	240
Query	313	VPGDPSSTAFPLVAALLVPGSDVITILNVLNPNTRTGLLITLQEMGADIEVINRLLAGGED	372
		VPGDPSSTAFPLVA L+VPGSD+TILNVLNPNTRTGLLITLQEMGA+IEV+N RLAGGED	
Sbjct	241	VPGDPSSTAFPLVAGLIVPGSDITILNVLNPNTRTGLLITLQEMGANIEVMNKRRLAGGED	300
Query	373	VADLRVRSSTLKGVTVPEDRAPPIDEYPI LAVAAAAFAEGATVMNGLEELRVKESDRLSA	432
		VADLRVR S LKGVTVPEDRAP MIDEYP+LAVAAAAFAEG TVMNGLEELRVKESDRLSA	
Sbjct	301	VADLRVRHSELKGVTVPEDRAPS MIDEYPVLAVAAAAFAEGTTVMNGLEELRVKESDRLSA	360
Query	433	VANGLKLNQVDCDEGETSLVVRGRPDGKGLGNASGA AVATHLDHRIAMSFLVMGLVSENP	492
		VA+GLKLNQVDCDEGE SLVVRGRP GKGLG SG V THLDHRIAMSFLVMGL SE+P	
Sbjct	361	VADGLKLNQVDCDEGEASLVVRGRPGGKGLKISGGQVKTHLDHRIAMSFLVMGLASEHP	420
Query	493	VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA	527
		VTVDDATMIATSFPEFM LM GLGAKIE ++ KAA	
Sbjct	421	VTVDDATMIATSFPEFMGLMTGLGAKIEEAENKAA	455

> [gi|113871287|ref|ZP\\_01411416.1|](#) 3-phosphoshikimate 1-carboxyvinyltransferase  
[Sinorhizobium medicae  
WSM419]

[gi|113725959|gb|EAU07051.1|](#) 3-phosphoshikimate 1-carboxyvinyltransferase  
[Sinorhizobium medicae  
WSM419]  
Length=456

Score = 745 bits (1923), Expect = 0.0, Method: Composition-based stats.  
Identities = 407/455 (89%), Positives = 425/455 (93%), Gaps = 0/455 (0%)

Query	73	MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG	132
		M HG++ RPATARKSS L GTVRIIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG	
Sbjct	1	MSHGSNPRPATARKSSDLKGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG	60
Query	133	KAMQAMGARIRKEGDTWIIDGVGNGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI	192
		+AMQAMGA+IRKEGDTWIIDGVGNGLLAPEAPLDFGNA TGCRMTMGLVGVYDFDSTFI	
Sbjct	61	RAMQAMGAKIRKEGDTWIIDGVGNGLLAPEAPLDFGNAGTGCRLTMGLVGVYDFDSTFI	120

Query	193	GDASLTKRPMGRVLNPLREMGVQVKSSEDGDRLPVTLRGPKTPPTPITYRVPMASAQVKS AV	252
Sbjct	121	GDASLTKRPMGRVL+PLREMGVQVKS +GDRLPVTLRGPKTP PITYRVPMASAQVKS AV	180
Query	253	LLAGLNTPGITTVIEPIMTCDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID	312
Sbjct	181	LLAGLNTPGITTV+EP+MT DHTEKMLQGFGLA+L+VETD DGVRTIRLEGRGKLTGQVID	240
Query	313	VPGDPSSTAFPLVAALLVPGSDVTILNVLNMPTRTGLILTLQEMGADIEVINRLRAGGED	372
Sbjct	241	VPGDPSSTAFPLVAALLVPGSD++I NVLMNPTRTGLILTLQEMGA IEV++ RLAGGED	300
Query	373	VADLRVRSSTLKGVTVPEDRAPPIMIDEYPI LAVAAAAFAEGATVMNGLEELRVKESDRLSA	432
Sbjct	301	VADLRVRSSTLKGVTVPEDRAPPIMIDEYPI LAVAAAAFAEGATVMNGLEELRVKESDRLSA	360
Query	433	VANGLKLNQVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP	492
Sbjct	361	VAEGLKLNQVDCDEGEASLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP	420
Query	493	VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA	527
Sbjct	421	VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA	455

>X [gi|115254522|emb|CAK05596.1](https://gi|115254522|emb|CAK05596.1) putative 3-phosphoshikimate 1-carboxyvinyltransferase (5-enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) [Rhizobium leguminosarum bv. viciae 3841]  
Length=452

Score = 733 bits (1893), Expect = 0.0, Method: Composition-based stats.  
Identities = 387/449 (86%), Positives = 416/449 (92%), Gaps = 0/449 (0%)

Query	73	MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG	132
Sbjct	1	MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG	60
Query	133	KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI	192
Sbjct	61	KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI	120
Query	193	GDASLTKRPMGRVLNPLREMGVQVKSSEDGDRLPVTLRGPKTPPTPITYRVPMASAQVKS AV	252
Sbjct	121	GDASLTKRPMGRVLNPLREMGVQV + +GDRLPVTLRGP TP+PI YRVPMASAQVKS AV	180
Query	253	LLAGLNTPGITTVIEPIMTCDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID	312
Sbjct	181	LLAGLNTPGITTVIEPIMTCDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID	240
Query	313	VPGDPSSTAFPLVAALLVPGSDVTILNVLNMPTRTGLILTLQEMGADIEVINRLRAGGED	372
Sbjct	241	VPGDPSSTAFPLVAAL+VPGSD+TI+NVLMNPTRTGLILTLQEMGADIEV+N RLAGGED	300
Query	373	VADLRVRSSTLKGVTVPEDRAPPIMIDEYPI LAVAAAAFAEGATVMNGLEELRVKESDRLSA	432
Sbjct	301	VADLRVRSSTLKGVTVPEDRAPPIMIDEYPI LAVAAAAFAEGATVMNGLEELRVKESDRLSA	360
Query	433	VANGLKLNQVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP	492
Sbjct	361	VADGLKLNQVDCDEGEDFLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP	420

```

Query 493  VTVDDATMIATSFPEFMDLMAGLGAKIEL 521
           VT+DDA MIATSFPEFM LM GLGAKIEL
Sbjct 421  VTIDDAAMIATSFPEFMQLMTGLGAKIEL 449

```

```

>X gi|61219221|sp|P0A2Y4|ARO_A_PSES2 3-phosphoshikimate 1-
carboxyvinyltransferase (5-enolpyruvylshikimate-3-phosphate
synthase) (EPSP synthase) (EPSPS)
gi|61219222|sp|P0A2Y5|ARO_A_CHSL 3-phosphoshikimate 1-carboxyvinyltransferase
(5-enolpyruvylshikimate-3-phosphate
synthase) (EPSP synthase) (EPSPS)
Length=449

```

Score = 697 bits (1798), Expect = 0.0, Method: Composition-based stats.  
Identities = 371/450 (82%), Positives = 403/450 (89%), Gaps = 3/450 (0%)

```

Query 73  MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 132
           M H AS +PATAR+S L+G +RIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG
Sbjct 1  MSHSASPKPATARRSEALTGEIRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60

Query 133 KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 192
           +AMQAMGA+IRKEGD WII+GVNG LL PEA LDFGNA TG RLTMGLVG YD ++FI
Sbjct 61  RAMQAMGAKIRKEGDVWIINGVNGCLLQPEAALDFGNAGTGARLTMGLVGTYDMKTSFI 120

Query 193 GDASLTKRPMGRVLNPLREMGVQVKSSEDGDRLPVTLRGPKTPTPI TYRVPMASAQVKS AV 252
           GDASL+KRPMGRVLNPLREMGVQV++ DGDR+P+TL GPKT PITYRVPMASAQVKS AV
Sbjct 121 GDASLSKRPMGRVLNPLREMGVQVEAADGDRMPLTLIGPKTANPITYRVPMASAQVKS AV 180

Query 253 LLAGLNTPGITTVIEPIIMTCDHTEKMLQGFANLTVETDADGVRTIRLEGRGKLTGQVID 312
           LLAGLNTPG+TTVIEP+MT DHTEKMLQGFGA+LTVETD DGVR IR+ G+GKL GQ ID
Sbjct 181 LLAGLNTPGVTTVIEPVMTRDHTKMLQGFADLTVETDKDGVRHIRITGQKLVGQTID 240

Query 313 VPGDPSSTAFPLVAALLVPGSDVTILNVLNPNTRTGLILTLQEMGADIEVINRLAGGED 372
           VPGDPSSTAFPLVAALLV GSDVTI NVLMNPNTRTGLILTLQEMGADIEV+N RLAGGED
Sbjct 241 VPGDPSSTAFPLVAALLVEGSDVTIRNVLNPNTRTGLILTLQEMGADIEVLNARLAGGED 300

Query 373 VADLRVRSSTLKGVTVPEDRAPPIDEYPI LA VAAAFAEGATVMNGLEELRVKESDRLSA 432
           VADLRVR+S LKGV VP +RAP MIDEYP+LA+AA+FAEG TVM+GL+ELRVKESDRL+A
Sbjct 301 VADLRVRASKLKGVVPPERAPSMIDEYPVLAIAASFAEGETVMDGLDELRVKESDRLAA 360

Query 433 VANGLKLVGDCDEGETSLVVRGRPDGKGLGNASGA AVATHLDHRIAMSFLVMGLVSEN P 492
           VA GL+ NGVDC EGE SL VRGRPDGKGLG G VATHLDHRIAMSFLVMGL +E P
Sbjct 361 VARGLEANGVDCTEGEMSLTVRGRPDGKGLG---GGTVATHLDHRIAMSFLVMGLAAEKP 417

Query 493  VTVDDATMIATSFPEFMDLMAGLGAKIELS 522
           VTVDD+ MIATSFPEFMD+M GLGAKIELS
Sbjct 418  VTVDDSNMIATSFPEFMDMMPGLGAKIELS 447

```

Blast 2o péptido

RID : 1159378658-14542-198966671817.BLASTQ4

5-enol-pyruvylshikimate-3-phosphate synthase class 2 precursor

[Agrobacterium sp. CP4]

Length=527

Score = 958 bits (2477), Expect = 0.0, Method: Composition-based stats.

Identities = 521/521 (100%), Positives = 521/521 (100%), Gaps = 0/521 (0%)

Query	1	MAQGIQTLNPNNSNFHKPQVPKSSSFLVFGSKLKNANSMLVLKKDSIFMQKFCFRISA	60
		MAQGIQTLNPNNSNFHKPQVPKSSSFLVFGSKLKNANSMLVLKKDSIFMQKFCFRISA	
Sbjct	7	MAQGIQTLNPNNSNFHKPQVPKSSSFLVFGSKLKNANSMLVLKKDSIFMQKFCFRISA	66
Query	61	SVATACMLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGE	120
		SVATACMLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGE	
Sbjct	67	SVATACMLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGE	126
Query	121	DVINTGKAMQAMGARIRKEGDTWIIDGVGNGLLAPLDFGNAATGCRLTMGLVGVYD	180
		DVINTGKAMQAMGARIRKEGDTWIIDGVGNGLLAPLDFGNAATGCRLTMGLVGVYD	
Sbjct	127	DVINTGKAMQAMGARIRKEGDTWIIDGVGNGLLAPLDFGNAATGCRLTMGLVGVYD	186
Query	181	FDSTFIGDASLTKRPMGRVLNPLREMGVQVKSEGDRLPVTLRGPKTPTPIYRVPMASA	240
		FDSTFIGDASLTKRPMGRVLNPLREMGVQVKSEGDRLPVTLRGPKTPTPIYRVPMASA	
Sbjct	187	FDSTFIGDASLTKRPMGRVLNPLREMGVQVKSEGDRLPVTLRGPKTPTPIYRVPMASA	246
Query	241	QVKSAVLLAGLNTPGITTVIEPIMTCDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKL	300
		QVKSAVLLAGLNTPGITTVIEPIMTCDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKL	
Sbjct	247	QVKSAVLLAGLNTPGITTVIEPIMTCDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKL	306
Query	301	TGQVIDVPGDPSSTAFPLVAALLVPGSDVTILNVLNMPTRTGLILTLQEMGADIEVINLR	360
		TGQVIDVPGDPSSTAFPLVAALLVPGSDVTILNVLNMPTRTGLILTLQEMGADIEVINLR	
Sbjct	307	TGQVIDVPGDPSSTAFPLVAALLVPGSDVTILNVLNMPTRTGLILTLQEMGADIEVINLR	366
Query	361	LAGGEDVADLRVRSSTLKGVTVPEDRAPPIMIDEYPI LAVAAAF AEGATVMNGLEELRVKE	420
		LAGGEDVADLRVRSSTLKGVTVPEDRAPPIMIDEYPI LAVAAAF AEGATVMNGLEELRVKE	
Sbjct	367	LAGGEDVADLRVRSSTLKGVTVPEDRAPPIMIDEYPI LAVAAAF AEGATVMNGLEELRVKE	426
Query	421	SDRLSAVANGLKLNQVDCDEGETSLVVRGRPDGKGLGNASGA AVATHLDHRIAMSFLVMG	480
		SDRLSAVANGLKLNQVDCDEGETSLVVRGRPDGKGLGNASGA AVATHLDHRIAMSFLVMG	
Sbjct	427	SDRLSAVANGLKLNQVDCDEGETSLVVRGRPDGKGLGNASGA AVATHLDHRIAMSFLVMG	486
Query	481	LVSENPVTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA	521
		LVSENPVTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA	
Sbjct	487	LVSENPVTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA	527

> [gi|18266432|gb|AAL67577.1|AF464188\\_1](#) CP4EPS [Glycine max]  
[gi|27549260|gb|AAO17037.1](#) CP4EPS protein [synthetic construct]  
Length=455

Score = 812 bits (2097), Expect = 0.0, Method: Composition-based stats.  
Identities = 452/455 (99%), Positives = 452/455 (99%), Gaps = 0/455 (0%)

Query	67	MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG	126
		MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG	
Sbjct	1	MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG	60
Query	127	KAMQAMGARIRKEGDTWIIDGVGNGLLAPLDFGNAATGCRLTMGLVGVYDFDSTFI	186
		KAMQAMGARIRKEGDTWIIDGVGNGLLAPLDFGNAATGCRLTMGLVGVYDFDSTFI	
Sbjct	61	KAMQAMGARIRKEGDTWIIDGVGNGLLAPLDFGNAATGCRLTMGLVGVYDFDSTFI	120
Query	187	GDASLTKRPMGRVLNPLREMGVQVKSEGDRLPVTLRGPKTPTPIYRVPMASAQVKS AV	246
		GDASLTKRPMGRVLNPLREMGVQVKSEGDRLPVTLRGPKTPTPIYRVPMASAQVKS AV	
Sbjct	121	GDASLTKRPMGRVLNPLREMGVQVKSEGDRLPVTLRGPKTPTPIYRVPMASAQVKS AV	180
Query	247	LLAGLNTPGITTVIEPIMTCDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID	306
		LLAGLNTPGITTVIEPIMTCDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID	
Sbjct	181	LLAGLNTPGITTVIEPIMTRDTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID	240

Query	307	VPGDPSSTAFPLVAALLVPGSDVTILNVLNPNPRTTGLILTLQEMGADIEVINRLAGGED	366
		VPGDPSSTAFPLVAALLVPGSDVTILNVLNPNPRTTGLILTLQEMGADIEVIN RLAGGED	
Sbjct	241	VPGDPSSTAFPLVAALLVPGSDVTILNVLNPNPRTTGLILTLQEMGADIEVINPRLAGGED	300
Query	367	VADLRVRSSTLKGVTVPEDRAPPIDEYPIILAVAAAFAGATVMNGLEELRVKESDRLSA	426
		VADLRVRSSTLKGVTVPEDRAP IDEYPIILAVAAAFAGATVMNGLEELRVKESDRLSA	
Sbjct	301	VADLRVRSSTLKGVTVPEDRAPSMIDEYPIILAVAAAFAGATVMNGLEELRVKESDRLSA	360
Query	427	VANGLKLNQVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP	486
		VANGLKLNQVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP	
Sbjct	361	VANGLKLNQVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP	420
Query	487	VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA	521
		VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA	
Sbjct	421	VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA	455

>X [gi|8469107|sp|Q9R4E4|AROA\\_AGRSC](#) 3-phosphoshikimate 1-carboxyvinyltransferase (5-enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS)  
[gi|114794058|pdb|2GG4|A](#) Chain A, Cp4 Epsp Synthase (Unliganded)  
[gi|114794059|pdb|2GG6|A](#) Chain A, Cp4 Epsp Synthase Liganded With S3p  
[gi|114794060|pdb|2GGA|A](#) Chain A, Cp4 Epsp Synthase Liganded With S3p And Glyphosate  
Length=455

Score = 809 bits (2090), Expect = 0.0, Method: Composition-based stats.  
Identities = 451/455 (99%), Positives = 451/455 (99%), Gaps = 0/455 (0%)

Query	67	MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG	126
		M HGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG	
Sbjct	1	MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG	60
Query	127	KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI	186
		KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI	
Sbjct	61	KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI	120
Query	187	GDASLTKRPMGRVLNPLREMGVQVKSEGDRLPVTLRGPKTPTPI TYRVPMASAQVKS AV	246
		GDASLTKRPMGRVLNPLREMGVQVKSEGDRLPVTLRGPKTPTPI TYRVPMASAQVKS AV	
Sbjct	121	GDASLTKRPMGRVLNPLREMGVQVKSEGDRLPVTLRGPKTPTPI TYRVPMASAQVKS AV	180
Query	247	LLAGLNTPGITTVIEPIMTCDHTEKMLQGFANLTVETDADGVRTIRLEGRGKLTGQVID	306
		LLAGLNTPGITTVIEPIMT DHTEKMLQGFANLTVETDADGVRTIRLEGRGKLTGQVID	
Sbjct	181	LLAGLNTPGITTVIEPIMTRDHTKMLQGFANLTVETDADGVRTIRLEGRGKLTGQVID	240
Query	307	VPGDPSSTAFPLVAALLVPGSDVTILNVLNPNPRTTGLILTLQEMGADIEVINRLAGGED	366
		VPGDPSSTAFPLVAALLVPGSDVTILNVLNPNPRTTGLILTLQEMGADIEVIN RLAGGED	
Sbjct	241	VPGDPSSTAFPLVAALLVPGSDVTILNVLNPNPRTTGLILTLQEMGADIEVINPRLAGGED	300
Query	367	VADLRVRSSTLKGVTVPEDRAPPIDEYPIILAVAAAFAGATVMNGLEELRVKESDRLSA	426
		VADLRVRSSTLKGVTVPEDRAP IDEYPIILAVAAAFAGATVMNGLEELRVKESDRLSA	
Sbjct	301	VADLRVRSSTLKGVTVPEDRAPSMIDEYPIILAVAAAFAGATVMNGLEELRVKESDRLSA	360
Query	427	VANGLKLNQVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP	486
		VANGLKLNQVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP	
Sbjct	361	VANGLKLNQVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP	420
Query	487	VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA	521
		VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA	
Sbjct	421	VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA	455

> [gi|114794061|pdb|2GGD|A](#) Chain A, Cp4 Epsp Synthase Ala100gly Liganded With S3p And Glyphosate  
Length=455

Score = 808 bits (2086), Expect = 0.0, Method: Composition-based stats.  
Identities = 450/455 (98%), Positives = 450/455 (98%), Gaps = 0/455 (0%)

Query	67	MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG	126
		M HGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG	
Sbjct	1	MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG	60
Query	127	KAMQAMGARIRKEGDTWIIDGVGNGLLAPLDFGNAATGCRLTMGLVGVYDFDSTFI	186
		KAMQAMGARIRKEGDTWIIDGVGNGLLAPLDFGNA TGCRMTMGLVGVYDFDSTFI	
Sbjct	61	KAMQAMGARIRKEGDTWIIDGVGNGLLAPLDFGNAAGTGCRLTMGLVGVYDFDSTFI	120
Query	187	GDASLTKRPMGRVLNPLREMGVQVKSEGDRLPVTLRGPKTPTPIYRVPMSAQVKSVA	246
		GDASLTKRPMGRVLNPLREMGVQVKSEGDRLPVTLRGPKTPTPIYRVPMSAQVKSVA	
Sbjct	121	GDASLTKRPMGRVLNPLREMGVQVKSEGDRLPVTLRGPKTPTPIYRVPMSAQVKSVA	180
Query	247	LLAGLNTPGITTVIEPIMTCDHTEKMLQGFANLTVETDADGVRTIRLEGRGKLTGQVID	306
		LLAGLNTPGITTVIEPIMT DHTEKMLQGFANLTVETDADGVRTIRLEGRGKLTGQVID	
Sbjct	181	LLAGLNTPGITTVIEPIMTRDTEKMLQGFANLTVETDADGVRTIRLEGRGKLTGQVID	240
Query	307	VPGDPSSTAFPLVAALLVPGSDVTILNVLNPNTRTGLLITLQEMGADIEVINRLRAGGED	366
		VPGDPSSTAFPLVAALLVPGSDVTILNVLNPNTRTGLLITLQEMGADIEVIN RLAGGED	
Sbjct	241	VPGDPSSTAFPLVAALLVPGSDVTILNVLNPNTRTGLLITLQEMGADIEVINRRLRAGGED	300
Query	367	VADLRVRSSTLKGVTVPEDRAPPIDEYPIAVAAAAFAEGATVMNGLEELRVKESDRLSA	426
		VADLRVRSSTLKGVTVPEDRAP MIDEYPIAVAAAAFAEGATVMNGLEELRVKESDRLSA	
Sbjct	301	VADLRVRSSTLKGVTVPEDRAPSMIDEYPIAVAAAAFAEGATVMNGLEELRVKESDRLSA	360
Query	427	VANGLKLNQVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP	486
		VANGLKLNQVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP	
Sbjct	361	VANGLKLNQVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP	420
Query	487	VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA	521
		VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA	
Sbjct	421	VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA	455

> [gi|15964006|ref|NP\\_384359.1](#)  3-phosphoshikimate 1-carboxyvinyltransferase  
[Sinorhizobium meliloti  
1021]

[gi|21362380|sp|Q92SV5|ARO\\_A\\_RHIME](#) 3-phosphoshikimate 1-carboxyvinyltransferase  
(5-enolpyruvylshikimate-3-phosphate  
synthase) (EPSP synthase) (EPSPS)

[gi|15073182|emb|CAC41690.1](#)  PUTATIVE 3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE PROTEIN  
[Sinorhizobium meliloti]  
Length=455

Score = 769 bits (1986), Expect = 0.0, Method: Composition-based stats.  
Identities = 410/455 (90%), Positives = 428/455 (94%), Gaps = 0/455 (0%)

Query	67	MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG	126
		M HG++ RPATARKSS L GT+RIPGDKSISHRSFMFGGLA+GETRITGLLEGEDVINTG	
Sbjct	1	MSHGNSRNPATARKSSDLKGTLRIPGDKSISHRSFMFGGLAAGETRITGLLEGEDVINTG	60
Query	127	KAMQAMGARIRKEGDTWIIDGVGNGLLAPLDFGNAATGCRLTMGLVGVYDFDSTFI	186
		KAMQAMGARIRKEGDTWIIDGVGNGLLAPLDFGNA TGCRMTMGLVGVYDFDSTFI	
Sbjct	61	KAMQAMGARIRKEGDTWIIDGVGNGLLAPLDFGNAAGTGCRLTMGLVGVYDFDSTFI	120

Query	187	GDASLTKRPMGRVLNPLREMGVQVKSEGDRLPVTLRGPKTPPTPITYRVPMASAQVKS AV	246
Sbjct	121	GDASLTKRPMGRVL+PLREMGVQVKS +GDRLPVTLRGPKTP PITYRVPMASAQVKS AV	180
Query	247	LLAGLNTPGITTVIEPIMTCDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID	306
Sbjct	181	LLAGLNTPGITTV+EP+MT DHTEKMLQGFGANLTVETDA+GVRTIRLEGRGKLTGQVID	240
Query	307	VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINRLLAGGED	366
Sbjct	241	VPGDPSSTAFPLVA L+VPGSD+TILNVLMNPTRTGLILTLQEMGA+IEV+N RLAGGED	300
Query	367	VADLRVRSSTLKGVTVPEDRAPMIDEYPI LAVAAAAFAEGATVMNGLEELRVKESDRLSA	426
Sbjct	301	VADLRVR S LKGVTVPEDRAP MIDEYP+LAVAAAAFAEG TVMNGLEELRVKESDRLSA	360
Query	427	VANGLKLNQVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP	486
Sbjct	361	VA+GLKLNQVDCDEGE SLVVRGRP GKGLG SG V THLDHRIAMSFLVMGL SE+P	420
Query	487	VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 521	
Sbjct	421	VTVDDATMIATSFPEFM LM GLGAKIE ++ KAA 455	

> [gi|113871287|ref|ZP\\_01411416.1|](#) 3-phosphoshikimate 1-carboxyvinyltransferase  
 [Sinorhizobium medicae  
 WSM419]  
[gi|113725959|gb|EAU07051.1|](#) 3-phosphoshikimate 1-carboxyvinyltransferase  
 [Sinorhizobium medicae  
 WSM419]  
 Length=456

Score = 743 bits (1919), Expect = 0.0, Method: Composition-based stats.  
 Identities = 407/455 (89%), Positives = 425/455 (93%), Gaps = 0/455 (0%)

Query	67	MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG	126
Sbjct	1	M HG++ RPATARKSS L GTVRIIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG	60
Query	127	KAMQAMGARIRKEGDTWIIDGVGNGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI	186
Sbjct	61	+AMQAMGA+IRKEGDTWIIDGVGNG LLAPEAPLDFGNA TGCRLTMGLVGVYDFDSTFI	120
Query	187	GDASLTKRPMGRVLNPLREMGVQVKSEGDRLPVTLRGPKTPPTPITYRVPMASAQVKS AV	246
Sbjct	121	GDASLTKRPMGRVL+PLREMGVQVKS +GDRLPVTLRGPKTP PITYRVPMASAQVKS AV	180
Query	247	LLAGLNTPGITTVIEPIMTCDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID	306
Sbjct	181	LLAGLNTPGITTV+EP+MT DHTEKMLQGFGLA+L+VETD DGVRTIRLEGRGKLTGQVID	240
Query	307	VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINRLLAGGED	366
Sbjct	241	VPGDPSSTAFPLVAALLVPGSD++I NVLMNPTRTGLILTLQEMGA IEV++ RLAGGED	300
Query	367	VADLRVRSSTLKGVTVPEDRAPMIDEYPI LAVAAAAFAEGATVMNGLEELRVKESDRLSA	426
Sbjct	301	VADLRVR S LKGVTVPE+RAP MIDEYP+LAVAAAAFAEGATVMNGLEELRVKESDRLSA	360
Query	427	VANGLKLNQVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP	486

```

                VA GLKLNQVDCDEGE SLVVRGRP GKGLGN SG V THLDHRIAMSFLVMGL SE P
Sbjct 361 VAEGGLKLNQVDCDEGEASLVVRGRPGGKGLGNDSSGGQVKTHLDHRIAMSFLVMGLASERP 420

Query 487 VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 521
                VTVDDATMIATSFPEFM LM GLGAKIE ++ KAA
Sbjct 421 VTVDDATMIATSFPEFMGLMTGLGAKIEETENKAA 455

```

> [gi|115254522|emb|CAK05596.1](#) putative 3-phosphoshikimate 1-carboxyvinyltransferase (5-enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) [Rhizobium leguminosarum bv. viciae 3841]  
Length=452

Score = 732 bits (1889), Expect = 0.0, Method: Composition-based stats.  
Identities = 387/449 (86%), Positives = 416/449 (92%), Gaps = 0/449 (0%)

```

Query 67 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 126
                ML+G++S+PATARKS+GL+G+VRIPGDKSISHRSFM GGLASGETRITGLLEGEDVINTG
Sbjct 1 MLNGSASKPATARKSAGLTGSVRIPGDKSISHRSFMIGGLASGETRITGLLEGEDVINTG 60

Query 127 KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 186
                +AMQAMGARIRKEG W+I+G GNG LLAP+APLDFGNA TG RLTMGLVG YDF STFI
Sbjct 61 RAMQAMGARIRKEGAQWVIEGTGNGALLAPDAPLDFGNAGTGVRLTMGLVGTVDYDFHSTFI 120

Query 187 GDASLTKRPMGRVLNPLREMGVQVKSEGDRLPVTLRGPKTPTPIYRVPMASAQVKSASV 246
                GDASL+KRPMGRVLNPLREMGVQV + +GDRLPVTLRGP TP+PI YRVPMASAQVKSASV
Sbjct 121 GDASLSKRPMGRVLNPLREMGVQVSASEGDRLPVTLRGPGTSPPIRVPMASAQVKSASV 180

Query 247 LLAGLNTPGITTVIEPIIMTCDHTEKMLQGFANLTVETDADGVRTIRLEGRGKLTGQVID 306
                LLAGLNTPG+TTVIEP+MT DHTEKMLQGFGA L+VETD DGVRTIRLEGRGKL GQVID
Sbjct 181 LLAGLNTPGVTTVIEPVMTRDHTKMLQGFGAALSVEVDGQVVRTIRLEGRGKLAGQVID 240

Query 307 VPGDPSSTAFPLVAALLVPGSDVTILNVLNPNTRTGLLITLQEMGADIEVINRLLAGGED 366
                VPGDPSSTAFPLVAAL+VPGSD+TI+NVLNPNTRTGLLITLQEMGADIEV+N RLAGGED
Sbjct 241 VPGDPSSTAFPLVAALIVPGSDITIVNVLNPNTRTGLLITLQEMGADIEVVNARLAGGED 300

Query 367 VADLRVRSSTLKGVTVPEDRAPPIMIDEYPILAVAAFAEGATVMNGLEELRVKESDRLSA 426
                VADLRVR S LKGVTVPEDRAP MIDEYPILAVAA FAEGATVM GLEELRVKESDRLSA
Sbjct 301 VADLRVRHSELKGVTVPEDRAPSMIDEYPILAVAAFAEGATVMKLEELRVKESDRLSA 360

Query 427 VANGLKLNQVDCDEGETSLVVRGRPDKGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 486
                VA+GLKLNQVDCDEGE L+VRGRPDKGKGLGNA+ V+THLDHRIAMSFLV+GL SE+
Sbjct 361 VADGLKLNQVDCDEGEDFLIVRGRPDKGKGLGNAADGRVSTHLDHRIAMSFLVLGLASEHA 420

Query 487 VTVDDATMIATSFPEFMDLMAGLGAKIEL 515
                VT+DDA MIATSFPEFM LM GLGAKIEL
Sbjct 421 VTIDDAAMIATSFPEFMQLMTGLGAKIEL 449

```

> [gi|61219221|sp|P0A2Y4|AROA\\_PSES2](#) 3-phosphoshikimate 1-carboxyvinyltransferase (5-enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS)  
[gi|61219222|sp|P0A2Y5|AROA\\_ACHSL](#) 3-phosphoshikimate 1-carboxyvinyltransferase (5-enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS)  
Length=449

Score = 696 bits (1796), Expect = 0.0, Method: Composition-based stats.  
Identities = 371/450 (82%), Positives = 403/450 (89%), Gaps = 3/450 (0%)

Query	67	MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG	126
		M H AS +PATAR+S L+G +RIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG	
Sbjct	1	MHSASPKPATARRSEALTGEIRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG	60
Query	127	KAMQAMGARIRKEGDTWIIDGVGNGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI	186
		+AMQAMGA+IRKEGD WII+GVGNG LL PEA LDFGNA TG RLTMGLVG YD ++FI	
Sbjct	61	RAMQAMGAKIRKEGDVWIINGVNGCLLQPEAALDFGNAAGTGARLTMGLVGTYDMKTSFI	120
Query	187	GDASLTKRPMGRVLNPLREMGVQVKSEGDRLPVTLRGPKTPTPITYRVPMASAQVKS AV	246
		GDASL+KRPMGRVLNPLREMGVQV++ DGDR+P+TL GKPT PITYRVPMASAQVKS AV	
Sbjct	121	GDASLSKRPMGRVLNPLREMGVQVEAADGDRMPLTLIGPKTANPITYRVPMASAQVKS AV	180
Query	247	LLAGLNTPGITTVIEPIMTCDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID	306
		LLAGLNTPG+TTVIEP+MT DHTEKMLQGFGA+LTVETD DGVR IR+ G+GKL GQ ID	
Sbjct	181	LLAGLNTPGVTTVIEPVMTRDHTEKMLQGFADLTVETDKDGVRRHIRTGQGLVGTQID	240
Query	307	VPGDPSSTAFPLVAALLVPGSDVTILNVLNMPTRTGLLITLQEMGADIEVINRLAGGED	366
		VPGDPSSTAFPLVAALLV GSDVTI NVLMNPTRTGLLITLQEMGADIEV+N RLAGGED	
Sbjct	241	VPGDPSSTAFPLVAALLVEGSDVTRNVLNMPTRTGLLITLQEMGADIEVLNARLAGGED	300
Query	367	VADLRVRSSTLKGVTVPEDRAPMIDEYPI LAVAAAFAEGATVMNGLEELRVKESDRLSA	426
		VADLRVR+S LKGV VP +RAP MIDEYP+LA+AA+FAEG TVM+GL+ELRVKESDRL+A	
Sbjct	301	VADLRVRASKLKGVVPPERAPSMIDEYPVLAIAASFAEGETVMDGLDELRVKESDRLAA	360
Query	427	VANGLKLNQVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSEN P	486
		VA GL+ NGVDC EGE SL VRGRPDGKGLG G VATHLDHRIAMSFLVMGL +E P	
Sbjct	361	VARGLEANGVDCTEGEMSLTVRGRPDGKGLG---GGTVATHLDHRIAMSFLVMGLAAEKP	417
Query	487	VTVDDATMIATSFPEFMDLMAGLGAKIELS	516
		VTVDD+ MIATSFPEFMD+M GLGAKIELS	
Sbjct	418	VTVDDSNMIATSFPEFMDMMPGLGAKIELS	447

Blast péptido 3

RID: 1159378878-19785-145962769428.BLASTQ4

5-enol-pyruvylshikimate-3-phospate synthase class 2 precursor  
[Agrobacterium sp. CP4]  
Length=527

Score = 881 bits (2277), Expect = 0.0, Method: Composition-based stats.  
Identities = 482/482 (100%), Positives = 482/482 (100%), Gaps = 0/482 (0%)

Query	1	MLVLKKSIFMQKFCFRISASVATACMLHGASSRPATARKSSGLSGTVRIPGDKSISHR	60
		MLVLKKSIFMQKFCFRISASVATACMLHGASSRPATARKSSGLSGTVRIPGDKSISHR	
Sbjct	46	MLVLKKSIFMQKFCFRISASVATACMLHGASSRPATARKSSGLSGTVRIPGDKSISHR	105
Query	61	SFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDGVGNGLLAPEAP	120
		SFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDGVGNGLLAPEAP	
Sbjct	106	SFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDGVGNGLLAPEAP	165
Query	121	LDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRVLNPLREMGVQVKSEGDRLP	180
		LDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRVLNPLREMGVQVKSEGDRLP	
Sbjct	166	LDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRVLNPLREMGVQVKSEGDRLP	225
Query	181	VTLRGPKTPTPITYRVPMASAQVKS AVLLAGLNTPGITTVIEPIMTCDHTEKMLQGFAN	240
		VTLRGPKTPTPITYRVPMASAQVKS AVLLAGLNTPGITTVIEPIMTCDHTEKMLQGFAN	
Sbjct	226	VTLRGPKTPTPITYRVPMASAQVKS AVLLAGLNTPGITTVIEPIMTCDHTEKMLQGFAN	285
Query	241	LTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVAALLVPGSDVTILNVLNMP	300

```

Sbjct 286 LTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVAALLVPGSDVTILNVL MNPT 345
Query 301 RTGLILTLQEMGADIEVINLRLAGGEDVADLRVRSSTLKGVTVPEDRAPP MIDEYPILAV 360
Sbjct 346 RTGLILTLQEMGADIEVINLRLAGGEDVADLRVRSSTLKGVTVPEDRAPP MIDEYPILAV 405
Query 361 AAFAEGATVMNGLEELRVKESDRLSAVANGLKLN GVDCDEGETSLVVRGRPDGKGLGNA 420
Sbjct 406 AAFAEGATVMNGLEELRVKESDRLSAVANGLKLN GVDCDEGETSLVVRGRPDGKGLGNA 465
Query 421 SGAAVATHLDHRIAMSFLVMGLVSEN PVTVDATMIATSFPEFMDLMAGLGAKIELSDTK 480
Sbjct 466 SGAAVATHLDHRIAMSFLVMGLVSEN PVTVDATMIATSFPEFMDLMAGLGAKIELSDTK 525
Query 481 AA 482
Sbjct 526 AA 527

```

```

> X gi|18266432|gb|AAL67577.1|AF464188\_1 CP4EPSPS [Glycine max]
gi|27549260|gb|AA017037.1 CP4EPSPS protein [synthetic construct]
Length=455

```

Score = 805 bits (2080), Expect = 0.0, Method: Composition-based stats.  
Identities = 452/455 (99%), Positives = 452/455 (99%), Gaps = 0/455 (0%)

```

Query 28 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 87
Sbjct 1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
Query 88 KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 147
Sbjct 61 KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
Query 148 GDASLTKRPMGRVLNPLREMGVQVKSEGDRLPVTLRGPKTPTPI TYRVPMASAQVKS AV 207
Sbjct 121 GDASLTKRPMGRVLNPLREMGVQVKSEGDRLPVTLRGPKTPTPI TYRVPMASAQVKS AV 180
Query 208 LLAGLNTPGITTVIEPIMTCDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 267
Sbjct 181 LLAGLNTPGITTVIEPIMTRDTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
Query 268 VPGDPSSTAFPLVAALLVPGSDVTILNVL MNPTRTGLILTLQEMGADIEVINLRLAGGED 327
Sbjct 241 VPGDPSSTAFPLVAALLVPGSDVTILNVL MNPTRTGLILTLQEMGADIEVINR LAGGED 300
Query 328 VADLRVRSSTLKGVTVPEDRAPP MIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 387
Sbjct 301 VADLRVRSSTLKGVTVPEDRAP MIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
Query 388 VANGLKLN GVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSEN P 447
Sbjct 361 VANGLKLN GVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSEN P 420
Query 448 VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 482
Sbjct 421 VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 455

```

> [gi|8469107|sp|Q9R4E4|AROA\\_AGRSC](#) 3-phosphoshikimate 1-carboxyvinyltransferase  
 (5-enolpyruvylshikimate-3-phosphate  
 synthase) (EPSP synthase) (EPSPS)  
[gi|114794058|pdb|2GG4|A](#) Chain A, Cp4 Epsp Synthase (Unliganded)  
[gi|114794059|pdb|2GG6|A](#) Chain A, Cp4 Epsp Synthase Liganded With S3p  
[gi|114794060|pdb|2GGA|A](#) Chain A, Cp4 Epsp Synthase Liganded With S3p And  
 Glyphosate  
 Length=455

Score = 803 bits (2073), Expect = 0.0, Method: Composition-based stats.  
 Identities = 451/455 (99%), Positives = 451/455 (99%), Gaps = 0/455 (0%)

Query	28	MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG	87
		M HGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG	
Sbjct	1	M SHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG	60
Query	88	KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI	147
		KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI	
Sbjct	61	KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI	120
Query	148	GDASLTKRPMGRVLNPLREMGVQVKSEGDRLPVTLRGPKTPTPI TYRVPMASAQVKS AV	207
		GDASLTKRPMGRVLNPLREMGVQVKSEGDRLPVTLRGPKTPTPI TYRVPMASAQVKS AV	
Sbjct	121	GDASLTKRPMGRVLNPLREMGVQVKSEGDRLPVTLRGPKTPTPI TYRVPMASAQVKS AV	180
Query	208	LLAGLNTPGITTVIEPIMTCDHTEKMLQGFANLTVETDADGVRTIRLEGRGKLTGQVID	267
		LLAGLNTPGITTVIEPIMT DHTEKMLQGFANLTVETDADGVRTIRLEGRGKLTGQVID	
Sbjct	181	LLAGLNTPGITTVIEPIMTRDHTKMLQGFANLTVETDADGVRTIRLEGRGKLTGQVID	240
Query	268	VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINRLAGGED	327
		VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVIN RLAGGED	
Sbjct	241	VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED	300
Query	328	VADLRVRSSTLKGVTVPEDRAPP MIDEYPI LAVAAAFAEGATVMNGLEELRVKESDRLSA	387
		VADLRVRSSTLKGVTVPEDRAP MIDEYPI LAVAAAFAEGATVMNGLEELRVKESDRLSA	
Sbjct	301	VADLRVRSSTLKGVTVPEDRAPSMIDEYPI LAVAAAFAEGATVMNGLEELRVKESDRLSA	360
Query	388	VANGLKLNQVDCDEGETSLVVRGRPDGKGLGNASGA AVATHLDHRIAMSFLVMGLVSEN P	447
		VANGLKLNQVDCDEGETSLVVRGRPDGKGLGNASGA AVATHLDHRIAMSFLVMGLVSEN P	
Sbjct	361	VANGLKLNQVDCDEGETSLVVRGRPDGKGLGNASGA AVATHLDHRIAMSFLVMGLVSEN P	420
Query	448	VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA	482
		VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA	
Sbjct	421	VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA	455

> [gi|114794061|pdb|2GGD|A](#) Chain A, Cp4 Epsp Synthase Ala100gly Liganded With  
 S3p And Glyphosate  
 Length=455

Score = 801 bits (2069), Expect = 0.0, Method: Composition-based stats.  
 Identities = 450/455 (98%), Positives = 450/455 (98%), Gaps = 0/455 (0%)

Query	28	MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG	87
		M HGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG	
Sbjct	1	M SHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG	60
Query	88	KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI	147
		KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNA TGCR LTMGLVGVYDFDSTFI	
Sbjct	61	KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAGTGCRLTMGLVGVYDFDSTFI	120
Query	148	GDASLTKRPMGRVLNPLREMGVQVKSEGDRLPVTLRGPKTPTPI TYRVPMASAQVKS AV	207

		GDASLTKRPMGRVNLNPLREMGVQVKSEGDRLPVTLRGPKTPTPITYRVPMASAQVKS AV	
Sbjct	121	GDASLTKRPMGRVNLNPLREMGVQVKSEGDRLPVTLRGPKTPTPITYRVPMASAQVKS AV	180
Query	208	LLAGLNTPGITTVIEPIMTCDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID	267
		LLAGLNTPGITTVIEPIMT DHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID	
Sbjct	181	LLAGLNTPGITTVIEPIMTRDHTKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID	240
Query	268	VPGDPSSTAFPLVAALLVPGSDVTILNVLNPNPRTGLILTLQEMGADIEVINRLR LAGGED	327
		VPGDPSSTAFPLVAALLVPGSDVTILNVLNPNPRTGLILTLQEMGADIEVIN RLAGGED	
Sbjct	241	VPGDPSSTAFPLVAALLVPGSDVTILNVLNPNPRTGLILTLQEMGADIEVINPRLAGGED	300
Query	328	VADLRVRSSTLKGVTVPEDRAPP MIDEY P ILAVAAAFAEGATVMNGLEELRVKESDRLSA	387
		VADLRVRSSTLKGVTVPEDRAP MIDEY P ILAVAAAFAEGATVMNGLEELRVKESDRLSA	
Sbjct	301	VADLRVRSSTLKGVTVPEDRAPSMIDEY P ILAVAAAFAEGATVMNGLEELRVKESDRLSA	360
Query	388	VANGLKLNQVDCDEGETSLVVRGRPDGKGLGNASGA AVATHLDHRIAMSFLVMGLVSENP	447
		VANGLKLNQVDCDEGETSLVVRGRPDGKGLGNASGA AVATHLDHRIAMSFLVMGLVSENP	
Sbjct	361	VANGLKLNQVDCDEGETSLVVRGRPDGKGLGNASGA AVATHLDHRIAMSFLVMGLVSENP	420
Query	448	VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA	482
		VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA	
Sbjct	421	VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA	455

> [gi|15964006|ref|NP\\_384359.1](#)  3-phosphoshikimate 1-carboxyvinyltransferase  
 [Sinorhizobium meliloti  
 1021]  
[gi|21362380|sp|Q92SV5|AROA\\_RHIME](#) 3-phosphoshikimate 1-carboxyvinyltransferase  
 (5-enolpyruvylshikimate-3-phosphate  
 synthase) (EPS synthase) (EPSPS)  
[gi|15073182|emb|CAC41690.1](#)  PUTATIVE 3-PHOSPHOSHIKIMATE 1-  
 CARBOXYVINYLTRANSFERASE PROTEIN  
 [Sinorhizobium meliloti]  
 Length=455

Score = 763 bits (1971), Expect = 0.0, Method: Composition-based stats.  
 Identities = 410/455 (90%), Positives = 428/455 (94%), Gaps = 0/455 (0%)

Query	28	MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG	87
		M HG++ RPATARKSS L GT+RIPGDKSISHRSFMFGGLA+GETRITGLLEGEDVINTG	
Sbjct	1	MSHGSNRPATARKSSDLKGTLRIPGDKSISHRSFMFGGLAAGETRITGLLEGEDVINTG	60
Query	88	KAMQAMGARIRKEGDTWIIDGVNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI	147
		KAMQAMGARIRKEGDTWIIDGVNG LLAPEAPLDFGNA TGCRLTMGLVGVYDFDSTFI	
Sbjct	61	KAMQAMGARIRKEGDTWIIDGVNGALLAPEAPLDFGNAGTGCRLTMGLVGVYDFDSTFI	120
Query	148	GDASLTKRPMGRVNLNPLREMGVQVKSEGDRLPVTLRGPKTPTPITYRVPMASAQVKS AV	207
		GDASLTKRPMGRVL+PLREMGVQVKS +GDRLPVTLRGPKT PITYRVPMASAQVKS AV	
Sbjct	121	GDASLTKRPMGRVLDPLREMGVQVKS AEGDRLPVTLRGPKT PNPITYRVPMASAQVKS AV	180
Query	208	LLAGLNTPGITTVIEPIMTCDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID	267
		LLAGLNTPGITTV+EP+MT DHTEKMLQGFGANLTVETA+GVRTIRLEGRGKLTGQVID	
Sbjct	181	LLAGLNTPGITTVIEPVMTRDHTKMLQGFGANLTVETAEGVRTIRLEGRGKLTGQVID	240
Query	268	VPGDPSSTAFPLVAALLVPGSDVTILNVLNPNPRTGLILTLQEMGADIEVINRLR LAGGED	327
		VPGDPSSTAFPLVA L+VPGSD+TILNVLNPNPRTGLILTLQEMGA+IEV+N RLAGGED	
Sbjct	241	VPGDPSSTAFPLVAGLIVPGSDITILNVLNPNPRTGLILTLQEMGANIEVMNKR LAGGED	300
Query	328	VADLRVRSSTLKGVTVPEDRAPP MIDEY P ILAVAAAFAEGATVMNGLEELRVKESDRLSA	387
		VADLRVR S LKGVTVPEDRAP MIDEY P +LAVAAAFAEG TVMNGLEELRVKESDRLSA	
Sbjct	301	VADLRVRHSELKGVTVPE DRAPS MIDEY PVLAVAAAFAEGTTVMNGLEELRVKESDRLSA	360

```

Query 388  VANGLKLNQVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 447
          VA+GLKLNQVDCDEGE SLVVRGRP GKGLG SG V THLDHRIAMSFLVMGL SE+P
Sbjct 361  VADGLKLNQVDCDEGEASLVVRGRPGGKGLGKISGGQVKTHLDHRIAMSFLVMGLASEHP 420

Query 448  VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 482
          VTVDDATMIATSFPEFM LM GLGAKIE ++ KAA
Sbjct 421  VTVDDATMIATSFPEFMGLMTGLGAKIEEAENKAA 455

```

```

>X gi|113871287|ref|ZP\_01411416.1 3-phosphoshikimate 1-carboxyvinyltransferase
[Sinorhizobium medicae
WSM419]
gi|113725959|gb|EAU07051.1 3-phosphoshikimate 1-carboxyvinyltransferase
[Sinorhizobium medicae
WSM419]
Length=456

```

Score = 737 bits (1903), Expect = 0.0, Method: Composition-based stats.  
Identities = 407/455 (89%), Positives = 425/455 (93%), Gaps = 0/455 (0%)

```

Query 28  MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 87
          M HG++ RPATARKSS L GTVRIIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG
Sbjct 1  MSHGSNPRPATARKSSDLKGTVRIIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60

Query 88  KAMQAMGARIRKEGDTWIIDGVGNGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 147
          +AMQAMGA+IRKEGDTWIIDGVGNG LLAPEAPLDFGNA TGCRLTMGLVGVYDFDSTFI
Sbjct 61  RAMQAMGAKIRKEGDTWIIDGVGNGALLAPEAPLDFGNAGTGCRLTMGLVGVYDFDSTFI 120

Query 148 GDASLTKRPMGRVNLPLREMGVQVKSSEDGDRLPVTLRGPKTPTPIYRVPMASAQVKS AV 207
          GDASLTKRPMGRVL+PLREMGVQVKS +GDRLPVTLRGPKTP PITYRVPMASAQVKS AV
Sbjct 121 GDASLTKRPMGRVLDPLREMGVQVKS AEGDRLPVTLRGPKTPNPITYRVPMASAQVKS AV 180

Query 208 LLAGLNTPGITTVIEPIIMTCDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 267
          LLAGLNTPGITTV+EP+MT DHTEKMLQGFGA+L+VETD DGVRTIRLEGRGKL GQVID
Sbjct 181 LLAGLNTPGITTVVEPVMTRDHTEKMLQGFADLSVETDRDGVRTIRLEGRGKLRGQVID 240

Query 268 VPGDPSSTAFPLVAALLVPGSDVTILNVLNPNPRTGLLITLQEMGADIEVINLR LAGGED 327
          VPGDPSSTAFPLVAALLVPGSD++I NVLMNPRTGLLITLQEMGA IEV++ RLAGGED
Sbjct 241 VPGDPSSTAFPLVAALLVPGSDLSIFNVLMNPRTGLLITLQEMGARIEVLSSRLAGGED 300

Query 328 VADLRVRSSTLKGVTVPEDRAPP MIDEYPI LAVAAAFAEGATVMNGLEELRVKESDR LSA 387
          VADLRVR S LKGVTVPE+RAP MIDEY+LAVAAAFAEGATVMNGLEELRVKESDR LSA
Sbjct 301 VADLRVRYSELKGVTVPEERAPSMIDEY PVLAVAAAFAEGATVMNGLEELRVKESDR LSA 360

Query 388  VANGLKLNQVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 447
          VA GLKLNQVDCDEGE SLVVRGRP GKGLGN SG V THLDHRIAMSFLVMGL SE P
Sbjct 361  VA EGLKLNQVDCDEGEASLVVRGRPGGKGLGNDSSGGQVKTHLDHRIAMSFLVMGLASERP 420

Query 448  VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 482
          VTVDDATMIATSFPEFM LM GLGAKIE ++ KAA
Sbjct 421  VTVDDATMIATSFPEFMGLMTGLGAKIEETENKAA 455

```

```

>X gi|115254522|emb|CAK05596.1 putative 3-phosphoshikimate 1-
carboxyvinyltransferase (5-enolpyruvylshikimate-3-phosphate
synthase) (EPSP synthase) [Rhizobium
leguminosarum bv. viciae 3841]
Length=452

```

Score = 725 bits (1872), Expect = 0.0, Method: Composition-based stats.  
Identities = 387/449 (86%), Positives = 416/449 (92%), Gaps = 0/449 (0%)

Query	28	MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG	87
Sbjct	1	MLNGSASKPATARKSAGLTGSVRIPGDKSISHRSFM GGLASGETRITGLLEGEDVINTG	60
Query	88	KAMQAMGARIRKEGDTWIIDGVGNGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI	147
Sbjct	61	+AMQAMGARIRKEG W+I+G GNG LLAP+APLDFGNA TG RLTMGLVG YDF STFI	120
Query	148	GDASLTKRPMGRVLNPLREMGVQVKSEGDRLPVTLRGPKTPTPIYRVPMASAQVKS	207
Sbjct	121	GDASLTKRPMGRVLNPLREMGVQV + +GDRLPVTLRGP TP+PI YRVPMASAQVKS	180
Query	208	LLAGLNTPGITTVIEPIMTCDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID	267
Sbjct	181	LLAGLNTPG+TTVIEP+MT DHTEKMLQGFGA L+VETD DGVRTIRLEGRGKL GQVID	240
Query	268	VPGDPSSTAFPLVAALLVPGSDVTILNVLNPNPRTGLILTLQEMGADIEVINRLAGGED	327
Sbjct	241	VPGDPSSTAFPLVAAL+VPGSD+TI+NVLNPNPRTGLILTLQEMGADIEV+N RLAGGED	300
Query	328	VADLRVRSSTLKGVTVPEDRAPPIDEYPIAVAAAFAEGATVMNGLEELRVKESDRLSA	387
Sbjct	301	VADLRVRHSELKGVTVPEDRAPSMIDEYPIAVAAACFAEGATVMKGLEELRVKESDRLSA	360
Query	388	VANGLKLNQVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSEN	447
Sbjct	361	VA+GLKLNQVDCDEGE L+VRGRPDGKGLGNA+ V+THLDHRIAMSFLV+GL SE+	420
Query	448	VTVDDATMIATSFPEFMDLMAGLGAKIEL	476
Sbjct	421	VT+DDA MIATSFPEFM LM GLGAKIEL	449

> [gi|61219221|sp|P0A2Y4|AROA\\_PSES2](#) 3-phosphoshikimate 1-carboxyvinyltransferase (5-enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS)  
[gi|61219222|sp|P0A2Y5|AROA\\_ACHSL](#) 3-phosphoshikimate 1-carboxyvinyltransferase (5-enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS)  
Length=449

Score = 690 bits (1781), Expect = 0.0, Method: Composition-based stats.  
Identities = 371/450 (82%), Positives = 403/450 (89%), Gaps = 3/450 (0%)

Query	28	MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG	87
Sbjct	1	M H AS +PATAR+S L+G +RIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG	60
Query	88	KAMQAMGARIRKEGDTWIIDGVGNGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI	147
Sbjct	61	+AMQAMGA+IRKEGD WII+GVGNG LL PEA LDFGNA TG RLTMGLVG YD ++FI	120
Query	148	GDASLTKRPMGRVLNPLREMGVQVKSEGDRLPVTLRGPKTPTPIYRVPMASAQVKS	207
Sbjct	121	GDASLTKRPMGRVLNPLREMGVQV++ DGDR+P+TL GKPT PITYRVPMASAQVKS	180
Query	208	LLAGLNTPGITTVIEPIMTCDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID	267
Sbjct	181	LLAGLNTPG+TTVIEP+MT DHTEKMLQGFGA+LTVETD DGVR IR+ G+GKL GQ ID	240
Query	268	VPGDPSSTAFPLVAALLVPGSDVTILNVLNPNPRTGLILTLQEMGADIEVINRLAGGED	327

```

Sbjct 241 VPGDPSSTAFPLVAALLV GSDVTI NVLMNPTRTGLILTLQEMGADIEV+N RLAGGED 300
          VPGDPSSTAFPLVAALLVEGSDVTIRNVLMNPTRTGLILTLQEMGADIEVLNARLAGGED
Query 328 VADLRVRSSTLKGVTVPEDRAPPIDEYPIILAVAAFAEGATVMNGLEELRVKESDRLSA 387
          VADLRVR+S LKGV VP +RAP MIDEYP+LA+AA+FAEG TVM+GL+ELRVKESDRL+A
Sbjct 301 VADLRVRASKLKGVVPPERAPSMIDEYPLVLAIAASFAEGETVMDGLDELRVKESDRLAA 360
          VADLRVRASKLKGVVPPERAPSMIDEYPLVLAIAASFAEGETVMDGLDELRVKESDRLAA
Query 388 VANGLKLNQVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 447
          VA GL+ NGVDC EGE SL VRGRPDGKGLG G VATHLDHRIAMSFLVMGL +E P
Sbjct 361 VARGLEANGVDCTEGEMSLTVRGRPDGKGLG---GGTVATHLDHRIAMSFLVMGLAAEKP 417
          VARGLEANGVDCTEGEMSLTVRGRPDGKGLG---GGTVATHLDHRIAMSFLVMGLAAEKP
Query 448 VTVDDATMIATSFPEFMDLMAGLGAKIELS 477
          VTVDD+ MIATSFPEFMD+M GLGAKIELS
Sbjct 418 VTVDDSNMIATSFPEFMDMMPGLGAKIELS 447
          VTVDDSNMIATSFPEFMDMMPGLGAKIELS

```

## Blast péptido 4

RID: 1159379249-3542-20574986494.BLASTQ2

Length=452

Score = 723 bits (1865), Expect = 0.0, Method: Composition-based stats.  
Identities = 387/449 (86%), Positives = 416/449 (92%), Gaps = 0/449 (0%)

```

Query 18 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 77
          ML+G++S+PATARKS+GL+G+VRIPGDKSISHRSFM GGLASGETRITGLLEGEDVINTG
Sbjct 1 MLNGSASKPATARKSAGLTGSVRIPGDKSISHRSFMIGGLASGETRITGLLEGEDVINTG 60
          MLNGSASKPATARKSAGLTGSVRIPGDKSISHRSFMIGGLASGETRITGLLEGEDVINTG
Query 78 KAMQAMGARIRKEGDTWIIDGVNGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 137
          +AMQAMGARIRKEG W+I+G GNG LLAP+APLDFGNA TG RLTMGLVG YDF STFI
Sbjct 61 RAMQAMGARIRKEGAQWVIEGTGNGALLAPDAPLDFGNAAGTGVRLTMGLVGTDFHSTFI 120
          RAMQAMGARIRKEGAQWVIEGTGNGALLAPDAPLDFGNAAGTGVRLTMGLVGTDFHSTFI
Query 138 GDASLTKRPMGRVLNPLREMGVQVKSSEDGDRLPVTLRGPKTPPTPIYRVPMASAQVKSASV 197
          GDASL+KRPMGRVLNPLREMGVQV + +GDRLPVTLRGP TP+PI YRVPMASAQVKSASV
Sbjct 121 GDASLSKRPMGRVLNPLREMGVQVSASEGDRLPVTLRGPSTPSPIRYRVPMASAQVKSASV 180
          GDASLSKRPMGRVLNPLREMGVQVSASEGDRLPVTLRGPSTPSPIRYRVPMASAQVKSASV
Query 198 LLAGLNTPGITTVIEPIMTCDHTEKMLQGFANLTVETDADGVRTIRLEGRGKLTGQVID 257
          LLAGLNTPG+TTVIEP+MT DHTEKMLQGFGA L+VETD DGVRTIRLEGRGKL GQVID
Sbjct 181 LLAGLNTPGVTTVIEPVMTRDHTKMLQGFGAALSVETDGDGVRTIRLEGRGKLAGQVID 240
          LLAGLNTPGVTTVIEPVMTRDHTKMLQGFGAALSVETDGDGVRTIRLEGRGKLAGQVID
Query 258 VPGDPSSTAFPLVAALLVPGSDVTIILNVLMNPTRTGLILTLQEMGADIEVINRLRAGGED 317
          VPGDPSSTAFPLVAAL+VPGSD+TI+NVLMNPTRTGLILTLQEMGADIEV+N RLAGGED
Sbjct 241 VPGDPSSTAFPLVAALIVPGSDITIVNVLMNPTRTGLILTLQEMGADIEVVNARLAGGED 300
          VPGDPSSTAFPLVAALIVPGSDITIVNVLMNPTRTGLILTLQEMGADIEVVNARLAGGED
Query 318 VADLRVRSSTLKGVTVPEDRAPPIDEYPIILAVAAFAEGATVMNGLEELRVKESDRLSA 377
          VADLRVR S LKGVTVPEDRAP MIDEYPIILAVAA FAEGATVM GLEELRVKESDRLSA
Sbjct 301 VADLRVRHSELKGVTVPEDRAPSMIDEYPIILAVAAFAEGATVMKGLEELRVKESDRLSA 360
          VADLRVRHSELKGVTVPEDRAPSMIDEYPIILAVAAFAEGATVMKGLEELRVKESDRLSA
Query 378 VANGLKLNQVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 437
          VA+GLKLNQVDCDEGE L+VRGRPDGKGLGNA+ V+THLDHRIAMSFLV+GL SE+
Sbjct 361 VADGLKLNQVDCDEGEDFLIVRGRPDGKGLGNAADGRVSTHLDHRIAMSFLVLGLASEHA 420
          VADGLKLNQVDCDEGEDFLIVRGRPDGKGLGNAADGRVSTHLDHRIAMSFLVLGLASEHA
Query 438 VTVDDATMIATSFPEFMDLMAGLGAKIEL 466
          VT+DDA MIATSFPEFM LM GLGAKIEL
Sbjct 421 VTIDDAAMIATSFPEFMQLMTGLGAKIEL 449
          VTIDDAAMIATSFPEFMQLMTGLGAKIEL

```

> [gi|61219221|sp|P0A2Y4|AROA\\_PSES2](#) 3-phosphoshikimate 1-carboxyvinyltransferase (5-enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS)

[gi|61219222|sp|P0A2Y5|ARO\\_ACHSL](#) 3-phosphoshikimate 1-carboxyvinyltransferase  
 (5-enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS)  
 Length=449

Score = 687 bits (1774), Expect = 0.0, Method: Composition-based stats.  
 Identities = 371/450 (82%), Positives = 403/450 (89%), Gaps = 3/450 (0%)

Query	18	MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG	77
		M H AS +PATAR+S L+G +RIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG	
Sbjct	1	MHSASPKPATARRSEALTGEIRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG	60
Query	78	KAMQAMGARIRKEGDTWIIDGVNGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI	137
		+AMQAMGA+IRKEGD WII+GVGNG LL PEA LDFGNA TG RLTMGLVG YD ++FI	
Sbjct	61	RAMQAMGAKIRKEGDVWIINGVNGCLLQPEAALDFGNAGTGARLTMGLVGTYDMKTSFI	120
Query	138	GDASLTKRPMGRVLNPLREMGVQVKSSEDGDRLPVTLRGPKTPTPIYRVPMASAQVKS AV	197
		GDASL+KRPMGRVLNPLREMGVQV++ DGDR+P+TL GPKT PITYRVPMASAQVKS AV	
Sbjct	121	GDASLSKRPMGRVLNPLREMGVQVEAADGDRMPLTLIGPKTANPITYRVPMASAQVKS AV	180
Query	198	LLAGLNTPGITTVIEPIMTCDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID	257
		LLAGLNTPG+TTVIEP+MT DHTEKMLQGFGA+LTVETD DGVR IR+ G+GKL GQ ID	
Sbjct	181	LLAGLNTPGVTTVIEPVMTRDHTEKMLQGFADLTVETDKDGVRRHIRITGQKLVGQTID	240
Query	258	VPGDPSSTAFPLVAALLVPGSDVTILNVLNPNTRTGLLITLQEMGADIEVINRLAGGED	317
		VPGDPSSTAFPLVAALLV GSDVTI NVLMNPNTRTGLLITLQEMGADIEV+N RLAGGED	
Sbjct	241	VPGDPSSTAFPLVAALLVEGSDVTIRNVLNPNTRTGLLITLQEMGADIEVLNARLAGGED	300
Query	318	VADLRVRSSTLKGVTVPEDRAPPIDEYPI LA VAAAFAEGATVMNGLEELRVKESDRLSA	377
		VADLRVR+S LKGV VP +RAP MIDEYP+LA+AA+FAEG TVM+GL+ELRVKESDRL+A	
Sbjct	301	VADLRVRASKLKGVVPPERAPSMIDEYPV LA IAASFAEGETVMDGLDELRVKESDRLAA	360
Query	378	VANGLKLNQVDCDEGETSLVVRGRPDGKGLGNASGA AVATHLDHRIAMSFLVMGLVSEN P	437
		VA GL+ NGVDC EGE SL VRGRPDGKGLG G VATHLDHRIAMSFLVMGL +E P	
Sbjct	361	VARGLEANGVDCTEGEMSLTVRGRPDGKGLG---GGTVATHLDHRIAMSFLVMGLAAEKP	417
Query	438	VTVDDATMIATSFPEFMDLMAGLGAKIELS	467
		VTVDD+ MIATSFPEFMD+M GLGAKIELS	
Sbjct	418	VTVDDSNMIATSFPEFMDMMPGLGAKIELS	447