

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 394 084**

51 Int. Cl.:

**C12N 15/82** (2006.01)

**C12N 9/02** (2006.01)

**C12N 15/62** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **19.06.2003 E 03760745 (4)**

97 Fecha y número de publicación de la solicitud europea: **15.06.2005 EP 1539969**

54 Título: **Péptido de direccionamiento plastidial**

30 Prioridad:

**21.06.2002 FR 0207729**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:

**17.01.2013**

73 Titular/es:

**GENOPLANTE-VALOR (100.0%)  
93, RUE HENRI ROCHEFORT  
91025 EVRY CEDEX, FR**

72 Inventor/es:

**MIRAS, STÉPHANE;  
SALVI, DANIEL;  
ROLLAND, NORBERT;  
JOYARD, JACQUES;  
FERRO, MYRIAM;  
GARIN, JÉROME y  
GRUNWALD, DIDIER**

74 Agente/Representante:

**PONS ARIÑO, Ángel**

**ES 2 394 084 T3**

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

## DESCRIPCIÓN

Péptido de direccionamiento plastidial

5 La presente invención se refiere a la producción de proteínas de interés en plantas, y en particular a su direccionamiento hacia el compartimento del plastidio.

10 Los plastos son orgánulos intracelulares de las plantas clorofílicas (algas, musgos, y plantas superiores). Existen diversos tipos principales de plástidos, según su contenido en pigmentos, y su naturaleza: los amiloplastos, ricos en almidón, los cloroplastos en los que los pigmentos principales son las clorofilas, y los cromoplastos cuyos pigmentos principales son los carotenoides. Estas tres clases de plástidos procedentes de precursores comunes, los proplastos, poseen una estructura básica común constituida por una doble membrana que encierra el estroma del plastidio. En los cloroplastos, existe un tercer sistema de membrana que forma unos sáculos en el interior del estroma denominados tilacoides.

15 Aparte de su papel primordial en la fotosíntesis, los cloroplastos también están implicados en reacciones de óxido-reducción, por ejemplo, la reducción de nitritos en amonio. Los plastos también desempeñan un papel esencial en la biosíntesis y/o almacenamiento de numerosas moléculas, entre las que se pueden citar el almidón, lípidos, carotenoides, la mayoría de aminoácidos, hormonas vegetales (ácido abscísico, precursores de giberelinas, del jasmonato, etc.).

20 A pesar de que los plastos tienen su propio genoma que codifica una parte de sus proteínas, una gran parte de las enzimas que intervienen en las diferentes funciones plastidiales están codificadas por el genoma nuclear y son importadas al interior de los plastos.

25 Esta importación se lleva a cabo por un mecanismo específico, que se ha estudiado más específicamente en el caso de los cloroplastos (para una revisión, cf. CHEN y SCHNELL, Trends Cell Biol. 9, 222 – 227, 1999; KEEGSTRA y CLINE, The Plant Cell 11, 557 – 570, 1999; SCHLEIFF y SOLL, Planta 211, 449 – 456, 2000; JACKSON-CONSTAN y KEEGSTRA, Plant Physiol. 125, 1567 – 1676, 2001). Este mecanismo implica la intervención de un sistema de importación en cada una de las dos membranas plastidiales: en la membrana externa, el complejo TOC (translocón en la membrana externa del cloroplasto) que comprende al menos tres proteínas: Toc 86, 75 y 34 (KESSLER et al., Science 266, 1035 – 1039, 1994; PERRY y KEEGSTRA, Plant Cell 6, 93 – 105, 1994); en la membrana interna, el complejo TIC (translocón en la membrana interna del cloroplasto) que comprende al menos cuatro proteínas: Tic 110, 55, 22 y 20 (KESSLER y BLOBEL, Proc. Natl. Acad. Sci. 93, 7684 – 7689, 1996; LÜBECK et al., EMBO J. 15, 4230 – 4238, 1996; CALIEBE et al., EMBO J. 16, 7342 – 7350, 1997; KOURANOV et al., J. Cell Biol., 143, 991 – 1002, 1998), y una proteína chaperona en el estroma: ClpC (AKITA et al., J. Cell Biol. 136, 983 – 994, 1997; NIELSEN et al., EMBO J. 16, 935 – 946, 1997).

40 Un elemento importante de este mecanismo es Toc75, que es la proteína más abundante en la membrana externa, y forma el poro central del canal de translocación situado en esta membrana (SCHNELL et al., Science 266, 1007 – 1012, 1994; TRANEL et al., EMBO J. 14, 2436 – 2446, 1995). Toc75 interactúa específicamente con una secuencia particular, denominada "péptido de direccionamiento" o "péptido de tránsito", que se encuentra en el extremo N-terminal de las proteínas importadas al interior de los plastos (MA et al., J. Cell Biol. 134, 315 – 327, 1996).

45 Muchos péptidos de direccionamiento han sido identificados en los precursores de proteínas dirigidas hacia el espacio intermembrana, la membrana interna, el estroma, y en el caso de los cloroplastos, hacia la membrana del tilacoides.

50 Entre las proteínas que se sabe que poseen un péptido de direccionamiento intraplastidial escindible se pueden citar en particular proteínas dirigidas al espacio intermembrana, (Tic22: KOURANOV et al., 1998, mencionado anteriormente; KOURANOV et al., J. Biol. Chem. 274, 25181 – 25194, 1999), proteínas dirigidas a la membrana interna (TPT (translocador de triosa-Pi/Pi): BRINK et al., J. Biol. Chem. 270, 20808 – 20815, 1995), proteínas dirigidas al estroma (subunidad pequeña de la ribulosa-1,5-bisfosfato carboxilasa (Rubisco): DE CASTRO SILVA FILHO et al., Plant Mol. Biol. 30, 769 – 780, 1996, anhidrasa carbónica), proteínas dirigidas a la membrana del tilacoides (LHCP (complejo de captación de luz): LAMPPA et al., J. Biol. Chem. 263, 14996 – 14999, 1988; Cfo-II: subunidad ATPasa) y al lumen del tilacoides (OEE1 (Oxygen Evolving Element 1): KO y CASHMORE, EMBO J. 8, 3187 – 3194, 1989).

60 Estos péptidos de direccionamiento en general contienen entre 40 y 100 aminoácidos, y en su mayor parte comprenden características comunes: están prácticamente desprovistos de aminoácidos cargados negativamente, como el ácido aspártico, el ácido glutámico, la asparagina o la glutamina; su región N-terminal está desprovista de aminoácidos cargados, y de aminoácidos como la glicina o la prolina; su región central contiene una proporción muy

elevada de aminoácidos básicos o hidroxilados, tales como la serina y la treonina; su región C-terminal es rica en arginina y tiene la capacidad de formar una estructura secundaria anfipática, en lámina beta.

- 5 En el caso de proteínas dirigidas al lumen del tilacoides, el péptido de direccionamiento consta de dos partes y contiene información adicional para atravesar la membrana del tilacoides (DE BOER y WEISBEEK, *Biochim. Biophys. Acta.* 1071, 221 – 253, 1991). En algunos casos, este péptido de direccionamiento que consta de dos partes también se puede encontrar en las proteínas dirigidas a la membrana del tilacoides (KARNAUCHOV y col., *J. Biol. Chem.* 269, 32.871 – 32.878, 1994).
- 10 En todos los casos, el péptido de direccionamiento se escinde después de su importación. Esta escisión se lleva cabo por proteasas específicas; se han descrito una proteasa localizada en el estroma (VANDERVERE et al, *Proc Natl. Acad. Sci.* 92, 7177 – 7181, 1995), y una proteasa localizada en el lumen del tilacoides (CHAAL et al, *J. Biol. Chem.* 273, 689 – 692, 1998).
- 15 Las proteínas dirigidas a la membrana externa generalmente no contienen el péptido señal escindible; la información de direccionamiento está contenida en la proteína madura (CLINE y HENRY, *Annu. Rev. Cell Dev. Biol.*, 12, 1 – 26, 1996); después de su síntesis en el citosol, estas proteínas se incorporan directamente a la membrana (VAN'T HOF et al, *FEBS Lett.* 291, 350 – 354, 1991 y *J. Biol. Chem.* 268, 4037 – 4042, 1993; PINADUWAGE y BRUCE, *J. Biol. Chem.* 271, 32907 – 32915, 1996) por medio de interacciones, cuya naturaleza aún no está clara, con la bicapa lipídica. La única excepción conocida hasta la fecha se refiere a la proteína Toc75 u (OEP75) en la que el direccionamiento hacia la membrana externa requiere de la presencia de un péptido de direccionamiento N-terminal escindible que consta de dos partes (TRANEL et al, 1995, anteriormente mencionada; TRANEL y KEEGSTRA, *Plant Cell* 8, 2093 – 2104, 1996).
- 20
- 25 Es sabido que es necesaria la utilización de péptidos de direccionamiento hacia los plastos para introducir en ellos proteínas de interés que permitan actuar sobre diversas funciones plastidiales, en particular, con el fin de mejorar las características de las plantas de interés agronómico, por ejemplo, la biosíntesis de lípidos, de almidón, de vitaminas, de hormonas o de proteínas por parte de dichas plantas, o su resistencia a enfermedades, insectos, o herbicidas. Por ejemplo, la Solicitud EP 189707, así como la Solicitud PCT WO 88/02402 proponen la utilización de péptidos de direccionamiento escindibles procedentes de precursores de proteínas cloroplásticas, y en particular del péptido de direccionamiento de la subunidad pequeña de la ribulosa-1,5-bisfosfato carboxilasa, para la importación de una proteína de interés hacia los cloroplastos; la Solicitud PCT WO 00/12732 propone la utilización de péptidos de direccionamiento de diferentes proteínas plastidiales para la importación de proteínas de interés hacia los plastos.
- 30
- 35 Las funciones plastidiales se pueden modificar de esta forma, y las características conferidas por estas modificaciones son muy diversas.

Con fines ilustrativos no limitantes, se pueden citar:

- 40 - el aumento de la resistencia a herbicidas, mediante la expresión del precursor de la acetolactato sintetasa (ALS), (LEE et al *EMBO J.*, 7, 1241 – 1248, 1988), de la acetolactato sintetasa mutada (PRESTON y POWLES, *Heredity* 88, 8 – 13, 2002); CHONG y CHOI, *Biochem. Biophys. Res. Commun.* 279, 462 – 467, 2000), o de la 3-enolpiruvilsiquimato-5-fosfato sintetasa (EPSP sintetasa) (KLEE et al, *Mol. Gen. Genet.*, 210, 437 – 442, 1987)
- 45 - el aumento de la resistencia a diversas situaciones de estrés, mediante la expresión de la zeaxantina epoxidasa (SEO et al, *Trends Plant Sci.*, 7, 41 – 48, 2002), de la colina monooxigenasa (SHEN et al, *Sheng Wu Gong Cheng Xue Bao*, 17, 1 – 6, 2001), del producto del gen ERD1\_ARATH (KIYOSUE et al, *Biochem. Biophys. Res. Commun.*, 15, 196, 1214 – 1220, 1993), de la ferroquelatasa (CHOW et al, *J. Biol. Chem.*, 31, 272, 27565 – 27571, 1997), de la ácido graso omega-3 desaturasa (IBA et al, *Tanpakushitsu Kakusan Koso*, 39, 2803 – 2813, 1994; MURAKAMI et al, *Science* 21, 287 (5452), 476 – 479, 2000), de la glutamina sintetasa (FUENTES et al, *J. Exp. Bot.*, 52, 1071 – 1081, 2001)
- 50
- 55 - la modificación del metabolismo de los plastidios, para aumentar la captura de energía lumínica (GAUBIER et al, *Mol. Gen. Genet.*, 1, 249, 58 – 64, 1995), la capacidad de fotosíntesis y de crecimiento (MIYAGAWA et al, *Nature Biotech.*, 19, 965 – 969, 2001), el contenido en carotenoides (HUGUENEY et al, *Eur. J. Biochem.*, 1, 209, 399 – 407, 1992; MANN et al, *Nature Biotech.*, 18, 888 – 892, 2000), o el contenido en diversas sustancias de interés, tales como el almidón (Solicitud PCT WO 00/11144), en aminoácidos esenciales (MUEHLBAUER et al, *Plant Physiol.*, 106, 1303 – 1312, 1994), en provitamina A (RÖMER et al, *Nature Biotech.*, 18, 666 – 669, 2000), en hormonas (JOYARD et al, *Plant Physiol.* 118, 715 – 723, 1998), etc.
- 60 - la sobreexpresión y el direccionamiento cloroplásticos de proteínas que se pueden utilizar para la biorremediación

(detoxificación o descontaminación de suelos contaminados), tales como la ferritina (LOBREAUX et al, Biochem. J., 15, 288 (Pt 3), 931 – 939, 1992), proteínas de la familia de las fitoquelatinas (CAZALE y CLEMENS, FEBS 507, 215 – 219, 2001; TSUJI et al, BBRC 293, 653 – 659, 2002), etc.

5 Todos los péptidos de direccionamiento intraplastidial conocidos en la técnica anterior permiten importar una proteína al interior de los plastos a través de los sistemas de importación de membrana TOC y TIC, como se ha indicado anteriormente. Se ha constatado que la utilización de estos péptidos para el direccionamiento de proteínas de interés al interior del cloroplasto podría presentar el inconveniente de saturar estos sistemas de importación, especialmente en el caso en el que la construcción péptido de interés / proteína de interés está situada bajo el control de un promotor fuerte tal como el promotor 35S, al entrar en competición con las proteínas dirigidas de forma natural hacia el cloroplasto. Por ello se producen "fugas" que se traducen, al cabo de unos días, en la presencia de la proteína de interés en otros compartimentos subcelulares como el citoplasma.

10 Sería deseable disponer de péptidos de direccionamiento intraplastidial, que no dependiesen del sistema de importación TOC/TIC, y permitiesen por tanto evitar los inconvenientes anteriormente mencionados.

15 En trabajos previos dirigidos a identificar, con un enfoque proteómico, las proteínas de preparaciones de membranas de cloroplastos de espinaca, los inventores han identificado, entre otros, péptidos que tienen una similitud de secuencia significativa con una proteína putativa de 41 kDa de *Arabidopsis* (número de acceso TrEMBL Q9SV68) (SEIGNEURIN-BERNY et al, Plant. J. 19, 217 – 228, 1999; FERRO et al, Electrophoresis 21, 3517 – 3526, 2000). Las secuencias de nucleótidos que codifican para esta proteína putativa de *Arabidopsis* también se mencionan en las solicitudes PCT WO 02/10210 y EP 1033405.

20 Prosiguiendo con sus trabajos con el fin de caracterizar mejor la proteína de *Arabidopsis*, y de su homóloga en espinacas, los inventores han constatado que aunque, de manera sorprendente, se trata de proteínas sintetizadas en el citoplasma e importadas a la membrana interna del cloroplasto, su importación se llevaría a cabo sin la escisión de un péptido de direccionamiento; además, el análisis de secuencias de estas proteínas revela que no existe ninguna secuencia que posea las características de los péptidos de direccionamiento plastidiales conocidos.

25 Las proteínas de la familia representada por la proteína de *Arabidopsis* de 41 kDa, y la proteína homóloga en espinaca se designan en lo sucesivo con la denominación IE41. (IE por *Inner Envelope* según la nomenclatura utilizada convencionalmente para este sistema de membrana).

30 La secuencia de la proteína IE41 de *Arabidopsis* está representada en el listado de secuencias anexo con la referencia SEQ ID NO: 1; la secuencia del ADNc que codifica la proteína IE41 de espinacas está representada en el listado de secuencias anexo con la referencia SEQ ID NO: 2, y la secuencia del polipéptido correspondiente está representada en el listado de secuencias anexo con la referencia SEQ ID NO: 3.

35 Los inventores han investigado cuáles son las regiones de la proteína IE41 implicadas en su direccionamiento plastidial, y han identificado una región de 41 aminoácidos (residuos 60 a 100) esencial para este direccionamiento.

40 La secuencia de esta región está representada en el listado de secuencias anexo con la referencia SEQ ID NO: 4 para la proteína IE41 de *Arabidopsis*, y con la referencia SEQ ID NO: 5 para la proteína IE41 de espinaca.

45 Además han constatado que cuando los fragmentos de IE41 que contienen esta región están fusionados por el extremo N-terminal de una proteína heteróloga, la proteína recombinante que resulta de esta fusión estaba dirigida hacia los cloroplastos de forma similar a la proteína IE41 completa.

50 La presente invención tiene por objeto un polipéptido de direccionamiento intraplastidial caracterizado porque está constituido por:

- un dominio A constituido por un polipéptido que presenta una identidad de al menos el 60 %, preferentemente de al menos el 70 %, de manera ventajosa de al menos el 80 %, y de manera más preferente de al menos el 90 % de identidad, o de al menos el 65 %, preferentemente de al menos el 75 %, de manera ventajosa de al menos el 85 %, y de manera más preferente de al menos el 95 % de similitud, con uno de los polipéptidos de las SEQ ID NO: 4 o SEQ ID NO: 5; y al menos un dominio seleccionado entre:

55 - un dominio B situado en el extremo N-terminal del dominio A, y constituido por un fragmento de uno de los polipéptidos de las SEQ ID NO: 1 o SEQ ID NO: 3 que comprende al menos los aminoácidos 49 a 59, preferentemente al menos los aminoácidos 39 a 59, de manera ventajosa al menos los aminoácidos 29 a 59, y de

manera más preferente al menos los aminoácidos 19 a 59, y de forma particularmente ventajosa al menos los aminoácidos 9 a 59 de dicho polipéptido, o bien constituido por un polipéptido que presenta una identidad de al menos el 60 %, preferentemente de al menos el 70 %, de manera ventajosa de al menos el 80 %, y de manera más preferente de al menos el 90 %, o una similitud de al menos el 65 %, preferentemente de al menos el 75 %, de manera ventajosa de al menos el 85 %, y de manera más preferente de al menos el 95 %, con dicho fragmento;

- un dominio C situado en el extremo C-terminal del dominio A, y constituido por un fragmento de uno de los polipéptidos de las SEQ ID NO: 1 o SEQ ID NO: 3 que comprende al menos los aminoácidos 101 a 111, preferentemente al menos los aminoácidos 101 a 121, de manera ventajosa al menos los aminoácidos 101 a 131, de manera más preferente al menos los aminoácidos 101 a 141, y como máximo los aminoácidos 101 a 151 de dicho polipéptido, o bien constituido por un polipéptido que presenta una identidad de al menos el 60 %, preferentemente de al menos el 70 %, de manera ventajosa de al menos el 80 %, y de manera más preferente de al menos el 90 %, o una similitud de al menos el 65 %, preferentemente de al menos el 75 %, de manera ventajosa de al menos el 85 %, y de manera más preferente de al menos el 95 %, con dicho fragmento.

Los porcentajes de identidad o similitud mencionados en el presente documento se determinan con la ayuda del *software* BLASTp (ALTSCHUL et al, Nucleic Acids Res. 25, 3389 – 3402, 1997), utilizando los parámetros por defecto.

Los dominios A, B y/o C definidos anteriormente pueden proceder de una misma proteína IE41; también pueden proceder de proteínas IE41 de orígenes diferentes.

La presente invención también tiene por objeto cualquier polipéptido quimérico que resulte de la fusión de un polipéptido de direccionamiento intraplástico de acuerdo con la invención con un polipéptido heterólogo. Dicho polipéptido heterólogo puede ser cualquier polipéptido de interés que se desea introducir en los plastos. Preferentemente, el polipéptido de direccionamiento intraplástico de acuerdo con la invención está situado en el extremo N-terminal del péptido heterólogo. No obstante también puede estar situado en su interior, o en su extremo C-terminal.

La presente invención también tiene por objeto la utilización de un polipéptido de direccionamiento intraplástico de acuerdo con la invención para la importación de una proteína de interés al interior de los plastos, y de manera ventajosa, para el direccionamiento de dicha proteína hacia la membrana plastidial interna.

De acuerdo con una forma de realización preferida de la presente invención, dicho polipéptido de direccionamiento intraplástico se utiliza para la importación de dicha proteína de interés al interior de los cloroplastos.

En particular, la presente invención tiene por objeto un procedimiento para la importación de una proteína de interés al interior de los plastos caracterizado porque comprende la expresión, en una célula vegetal que contiene dichos plastos, de un polipéptido quimérico que resulta de la fusión de un polipéptido de direccionamiento intraplástico de acuerdo con la invención con un polipéptido heterólogo.

La presente invención también tiene por objeto:

- cualquier polinucleótido que codifica un polipéptido de direccionamiento intraplástico o un polipéptido quimérico de acuerdo con la invención;

- cualquier casete de expresión recombinante que comprenda un polinucleótido de acuerdo con la invención situado bajo el control de secuencias de regulación de la transcripción adecuadas (en particular, promotores y terminadores de la transcripción);

- cualquier vector recombinante que resulte de la inserción, en un vector hospedador adecuado, de un polinucleótido o de un casete de expresión de acuerdo con la invención.

La presente invención también tiene por objeto células hospedadoras que alberguen un polinucleótido, un casete de expresión, o un vector recombinante de acuerdo con la invención.

La presente invención también comprende plantas transgénicas genéticamente modificadas por un polinucleótido o un casete de expresión de acuerdo con la invención, así como los descendientes de estas plantas. La invención también comprende células y tejidos vegetales, así como los órganos o partes de plantas, incluyendo hojas, tallos, raíces, flores, frutos y/o semillas obtenidas a partir de estas plantas.

En la puesta en práctica de la presente invención se pueden utilizar las técnicas convencionales de construcción de vectores recombinantes, de transformación de células o de organismos hospedadores, y de producción de proteínas recombinantes.

5 La selección del vector hospedador, y de las secuencias de regulación de la expresión particularmente se llevará a cabo en función del procedimiento de transformación y de la planta hospedadora seleccionados, y/o del tipo de célula o de tejido en el que se desea obtener la expresión.

10 Se conocen muchos promotores que se pueden utilizar para la expresión en células vegetales. A modo de ejemplo, se puede seleccionar un promotor constitutivo, como el promotor 35S de CaMV o sus derivados, o el promotor de la actina o de la ubiquitina, etc. También se puede seleccionar un promotor inducible o un promotor específico de tejido, con el fin de producir el direccionamiento plastidial de la proteína de interés sólo en ciertas fases del desarrollo de la planta, en ciertas condiciones medioambientales, o en ciertos tejidos diana.

15 Por ejemplo, si se desea obtener preferentemente un direccionamiento de la proteína de interés hacia los cloroplastos, se expresará un polipéptido quimérico de acuerdo con la invención bajo el control de un promotor específico de tejidos u órganos ricos en plastos. A modo de ejemplo, los promotores del virus de la *Chlorella* que regulan la expresión génica de la adenina metiltransferasa (MITRA y HIGGINS, Plant Mol. Biol. 26, 85 – 93, 1994) o el del virus del mosaico de la yuca (VERDAGUER et al, Plant Mol. Biol. 37, 1055 – 1067, 1998) se expresan principalmente en los tejidos verdes. Los elementos reguladores del promotor del gen 2A11 del tomate permiten una expresión específica en los frutos (VAN HAAREN y HOUCK, Plant Mol. Biol. 17, 615 – 630, 1991), etc.

20 Son muy conocidos los procedimientos de transformación de células vegetales o de plantas completas: a modo de ejemplos no limitantes, se citará la transformación de protoplastos en presencia de polietilenglicol, mediante electroporación, utilización de una pistola de partículas, microinyección citoplasmática o nuclear, o la transformación a través de *Agrobacterium*.

25 La presente invención se puede poner en práctica en las aplicaciones habituales de los péptidos de direccionamiento hacia plastos, y especialmente en las aplicaciones mencionadas anteriormente, para actuar sobre diversas funciones plastidiales. Se trata en particular de la modificación de las funciones propias de la membrana interna de la envuelta, por ejemplo, la biosíntesis de pigmentos, de quinonas, de ácidos grasos, de vitaminas y de precursores de hormonas vegetales, pero también, de la importación de todos los iones y metabolitos hacia los plastos.

30 Las características de los péptidos de direccionamiento plastidiales de acuerdo con la invención, que son muy diferentes de las de los péptidos de direccionamiento plastidiales conocidos permiten suponer que los péptidos de direccionamiento de acuerdo con la invención utilizan un sistema de importación diferente de aquel que implica a las proteínas TOC y TIC.

35 Por lo tanto, las proteínas de interés dirigidas hacia los plastos con la ayuda de un péptido de direccionamiento de acuerdo con la invención no compiten con las proteínas dirigidas de manera natural hacia el plasto a través del sistema TOC y TIC, y no lo saturan. En particular, esto permitiría evitar las fugas hacia los otros compartimentos subcelulares, y conservar las proteínas de interés dentro del cloroplasto, incluso después de varios días de expresión. Esto también permitiría dirigir proteínas hacia los plastos, para las cuales no sería funcional la importación convencional por medio de una secuencia de direccionamiento escindible en el N-término y la utilización del sistema TIC/TOC.

40 La presente invención se comprenderá mejor con ayuda de la descripción siguiente, que se refiere a ejemplos no limitantes que ilustran la obtención y caracterización de péptidos de direccionamiento plastidiales de acuerdo con la invención y su puesta en práctica para la importación de proteínas heterólogas al interior de los cloroplastos.

#### **EJEMPLO 1: CARACTERIZACIÓN Y CLONACIÓN DE LA PROTEÍNA IE41 DE *ARABIDOPSIS THALIANA***

45 En trabajos previos dirigidos a identificar las proteínas más hidrófobas de preparaciones de membranas de cloroplastos de espinaca (SEIGNEURIN-BERNY et al, Plant. J., 19, p. 217 – 228, 1999; FERRO et al, Electrophoresis, 21, 3517 – 3526, 2000) se han puesto de manifiesto varios péptidos que derivan de una proteína de 41 kDa que presenta una gran similitud de secuencia con una proteína putativa de *Arabidopsis* (número de acceso TrEMBL Q9SV68).

50 Sin embargo, el análisis de la secuencia primaria de esta proteína de 41 kDa con la ayuda de la aplicación TMPred (HOFMANN y STOFFEL, Biol. Chem. Hoppe-Seyler., 347, 166, 1993), no ha permitido detectar segmentos

transmembrana que puedan asegurar el anclaje de la proteína en una bicapa lipídica.

Para confirmar la localización de esta proteína en la envuelta del cloroplasto, se clonó el ADNc correspondiente y la proteína recombinantes se sobre-expresó en *E. coli* con el fin de obtener anticuerpos policlonales dirigidos contra esta proteína.

### **Expresión en *E. coli***

El ADNc que codifica para la proteína de *Arabidopsis* de 41 kDa se obtuvo por PCR, a partir de una librería de ADNc de *Arabidopsis*, utilizando los siguientes cebadores:

TCA CATATG GCTGGAAACTCAATGCAC (SEQ ID NO: 10)  
que permite la introducción de un sitio de restricción *NdeI* (subrayado) en el extremo 5' del ADNc;

ATG GATCC AACGCTCTTATGGCTCGAC (SEQ ID NO: 11)  
que permite la introducción de un sitio de restricción *BamHI* (subrayado) en el extremo 3' del ADNc.

El fragmento de amplificación se clonó en el plásmido pBluescript KS-. El inserto se digirió con las enzimas de restricción *NdeI* y *BamHI*, y se insertó en el vector de expresión pET-15b (Novagen).

El vector resultante permite la expresión de una proteína recombinante que tiene una cola de polihistidina (His-tag) en su extremo N-terminal [(His-tag)-P41].

Este vector se utiliza para transformar la cepa BLR(DE3) de bacterias *E. coli*.

Las bacterias recombinantes se cultivaron en 500 ml de medio LB que contiene 100 µg/ml de ampicilina a 37 °C. Cuando la densidad óptica a 600 nm (DO<sub>600</sub>) del cultivo de *E. coli* alcanzó 0,6, se añadió IPTG (isopropil-β-D-galactotipiranósido) a una concentración final de 1 mM para inducir la expresión de la proteína de 41 kDa.

Después de 3 horas de cultivo, las células se centrifugaron durante 2 min a 13.000 rpm (Eppendorf 5415D).

El sedimento se resuspendió en 20 ml de tampón de lisis (NaH<sub>2</sub>PO<sub>4</sub> 50 mM, NaCl 300 mM, imidazol 10 mM, pH 8) y las bacterias se lisaron por sonicación (sonicador OMRON, tipo STP.YM.220.VAC, 6 x 1 min, 0 °C).

Después de la sonicación, el extracto de proteínas total se analizó por SDS-PAGE al 12 %. Los geles se tiñeron con azul de Coomassie (R-250, Biorad).

Los resultados se muestran en la Figura 1A: Leyenda para la Figura 1A:

pET-15b = bacterias transformadas con el plásmido no recombinante (control negativo)  
pET-15b + inserto = bacterias transformadas con el plásmido recombinante  
- = Sin inducción con IPTG  
+ = Inducción con IPTG  
\* = Banda correspondiente a la proteína recombinante de 41 kDa

Después de la inducción con IPTG, la proteína recombinante se expresa intensamente en las bacterias transformadas con el plásmido pET-15b que comprende el inserto de ADNc de *Arabidopsis*.

### **Solubilización de la proteína recombinante**

El sedimento bacteriano se suspendió en 1 ml de tampón Tris/HCl 20 mM, pH 6,8 y se lisó por sonicación (sonicador OMRON, tipo STP.YM.220.VAC, 6 x 1 min, 0 °C).

Después de la sonicación, una primera centrifugación (2 min, 13.000 rpm) del extracto de proteínas total permite aislar las proteínas solubles del sobrenadante. Las proteínas insolubles del sedimento se suspendieron en 1 ml de un tampón que contiene detergente (Tris/HCl 20 mM a pH 6,8, 0,5 % de Triton X-100). Una segunda centrifugación permite aislar las proteínas de membrana solubilizadas por el Triton X-100. Las proteínas no solubilizadas se resuspendieron en Tris/HCl 20 mM, pH 6,8 y se analizaron. Las diferentes fracciones se analizaron mediante SDS-PAGE al 12 % como se ha descrito anteriormente.

Los resultados se muestran en la Figura 1B: Leyenda para la Figura 1B:

pET-15b = bacterias transformadas con el plásmido no recombinante (control negativo)

pET-15b + inserto = bacterias transformadas con el plásmido recombinante

S = proteínas solubles

5 D = proteínas solubilizadas en Triton X-100

1 = proteínas no solubilizadas en Triton X-100

\* = Banda correspondiente a la proteína recombinante de 41 kDa

10 Observamos que la proteína recombinante (\*) se encuentra en 2 fracciones distintas: una fracción hidrosoluble presente en el citosol de *E. coli* y una fracción insoluble, que no se solubiliza en Triton X-100, lo que indica que es probable que se agregue en forma de cuerpos de inclusión.

### **Purificación de la proteína recombinante**

15 La fracción soluble de la proteína (His\_tag)-P41 recombinante se purifica por cromatografía de afinidad.

Después de la centrifugación (10 min a 6.000 rpm, Eppendorf 5415D), el sobrenadante se aplicó a una columna de afinidad "His-bind resin" (NOVAGEN) de 2,5 ml cargada con 13 ml de tampón de carga (NiSO<sub>4</sub> 50 mM) y se equilibró con 5 ml de tampón de equilibrio (Tris/HCl 20 mM a pH 7,9, imidazol 5 mM, NaCl 0,5 M).

20 La columna se lavó con 2 volúmenes que contienen imidazol 35 mM, y 2 volúmenes de tampón de lisis que contiene imidazol 60 mM (L2). La proteína se eluyó con 6 volúmenes de tampón de lisis que contiene imidazol 250 mM. Las diferentes fracciones se analizaron por SDS-PAGE al 12 % y se revelaron con azul de Coomassie.

25 Los resultados se muestran en la Figura 2. Leyenda de la Figura 2:

C = sedimento bacteriano diluido en tampón de lisis (10 µl)

S = proteínas bacterianas solubles (10 µl)

P = proteínas no unidas (10 µl)

30 L, L1, L2 = lavados con el tampón de lisis (imidazol 35 y 60 mM) (15 µl)

Elución = fracción eluida en presencia de imidazol 250 mM.

### **Producción de anticuerpos policlonales**

35 La proteína recombinante (His\_tag)-P41 purificada, se desaló (columna Sephadex G25) y se almacenó a -80 °C.

Esta proteína recombinante se utiliza para inmunizar conejos con el fin de producir anticuerpos policlonales dirigidos contra la proteína de 41 kDa de *Arabidopsis*.

### **EJEMPLO 2: LOCALIZACIÓN DE LA PROTEÍNA DE 41 KDA EN LOS CLOROPLASTOS**

Durante el procedimiento de purificación, la proteína de 41 kDa se comporta como una proteína hidrosoluble y débilmente hidrófoba, lo que plantea la cuestión de su asociación eficaz con la envuelta del cloroplasto.

45 Su localización subcelular se ha verificado mediante el análisis de las diferentes fracciones del cloroplasto.

Se obtuvieron cloroplastos en bruto a partir de 3 – 4 kg de hojas de espinaca (*Spinacia oleracea L.*) y se purificaron por centrifugación isopícnica en gradiente de Percoll (DOUCE y JOYARD, *Methods in chloroplast Molecular Biology*. Edelman, M., Hallick, R. y Chua, N.-H., eds. (Amsterdam: Elsevier Science Publishers BV), 239 – 256, 1982). En esta etapa, se añadieron inhibidores de proteasas (PMSF 1 mM, benzamidina 1 mM y ácido aminocaproico 0,5 mM) para evitar la degradación de las proteínas. Los cloroplastos purificados se lisaron en medio hipotónico, y las membranas de la envuelta se purificaron a partir del lisado por centrifugación en un gradiente de sacarosa (DOUCE Y JOYARD, 1982, véase más arriba). Se obtuvieron subfracciones de la envuelta enriquecidas respectivamente en las membranas externas e internas de acuerdo con el protocolo descrito por BLOCK et al (*J. Biol. Chem.*, 258, 13273 – 13280, 1983).

60 Todas las etapas anteriores se llevan a cabo entre 0 y 5 °C. Las fracciones obtenidas se almacenaron en nitrógeno líquido en MOPS-NaOH 50 mM, pH 7,8, en presencia de inhibidores de proteasas (benzamidina 1 mM y ácido aminocaproico 0,5 mM).

**Análisis de las fracciones del cloroplasto**

El análisis por SDS-PAGE de los cloroplastos totales, o de sus fracciones de membranas de la envuelta, estroma, membranas de los tilacoides (15 µg), así como de sub-fracciones de la envuelta de los cloroplastos (15 µg) se realizaron como se describe por CHUA (Methods Enzymol., 69, 434 – 436, 1980). Los geles de resolución y de concentración (12 – 15 % de acrilamida), como tampón de migración, contienen el 0,1 % de SDS. Los péptidos se revelan con azul de Coomassie (MANIATIS et al, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982) o con nitrato de plata (MERRIL et al, Anal. Biochem., 110, 201 – 207, 1981).

Para los análisis de Western blot, se detectó la proteína de 41 kDa usando anticuerpos policlonales marcados con fosfatasa alcalina y dirigidos contra la proteína recombinante de *Arabidopsis* producida como se ha descrito en el

**Ejemplo 1.**

Los resultados se muestran en la Figura 3A.  
Leyenda de la Figura 3A:

C = proteínas del cloroplasto  
E = proteínas de membrana de la envuelta  
S = proteínas del estroma  
T = proteínas de las membranas de los tilacoides

Estos resultados muestran que la proteína de 41 kDa se asocia únicamente con la envuelta del cloroplasto y no se detecta ni en el estroma ni en los tilacoides.

Los cloroplastos intactos purificados están desprovistos de proteínas citosólicas que puedan contaminar la preparación (DOUCE y JOYARD, 1980, véase más arriba). Sin embargo, la proteína de 41 kDa podría interactuar específicamente con la membrana externa de la envuelta del cloroplasto y así se podría purificar simultáneamente con las preparaciones de la envuelta. Para descartar esta posibilidad, se trataron 20 µg de envuelta derivada de cloroplastos intactos con termolisina de *Bacillus thermoproteolyticus* (0, 20, 50 y 100 µg/ml) con el fin de digerir los polipéptidos localizados en la superficie externa de la membrana externa de la envuelta (JOYARD et al, J. Biol. Chem., 258, 10000 – 10006, 1983). A modo de control, las proteínas de la envuelta solubilizadas se sometieron al mismo tratamiento proteolítico.

La presencia de la proteína de 41 kDa se detectó por Western blot como se ha descrito anteriormente.

Los resultados se muestran en las Figuras 3B y 3C.  
Leyenda de las figuras 3B y 3C:

0 = ausencia de termolisina  
20 = termolisina a 20 µg/ml  
50 = termolisina a 50 µg/ml  
100 = termolisina a 100 µg/ml

La proteína de 41 kDa no se ve afectada por el tratamiento con la termolisina (Figura 3B), mientras que el mismo procesamiento proteolítico realizado en proteínas de la envuelta solubilizadas muestran la sensibilidad de la proteína de 41 kDa al tratamiento con termolisina (Figura 3C). Este resultado descarta la hipótesis de que la proteína de 41 kDa esté localizada en la superficie externa de la membrana externa. De hecho, la proteína de 41 kDa no es una proteína del citosol que contamine preparaciones de la envuelta del cloroplasto.

Se utilizan subfracciones de la envuelta enriquecidas respectivamente en las membranas externas e internas para precisar la sub-localización de la proteína de 41 kDa en la membrana de la envuelta del cloroplasto. La naturaleza de estas subfracciones de la envuelta se confirmó mediante el uso de los marcadores IE18 y OEP24, que son respectivamente proteínas intrínsecas de la membrana interna y externa de la envuelta del cloroplasto. Las proteínas IE18 y OEP24 se detectan utilizando anticuerpos policlonales dirigidos específicamente contra cada una de estas proteínas.

Los resultados se muestran en la Figura 3D.  
Leyenda de la Figura 3D:

E = proteínas de las membranas de la envuelta  
OM = proteínas de la membrana externa  
IM = proteínas de la membrana interna

5 La proteína de 41 kDa se encuentra, al igual que la proteína IE18, únicamente en las preparaciones de las envueltas del cloroplasto enriquecidas en la membrana interna.

Todos los resultados representados en las Figuras 3A – 3D muestran que la proteína de 41 kDa está localizada en la membrana interna de la envuelta del cloroplasto.

10 De acuerdo con la nomenclatura convencional, esta proteína de la envuelta interna, que tiene una masa teórica mediante electroforesis en gel de poliacrilamida de 41 kDa, se denomina IE41 (por "Inner Envelope Protein of 41 kDa").

### 15 **Análisis de las interacciones entre la proteína IE41 y la membrana interna de la envuelta del cloroplasto**

Para analizar con más precisión la forma de interaccionar de la proteína IE41 con la membrana interna de los cloroplastos, se sometieron a prueba diferentes hipótesis:

20 1) La proteína IE41 es una proteína soluble localizada en el espacio intermembrana situado entre las membranas externa e interna de la envuelta del cloroplasto. Esta proteína se purificaría simultáneamente con las preparaciones de la envuelta, y secuestrada en las vesículas de la envuelta. En este caso, la sonicación de las vesículas de la envuelta permitiría la liberación de IE41.

25 Para someter a prueba esta primera hipótesis, se solubilizaron proteínas de la envuelta (500 µg) en MOPS 50 mM, pH 7,8 (500 µl). Las vesículas de la envuelta se sometieron a sonicación durante 10 segundos y se centrifugaron (20 min a 72.000 g, Beckman L2 65B, rotor SW28) para separar las proteínas solubles y de membrana. Cada fracción (20 µl) se analizó por SDS-PAGE al 12 % (revelado: azul de Coomassie) y Western blot.

30 Los resultados se muestran en la Figura 4A. Leyenda de la Figura 4A:

- = Sin sonicación  
+ = Sonicación  
S = proteínas solubles  
35 I = proteínas de membrana

Los análisis por SDS-PAGE y Western blot de las fracciones tratadas de la envuelta muestran que la proteína IE41 no se solubiliza después de la sonicación de las vesículas de la envuelta. Por contra, las proteínas principales solubles del estroma (Rbcl) que quedan secuestradas en las vesículas de la envuelta, y que se sabe que contaminan las fracciones de la envuelta, se solubilizan con este tratamiento. Esto demuestra que la proteína IE41 ni es una proteína soluble del espacio intermembrana, ni una proteína soluble del estroma que contamina la fracción purificada de la envuelta.

45 2) La proteína IE41 puede estar unida a la membrana interna:

- mediante el anclaje a la bicapa lipídica o mediante su inserción parcial en la misma; en este caso, sólo el uso de detergente puede permitir la solubilización de la proteína IE41;

50 - mediante interacciones electrostáticas con una o más proteínas de membrana o con la superficie polar de la bicapa lipídica; en este caso, estas interacciones se pueden romper, y la proteína IE41 se puede solubilizar, mediante un tratamiento alcalino o con elevadas concentraciones salinas.

Para determinar el tipo de interacciones implicadas en la unión de IE41 con la membrana interna, se llevaron a cabo los siguientes experimentos:

55 a) solubilización por Triton X-100:

Se diluyeron vesículas de la envuelta (0,8 mg) en 1 ml de MOPS 50 mM, pH 7,8 que contiene 0,05, 0,1 ó 0,2 % (v/v) de Triton X-100. Después de incubar durante 30 min a 4 °C, la mezcla se centrifugó para separar las proteínas solubles y de membrana. Todas las fracciones (20 µl) se analizaron mediante SDS-PAGE al 12 % (revelado: azul de Coomassie) y Western blot.

Los resultados se muestran en la Figura 4C. Leyenda de la Figura 4C:

- 5  
 Mix = vesículas de la envuelta purificadas  
 M = proteínas de membrana de la envuelta  
 S = proteínas solubles de la envuelta

Estos resultados muestran que la proteína IE41 se puede solubilizar completamente con concentraciones de Triton X-100 (0,2 %), muy inferiores a las necesarias (aproximadamente del 2 %) para solubilizar proteínas intrínsecas.

- 10  
 b) solubilización por tratamiento alcalino o tratamiento con una disolución salina.

Se incubaron vesículas de la envuelta purificadas (500 µg) durante 30 min a 4 °C en diferentes medios (500 µl):

- 15  
 1) MOPS 10 mM, pH 7,8 + NaCl 0,5 M;  
 2) MOPS 10 mM, pH 7,8 + KI 0,5 M;  
 3) Na<sub>2</sub>CO<sub>3</sub> 0,1 M, pH 11;  
 4) NaOH 0,1 N,

- 20  
 y a continuación se sometieron a sonicación y se centrifugaron como se ha descrito anteriormente con el fin de separar las proteínas solubles y de membrana.

- 25  
 Todas las fracciones se analizaron (20 µl) mediante SDS-PAGE al 12 % (revelado: azul de Coomassie) y Western blot (detección con anticuerpos policlonales dirigidos contra la proteína IE41 de *Arabidopsis* (dilución 1/5.000) o dirigidos contra la proteína IE18 (dilución 1/5000).

Los resultados se muestran en la Figura 4B.  
 Leyenda de la Figura 4B:

- 30  
 + = sonicación (10 segundos)  
 NaCl 0,5 M = tratamiento 1  
 KI 0,5 M = tratamiento 2  
 Na<sub>2</sub>CO<sub>3</sub> 0,1 M, pH 11 = tratamiento 3  
 NaOH 0,1 N = tratamiento 4  
 35  
 S = fracción de proteínas solubles  
 I = fracción de proteínas insolubles

- 40  
 Estos resultados muestran que la proteína IE41 se solubiliza al menos parcialmente con los tratamientos salinos (KI, NaCl) o alcalinos moderados (Na<sub>2</sub>CO<sub>3</sub>), que no tienen ningún efecto sobre la proteína intrínseca IE18, esta última que sólo se puede solubilizar mediante un tratamiento alcalino fuerte (NaOH).

Los resultados presentados en las figuras 4A, 4B y 4C muestran que la proteína IE41 es una proteína extrínseca cuya unión a la membrana interna de la envuelta del cloroplasto implica interacciones electrostáticas.

45 **EJEMPLO 3: PURIFICACIÓN Y CARACTERIZACIÓN DE LA PROTEÍNA IE41 DE ESPINACA**

- Sorprendentemente, la proteína IE41 purificada a partir de cloroplastos de espinaca y la proteína recombinante de *Arabidopsis* tienen un tamaño similar en SDS-PAGE y Western blot, lo que sugiere la posibilidad de que IE41 pueda ser dirigida a la membrana interna de la envuelta sin necesidad de la escisión de una secuencia de importación N-terminal.

Para probar esta hipótesis, la proteína IE41 presente en la envuelta de los cloroplastos de espinaca se purificó con el fin de secuenciar y comparar su secuencia con la del ADNc correspondiente.

55 **Inmunopurificación de la proteína IE41 de espinaca**

- Se disolvieron las proteínas de la envuelta de cloroplastos (1 mg) en 1 ml de tampón Tris/HCl 50 mM, NaCl 150 mM, CHAPS 6 mM y se centrifugó (20 min, 72.000 g, Beckman L2 65B, rotor SW28). Las proteínas solubles se incubaron durante 1 hora a 4 °C con 33 µl de suero policlonal dirigido contra la proteína recombinante IE41 de *Arabidopsis*. A continuación se añadieron 50 µg de agarosa-proteína A (Boehringer) y la mezcla se incubó durante 3 horas a 4 °C. Después de tres lavados sucesivos por centrifugación (Eppendorf 5415D, 12.000 rpm, 20 min, 4 °C) y la

resuspensión del sedimento en 1 ml de tampón de solubilización (MOPS 20 mM, NaCl 150 mM, CHAPS 6 mM, pH 7,8), se añadió un exceso (50 µg) de la proteína recombinante (His-tag)-IE41 en 200 µl de tampón de solubilización. La mezcla se incubó durante 1 hora a 4 °C, y se centrifugó durante 20 min a 12.000 rpm (Eppendorf 5415D). El sobrenadante se incubó durante 1 h con la resina Ni-NTA (QIAGEN) previamente equilibrada en tampón de solubilización, con el fin de eliminar la mayor parte de la proteína recombinante (His-tag)-IE41.

Cada fracción (20 µl) se analizó por SDS-PAGE al 12 % (revelado con nitrato de plata) y Western blot (usando los anticuerpos policlonales de conejo dirigidos contra IE41 descritos en el Ejemplo 1).

Los resultados se muestran en la Figura 5. Leyenda de la Figura 5:

- A: análisis por SDS-PAGE;
- B: Western blot;
- Mix = proteínas de la envuelta solubilizadas;
- C = proteínas insolubles;
- S = proteínas solubles;
- L1, L2, L3 = fracciones recogidas durante los tres lavados sucesivos;
- E1 = fracción eliminada por incubación con la resina Ni-NTA;
- E2 = proteína IE41 de espinaca purificada (So) + (His-tag)-IE41.

La fracción E2 que comprende la proteína IE41 natural de espinaca purificada (So) sigue estando contaminada por la proteína (His-tag)-IE41 recombinante de *Arabidopsis*.

La diferencia de tamaño entre estas dos proteínas corresponde a la cola de polihistidina (His-tag) presente en el extremo N-terminal de la proteína recombinante.

#### **EJEMPLO 4: OBTENCIÓN DEL ADNc QUE CODIFICA PARA LA PROTEÍNA IE41 DE ESPINACA**

La secuenciación parcial de la proteína IE41 de espinaca eluida a partir del gel de electroforesis permitió obtener 9 secuencias peptídicas diferentes. Estas secuencias se utilizaron para definir los cebadores degenerados que han permitido aislar el ADNc que codifica la proteína IE41.

La secuencia completa de este ADNc (SEQ ID NO: 2), así como la secuencia polipeptídica deducida (SEQ ID NO: 3) se muestran en la Figura 6. El codón de iniciación de la traducción ATG está indicado en negrita y el codón de detención TAA está subrayado. Las 9 secuencias peptídicas obtenidas por secuenciación directa de la proteína IE41 de espinaca están resaltadas en gris. La correspondencia entre los péptidos obtenidos con la secuencia deducida del ADNc de la proteína IE41 de espinaca, particularmente en la región N-terminal, así como la presencia de un codón de detención en la posición 3' de la metionina de iniciación y en el mismo marco de lectura, demuestran que el ADNc predicho está completo y que esta proteína no está sometida a maduración post-traducciona durante su direccionamiento hacia la membrana interna de la envuelta del cloroplasto.

La proteína IE41 de espinaca presenta una identidad del 75,1 % y una similitud del 88,8 % con la proteína IE41 de *Arabidopsis*. Esta gran similitud, y el hecho de que *Arabidopsis* contiene un sólo gen *ie41* por genoma haploide, permiten llegar a la conclusión de que estas proteínas están codificadas por los genes ortólogos de *Arabidopsis* y de espinaca.

Las proteínas IE41 de *Arabidopsis* y de espinaca se alinearon con proteínas homólogas de bacterias, de levaduras y de animales.

Los resultados se muestran en la Figura 7.

*Arabidopsis thaliana*: IE41 ATH (SEQ ID NO: 1)  
Espinaca: IE41 SOL (SEQ ID NO: 3)

Proteínas homólogas:

- Escherichia coli*: QORECOLI (SEQ ID NO: 6),
- Saccharomyces cerevisiae*: QORYEAST (SEQ ID NO: 7),
- Cavia porcellus*: QORCAVPO (SEQ ID NO: 8),
- Ratón: QORMOUSE (SEQ ID NO: 9).

Los restos conservados en las 6 secuencias peptídicas están resaltados en color gris oscuro. Los restos

conservados en la secuencia de IE41 y en al menos otra secuencia de proteínas homólogas están resaltados en gris. Las similitudes entre los restos se basan en los siguientes grupos: ASPTG, ILMV, KRH, NQ, DE, YWF y C.

5 Las búsquedas de homología indican que la proteína IE41 pertenece a la superfamilia de las deshidrogenasas, y más particularmente al grupo de las quinonas oxidoreductasas de tipo  $\xi$ -cristalino (JÖRNVALL et al, FEBS 3, 240 – 244, 1993). Además, la comparación de secuencias entre las proteínas IE41 y otras proteínas de la misma familia revela que los primeros 50 restos de la región N-terminal de estas proteínas están altamente conservados en bacterias, plantas y animales. Esta observación sugiere que la región N-terminal de la proteína IE41 de plantas no estaría implicada en el direccionamiento hacia el cloroplasto, y lo más probable es que estuviera conservada a causa de la presión selectiva ejercida durante el transcurso de la evolución sobre el dominio catalítico de la proteína.

#### **EJEMPLO 5: ANÁLISIS DEL DIRECCIONAMIENTO PLASTIDIAL DE LA PROTEÍNA IE41 DE ARABIDOPSIS EN CÉLULAS DE ARABIDOPSIS Y DE TABACO**

15 Para definir el dominio esencial para la importación de la proteína IE41, se expresan diferentes construcciones que codifican formas truncadas de esta proteína fusionadas a la GFP en células de Arabidopsis y de tabaco.

#### **Construcción de vectores de expresión**

20 Se han descrito anteriormente por CHIU et al (Curr. Biol., 6, 325 – 330, 1996) el plásmido [35 $\Omega$ -sGFP(S65T)] usado para estas construcciones, que comprende la secuencia codificante de la GFP bajo el control del promotor 35S, y el plásmido [35 $\Omega$ -TP-sGFP(S65T)], que comprende la secuencia que codifica el péptido de direccionamiento (TP) de la subunidad pequeña de la ribulosa-1,5-bisfosfato carboxilasa, fusionada a la secuencia que codifica la GFP.

25 La secuencia que codifica la proteína IE41 de Arabidopsis se amplifica por PCR usando los siguientes dos cebadores:

*Xho*I-N-ter CCTCTCGAGATGGCTGGAAACTCATGCAC (SEQ ID NO: 12),  
y  
30 *Nco*I-C-ter CAACCCATGGATGGCTCGACAATGATCTTC (SEQ ID NO: 13), que introducen respectivamente un sitio *Xho*I y un sitio *Nco*I (subrayados).

El producto de la PCR se clonó con los extremos romos en el vector pBluescript KS (Stratagene). El fragmento *Xho*I-*Nco*I escindido del plásmido así obtenido se insertó en el plásmido 35 $\Omega$ -sGFP(S65T) previamente digerido con *Sal*I-*Nco*I para crear el vector 35 $\Omega$ -IE41-sGFP(S65T), que comprende la región codificante de la proteína IE41 de Arabidopsis fusionada a la GFP. Se utilizó un protocolo similar para otras construcciones:

\* La secuencia que codifica la proteína IE41 de Arabidopsis desprovista de los primeros 31 aminoácidos se obtiene mediante amplificación por PCR usando los siguientes dos cebadores:

40 *Sal*I-N-ter CGGTTGTCGACATGAAGAGTAATGAGGTTTGCCTG (SEQ ID NO: 14)  
*Nco*I-C-ter CAACCCATGGATGGCTCGACAATGATCTTC (SEQ ID NO: 13).

El plásmido 35 $\Omega$ - $\Delta$ (1-31)IE41-sGFP(S65T) se obtiene mediante la inserción de esta secuencia en el plásmido 35 $\Omega$ -sGFP(S65T).

\* La secuencia que codifica la proteína IE41 de Arabidopsis desprovista de los primeros 59 aminoácidos se amplifica por PCR usando los siguientes dos cebadores:

50 *Sal*I-N-ter GAATGGTTCGACATGTTTCTGCCCGCAAGTTC (SEQ ID NO: 15), y  
*Nco*I-C-ter CAACCCATGGATGGCTCGACAATGATCTTC (SEQ ID NO: 13).

El plásmido 35 $\Omega$ - $\Delta$ (1-59)IE41-sGFP(S65T) se obtiene mediante la inserción de esta secuencia en el plásmido 35 $\Omega$ -sGFP(S65T).

\* La secuencia que codifica la proteína IE41 de Arabidopsis desprovista de los primeros 99 aminoácidos se amplifica por PCR usando los siguientes dos cebadores:

60 *Sal*I-N-ter GGTGTCGACATGCTAGGTGGAGGTGGACTTG (SEQ ID NO: 16)  
*Nco*I-C-ter CAACCCATGGATGGCTCGACAATGATCTTC (SEQ ID NO: 13).

El plásmido 35Ω-Δ(1-99)IE41-sGFP(S65T) se obtiene mediante la inserción de esta secuencia en el plásmido 35Ω-sGFP(S65T).

5 \* La secuencia que codifica los aminoácidos 6-100 de la proteína IE41 de *Arabidopsis* se amplifica por PCR usando los siguientes dos cebadores:

*XhoI-N-ter* CCTCTCGAGATGGCTGGAAAACTCATGCAC (SEQ ID NO: 17)

*NcoI-C-ter* ACCCATGGCTAGATGGCTAAGAACCGCTAC (SEQ ID NO: 18).

10 El cebador SEQ ID NO: 17 comprende un nucleótido adicional con respecto al cebador SEQ ID NO: 15, que crea un desplazamiento en el marco de lectura en el producto de amplificación, cuya traducción se inicia en el codón ATG correspondiente a la metionina en posición 6 de la proteína IE41.

15 El plásmido 35Ω-(6-100)IE41-sGFP(S65T) se obtiene mediante la inserción de esta secuencia en el plásmido 35Ω-sGFP(S65T).

\* La secuencia que codifica los aminoácidos 60-100 de la proteína IE41 de *Arabidopsis* se amplifica por PCR usando los siguientes dos cebadores:

20 *Sall-N-ter* GAATGGTCCGACATGTTTCTGCCCGCAAGTTC (SEQ ID NO: 15)

*NcoI-C-ter* ACCCATGGCTAGATGGCTAAGAACCGCTAC (SEQ ID NO: 18).

25 El plásmido 35Ω-(60-100)IE41-sGFP(S65T) se obtiene mediante la inserción de esta secuencia en el plásmido 35Ω-sGFP(S65T).

Estas distintas construcciones se muestran en la Figura 8.

Leyenda de la Figura 8:

30 IE41 = plásmido 35Ω-IE41-sGFP(S65T)  
 Δ(1-31)IE41 = plásmido 35Ω-Δ(1-31)IE41-sGFP(S65T)  
 Δ(1-59)IE41 = plásmido 35Ω-Δ(1-59)IE41-sGFP(S65T)  
 Δ(1-99)IE41 = plásmido 35Ω-Δ(1-99)IE41-sGFP(S65T)  
 (6-100)IE41 = plásmido 35Ω-(6-100)IE41-sGFP(S65T)  
 35 (60-100)IE41 = plásmido 35Ω-(60-100)IE41-sGFP(S65T)

### **Bombardeo de células de *Arabidopsis* y de tabaco**

40 Las células de *Arabidopsis* se cultivan con luz durante 3 días en medio B5 de Gamborg (Sigma, pH 5,8) suplementado con sacarosa al 1,5 % y ANA 1 μM (ácido naftalenacético). Se transfirieron 15 ml de suspensión de células (correspondiente a 0,5 g aproximadamente) a placas Petri que contenían el mismo medio de crecimiento suplementado con Bacto-agar al 0,8 % y se incubaron durante 18 – 36 h a la luz.

45 Se cultivaron células BY2 de tabaco durante 5 días a 27 °C en medio de Murashige y Skoog (medio MS, Duchefa, pH 5,8) suplementado con sacarosa al 3 %, 0,2 % de KH<sub>2</sub>PO<sub>4</sub>, 0,2 % de mioinositol, 2.4D 1 μM (ácido 2,4-diclorofenoxiacético) y tiamina 3 μM. Se transfirieron 2,5 ml de suspensión de células (correspondiente a 0,3 g aproximadamente) a placas Petri que contenían el mismo medio de crecimiento suplementado con Bacto-agar al 1 % y se pusieron a 27 °C durante 18 – 24 h.

50 Los plásmidos que contienen las construcciones a probar utilizadas para el bombardeo tisular se preparan usando el "QIAfilter Plasmid Midi Kit" (Qiagen, Alemania).

El plásmido [35Ω-sGFP(S65T)] (GFP) y el plásmido [35Ω-TP-sGFP(S65T)] (TP-GFP) se utilizan respectivamente como control negativo y control positivo.

55 Los plásmidos (1 μg) se introdujeron en las células utilizando una pistola de partículas neumática (PDS-1000/He, BioRad). Las condiciones de bombardeo son las siguientes: se utilizó una presión de helio de 1.350 psi, discos de ruptura de 1.100 psi (Biorad); una distancia al objetivo de 10 cm; y microportas de oro de 1 μm (Biorad). Después del bombardeo, las células se incubaron en placas de Petri durante 18 – 36 h (a la luz para las células de *Arabidopsis*), y a continuación se transfirieron a portaobjetos de vidrio antes de la microscopía de fluorescencia.

60

**Microscopía de fluorescencia**

La localización de la GFP y de los péptidos de fusión a la GFP se analizó en las células transformadas por microscopía de fluorescencia utilizando un microscopio de fluorescencia Zeiss Axioplan 2 y una cámara digital CCD (Hamamatsu). Los juegos de filtros utilizados son: Juego de filtros Zeiss 13, 488013 – 0000 (excitador BP 470/20, divisor de haz FT 493, emisor BP 505 – 530) y juego de filtros Zeiss 15, 488015 – 0000 (excitador BP 546/12, divisor de haz FT 580, emisor LP 590) para la GFP y la autofluorescencia de las clorofilas, respectivamente.

En estas condiciones, la presencia de clorofila (específicamente localizada en los cloroplastos) y la localización de la GFP en la célula se visualizan por una fluorescencia intensa.

En las células de *Arabidopsis* transformadas con las construcciones GFP,  $\Delta(1-99)$ IE41, y (60-100)IE41, la fluorescencia de la GFP es difusa y localizada en el citosol y el núcleo, no se observa la co-localización con la clorofila.

En las células de *Arabidopsis* transformadas con las construcciones IE41,  $\Delta(1-31)$ IE41,  $\Delta(1-59)$ IE41, (6-100)IE41, y con el control positivo de localización TP-GFP, por el contrario se observa una co-localización, en los cloroplastos, entre la fluorescencia de GFP y la autofluorescencia de la clorofila.

Los resultados son similares en las células BY2 del tabaco no clorofílicas: los marcajes fluorescentes observados con las construcciones IE41,  $\Delta(1-59)$ IE41 y (6-100)IE41, y con el control positivo de localización TP-GFP, corresponden a una localización plastidial; sin embargo, los marcajes fluorescentes observados con las construcciones GFP,  $\Delta(1-99)$ IE41, y (60-100)IE41 corresponden a la localización citosólica y nuclear.

Estos experimentos muestran que el direccionamiento también es eficaz en plástidos no clorofílicos.

Todos los resultados anteriores demuestran que:

- la proteína IE41 completa fusionada a la GFP está dirigida hacia el cloroplasto;
- los 59 restos localizados en el extremo N-terminal no son esenciales para la importación;
- los 99 restos localizados en la región N-terminal contienen una región esencial para la importación;
- una secuencia de 94 restos, que corresponde a los aminoácidos N-terminales 6-100 es suficiente para catalizar la importación, los 223 restos (101-323) C-terminales no son esenciales para la importación.

La secuencia interna de 40 aminoácidos que se extiende entre los restos 60 y 100, corresponde al dominio esencial para la importación. Sin embargo, este dominio, que debe estar presente para dirigir la proteína a los plástidos, no es suficiente para un buen direccionamiento.

**EJEMPLO 6: ANÁLISIS EN LA PLANTA DEL DIRECCIONAMIENTO PLASTIDIAL DE LA PROTEÍNA IE41**

Los plásmidos 35 $\Omega$ -IE41-sGFP(S65T), 35 $\Omega$ - $\Delta(1-31)$ IE41-sGFP(S65T), 35 $\Omega$ - $\Delta(1-59)$ IE41-sGFP(S65T), 35 $\Omega$ - $\Delta(1-99)$ IE41-sGFP(S65T), 35 $\Omega$ -(6-100)IE41-sGFP(S65T), y 35 $\Omega$ -(60-100)IE41-sGFP(S65T), así como los plásmidos control 35 $\Omega$ -sGFP(S65T) y 35 $\Omega$ -TP-sGFP(S65T), se digirieron con *EcoRI*/*HindIII* para recuperar los casetes de expresión.

Éstos se insertaron en el plásmido binario pEL103 (derivado del plásmido pBI121 (AF485783), que contiene un gen de resistencia a la kanamicina), y el plásmido resultante se utilizó para transformar la cepa C58 de *Agrobacterium tumefaciens* mediante electroporación. Las bacterias transformadas se utilizaron para transformar plantas de *Arabidopsis* WS mediante la técnica de "inmersión floral" (The Plant Journal 1998; 16: 735 – 743). Las plantas transgénicas se seleccionaron en base a su resistencia a la kanamicina.

Para analizar la expresión de las proteínas de fusión en las plantas transgénicas, se extrajeron las proteínas totales a partir de 10 mg de hojas de cada una de las plantas sometidas a ensayo, y se disolvieron en el tampón siguiente: pirofosfato tetrasódico (13,4 g/l), Tris-HCl pH 6,8 (50 mM), SDS (1 %).

El extracto de proteínas se analizó por SDS-PAGE (12 % de acrilamida) y transferencia de Western utilizando un anticuerpo dirigido contra GFP (anticuerpo 2A5 (Euromedex) a 1/4000<sup>e</sup> en TBST/ 5% de leche; anticuerpo secundario: IgG dirigido contra ratón conjugado con fosfatasa alcalina (Promega) diluido a 1/10000<sup>e</sup> en TBS-Triton), o anticuerpos policlonales de conejo dirigidos contra IE41 descritos en el Ejemplo 1 (a 1/5000<sup>e</sup> en TBS-Triton/5 % de leche; anticuerpo secundario IgG dirigido contra conejo conjugado con fosfatasa alcalina (Promega) diluido a 1/10000<sup>e</sup> en tampón TBS-Triton).

Los resultados de estos análisis se muestran en la Figura 9;

Leyenda de la Figura 9:

- 5        - A: análisis por SDS-PAGE;  
 - B: transferencia de Western con anticuerpos dirigidos contra GFP: las flechas negras indican la presencia de la proteína GFP en las fusiones expresadas en *Arabidopsis*;  
 - C: transferencia de Western con un anticuerpo dirigido contra IE41: las flechas negras indican la presencia de la proteína IE41 en las fusiones expresadas en *Arabidopsis*, el diamante blanco indica la posición de la proteína IE41 natural presente en todos los extractos;
- 10        - WT = planta no transformada;  
 - M = marcadores de pesos moleculares;  
 - GFP = plásmido 35Ω-sGFP(S65T);  
 - TP GFP = plásmido 35Ω-TP-sGFP(S65T);
- 15        - IE41 GFP = plásmido 35Ω-IE41-sGFP(S65T);  
 - Δ(1-59)IE41 GFP = plásmido 35Ω-Δ(1-59)IE41-sGFP(S65T);  
 - Δ(1-99)IE41 GFP = plásmido 35Ω-Δ(1-99)IE41-sGFP(S65T);  
 - (6-100)IE41 GFP = plásmido 35Ω-(6-100)IE41-sGFP(S65T);  
 - (60-100)IE41 GFP = plásmido 35Ω-(60-100)IE41-sGFP(S65T).

20        Estos resultados muestran que las proteínas de fusión se expresan en todas las plantas transformadas.

La localización subcelular de las proteínas expresadas por las diferentes construcciones se visualizó por microscopía de fluorescencia como se describe en el Ejemplo 5 anterior.

25        Los resultados se muestran en la Figura 10; leyenda de la Figura 10:

- 30        - GFP = plásmido 35Ω-sGFP(S65T);  
 - TP-RBCS GFP = plásmido 35Ω-TP-sGFP(S65T);  
 - IE41 GFP = plásmido 35Ω-IE41-sGFP(S65T);  
 - (6-100)IE41 GFP = plásmido 35Ω-(6-100)IE41-sGFP(S65T);

35        Sucede que en las condiciones de expresión utilizadas anteriormente, es la región N-terminal de la proteína IE41 (restos 6-100) la que confiere una mayor especificidad de direccionamiento al plástido. De hecho, la construcción (6-100) IE41-GFP, que sólo expresa esta región, permite el direccionamiento sistemático de toda la fluorescencia hacia los plástidos. Por contra, la proteína IE41 completa (construcción IE41-GFP) induce un direccionamiento plastidial menos específico. En estas condiciones, la proteína IE41 también parece dirigirse a otros compartimentos intracelulares.

40        **LISTADO DE SECUENCIAS**

< 110 > GENOPLANTE-VALOR  
 MIRAS, Stéphane  
 SALVI, Daniel

45        ROLLAND, Norbert  
 JOYARD, Jacques  
 FERRO, Myriam  
 GARIN, Jérôme  
 GRUNWALD, Didier

50        < 120 > PEPTIDO DE DIRECCIONAMIENTO PLASTIDIAL  
 < 130 > MJPbv1516/7  
 < 150 > FR 02 07729  
 < 151 > 2002-06-21

55        < 160 > 18  
 < 170 > PatentIn version 3.1  
 < 210 > 1  
 < 211 > 329

60        < 212 > PRT  
 < 213 > *Arabidopsis thaliana*

< 400 > 1

```

Met Ala Gly Lys Leu Met His Ala Leu Gln Tyr Asn Ser Tyr Gly Gly
1           5           10           15
Gly Ala Ala Gly Leu Glu His Val Gln Val Pro Val Pro Thr Pro Lys
           20           25           30
Ser Asn Glu Val Cys Leu Lys Leu Glu Ala Thr Ser Leu Asn Pro Val
           35           40           45
Asp Trp Lys Ile Gln Lys Gly Met Ile Arg Pro Phe Leu Pro Arg Lys
50           55           60
    
```



```

ccatcctaatac gactcact atagggctcg agcggccgcc cgggcaggtc aaactgtggt      60
aagataatac agtaccatta ccatctgacg cgcaaatggc tgctaagcta atgcatgcca      120
ttcaatattc tggctatggt ggtggaactg atgctttaa gcatggtgaa gttgctgttc      180
ctgatccaaa gtctgatgag ttattgctta aaattgaggc tgcaactttg aaccaattg      240
attggaagat tcagaagggt gtacttcgtc ccctcttacc cgcgaagttc cctactatac      300
ctggaactga tgttgctggg gaggtagtcc aggctggatc tgctgtaa at aggtttaaaa      360
ctggtgacaa agtcgtggcc gtgcttagtc atgctactgg ggggtgacta gctgaatatg      420
ccgtggcgaa ggagaacctg acagttgcta gaccaccaga agtatcagca gcagaagggtg      480
ctgccttacc tgttgctgcc ctcaaggctc accaagctct caccagttt gccaacatca      540
agctcgatgg aagtggtgaa aggaagaaca tattgatcac ggctgcatca gggggtgtgg      600
gccactatgc ggtccagctg gcaaagctcg ggaacacgca tgtaacagca acatgtggag      660
cccgcaacct agatttcgtg aaaggcttgg gtgccgatga ggttcttgac tacaaaacac      720
ctgaaggggc gtccttgaca agcccgtcag gaaagaaata tgactacgta gtccacgggtg      780
caagcggaat cccttgggtc acctttgagc ccaatttgag tgaagcaggt aaggtaatag      840
atgtgactcc tggcccaact gcaatgatga catttgcttg gaaaaagcta acattctcca      900
aaaagcagct ggtgcctctg cttttgatac caaagatccc caactttgaa tatgttgatga      960
atgttgtaaa ggaaaagaag cttaaaacag tcatagactc taaacatccc ttgagtaaag     1020
gtgaagatgc ttggagtagg ataatgggtg gtcatgctac agggaagatt ataatcgagc     1080
cttgaataga aatatattgat gcagaccgcg tatatattgc ttgaaggta caaactttta     1140
agtttatagt acttgagtta tactttccta gttgtaaaca ttcaagtatt tcataatggt     1200
atactttcct agtttcctcc aaaaaaaaaa                                     1228

```

- < 210 > 3
- < 211 > 329
- < 212 > PRT
- 5 < 213 > Spinacia oleracea
- < 400 > 3

```

Met Ala Ala Lys Leu Met His Ala Ile Gln Tyr Ser Gly Tyr Gly Gly
 1           5           10          15
Gly Thr Asp Ala Leu Lys His Val Glu Val Ala Val Pro Asp Pro Lys
          20          25          30
Ser Asp Glu Leu Leu Leu Lys Ile Glu Ala Ala Thr Leu Asn Pro Ile
          35          40          45

```

Asp Trp Lys Ile Gln Lys Gly Val Leu Arg Pro Leu Leu Pro Arg Lys  
 50 55 60  
 Phe Pro Thr Ile Pro Gly Thr Asp Val Ala Gly Glu Val Val Gln Ala  
 65 70 75 80  
 Gly Ser Ala Val Asn Arg Phe Lys Thr Gly Asp Lys Val Val Ala Val  
 85 90 95  
 Leu Ser His Ala Thr Gly Gly Ala Leu Ala Glu Tyr Ala Val Ala Lys  
 100 105 110  
 Glu Asn Leu Thr Val Ala Arg Pro Pro Glu Val Ser Ala Ala Glu Gly  
 115 120 125  
 Ala Ala Leu Pro Val Ala Ala Leu Thr Ala His Gln Ala Leu Thr Gln  
 130 135 140  
 Phe Ala Asn Ile Lys Leu Asp Gly Ser Gly Glu Arg Lys Asn Ile Leu  
 145 150 155 160  
 Ile Thr Ala Ala Ser Gly Gly Val Gly His Tyr Ala Val Gln Leu Ala  
 165 170 175  
 Lys Leu Gly Asn Thr His Val Thr Ala Thr Cys Gly Ala Arg Asn Leu  
 180 185 190  
 Asp Phe Val Lys Gly Leu Gly Ala Asp Glu Val Leu Asp Tyr Lys Thr  
 195 200 205  
 Pro Glu Gly Ala Ser Leu Thr Ser Pro Ser Gly Lys Lys Tyr Asp Tyr  
 210 215 220  
 Val Val His Gly Ala Ser Gly Ile Pro Trp Ser Thr Phe Glu Pro Asn  
 225 230 235 240  
 Leu Ser Glu Ala Gly Lys Val Ile Asp Leu Thr Pro Gly Pro Thr Ala  
 245 250 255  
 Met Met Thr Phe Ala Trp Lys Lys Leu Thr Phe Ser Lys Lys Gln Leu  
 260 265 270  
 Val Pro Leu Leu Leu Ile Pro Lys Ile Pro Asn Phe Glu Tyr Val Val  
 275 280 285  
 Asn Leu Val Lys Glu Lys Lys Leu Lys Thr Val Ile Asp Ser Lys His  
 290 295 300  
 Pro Leu Ser Lys Gly Glu Asp Ala Trp Ser Arg Ile Met Gly Gly His  
 305 310 315 320  
 Ala Thr Gly Lys Ile Ile Ile Glu Pro  
 325

ES 2 394 084 T3

< 211 > 41  
 < 212 > PRT  
 < 213 > Arabidopsis thaliana  
 < 400 > 4

Phe Leu Pro Arg Lys Phe Pro Cys Ile Pro Ala Thr Asp Val Ala Gly  
 1 5 10 15

Glu Val Val Glu Val Gly Ser Gly Val Lys Asn Phe Lys Ala Gly Asp  
 20 25 30

Lys Val Val Ala Val Leu Ser His Leu  
 35 40

5

< 210 > 5  
 < 211 > 41  
 < 212 > PRT  
 < 213 > Spinacia oleracea  
 < 400 > 5

10

Leu Leu Pro Arg Lys Phe Pro Thr Ile Pro Gly Thr Asp Val Ala Gly  
 1 5 10 15

Glu Val Val Gin Ala Gly Ser Ala Val Asn Arg Phe Lys Thr Gly Asp  
 20 25 30

Lys Val Val Ala Val Leu Ser His Ala  
 35 40

< 210 > 6  
 < 211 > 327  
 < 212 > PRT  
 < 213 > Escherichia coli  
 < 400 > 6

15

Met Ala Thr Arg Ile Glu Phe His Lys His Gly Gly Pro Glu Val Leu  
 1                                   5                                   10                                   15  
 Gln Ala Val Glu Phe Thr Pro Ala Asp Pro Ala Glu Asn Glu Ile Gln  
                                   20                                   25                                   30  
 Val Glu Asn Lys Ala Ile Gly Ile Asn Phe Ile Asp Thr Tyr Ile Arg  
                                   35                                   40                                   45  
 Ser Gly Leu Tyr Pro Pro Pro Ser Leu Pro Ser Gly Leu Gly Thr Glu  
                                   50                                   55                                   60  
 Ala Ala Gly Ile Val Ser Lys Val Gly Ser Gly Val Lys His Ile Lys  
 65                                   70                                   75                                   80  
 Ala Gly Asp Arg Val Val Tyr Ala Gln Ser Ala Leu Gly Ala Tyr Ser  
                                   85                                   90                                   95  
 Ser Val His Asn Ile Ile Ala Asp Lys Ala Ala Ile Leu Pro Ala Ala  
                                   100                                   105                                   110  
 Ile Ser Phe Glu Gln Ala Ala Ala Ser Phe Leu Lys Gly Leu Thr Val  
                                   115                                   120                                   125  
 Tyr Tyr Leu Leu Arg Lys Thr Tyr Glu Ile Lys Pro Asp Glu Gln Phe  
                                   130                                   135                                   140  
 Leu Phe His Ala Ala Ala Gly Gly Val Gly Leu Ile Ala Cys Gln Trp  
 145                                   150                                   155                                   160  
 Ala Lys Ala Leu Gly Ala Lys Leu Ile Gly Thr Val Gly Thr Ala Gln  
                                   165                                   170                                   175  
 Lys Ala Gln Ser Ala Leu Lys Ala Gly Ala Trp Gln Val Ile Asn Tyr  
                                   180                                   185                                   190  
 Arg Glu Glu Asp Leu Val Glu Arg Leu Lys Glu Ile Thr Gly Gly Lys  
                                   195                                   200                                   205  
 Lys Val Arg Val Val Tyr Asp Ser Val Gly Arg Asp Thr Trp Glu Arg  
                                   210                                   215                                   220  
 Ser Leu Asp Cys Leu Gln Arg Arg Gly Leu Met Val Ser Phe Gly Asn  
 225                                   230                                   235                                   240  
 Ser Ser Gly Ala Val Thr Gly Val Asn Leu Gly Ile Leu Asn Gln Lys  
                                   245                                   250                                   255  
 Gly Ser Leu Tyr Val Thr Arg Pro Ser Leu Gln Gly Tyr Ile Thr Thr  
                                   260                                   265                                   270  
 Arg Glu Glu Leu Thr Glu Ala Ser Asn Glu Leu Phe Ser Leu Ile Ala  
                                   275                                   280                                   285

Ser Gly Val Ile Lys Val Asp Val Ala Glu Gln Gln Lys Tyr Pro Leu  
 290 295 300  
 Lys Asp Ala Gln Arg Ala His Glu Ile Leu Glu Ser Arg Ala Thr Gln  
 305 310 315 320  
 Gly Ser Ser Leu Leu Ile Pro  
 325

< 210 > 7  
 < 211 > 334  
 < 212 > PRT  
 5 < 213 > Saccharomyces cerevisiae  
 < 400 > 7

Met Lys Cys Thr Ile Pro Glu Gln Gln Lys Val Ile Leu Ile Asp Glu  
 1 5 10 15  
 Ile Gly Gly Tyr Asp Val Ile Lys Tyr Glu Asp Tyr Pro Val Pro Ser  
 20 25 30  
 Ile Ser Glu Glu Glu Leu Leu Ile Lys Asn Lys Tyr Thr Gly Val Asn  
 35 40 45  
 Tyr Ile Glu Ser Tyr Phe Arg Lys Gly Ile Tyr Pro Cys Glu Lys Pro  
 50 55 60  
 Tyr Val Leu Gly Arg Glu Ala Ser Gly Thr Val Val Ala Lys Gly Lys  
 65 70 75 80  
 Gly Val Thr Asn Phe Glu Val Gly Asp Gln Val Ala Tyr Ile Ser Asn  
 85 90 95  
 Ser Thr Phe Ala Gln Tyr Ser Lys Ile Ser Ser Gln Gly Pro Val Met  
 100 105 110  
 Lys Leu Pro Lys Gly Thr Ser Asp Glu Glu Leu Lys Leu Tyr Ala Ala  
 115 120 125  
 Gly Leu Leu Gln Val Leu Thr Ala Leu Ser Phe Thr Asn Glu Ala Tyr  
 130 135 140  
 His Val Lys Lys Gly Asp Tyr Val Leu Leu Phe Ala Ala Ala Gly Gly  
 145 150 155 160  
 Val Gly Leu Ile Leu Asn Gln Leu Leu Lys Met Lys Gly Ala His Thr  
 165 170 175  
 Ile Ala Val Ala Ser Thr Asp Glu Lys Leu Lys Ile Ala Lys Glu Tyr  
 180 185 190

Gly Ala Glu Tyr Leu Ile Asn Ala Ser Lys Glu Asp Ile Leu Arg Gln  
 195 200 205  
 Val Leu Lys Phe Thr Asn Gly Lys Gly Val Asp Ala Ser Phe Asp Ser  
 210 215 220  
 Val Gly Lys Asp Thr Phe Glu Ile Ser Leu Ala Ala Leu Lys Arg Lys  
 225 230 235 240  
 Gly Val Phe Val Ser Phe Gly Asn Ala Ser Gly Leu Ile Pro Pro Phe  
 245 250 255  
 Ser Ile Thr Arg Leu Ser Pro Lys Asn Ile Thr Leu Val Arg Pro Gln  
 260 265 270  
 Leu Tyr Gly Tyr Ile Ala Asp Pro Glu Glu Trp Lys Tyr Tyr Ser Asp  
 275 280 285  
 Glu Phe Phe Gly Leu Val Asn Ser Lys Lys Leu Asn Ile Lys Ile Tyr  
 290 295 300  
 Lys Thr Tyr Pro Leu Arg Asp Tyr Arg Thr Ala Ala Ala Asp Ile Glu  
 305 310 315 320  
 Ser Arg Lys Thr Val Gly Lys Leu Val Leu Glu Ile Pro Gln  
 325 330

- < 210 > 8
- < 211 > 329
- < 212 > PRT
- 5 < 213 > *Cavia porcellus*
- < 400 > 8

Met Ala Thr Gly Gln Lys Leu Met Arg Ala Ile Arg Val Phe Glu Phe  
 1 5 10 15  
 Gly Gly Pro Glu Val Leu Lys Val Gln Ser Asp Val Ala Val Pro Ile  
 20 25 30  
 Pro Lys Asp His Gln Val Leu Ile Lys Val His Ala Cys Gly Ile Asn  
 35 40 45  
 Pro Val Glu Thr Tyr Ile Arg Ser Gly Thr Tyr Thr Arg Ile Pro Leu  
 50 55 60  
 Leu Pro Tyr Thr Pro Gly Thr Asp Val Ala Gly Val Val Glu Ser Ile  
 65 70 75 80  
 Gly Asn Asp Val Ser Ala Phe Lys Lys Gly Asp Arg Val Phe Thr Thr  
 85 90 95

```

Ser Thr Ile Ser Gly Gly Tyr Ala Glu Tyr Ala Leu Ala Ser Asp His
      100                      105                      110
Thr Val Tyr Arg Leu Pro Glu Lys Leu Asp Phe Arg Gln Gly Ala Ala
      115                      120                      125
Ile Gly Ile Pro Tyr Phe Thr Ala Cys Arg Ala Leu Phe His Ser Ala
      130                      135                      140
Arg Ala Lys Ala Gly Glu Ser Val Leu Val His Gly Ala Ser Gly Gly
145                      150                      155                      160
Val Gly Leu Ala Ala Cys Gln Ile Ala Arg Ala Tyr Gly Leu Lys Val
      165                      170                      175
Leu Gly Thr Ala Gly Thr Glu Glu Gly Gln Lys Val Val Leu Gln Asn
      180                      185                      190
Gly Ala His Glu Val Phe Asn His Arg Asp Ala His Tyr Ile Asp Glu
      195                      200                      205
Ile Lys Lys Ser Ile Gly Glu Lys Gly Val Asp Val Ile Ile Glu Met
      210                      215                      220
Leu Ala Asn Val Asn Leu Ser Asn Asp Leu Lys Leu Leu Ser Cys Gly
225                      230                      235                      240
Gly Arg Val Ile Ile Val Gly Cys Arg Gly Ser Ile Glu Ile Asn Pro
      245                      250                      255
Arg Asp Thr Met Ala Lys Glu Ser Thr Ile Ser Gly Val Ser Leu Phe
      260                      265                      270
Ser Ser Thr Lys Glu Glu Phe Gln Gln Phe Ala Ser Thr Ile Gln Ala
      275                      280                      285
Gly Met Glu Leu Gly Trp Val Lys Pro Val Ile Gly Ser Gln Tyr Pro
      290                      295                      300
Leu Glu Lys Ala Ser Gln Ala His Glu Asn Ile Ile His Ser Ser Gly
305                      310                      315                      320
Thr Val Gly Lys Thr Val Leu Leu Met
      325

```

< 210 > 9  
 < 211 > 331  
 < 212 > PRT  
 < 213 > Mus musculus  
 < 400 > 9

5





< 223 > Cebador de la PCR  
< 400 > 16  
ggtgtcgac atgctaggtg gaggtggact tg 32  
5 < 210 > 17  
< 211 > 31  
< 212 > ADN  
< 213 > Secuencia artificial  
< 220 >  
10 < 223 > Cebador de la PCR  
< 400 > 17  
cctctcgaga tggctggaaa aactcatgca c 31  
< 210 > 18  
< 211 > 30  
< 212 > ADN  
15 < 213 > Secuencia artificial  
< 220 >  
< 223 > Cebador de la PCR  
< 400 > 18  
20 acctatggct agatggctaa gaaccgctac 30

## REIVINDICACIONES

1. Polipéptido de direccionamiento intraplasmático, **caracterizado porque** está constituido por:
- 5 - un dominio A constituido por un polipéptido que presenta una identidad de al menos el 60 %, o una similitud de al menos el 65 %, con uno de los polipéptidos de las SEQ ID NO: 4 o SEQ ID NO: 5;
- y al menos un dominio seleccionado entre:
- 10 - un dominio B situado en el extremo N-terminal del dominio A, y constituido por un fragmento de uno de los polipéptidos de las SEQ ID NO: 1 o SEQ ID NO: 3 que comprende al menos los aminoácidos 49 a 59 de dicho polipéptido, o alternativamente de un polipéptido que presenta una identidad de al menos el 60 %, o una similitud de al menos el 65 %, con dicho fragmento;
- 15 - un dominio C situado en el extremo C-terminal del dominio A, y constituido por un fragmento de uno de los polipéptidos de las SEQ ID NO: 1 o SEQ ID NO: 3 que comprende al menos los aminoácidos 101 a 111, y como máximo los aminoácidos 101 a 151 de dicho polipéptido, o alternativamente de un polipéptido que presenta una identidad de al menos el 60 %, o una similitud de al menos el 65 %, con dicho fragmento.
- 20 2. Polipéptido de acuerdo con la reivindicación 1, **caracterizado porque** el dominio B está constituido de un fragmento que comprende al menos los aminoácidos 39 a 59 de los polipéptidos de las SEQ ID NO: 1 o SEQ ID NO: 3, o alternativamente de un polipéptido que presenta una identidad de al menos el 60 %, o una similitud de al menos el 65 %, con dicho fragmento.
- 25 3. Polipéptido de acuerdo con una cualquiera de la reivindicación 1 o la reivindicación 2, **caracterizado porque** el dominio C está constituido por un fragmento que comprende al menos los aminoácidos 101 a 121 de los polipéptidos de las SEQ ID NO: 1 o SEQ ID NO: 3, o alternativamente de un polipéptido que presenta una identidad de al menos el 60 %, o una similitud de al menos el 65 %, con dicho fragmento.
- 30 4. Polipéptido quimérico, que comprende un polipéptido de direccionamiento intraplasmático de acuerdo con una cualquiera de las reivindicaciones 1 a 3 fusionado con una proteína heteróloga.
5. Polipéptido quimérico de acuerdo con la reivindicación 4, **caracterizado porque** el polipéptido de direccionamiento intraplasmático está situado en el extremo N-terminal de la proteína heteróloga.
- 35 6. Uso de un polipéptido de direccionamiento intraplasmático de acuerdo con la reivindicación 1 para la importación de una proteína de interés hacia plástidos.
7. Uso de acuerdo con la reivindicación 6, **caracterizado porque** dicho polipéptido de direccionamiento intraplasmático se utiliza para importar dicha proteína de interés hacia los cloroplastos.
- 40 8. Procedimiento para la importación de una proteína de interés hacia plástidos, **caracterizado porque** comprende la expresión, en una célula vegetal que contiene dichos plástidos, de un polipéptido quimérico que resulta de la fusión de un polipéptido de direccionamiento intraplasmático de acuerdo con la reivindicación 1 con dicha proteína de interés.
- 45 9. Polinucleótido codificante para un polipéptido de acuerdo con una cualquiera de las reivindicaciones 1 a 5.
- 50 10. Casete de expresión que comprende un polinucleótido de acuerdo con la reivindicación 9 situado bajo el control de secuencias para regular la transcripción.
11. Vector recombinante que resulta de la inserción de un polinucleótido de acuerdo con la reivindicación 9 o de un casete de expresión de acuerdo con la reivindicación 10 en un vector hospedador.
- 55 12. Planta transgénica transformada mediante, y que comprende, un polinucleótido de acuerdo con la reivindicación 9 o un casete de expresión de acuerdo con la reivindicación 10.

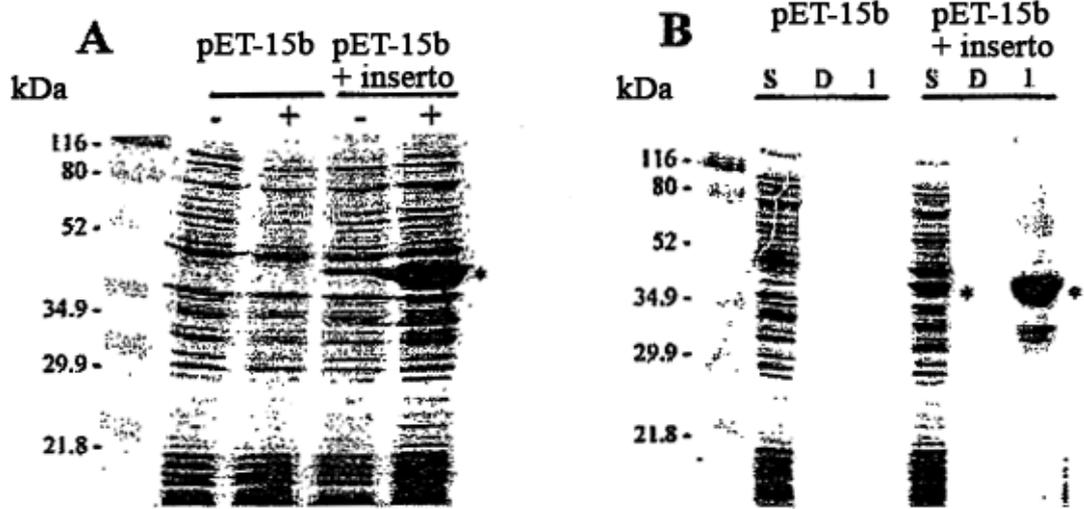


Fig . 1

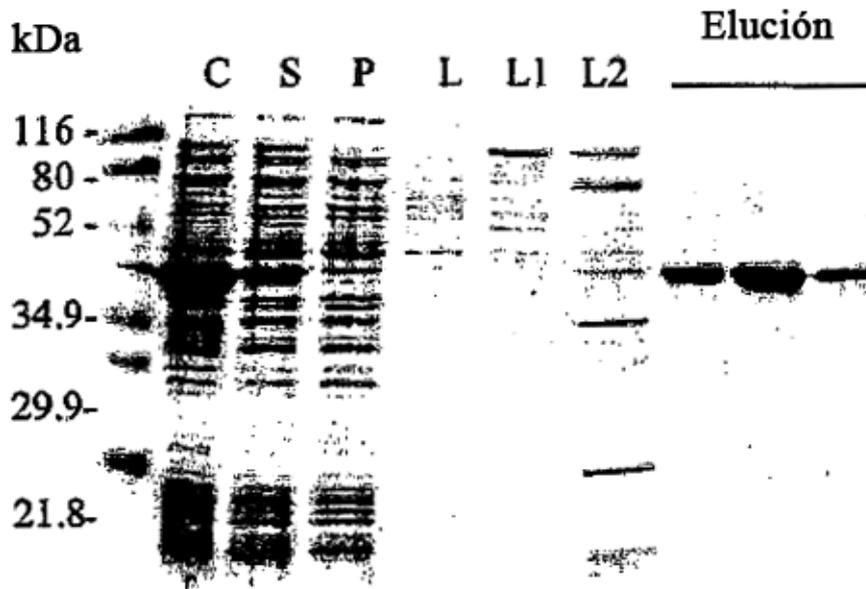


Fig . 2

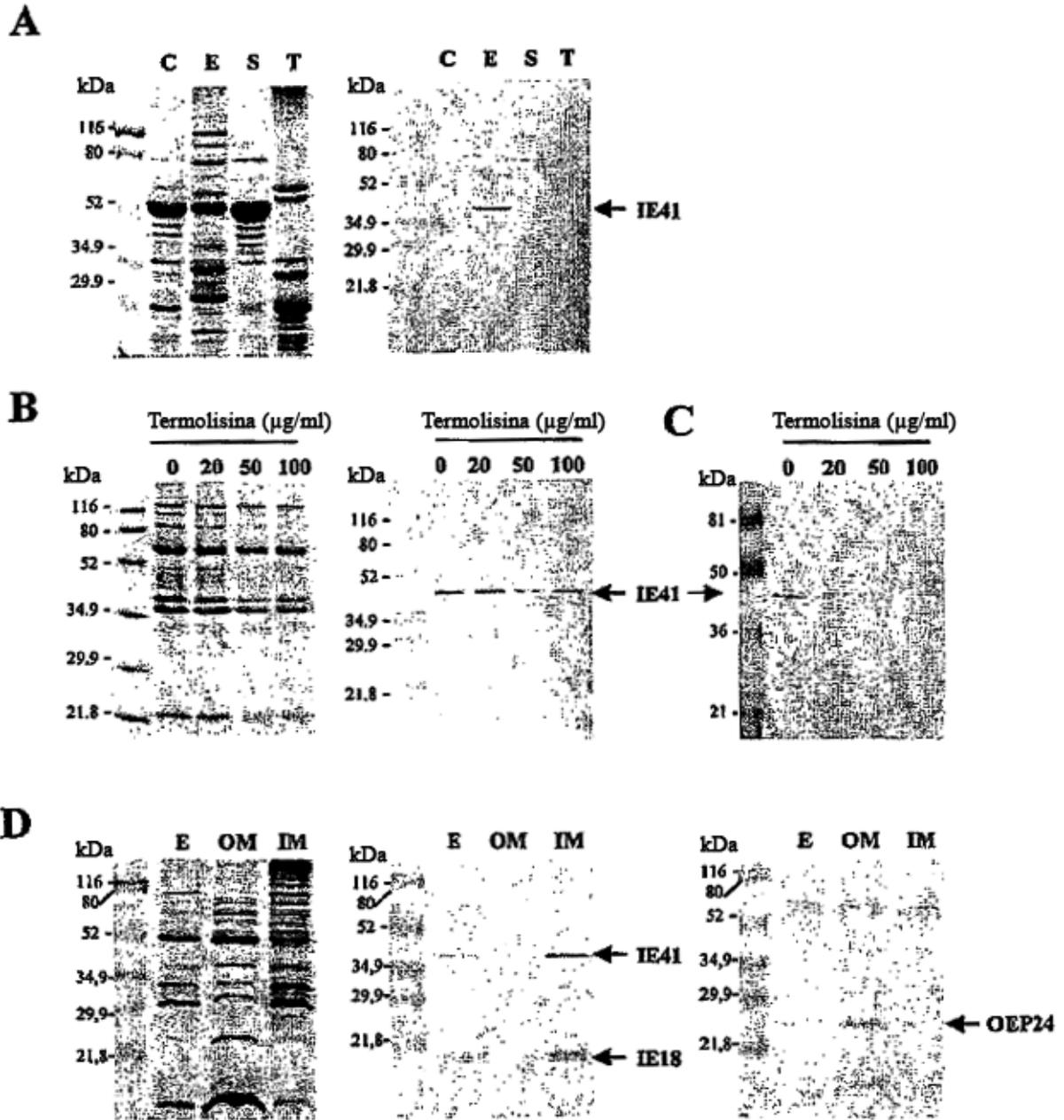


Fig . 3

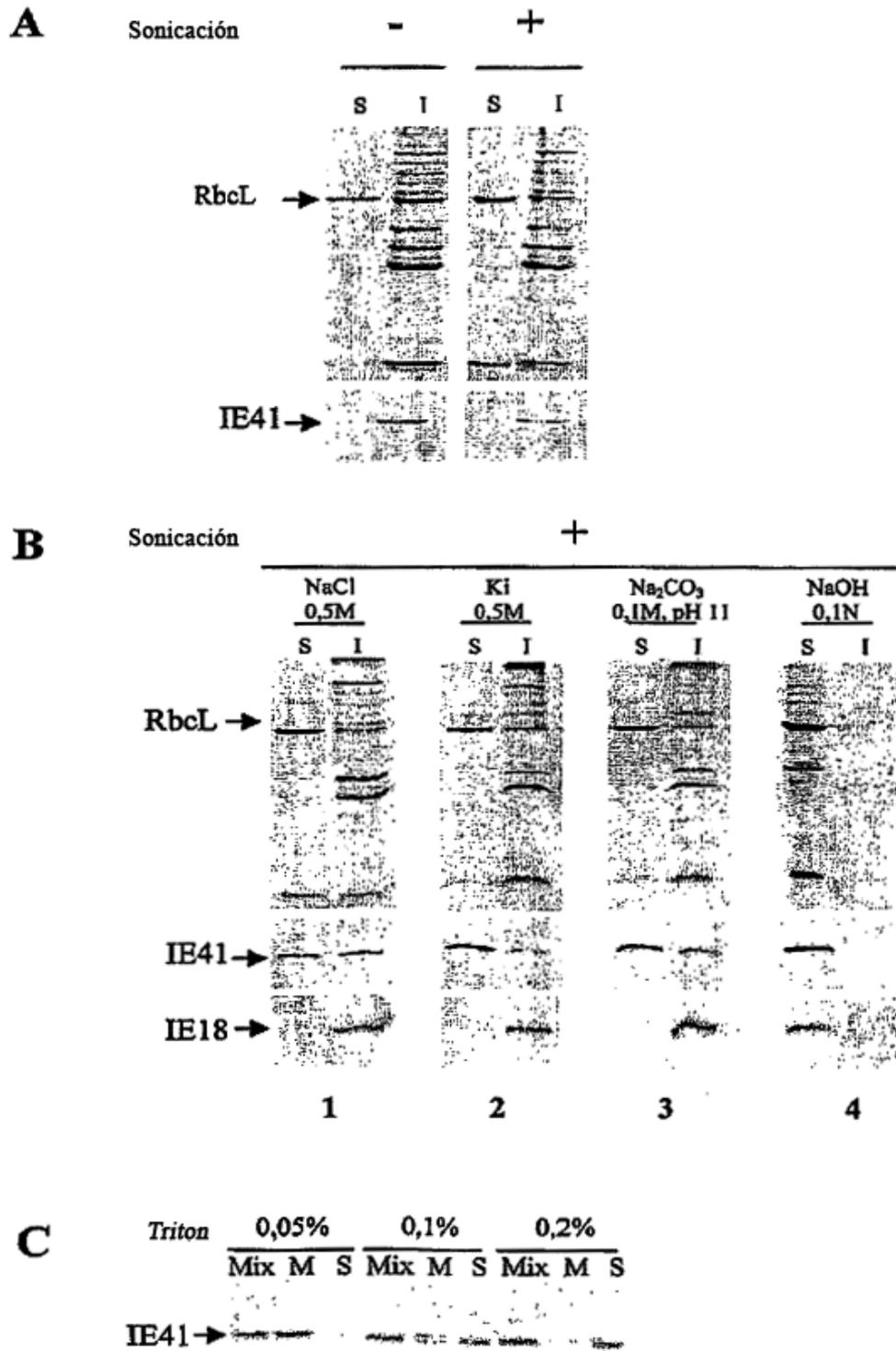


Fig. 4

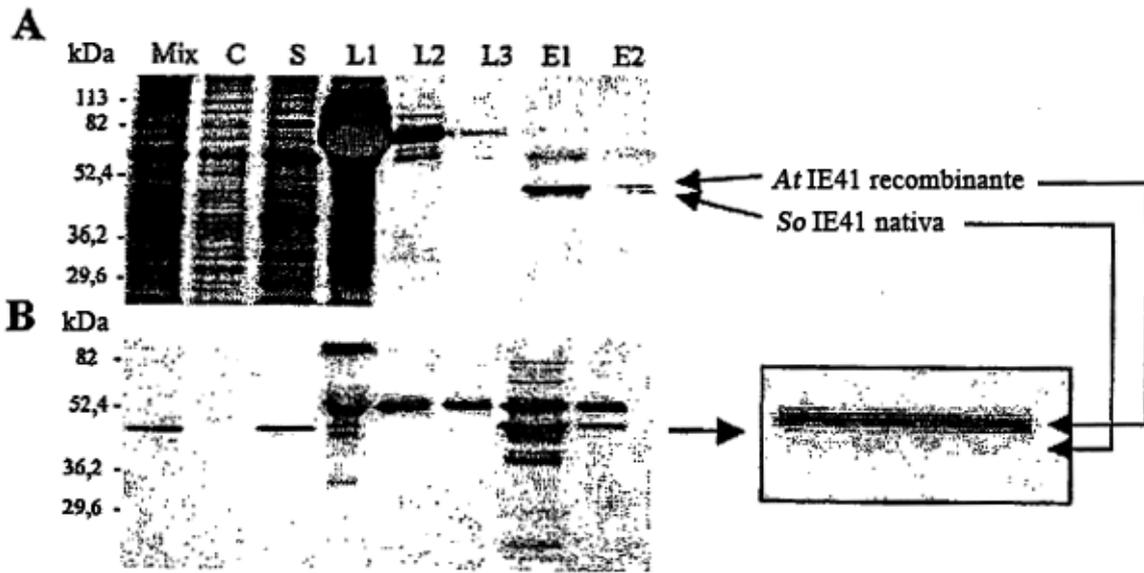


Fig . 5

	ccatcctaata	8
cgactcactatagggctogagcggcgcgcccgggcaggtcaaactgtggtaagataataca		68
gtaccattaccatctgacgcgcaaatggctgctaagctaatgcatgogattcaatattct		128
	M A A K L M H A I O Y S	12
ggctatgggtggggaactgatgctttaagcatggtgaagttgctgttctctgatccaaag		188
G Y G G G T D A L K H V E V A V P D P K		32
tctgatgagttattgcttaaaattgaggctgcaactttgaaccaattgattggaagatt		248
S D E L L L K L E A A T L N P I D W K I		52
cagaaggggtgtacttcgtccctcttaccgccgaagttcctactatacctggaactgat		308
Q K G V L R F D L P R K F P T I P G T D		72
gttgctggggaggttagtccaggctggatctgctgtaaataggtttaaaactggtgacaaa		368
V A G E V V Q A G S A V N R F K T G D K		92
gtcgtggccgtgcttagtcatgctactgggggtgcactagctgaatatgccgtggcgaag		428
V V A V L S H A T G G A L A E Y A V A K		112
gagaacctgacagttgctagaccaccagaagtatcagcagcagaaggtgctgccttacct		488
E N L T V A R P P E V S A A E G A A L P		132
gttgctgccctcacggctcaccaagctctcaccagtttgccaacatcaagctcgatgga		548
V A A L T A H Q A L T Q F A N I K L D G		152
agtggtgaaaggaagaacatattgatcacggctgcatcaggggggtggtgggccactatgcg		608
S G E R K N I L L T A A S G G V G H Y A		172
gtccagctggcaaaagctcgggaacacgcatgtaacagcaacatgtggagcccgcaaccta		668
V O L A R L G N T H V T A T C G A R N L		192
gatttcgtgaaaggcttgggtgccgatgaggttcttgactacaaaacacctgaagggcg		728
D F V K G L G A D E V L D Y K T P E G A		212
tccttgacaagcccgtcaggaaagaatatgactacgtagtccacgggtgcaagcgggaatc		788
S L T S P S G K K Y D Y V V H G A S G I		232
ccttgggtccaccttgagcccaatttgagtgaagcaggttaaggtaatagatttgactcct		848
P W S T F E P N L S E A G K V I D L T P		252
ggcccaactgcaatgatgacatttgcttggaaaaagctaacattctccaaaaagcagctg		908
G P T A M M T F A W K K L T F S K K Q L		272
gtgctctgcttttgataccaaagatccccaactttgaatatggttgtgaatttggtaaag		968
V P L L L L I P K I P N F E Y V V N L V K		292
gaaagaagcttaaaacagtcatagactctaaacatcccttgagtaaaggtgaagatgct		1028
E K K L K T V I D S K H P L S K G E D A		312
tggagtaggataatgggtggtcatgctacaggggaagattataatcgagccttgaatagaa		1088
W S R I M G G H A T G K I I E P *		329
aatattgatgcagaccgctatatattgcttgaaggttacaaacttttaagtttatagta		1148
cttgagttatactttcctagttgtaaacattcaagtatctcataatggttatactttccta		1208
gtttctccaaaaaaa		1225

Fig . 6

```

QORECOLI  ---EFHKHGG---VVOAV---EPADRAENEDQENKAI---ENF
QORYEAST MKC IPEQV LIDEIGY---VVKYE---YVPSISEEELTKNKYTVNY
QORCAVPO  ---GONM---RVFEG---VVKVQS---VVEIKDHO---VHACGNEVE
QORMOUSE  ---GOMBA---RVFEG---VVKLQSDV---VVSQSHO---VHACGNEVE
IE41 ATH  ---MHA---QNSYGGGAG---EHV---QVVE---TEKSNE---LEAT---P
IE41 SOL  ---MHA---QYSGYGGT---A---KHV---VAVRDEKSD---LEA---NE

```

```

QORECOLI  TYRSL-Y---PSL---SGLSTEAG---SKVCSG---KHICASRWV-Y---QAL
QORYEAST  SYFRGL-Y---CEK---YVLR---EASGT---VAKKKG---TNEVEPOV---AYLNE---TF
QORCAVPO  TYR---SGTYRT---LLYTP---VAGV---ESIEND---SARKKEDK---FTS---IS
QORMOUSE  TYR---SGAYRK---ALYTP---SVVCI---ESVGDK---SARKKEDK---FCYS---IVS
IE41 ATH  WK---QGMIR---FLRKF---CIP---VAGE---VEVSG---KN---AVV---HLGG
IE41 SOL  WK---QGLR---LLRKF---ETP---VAGE---VQASA---NR---ST---VAVL---HTIG

```

```

QORECOLI  ---SVHNIADK---AILE---AIS---F---Q---SFKGL---VYLL---RKY---Y---
QORYEAST  ---OYKI---SOG---MKL---KGTSD---ELKLY---AGL---QVLA---LSTNE---YHV---
QORCAVPO  ---SYEY---LASDH---NYRL---E---KL---FROD---IG---PYFT---CRA---FH---ARA---
QORMOUSE  ---SYEY---LAADD---NYPL---E---TLN---FROD---LG---PYFT---CRA---FH---ARA---
IE41 ATH  ---LET---VATEKL---KRPOEVGA---S---L---P---AGL---LOA---TNEAGL---LD
IE41 SOL  ---LEY---VAKENL---ARPEVSAA---S---L---P---AAL---HOA---TOFANI---LD

```

```

QORECOLI  ---ED---QF---FH---GG---GLI---C---W---ALGE---I---V---AQAQSAKAC---WO
QORYEAST  ---K---Y---F---AGG---LILN---IL---MKG---TI---VA---DEK---KIAKEY---CAEY
QORCAVPO  ---S---H---S---GLA---C---RAYGL---L---A---EEGQKVVLON---HE
QORMOUSE  ---S---H---S---GLA---C---RAYGL---L---A---EEGKKLVLON---HE
IE41 ATH  G---KKAN---T---S---G---HY---V---LANA---T---C---ARN---TEFV---SL---DE
IE41 SOL  G---RKN---T---S---G---HY---V---L---GNT---T---C---ARN---LDFV---KGL---DE

```

```

QORECOLI  I---NYHEEDLVER---KEITG---K---VRV---YDSV---RDTWERSLDC---ORR---L---S---FN
QORYEAST  I---NASKEDILRQVLKFTNG---GV---ASFDSV---KDT---FELSIAA---KRK---V---S---FN
QORCAVPO  F---NH---DAHYIDE---KKSIE---GV---V---IEM---L---N---V---L---S---NDLKL---SCG---R---IV---
QORMOUSE  F---NH---EANYIDK---KMSVD---NDKGV---V---IEM---L---N---V---L---S---NDLKL---SHG---R---V---W---
IE41 ATH  LDY---TPEGAA---K---S---G---K---Y---A---VHCANG---PFSV---FEFN---SENSK---DT---P
IE41 SOL  LDY---TPEGAS---T---S---S---G---K---Y---Y---VHGA---GLP---W---STFEFN---SEACK---DL---P

```

```

QORECOLI  SS---T---GVN---GILNOKGSLYVTRPSLQ---YTT---E---E---L---TEASNE---FSL---ASGV
QORYEAST  AS---L---PPE---ITEL---S---NITLVRPO---Y---Y---ADH---E---W---KY---SDEFFGLN---SK---
QORCAVPO  CR---E---INE---D---M---M---ESTISGVS---F---S---E---F---F---Q---Q---ASTF---QAG---ELGW
QORMOUSE  CR---E---INE---D---M---M---ETSIIIGVS---S---S---E---F---F---Q---Q---AGL---QAG---EKGW
IE41 ATH  GRN---W---TY---VK---I---MS---K---Q---V---L---L---L---I---R---A---N---L---E---N---T---KEGK
IE41 SOL  GP---M---M---T---W---K---L---F---S---K---Q---V---L---L---L---I---P---N---F---E---N---V---N---L---KEKKE

```

```

QORECOLI  KVD---AEQKY---PKD---QR---HE---I---E---S---R---A---Q---G---S---M---I---P
QORYEAST  LN---KIYKY---R---DYRT---AAD---E---R---K---V---E---L---G---E---I---P
QORCAVPO  PV---G---SQ---Y---E---K---S---Q---H---E---M---I---H---S---G---I---V---E---T---M---L
QORMOUSE  PV---G---SE---Y---E---K---A---Q---H---E---D---I---H---S---G---K---T---M---L
IE41 ATH  TV---D---S---K---H---S---K---E---D---W---A---K---S---D---H---A---S---K---I---E---P
IE41 SOL  TV---D---S---K---H---S---K---E---D---W---S---R---I---G---H---A---S---K---I---E---P

```

Fig . 7

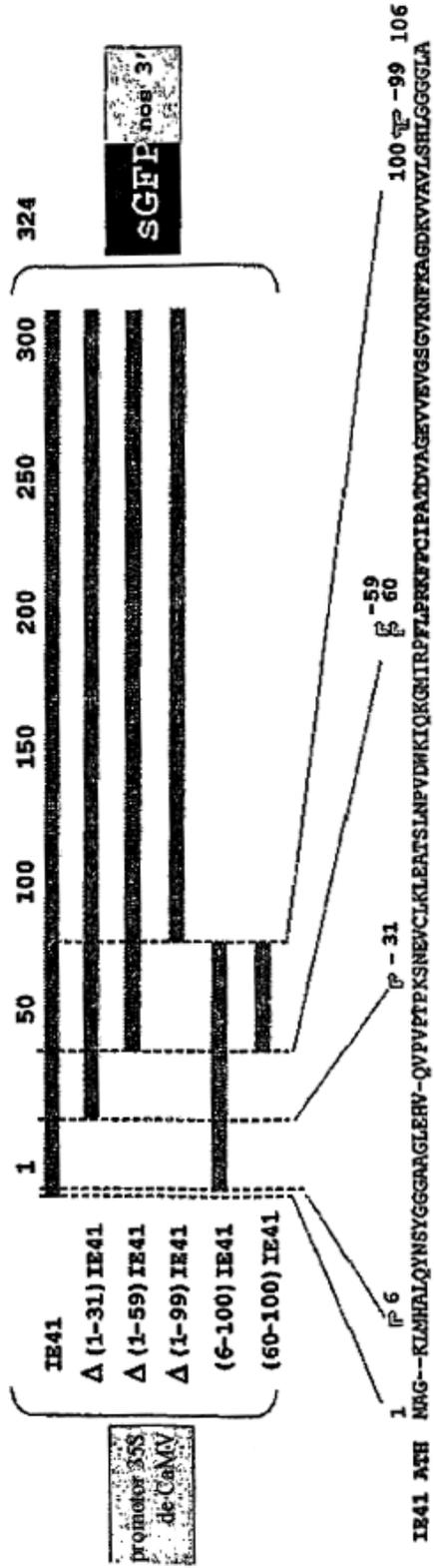
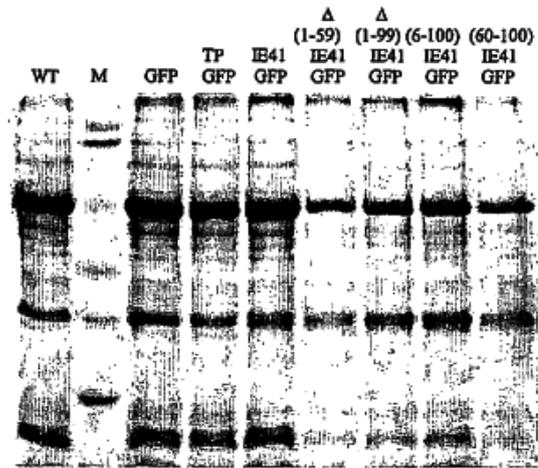
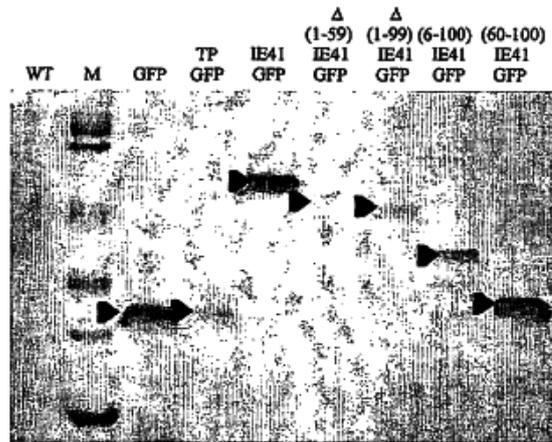


Fig . 8

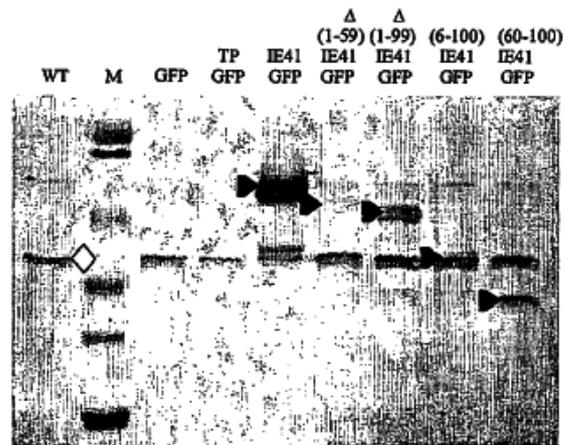
**A**



**B**



**C**



**Fig. 9**

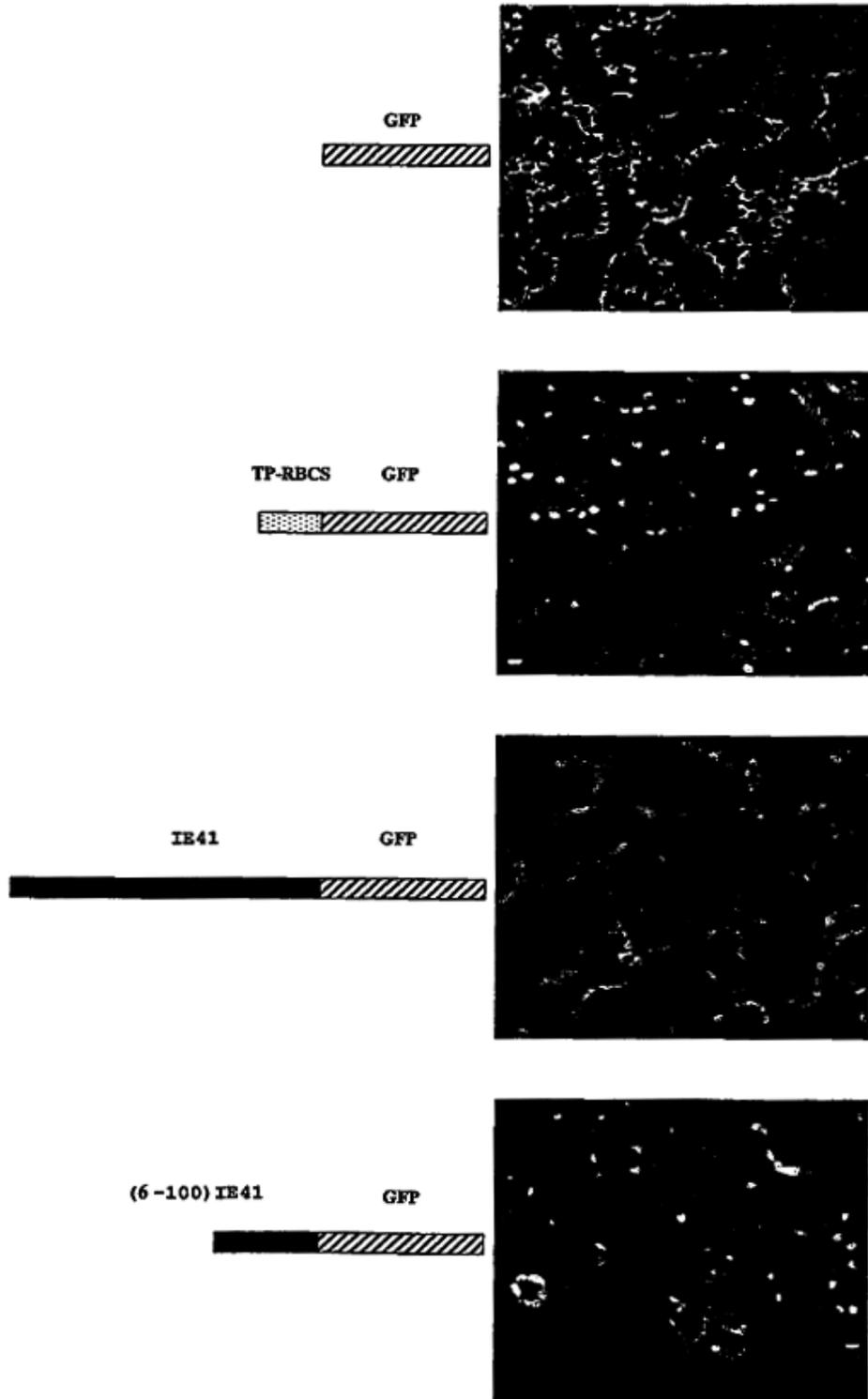


Fig . 10