



19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 304 792**

51 Int. Cl.:

**C12N 15/53** (2006.01)

**C12N 9/04** (2006.01)

**C12P 19/02** (2006.01)

**C12N 1/21** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Número de solicitud europea: **98115231 .7**

86 Fecha de presentación : **13.08.1998**

87 Número de publicación de la solicitud: **0897984**

87 Fecha de publicación de la solicitud: **24.02.1999**

54 Título: **Gen D-sorbitol deshidrogenasa.**

30 Prioridad: **21.08.1997 EP 97114432**

45 Fecha de publicación de la mención BOPI:  
**16.10.2008**

45 Fecha de la publicación del folleto de la patente:  
**16.10.2008**

73 Titular/es: **DSM IP Assets B.V.**  
**Het Overloon 1**  
**6411 TE Heerlen, NL**

72 Inventor/es: **Hoshino, Tatsuo;**  
**Ojima, Setsuko;**  
**Shinjoh, Masako;**  
**Tomiyama, Noribumi y**  
**Miyazaki, Taro**

74 Agente: **Lehmann Novo, María Isabel**

ES 2 304 792 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

Gen D-sorbitol deshidrogenasa.

5 La presente invención se refiere a un nuevo código de ADN para sorbitol deshidrogenasa de un microorganismo perteneciente a las bacterias de ácido acético incluyendo el género *Gluconobacter* y el género *Acetobacter*, un vector de expresión conteniendo dicho ADN, y organismos recombinantes conteniendo dicho vector de expresión. Además la presente invención se refiere al proceso para producir D-sorbitol deshidrogenasa recombinante y el proceso para producir L-sorbosa utilizando dichas enzimas recombinantes u organismos recombinantes conteniendo dicho vector de expresión.

10 L-Sorbosa es un intermediario importante en el proceso actual industrial de producción de la vitamina C, el cual es principalmente practicado por el método de Reichstein (Helvetica chimica Acta, 17: 311, 1934). En el proceso, la única transformación microbiana es la producción de L-sorbosa a partir de D-sorbitol por cepas de *Gluconobacter* o *Acetobacter*. La transformación se considera que es llevada a cabo por la D-sorbitol deshidrogenasa (SLDH) independiente de NAD/NADP. L-Sorbosa es también un sustrato bien conocido en el arte para producir microbiológicamente el ácido 2-ceto-L-gulónico, el cual es un intermediario útil en la producción de vitamina C.

20 Es conocido que existen D-sorbitol deshidrogenasa independiente de NAD/NADP, las cuales catalizan la oxidación de D-sorbitol a L-sorbosa. Dicha D-sorbitol deshidrogenasa fue aislada y caracterizada a partir de *Gluconobacter suboxydans* var  $\alpha$  IFO 3254 (E. Shinagawa y otros, Agric Biol. Chem., 46: 135-141, 1982), y se encontró que está compuesta de tres subunidades con pesos moleculares de 63 kDa, 51 kDa y 17 kDa; la mayor subunidad es la deshidrogenasa conteniendo a FAD como un cofactor, la segunda de ellas es citocromo c AB/vs/26.6.98 y la más pequeña es una proteína con función desconocida; y muestra su pH óptimo a 4.5. Tal SLDH también fue purificada y caracterizada a partir de *G. suboxydans* ATCC 621 (KCTC 2111) (E-S Choi y otros, FEMS Microbiol. Lett. 125: 45-50, 1995) y se encontró que está compuesta de tres subunidades con pesos moleculares de 75 kDa, 50 kDa y 14 kDa; la mayor subunidad es la deshidrogenasa conteniendo a la pirroloquinolina quinona (PQQ) como un cofactor, la segunda de ellas es citocromo c y la más pequeña es una proteína con función desconocida. Los inventores también purificaron y caracterizaron la SLDH NAD/NADP-independiente a partir de *G. suboxydans* IFO 3255(T. Hoshino y otros, EP 728840; la SLDH está compuesta de un tipo de subunidad con peso molecular de 79.0 +/- 0.5 kDa y muestra su pH óptimo de 6 a 7 y muestra actividad deshidrogenasa en manitol y glicerol así como en D-sorbitol.

30 Aunque diversas SLDHs han sido purificadas, sus genes no han sido clonados aún. Es útil clonar el gen de SLDH para una producción eficiente de la enzima SLDH y para construir organismos recombinantes que tengan una actividad SLDH intensificada para mejorar el rendimiento productivo de L-sorbosa. Es útil también introducir el llamado gen SLDH dentro de organismos deseados, por ejemplo *Gluconobacter* transformando L-sorbosa en ácido 2-ceto-L gulónico para construir microorganismos recombinantes los cuales producen directamente el ácido 2-ceto-L gulónico a partir del D-sorbitol.

40 La presente invención proporciona el código de secuencia nucleótida (gen) para D-sorbitol deshidrogenasa (SLDH) originando a partir de un microorganismo perteneciente a las bacterias de ácido acético incluyendo el género *Gluconobacter* y el género *Acetobacter*; una molécula de ADN que comprende la secuencia nucleótida así como la llamada combinación de dicho ADN con un ADN que comprende la secuencia nucleótida de una proteína funcional activando *in vivo* la llamada SLDH, vectores de expresión que portan el ADN que comprende la secuencia nucleótida SLDH o dicha combinación de ADNs; organismos recombinantes que portan los vectores de expresión; un proceso para producir el SLDH recombinante; y un proceso para producir L-sorbosa utilizando el SLDH recombinante y una proteína la cual es funcional activando *in vivo* SLDH.

50 Además la presente invención está dirigida a secuencias de ADN que codifican los polipéptidos con actividad SLDH y polipéptidos funcionales activando *in vivo* dicha SLDH, como es divulgado por ejemplo en la lista de secuencia como SEQ ID NO:2 y NO:3 así como sus cadenas complementarias, o aquellas en las cuales incluye estas secuencias, secuencias de ADN las cuales se hibridan bajo condiciones rigurosas con tales secuencias.

55 Un hombre especializado en el arte está familiarizado con la hibridación rigurosa y las condiciones rigurosas de lavado, las cuales se describen, por ejemplo en Sambrook y otros, Molecular Cloning, Cold Spring Harbor Laboratory Press, New York, USA, second edition, 1989.

Seguidamente una breve descripción de los dibujos se presenta.

60 La Figura 1 ilustra las secuencias parciales de aminoácidos de polipéptidos SLDH y oligonucleótidos. Las secuencias de aminoácidos en negritas en la figura fueron usadas para sintetizar dos secuencias de oligonucleótidos (S7 y S6R) por PCR. Las flechas muestran la dirección de síntesis de ADN. Todos los cebadores fueron mezclas de ADN degenerado que tienen errores para el codón de uso de *Gluconobacter*.

65 La Figura 2 ilustra un mapa de restricción del gen SLDH clonado en la presente invención y la estructura génica de la región de ADN que codifica para SLDH y ORF2.

## ES 2 304 792 T3

La Figura 3 ilustra la secuencia nucleótida que codifica para SLDH y ORF2 con secuencias aguas arriba y aguas abajo e ilustra la secuencia de aminoácidos inferida de SLDH y ORF2. La Figura 3 ilustra también los supuestos sitios de enlace ribosomal (secuencias SD) de los genes SLDH y ORF2 y las posibles secuencias terminales de la transcripción.

5

La Figura 4 ilustra los pasos para la construcción de los plásmidos pTNB114, pTNB115, y pTNB116 para la expresión de los genes de SLDH y/o ORF2 en *E. coli* bajo el control del promotor lac.

La Figura 5 ilustra los pasos para la construcción del plásmido pTNB110 el cual porta los genes de SLDH y/o ORF2 bajo el control de un promotor común pA (descrito abajo en el ejemplo 6) y pTNB143 el cual porta el gen de ORF2 bajo el control de pA localizado contiguo al gen en la secuencia aguas arriba y el gen SLDH bajo el control de otro promotor pA localizado contiguo al gen SLDH en su secuencia aguas arriba.

Los inventores han aislado el gen de SLDH junto con el gen funcional que desarrolla *in vivo* la actividad de SLDH por técnicas de ADN recombinante y determinaron la secuencia nucleótida.

La presente invención proporciona las secuencias de ADN que codifican para SLDH de *Gluconobacter* y un gen funcional que desarrolla *in vivo* actividad de SLDH después indicada aquí como gen ORF2, los vectores de expresión conteniendo dicho ADN para SLDH y las células huésped portando dichos vectores de expresión.

20

Brevemente, el(los) gen(es) de SLDH y/o ORF2, la molécula de ADN conteniendo dicho(s) gen(es), el vector de expresión recombinante y el organismo recombinante utilizado en la presente invención pueden obtenerse por los siguientes pasos:

(1) Aislamiento del ADN cromosómico a partir de los microorganismos los cuales pueden proporcionar actividad de SLDH que transforma D-sorbitol en L-sorbosa y construcción de la biblioteca genómica con el ADN cromosómico en una célula huésped apropiada, por ejemplo *E. coli*.

25

(2) Clonación del(de los) gen(es) SLDH y/o ORF2 a partir del ADN cromosómico por hibridación de colonias, placas, o Southern-blot, clonación por PCR (Reacción en cadena de la polimerasa), análisis Western-blot y similares.

30

(3) Determinación de la secuencia nucleótida del (de los) gen(es) SLDH y/o ORF2 obtenidos anteriormente por métodos convencionales para seleccionar la molécula de ADN conteniendo dicho(s) gen(es) SLDH y/o ORF2 y construcción del vector de expresión recombinante en el cual el(los) gen(es) SLDH y/o ORF2 pueden expresarse eficientemente.

35

(4) Construcción de organismos recombinantes que portan el(los) gen(es) SLDH y/o ORF2 por un método apropiado para introducir el ADN dentro de la célula huésped, por ejemplo, transformación, transducción, transconjugación y/o electroporación.

40

Los materiales y técnicas empleadas en el aspecto anterior de la presente invención son ejemplificados en detalle como sigue:

El ADN cromosómico total puede ser purificado por un procedimiento bien conocido en el arte. El gen objetivo puede ser clonado en cualquiera de los vectores plásmidos o fagos a partir del ADN cromosómico total típicamente por cualquiera de los siguientes métodos ilustrados:

45

(i) La secuencia parcial de aminoácidos es determinada a partir de las proteínas purificadas o fragmentos de péptidos de las mismas. Tal proteína entera o fragmentos de péptidos pueden ser preparados por el aislamiento de tal proteína entera o por tratamiento con peptidasas a partir del gel después de la electroforesis en gel SDS-poliacrilamida. De este modo la proteína obtenida o los fragmentos de ella son aplicados al secuenciador de proteína como el Applied Biosystems secuenciador automático de fase gaseosa 470A. La secuencia de aminoácidos puede ser utilizada para diseñar y preparar sondas de oligonucleótidos y/o cebadores con el sintetizador de ADN como el Applied Biosystems secuenciador automático de ADN 381A. Dichas sondas pueden ser empleadas para el aislamiento de clones que portan el gen objetivo a partir de la biblioteca genómica de la cepa que porta el gen objetivo con la ayuda de la hibridación de placas, colonias, o Southern-blot.

50

55

(ii) Alternativamente, para el propósito de selección de clones que expresan la proteína objetivo a partir de la biblioteca genómica, métodos inmunológicos con anticuerpos preparados contra la proteína objetivo pueden ser aplicados.

60

(iii) El fragmento de ADN del gen objetivo puede ser amplificado a partir del ADN cromosómico por el método PCR con un conjunto de cebadores, es decir dos oligonucleótidos sintetizados de acuerdo con la secuencia de aminoácidos determinada como anteriormente. Luego el clon que porta el gen objetivo entero puede ser aislado a partir de la biblioteca genómica construida, por ejemplo en *E. coli* por hibridación de placas, colonias, o Southern-blot con el producto PCR obtenido anteriormente como la sonda.

65

## ES 2 304 792 T3

(iv) El camino adicional alternativo a la clonación es el tamizaje de clones complementando cepas deficientes de SLDH construidas por mutación convencional con mutagénesis química o por técnicas de ADN recombinante, por ejemplo con transposon Tn5 para afectar el gen objetivo.

5 Secuencias de ADN las cuales pueden ser logradas por la reacción en cadena de la polimerasa empleando cebadores diseñados sobre la base de las secuencias de ADN divulgadas allí por métodos conocidos en el arte también son un propósito de la presente invención. Está implícito que la secuencia de ADN de la presente invención también puede ser lograda sintéticamente como es descrito, por ejemplo en EP 747 483.

10 El anticuerpo mencionado anteriormente puede ser preparado con proteína SLDH purificada o sus fragmentos de péptidos como antígeno por tal método descrito en *Methods in Enzymology*, vol. 73, p 46, 1981.

15 Una vez que el clon que porta el gen deseado es obtenido, la secuencia nucleótida del gen objetivo puede ser determinada por el método bien conocido tal como el método terminador de cadena dideoxi con fago M13 (Sanger F. y otros, *Proc. Natl. Acad. Sci. USA*, 74:5463-5467, 1977).

20 El gen deseado que expresa actividad de D-sorbitol deshidrogenasa de la presente invención es ilustrado en la Figura 2 y la Figura 3. Este gen específico codifica la enzima SLDH que tiene 716 residuos de aminoácidos junto con un péptido señal de 24 residuos de aminoácidos. Los inventores encontraron un marco de lectura abierta justo en la secuencia aguas arriba de la anterior estructura génica del SLDH y designada ésta como ORF2. Este gen ORF2 codifica una proteína que tiene 126 residuos de aminoácidos y fue sugerido ser funcional en proporción a la actividad enzimática deseada para el SLDH recombinante de la presente invención, en particular cuando dicho SLDH es expresado en células huésped de diferentes géneros a partir de bacterias de ácido acético.

25 Para expresar el gen deseado/secuencia nucleótida aislada a partir de bacterias de ácido acético incluyendo el género *Gluconobacter* y el género *Acetobacter* eficientemente, varios promotores pueden ser empleados; por ejemplo, el promotor original del gen, promotores de genes de resistencia a antibióticos tales como el gen resistente a la kanamicina de Tn5 (D.E. Berg, and C.M. Berg. 1983. *Bio/Technology* 1:417-435), gen resistente a la ampicilina de pBR322, y  $\beta$ -galactosidasa de *E. coli* (lac), promotor trp-, tac-, trc-, promotores de fago lambda y cualquiera de los  
30 promotores que puedan ser funcionales en el organismo huésped. Para este propósito, el organismo huésped puede ser seleccionado a partir de de microorganismos que incluyen bacterias tales como *Escherichia coli*, *Pseudomonas putida*, *Acetobacter xylinum*, *Acetobacter paterianus*, *Acetobacter aceti*, *Acetobacter hansenii*, y *Gluconobacter albidus*, *Gluconobacter capsulatus*, *Gluconobacter cerinus*, *Gluconobacter dioxyaceticus*, *Gluconobacter gluconicus*, *Gluconobacter industrius*, *Gluconobacter melanogenus*, *Gluconobacter nonoxygluconicus*, *Gluconobacter oxydans*, por  
35 ejemplo, *Gluconobacter oxydans* DSM 4025, el cual ha sido depositado como DSM 4025 en Marzo 17, 1987 bajo las condiciones del Tratado de Budapest para Deutsche Sammlung von Mikroorganismen und Zellculturen GMBH, Braunschweig, BRD, *Gluconobacter oxydans* subespecie *sphaericus*, *Gluconobacter roseus*, *Gluconobacter rubiginosus*, *Gluconobacter suboxydans*, células mamíferas y células vegetales.

40 Para la expresión, otros elementos reguladores, tales como la secuencia Shine-Dalgarno (SD) (por ejemplo, AG-GAGG etc. incluyendo secuencias factibles naturales y sintéticas en la célula huésped) y un terminador transcripcional (estructura repetitiva invertida que incluye cualquier secuencia factible natural y sintética en la célula huésped) los cuales son factibles en la célula huésped dentro de la cual la secuencia que codifican será introducida pueden ser empleados con los promotores descritos anteriormente.

45 Para la expresión de polipéptidos unidos a la membrana, similar a la proteína SLDH de la presente invención, un péptido señal, conteniendo usualmente de 15 a 50 residuos de aminoácidos y totalmente hidrofóbico, es de preferencia asociado. Un ADN que codifica un péptido señal puede ser seleccionado a partir de cualquier secuencia factible natural y sintética en la célula huésped deseada.

50 Una amplia variedad de combinaciones huésped/vectores de clonación pueden ser usadas en la clonación de la doble cadena de ADN. El vector de clonación es generalmente un plásmido o fago que contiene el origen de replicación, elementos reguladores, un sitio de clonación incluyendo un sitio de multi-clonación y marcadores selectivos tales como genes de resistencia a antibióticos incluyendo genes de resistencia para ampicilina, tetraciclina, kanamicina, estreptomycin, gentamicina, espectinomycin, etc.

55 De preferencia el vector para la expresión del gen de la presente invención en *E. coli* es seleccionado a partir de algunos vectores empleados en *E. coli*, tales como pBR322 o sus derivados incluyendo pUC18 y pBluescript II (Stratagene Cloning Systems, CA, USA), pACYC177 y pACYC184 (*J. Bacteriol.*, 134:1141-1156, 1978) y sus derivados, y un vector derivado a partir una amplia gama de plásmidos huéspedes tales como RK2 (C.M. Thomas, *Plasmid* 5: 10, 1981) y RSF1010 (P. Guerry y otros, *J. Bacteriol.* 117: 619-630, 1974). El vector preferido para la expresión de la secuencia nucleótida de la presente invención en bacterias incluyendo *Gluconobacter*, *Acetobacter* y *P. putida* es seleccionado a partir de algunos vectores los cuales pueden replicarse en *Gluconobacter*, *Acetobacter* y/o *P. putida*, así como en el organismo de clonación preferido tal como *E. coli*. El vector preferido es un vector de amplia gama huésped como el vector cósmido similar a pVK100 (V.C. Knauf y otros, *Plasmid*. 8: 45-54, 1982) y sus derivados y RSF1010. El número de copias y la estabilidad del vector deberían ser cuidadosamente considerados para una estable y eficiente expresión del gen clonado y también para un eficiente cultivo de la célula huésped que porta el gen clonado. Moléculas de ADN conteniendo elementos transponibles tales como el Tn5 que también pueden ser empleadas como

vector para introducir el gen deseado dentro del huésped preferido, especialmente en un cromosoma. Moléculas de ADN conteniendo algunos ADNs aislados a partir del huésped preferido junto con el gen de la presente invención también es útil para introducir este gen dentro del huésped preferido aplicando alguno de los métodos convencionales, por ejemplo, transformación, transducción, transconjugación o electroporación, los cuales son bien conocidos para aquellos especializados en el arte, considerando la naturaleza del huésped y la molécula de ADN.

Huéspedes útiles pueden incluir microorganismos, células de mamíferos, células vegetales y similares. Dentro de los microorganismos preferidos, aquí pueden ser mencionadas las bacterias tales como *E. coli*, *P. putida*, *A. xylinum*, *A. patetianus*, *A. aceti*, *A. hansenii*, *Gluconobacter albidus*, *Gluconobacter capsulatus*, *Gluconobacter cerinus*, *Gluconobacter dioxyaceticus*, *Gluconobacter gluconicus*, *Gluconobacter industrius*, *Gluconobacter melanogenus*, *Gluconobacter nonoxygluconicus*, *Gluconobacter oxydans* y *Gluconobacter oxydans* subespecie *sphaericus*, *Gluconobacter roseus*, *Gluconobacter rubiginosus*, *Gluconobacter suboxydans* y algunas bacterias capaces de expresar SLDH recombinante y/o gen(es) ORF2. Equivalentes funcionales, subcultivos, mutantes y variantes de dichos microorganismos también pueden ser empleados en la presente invención. La cepa preferida es la *E. coli* K12 o sus derivadas, *P. putida*, *Gluconobacter* o cepas *Acetobacter*.

Los genes y secuencias nucleótidas de SLDH y/o ORF2 proporcionados en esta invención están ligados dentro de un vector adecuado conteniendo regiones reguladoras tales como un promotor, un sitio de enlace ribosomal y un terminador transcripcional factible en las células huéspedes descritas anteriormente con métodos bien conocidos en el arte para producir un vector de expresión. Cuando los genes de SLDH y ORF2 son clonados en combinación, cualquiera de los dos genes puede ser clonado en serie o separadamente en el mismo plásmido también en el ADN cromosómico. Las Figuras 4 y 5 ejemplifican la forma de la clonación combinada de los genes SLDH y ORF2 en el plásmido. También se puede clonar un gen en el plásmido y el otro en el ADN cromosómico.

Para construir un microorganismo recombinante que porte un vector de expresión recombinante, varios métodos de transferencia de genes incluyendo transformación, transducción, apareamiento conyugal (Capítulos 14 y 15, Methods for general and molecular bacteriology, Philipp Gerhardt *et al.* ed., American Society for Microbiology, 1994) y electroporación pueden ser empleados. Los métodos para construir el organismo recombinante pueden ser seleccionados a partir de métodos bien conocidos en el campo de la biología molecular. Sistemas de transformación usuales pueden ser empleados para *E. coli*, *Pseudomonas*, *Gluconobacter* y *Acetobacter*. El sistema de transducción puede ser empleado para *E. coli*. El sistema de apareamiento conyugal puede ser ampliamente empleado en bacterias Gram positivas y Gram negativas incluyendo *E. coli*, *P. putida*, y *Gluconobacter*. El apareamiento conyugal preferido es divulgado en WO89/06688. La conjugación puede ocurrir en medio líquido o en una superficie sólida. El receptor preferido para la producción de SLDH y/o ORF2 se selecciona a partir de *E. coli*, *P. putida*, *Gluconobacter* y *Acetobacter*. Al receptor para el apareamiento conyugal, un marcador selectivo es usualmente adicionado; por ejemplo, la resistencia contra el ácido nalidíxico o la rifampicina es usualmente seleccionada. La resistencia natural también puede ser empleada; por ejemplo resistencia contra polimixina B es útil para muchas *Gluconobacters*.

La presente invención proporciona un proceso para producir SLDH recombinante. Se puede aumentar el rendimiento productivo de la enzima SLDH introduciendo el gen de SLDH proporcionado en la presente invención dentro de organismos incluyendo bacterias de ácido acético del género *Gluconobacter* y el género *Acetobacter*. También se puede producir SLDH activa en microorganismos además de *Gluconobacter* empleando el gen de la SLDH de la presente invención en combinación con el gen de ORF2 de la presente invención. El recombinante de SLDH puede ser inmovilizado en un portador sólido para una reacción enzimática en fase sólida. La presente invención también proporciona organismos recombinantes. Se puede producir L-sorbosa a partir de D-sorbitol con los organismos recombinantes. También se puede producir L-sorbosa a partir de D-sorbitol incluso en un organismo huésped además de las bacterias de ácido acético introduciendo el gen SLDH en combinación con el gen ORF2 de la presente invención.

## 50 Ejemplos

### Ejemplo 1

#### Determinación de la secuencia de aminoácidos de SLDH

55 (1) Análisis de la secuencia de aminoácidos N- terminal de polipéptidos SLDH tratados con lisilendopeptidasas

Secuencias parciales de aminoácidos de los péptidos No.1, 3 y 8 preparadas a partir de la proteína SLDH fueron determinadas (Fig. 1; SEQ ID NO:4 a 6). SLDH purificada de *G. suboxydans* IFO 3255 (T. Hoshino y otros, EP 728840) fue digerida con lisilendopeptidasa; la mezcla de reacción (25 ml) contenía 1.2 mM de lisilendopeptidasa y 3 nmol de la proteína SLDH en buffer Tris-HCl 100 mM, pH 9.0, y la reacción fue llevada a cabo a 37°C durante 15 horas. Los fragmentos de péptido resultantes fueron separados por HPLC (columna: protein C-4, VYDAC, California, USA) con gradiente de acetonitrilo/isopropanol en TFA 0.1% a un velocidad de flujo de 1.0 ml/min. La elusión de los péptidos fue monitoreada por absorbancia UV a 214 nm, y las fracciones de los picos fueron colectadas manualmente y sometidas a análisis de secuencia de aminoácidos N- terminal con un secuenciador de aminoácidos (Applied Biosystems model 65 470A, The Perkin Elmer Corp., Conn., USA).

## Ejemplo 2

*Clonación del gen parcial de SLDH por PCR*

5 La PCR fue conducida con el ADN cromosómico de *G. suboxydans* IFO 3255 y cebadores de oligonucleótidos de ADN degenerados, s7 y s6R cuyas secuencias son mostradas en la Fig. 1. La amplificación por PCR fue llevada a cabo con Amplitaq polimerasa termoestable (Perkin-Elmer, Roche Molecular Systems Inc., NJ, USA), empleando un termociclador (ZYMOREACTOR II AB-1820, ATTO, Tokio, Japan). Seguidamente la mezcla de reacción (50  $\mu$ l) fue empleada para PCR: 200  $\mu$ M de dNTPs, 10-20 pmol de cada cebador (degeneración 54-288), 6 ng de ADN cromosómico de *G. suboxydans* IFO 3255 y 0.5 unidades de ADN polimerasa en el buffer proporcionado por el proveedor. La reacción constó de 30 ciclos de 1) paso de desnaturalización a 94°C durante 1 minuto; 2) paso de templado a 48°C durante 1 minuto; 3) paso de síntesis a 72°C durante 2 minutos. Por consiguiente un ADN de 1.6 kb fue amplificado y clonado en un vector de *E. coli*, pUC57/T (MBI Fermentas, Vilnius, Lithuania), el cual tiene un extremo terminal 3'-ddT para ligamento directo de un fragmento de ADN amplificado y obtener el plásmido recombinante pMT20. El ADN clonado fue sometido a la secuenciación de nucleótidos por el método de terminador de cadena dideoxi (F. Sanger y otros, Proc. Natl. Acad. Sci. USA, 74:5463-5467, 1977); el fragmento de 1.6 kb que codifica los péptidos No.3 y No.8.

## 20 Ejemplo 3

*Clonación completa del gen de SLDH*(1) *Construcción de la biblioteca genómica de G. suboxydans IFO 3255*

25 El ADN cromosómico de *G. suboxydans* IFO 3255 fue preparado a partir de células cultivadas en placas de agar MB durante 2 días. El ADN cromosómico (160  $\mu$ g) fue parcialmente digerido con 20 unidades de *Eco* RI en 500  $\mu$ l de mezcla de reacción. Porciones de muestras fueron retiradas a los 5, 10, 15, 30 y 60 minutos y el grado de digestión fue detectado por electroforesis en gel de agarosa. Las primeras cuatro porciones que contenían fragmentos de ADN parcialmente digeridos fueron combinadas y sometidas a una electroforesis en gel preparativo (agarosa: 0.6%). Fragmentos de 15-35 kb fueron recortados y electroeluidos a partir del gel. El eluato fue filtrado y precipitado con acetato de sodio y etanol a -80°C. Los fragmentos de ADN fueron colectados por centrifugación y suspendidos en 200  $\mu$ l de buffer Tris-HCl 10 mM, pH 8.0, conteniendo EDTA 1 mM.

35 En paralelo, 1.8  $\mu$ g de vector cósmido pVK100 fue completamente digerido con *Eco* RI y tratado con fosfatasa alcalina de intestino de ternero. El pVK100 linealizado y defosforilado fue ligado con fragmentos *Eco* RI de 15-35 kb del ADN cromosómico de *G. suboxydans* IFO 3255 (5  $\mu$ g) con el kit de ligamento (Takara Shuzo, Kyoto, Japan) en 20 tubos separados bajo las condiciones recomendadas por el proveedor para la obtención del ADN altamente polimerizado. El ADN ligado fue entonces empleado para el empaquetamiento *in vitro* de acuerdo con el método descrito por el proveedor (Amersham Japan). Las partículas de fago resultantes fueron empleadas para infectar *E. coli* ED8767, el huésped para la biblioteca genómica. Por consiguiente, 4,271 colonias fueron obtenidas y de todas las colonias probadas (20 colonias) poseían los insertos de ADNs con la talla promedio de aproximadamente 25 kb.

45 (2) *Hibridación de colonia para obtención del gen SLDH completo*

La biblioteca de cósmidos descrita anteriormente fue tamizada para aislar el clon portando el gen SLDH completo por el método de hibridación de colonias con ADN de pMT20 de 1.6 kb marcados con P<sup>32</sup>. Un clon fue aislado y designado como pSLII, el cual porta un inserto de aproximadamente 25 kb en el vector pVK100. A partir de pSLII, el fragmento *Pst* I de 6.2 kb fue aislado y clonado dentro de pUC18 para obtener el plásmido pUSLIIP. El fragmento *Pst* I de 6.2 kb conteniendo los genes SLDH y ORF2 fue aislado a partir de pUSLIIP y subclonado dentro del vector pCRII (Invitrogen Corporation, CA, USA) para producir pCRSLIIP. A partir del plásmido resultante, el fragmento de ADN conteniendo el fragmento *Pst* I de 6.2 kb fue aislado como fragmento *Hind* III-*Xho* I y clonado entre los sitios *Hind* III y *Xho* I de pVK100 para obtener pVKSIIIP.

55 (3) *Expresión del gen SLDH en E. coli*

Para confirmar si el fragmento *Pst* I de 6.2 kb codifica el SLDH objetivo, células de *E. coli* JM109 portando pUSLIIP fueron sometidas a análisis por Western-blot con anticuerpos anti-SLDH como descrito anteriormente. Proteínas inmuno-positivas con peso molecular de aproximadamente 80 kDa fueron observadas en el transformante, indicando que el fragmento *Pst* I codifica los polipéptidos con el peso molecular del SLDH intacto (79 kDa +/- 0.5 kDa).

65 (4) *Construcción del Gluconobacter deficiente de SLDH, cepa 26A11, como la cepa prueba para la complementación de la actividad SLDH*

La mutagénesis del transposon Tn5 fue realizada con *G. melanogenus* IFO 3293 como la madre. Tn5, un elemento transponible y codifica para Kmr, causa mutaciones nulas al azar en el ADN de su organismo huésped y ampliamente empleado como mutágeno en bacterias Gram-negativas. La IFO 3293 fue seleccionada como madre por las razones

## ES 2 304 792 T3

siguientes: (i) produjo L-sorbose a partir de D-sorbitol, (ii) mostró polipéptidos inmuno-positivos con peso molecular de aproximadamente 80 kDa en análisis por Western-blot con los anticuerpos preparados contra SLDH purificada a partir de *G. suboxydans* IFO 3255 y (iii) su frecuencia para generar mutantes Tn5 fue mucho mayor que la de *G. suboxydans* IFO 3255.

5

*G. melanogenus* IFO 3293 fue cultivada en tubos de prueba conteniendo 5 ml de medio MB, conteniendo 25 g/l de manitol, 5 g/l de extracto de levadura (Difco Laboratorios), 3 g/l de Bactopeptona (Difco) a 30°C toda la noche. *E. coli* HB101 (pRK2013) [D.H. Figurski, Proc. Natl. Acad. Sci. USA 76: 1648-1652, 1979] y *E. coli* HB101 (pSUP2021) [R. Simon, y otros, BIO/TECHNOL. 1: 784-791, 1983] fueron cultivadas en tubos de prueba conteniendo 5 ml de medio LB con 50 µg/ml de kanamicina a 37°C toda la noche. Las células fueron colectadas separadamente por centrifugación y suspendidas en la mitad del volumen del medio MB. Cada suspensión celular fue mezclada en la proporción 1:1:1 y la mezcla fue colocada en filtro de nitrocelulosa sobre la superficie de la placa de agar MB. Las placas fueron incubadas a 27°C toda la noche y las células resultantes en el filtro fueron raspadas, suspendidas en un volumen adecuado de medio MB y extendidas en la placa elegida (placa MPK), MB conteniendo 10 µg/ml de polimixina B y 50 µg/ml de kanamicina. La placa MPK fue incubada a 27°C de 3 a 4 días.

10

Los 3,436 mutantes Tn5 resultantes fueron sometidos al tamizaje por inmuno-dot blot con el anticuerpo anti-SLDH. Las células de cada cepa fueron suspendidas independientemente en 50 µl de buffer Laemmli compuesto de Tris-HCl 62.5 mM (pH6.5), glicerol 10%, SDS 2% y β- mercaptoetanol 5% en placas de microtitulación de 96-pocillos e incubada a 60°C durante 2 horas. Los Extractos Libres de Células (CFEs) fueron estampados en filtros de nitrocelulosa y las muestras inmuno-positivas en los CFEs fueron tamizadas con el kit de sustrato conjugado AP (Bio-RAD Laboratorios, Richmond, Calif., USA). Como resultado, solo una cepa 26A11 sin señal positiva en el tamizaje por inmuno-dot blot fue obtenida.

20

La deficiencia de SLDH en la cepa 26A11 fue confirmada por análisis de Western-blot; este expresó al menos la cantidad de 1/500 de SLDH comparado con su cepa madre. 26A11 no fue una cepa deficiente del SLDH completo, sino la cepa con el gen SLDH reprimido por la inserción de Tn5; el sitio de inserción fue encontrado siendo próximo del C- terminal por determinación de la secuencia nucleótida alrededor del punto de inserción Tn5. A continuación, la reacción celular restante fue conducida para examinar la actividad SLDH total en 26A11 y de las cepas *Gluconobacter* salvajes. En el buffer fosfato de potasio 100 mM (pH7.0) conteniendo D-sorbitol 2%, 26A11 convirtió ligeramente D-sorbitol a L-sorbose, mientras las cepas salvajes IFO 3293 y IFO 3255 lo hicieron completamente en 39.5 horas a 30°C.

25

30

### (5) Expresión del gen SLDH en la cepa 26A11

35

Para confirmar la actividad SLDH de los clones SLDH obtenidos, la prueba de complementación fue conducida. Los plásmidos pSLII y pVKSLIIP fueron introducidos dentro de 26A11 por el apareamiento conyugal. El trans-conjugante portando pSLII o pVKSLIIP regeneró la actividad SLDH en la reacción celular mini-restante y mostró polipéptidos inmuno-reactivos de aproximadamente 80 kDa en el análisis por Western-blot.

40

### (6) Secuenciación de nucleótidos del gen SLDH

El plásmido pUSLIIP fue empleado para la secuenciación de nucleótidos de los genes de SLDH y ORF2. La secuencia nucleótida determinada (SEQ ID NO: 1; 3,481 pb) reveló que el ORF del gen SLDH (2,223 pb, nucleótido No.572 a 2794 en SEQ ID NO: 1) que codifica el polipéptido de 740 residuos de aminoácidos (SEQ ID NO: 2), en la cual estuvieron tres secuencias de aminoácidos (Péptidos No.1, 3 y 8 mostrados en SEQ ID NO: 4 al 6) determinadas a partir del polipéptido SLDH purificado. Adicionalmente al ORF SLDH, uno más ORF, ORF2, fue encontrado justo aguas arriba del ORF SLDH como es ilustrado en las Figuras 2 y 3. El ORF y ORF2 (381 pb, nucleótido No.192 a 572 en SEQ ID NO: 1) codifican el polipéptido de 126 residuos de aminoácidos (SEQ ID NO: 3).

50

La 4<sup>ta</sup> secuencia de aminoácidos del polipéptido No.1 fue determinada como Glu por el secuenciador de aminoácidos, pero esta fue Ala de acuerdo con la secuencia de ADN. La 11<sup>na</sup> secuencia de aminoácidos del polipéptido No.3 fue determinada como Gln por el secuenciador de aminoácidos, pero esta fue Pro de acuerdo con la secuencia de ADN. La región similar al péptido señal (SEQ ID NO: 8) está posiblemente incluida en la secuencia de aminoácidos deducida: esta contiene (i) muchos residuos hidrofóbicos, (ii) residuos positivamente cargados cerca del N-terminal, y (iii) el sitio Ala-Xaa-Ala como una señal clave. La verdadera secuencia señal fue determinada como es descrito en el ejemplo 3 (7). El supuesto sitio del enlace ribosomal (Shine-Dalgarno, SD, secuencia) para el gen SLDH fue localizado a 8 pb aguas arriba del codón de iniciación (AGAGGAG en nucleótido No.558-564 de Seq ID NO: 1). La supuesta secuencia SD para el gen ORF2 fue localizada a 10 pb aguas arriba del codón de iniciación (GGGAGG en el nucleótido No.177-182 de Seq ID NO: 1). Estuvieron algunas secuencias repetitivas invertidas inmediatamente aguas abajo de la estructura génica de SLDH (secuencia nucleótida del No.2803-2833 y del No.2838-2892) como es ilustrado en la Figura 3; las mismas funcionan como lazos terminadores de la transcripción para el gen SLDH. Para el gen ORF2, la secuencia repetitiva invertida fue encontrada en el No.684-704.

60

Búsquedas homólogas para SLDH y ORF2 fueron realizadas con los programas mpBlast (NCBI, Bethesda, Md. USA) y Motifs in GCG (Genetics Computer Group, University Research Park, Md. USA). El polipéptido SLDH tuvo una secuencia comúnmente conservada en quinoproteína en la región cerca del C-terminal (dentro de los residuos de aminoácidos No.632 al 692 de la SEQ ID No.2) y otra secuencia identificada como una quinoproteína motivo (Prosit

65

## ES 2 304 792 T3

No. PS00363: [DN]W.{3}G[RK].{6}[FY]S.{4}[LIVM]N.{2}NV.{2}L[RK]; residuo de aminoácido No.79 al 107 de la SEQ ID No.2).

ORF2 mostró homología con la región N-terminal de la D-glucosa deshidrogenasa (GDH) dependiente del PQQ unida a la membrana de *G. oxydans*, *E. coli*, *A. calcoaceticus*, la cual es conocida como una región extendida de membrana para la unión del GDH a la membrana. Las identidades de ORF2 para la región N-terminal de las GDHs de *G. oxydans*, *E. coli*, *A. calcoaceticus* fueron 30%, 32%, y 37%, respectivamente. La proteína ORF2 funciona como una proteína áncora para lograr unión tipo del SLDH a la membrana.

### Ejemplo 4

#### *Determinación de la secuencia N-terminal y C-terminal del polipéptido SLDH maduro*

La secuenciación directa del N-terminal no dio resultado, indicando que el N-terminal está bloqueado. Luego, el polipéptido SLDH fue tratado con la endoproteinasa Lys-C (Wako, Osaka Japan) en Tris-HCl 0.1 M a 37°C por 20 horas con una proporción sustrato para enzima de 20:1 (p/p). El total digerido fue analizado por HPLC fase reversa (RP300, 1 mm x 25 cm, Applied Biosystems, Foster City, CA) y cada pico fue sometido a espectrometría de masa (instrumento de triple cuádruplo TSQ700, Finningan-MAT, San Jose, California) para determinar el peso molecular. Uno de los digeridos descrito como SEQ ID No.7 fue asignado como la secuencia N-terminal por análisis de espectrometría de masa y análisis de la composición de aminoácidos junto con la composición de aminoácidos pronosticada a partir de la secuencia nucleótida determinada que se muestra como SEQ ID NO:1. Análisis adicional con la disociación inducida colisional (CID) fue llevado a cabo para confirmar la identidad del péptido con secuencia N-terminal. El N-terminal fue determinado siendo un residuo pirrolglutamil.

Desde que el N-terminal de SLDH fue determinado siendo Gln-Phe-Ala-Pro-Ala-Gly-Ala-Gly-Gly-Glu-Pro-Ser-Ser-Ser-Val-Pro-Gly-Pro-Gly-Asn-Ala-Ser-Glu-Pro-Thr-Glu-Asn-Ser-Pro-Lys como se muestra en SEQ ID NO.7, la secuencia señal fue confirmada siendo largos residuos de 24 aminoácidos con la secuencia de Met-Arg-Arg-Pro-Tyr-Leu-Leu-Ala-Thr-Ala-Ala-Gly-Leu-Ala-Leu-Ala-Cys-Ser-Pro-Leu-Ile-Ala-His-Ala listada como SEQ ID NO:8. La secuencia C-terminal también fue determinada empleando los péptidos recuperados a partir de la digestión con proteasa V8 siendo Pro-Asp-Ala-Ile-Lys-Gln (SEQ ID NO:9).

### Ejemplo 5

#### *Expresión de los genes SLDH y/o ORF2 en E. coli*

A partir del pCRSLIIP descrito en el ejemplo 3 (2), los plásmidos que portan el gen SLDH con y sin el gen ORF2 bajo el control del promotor lac fueron construidos como es ilustrado en la Figura 4. Los tres plásmidos resultantes son pTNB114 portando los genes de SLDH y ORF2, pTNB115 portando el gen SLDH y el gen ORF2 truncados en su N-terminal conteniendo el sitio de enlace ribosomal y el codón de iniciación (ATG) y pTNB116 portando el gen ORF2 truncado en su mayoría y el gen SLDH intacto. Estos tres plásmidos fueron introducidos dentro de *E. coli* por transformación convencional. La producción del polipéptido SLDH fue detectada por análisis Western-blot con extractos libres de células de los transformantes resultantes y la actividad de SLDH fue ensayada con las células restantes. La reacción celular restante fue llevada a cabo en la mezcla de reacción compuesta de NaCl 0.3%, CaCO<sub>3</sub> 1%, D-sorbitol 4% y PMS 1 mM con y sin 1 µg/ml de PQQ a temperatura ambiente por 17 horas. La actividad de SLDH para producir L-sorbosa fue analizada por ensayo TLC con gel de sílice 60 F<sub>254</sub>, 0.25 mm, Merk con el solvente revelador compuesto de *n*-propanol-H<sub>2</sub>O-H<sub>3</sub>PO<sub>4</sub>-HCOOH 1% (400:100:10:1) y reactivo en aerosol de naftoresorcinol. Por consiguiente, el polipéptido SLDH fue detectado en todos los transformantes, incluso sin expresión génica de ORF2, en análisis por Western-blot, pero la actividad de SLDH para producir L-sorbosa fue detectada solo en el transformante portando pTNB114 conteniendo el gen ORF2 intacto bajo condiciones de reacción celular restante en presencia de PQQ.

### Ejemplo 6

#### *Expresión de los genes SLDH y/o ORF2 en E. coli*

La Figura 5 ilustra los pasos para la construcción de pTNB110 y pTNB143. El plásmido pTNB110 portando los genes ORF2 y SLDH bajo el control del promotor del gen de la enzima A (pA) de DSM4025 (T. Hoshino y otros, European Patent Application No.9611500.8) fue construido por la inserción del fragmento *Hind* III-*Kpn* I conteniendo pA, fragmentos *Kpn* I-*Xho* I a partir de pTNB114 entre los fragmentos de *Hind* III-*Xho* I de pUC18. El plásmido pTNB143 portando los genes ORF2 y SLDH como unidades de expresión independiente, gen ORF2 con pA y gen SLDH con pA, fue construido por inserción del fragmento *Sal* I de 0.9 kb a partir de pTNB141 y el fragmento *Hind* III-*Xho* I de 3.2 kb a partir de pTNB135 dentro del vector híbrido pUC57/pCRII (fragmento *Sca* I-*Hind* III con plac a partir de pUC57 y el fragmento *Xho* I a *Sca* I sin plac a partir de pCRII) como se muestra en la Figura 5. Los plásmidos, pTNB110 y pTNB143, fueron introducidos dentro de *E. coli* por transformación convencional. Los transformantes resultantes fueron sometidos a análisis por Western-blot y a la reacción celular restante como descrito en el ejemplo

## ES 2 304 792 T3

5. Por consiguiente, ambos transformantes portando pTNB110 y pTNB143 produjeron proteína SLDH y mostraron actividad SLDH para producir L-sorbosa a partir de D-sorbitol en presencia de PQQ.

### 5 Ejemplo 7

#### *Expresión del gen SLDH en G. oxydans DSM 4025*

10 El plásmido pTNB136 para la expresión del gen SLDH en *G. oxydans* DSM 4025 tiene intensas actividades L-sorbosa y L-sorbosona deshidrogenasa junto con débil actividad D-sorbitol deshidrogenasa para producir L-sorbosa, fue construido por la inserción del fragmento *Hind* III-*Xho* I a partir de pTNB135 (Figura 5) entre los sitios de *Hind* III y *Xho* I de pVK100. El plásmido pTNB136 y su vector pVK100 fueron introducidos por apareamiento conyugal dentro de la cepa GOBΔK, la cual es un mutante de *G. oxydans* DSM 4025 cuyo gen de la enzima B (EP 832 974) tiene D-sorbitol deshidrogenasa para producir L-sorbosa es eliminado por sustitución de dos fragmentos *Eco* RI conteniendo 15 gen de la Enzima B con el casete génico de resistencia a kanamicina (fragmento *Eco* RI 1.28 kb de pUC4K; Pharmacia Uppsala, Sweden). La afectación génica fue conducida por técnicas de ADN recombinante bien conocidas en el arte. El transconjugante resultante, GOBΔK portando pTNB136 o pVK100, produjo 5 g/L o por debajo de 2 g/l de 2KGA en D-sorbitol 10%, respectivamente, por fermentación en frasco conducida a 30°C durante 4 días (medio: D-sorbitol 10%, urea 1.6%, glicerol 0.05%, MgSO<sub>4</sub>·7H<sub>2</sub>O 0.25%, licor de maíz en remojo 3%, células húmedas de levadura panadera 20 6.25% y CaCO<sub>3</sub> 1.5%). Análisis por Western-blot con anticuerpos anti-SLDH revelaron que los transconjugantes portando pTNB136 expresaron los polipéptidos SLDH inmuno-reactivos.

25

30

35

40

45

50

55

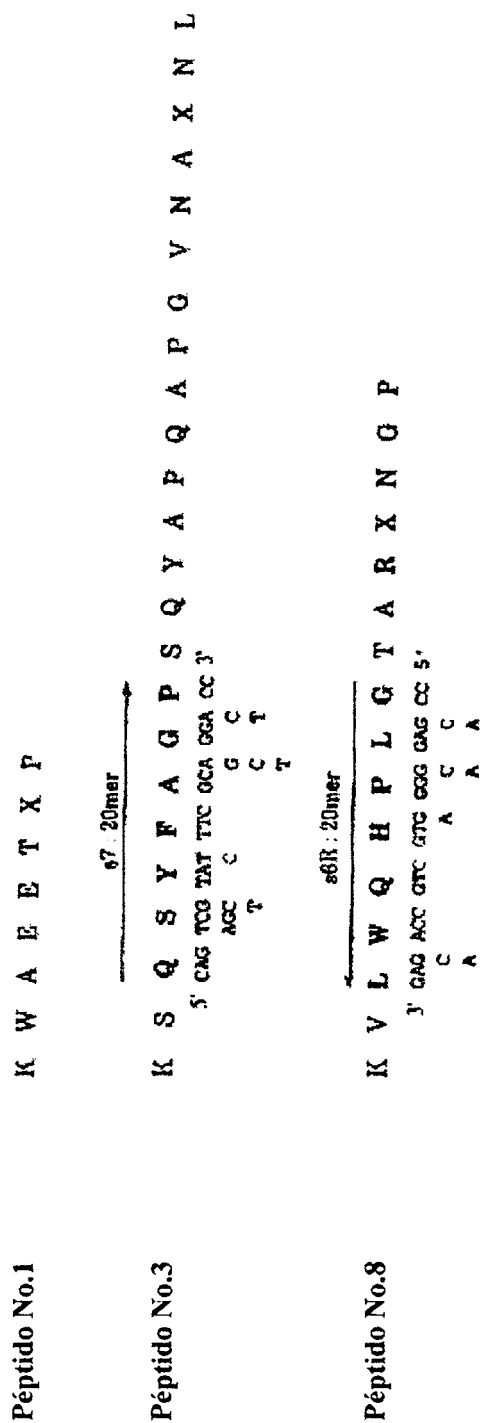
60

65

## ES 2 304 792 T3

### REIVINDICACIONES

- 5 1. ADN que codifica una proteína que tiene actividad D-sorbitol deshidrogenasa cuando es clonado dentro de una bacteria de ácido acético, hibridable bajo condiciones rigurosas, con un ADN de acuerdo con la SEQ ID NO:1.
2. ADN que codifica una proteína funcional activando *in vivo* D-sorbitol deshidrogenasa, hibridable bajo condiciones rigurosas, con un ADN de acuerdo con la SEQ ID NO:1.
- 10 3. Un vector de expresión comprendiendo un ADN de acuerdo con las reivindicaciones 1 y/o 2.
4. Célula huésped transformada con un ADN de acuerdo con las reivindicaciones 1 y/o 2 o un vector de acuerdo a la reivindicación 3.
- 15 5. Célula huésped de acuerdo con la reivindicación 4 seleccionada a partir de un grupo compuesto por bacterias de ácido acético incluyendo el género *Gluconobacter* y el género *Acetobacter*.
- 20 6. Un proceso para producir D-sorbitol deshidrogenasa recombinante que comprende los pasos de cultivar la célula huésped que comprende el ADN de acuerdo con las reivindicaciones 1 y 2 en un medio apropiado, y recobrar una proteína con actividad de D-sorbitol deshidrogenasa a partir del cultivo.
7. Un proceso el cual comprende la conversión de D-sorbitol en L-sorbosa con la ayuda de una proteína codificada por un ADN de acuerdo con las reivindicaciones 1 y 2 o con la célula huésped comprendiendo un ADN de acuerdo con las reivindicaciones 1 y 2.
- 25 8. Un proceso de acuerdo con la reivindicación 7 en el que dicha L-sorbosa es además convertida en 2-KGA y/o vitamina C.
- 30
- 35
- 40
- 45
- 50
- 55
- 60
- 65



Las secuencias de aminoácidos en negritas fueron empleadas para sintetizar dos secuencias de oligonucleótidos (S7 y S6R) para PCR. Las flechas muestran la dirección de síntesis de ADN. Los cebadores fueron mezclas de ADN degenerado que tienen errores para el codón de uso de *Gluconobacter*.

Figura 1

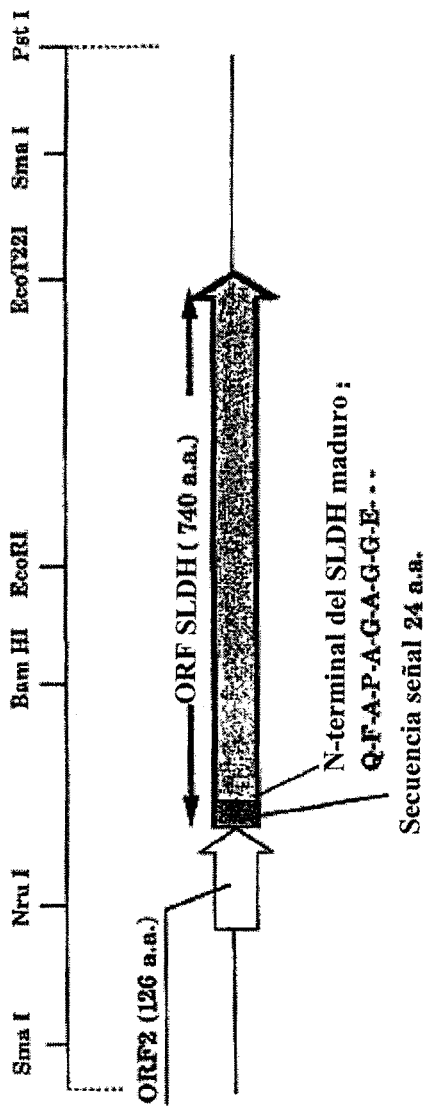


Figura 2

ES 2 304 792 T3

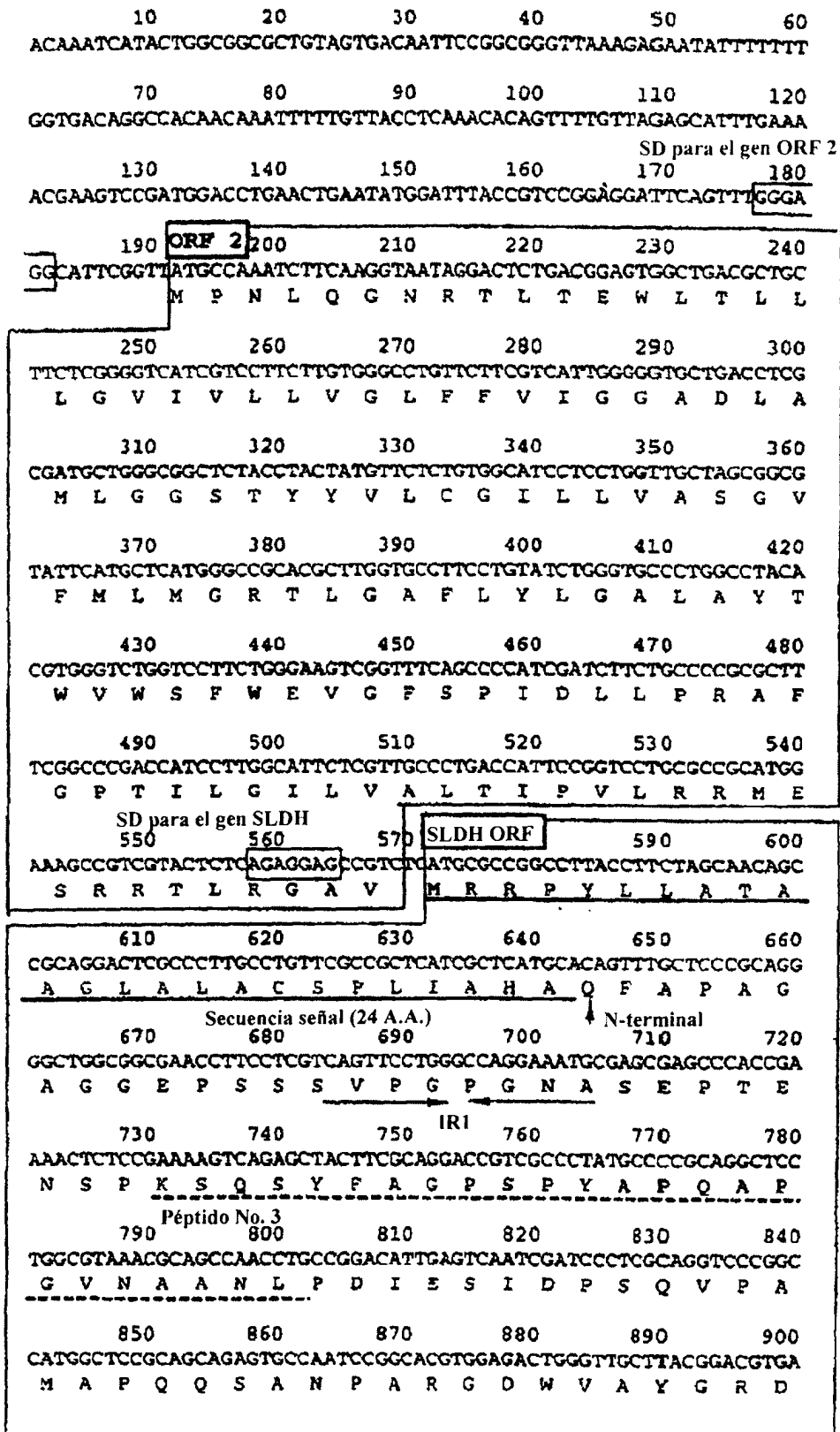


Figura 3a

ES 2 304 792 T3

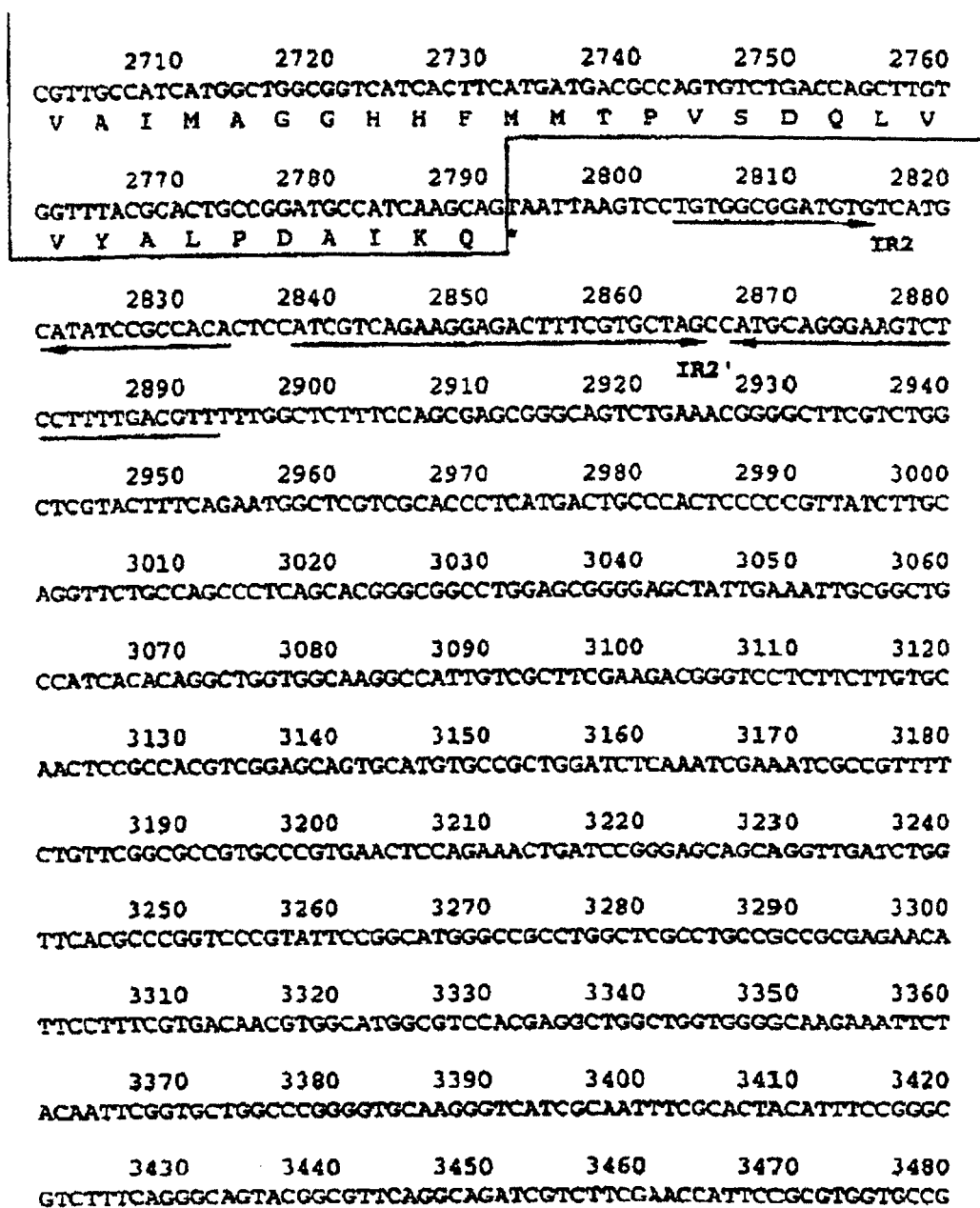
910	920	930	940	950	960
CGATCATCAGACGGGATACTCTCCGCTTTCGGAAATCACGCCTGAGAACGCAAGCAAGCT					
D H Q T R Y S P L S E I T P E N A S K L					
970	980	990	1000	1010	1020
CAAGGTCGCTTTCGCTTACCACACGGGGAGTTATCCGCGTCCGGGACAGGTGAACAAATG					
K V A F V Y H T G S Y P R P G Q V N K W					
					Péptido No. 1
1030	1040	1050	1060	1070	1080
GGCCGCCGAAACCACGCCGATCAAGGTTGGTGACGGTCTCTACACATGTTCCGCCATGAA					
A A E T T P I K V G D G L Y T C S A M N					
1090	1100	1110	1120	1130	1140
CGACATCATCAAGCTGGATCCGGCTACGGGTAAGCAGATCTGGCGTCGGAACGTGGATGT					
D I I K L D P A T G K Q I W R R N V D V					
1150	1160	1170	1180	1190	1200
CAAATACCACTCCATTCCCTATAACCGCTGCCTGTAAGGGTGTGACGTATTTACGTCCTC					
K Y H S I P Y T A A C K G V T Y F T S S					
1210	1220	1230	1240	1250	1260
CGTGGTCCCGGAAGGCCAGCCCTGCCACAATCGCCTTATCGAAGGCACGCTGGATATGGC					
V V P E G Q P C H N R L I E G T L D M R					
1270	1280	1290	1300	1310	1320
TCTGATGCGGTTGACCGCGGAGACAGGGGATTTCTGCCCTAATTTCCGGTCAATGGTGGTCA					
L I A V D A E T G D F C P N F G H G G Q					
1330	1340	1350	1360	1370	1380
GGTCAACCTGATGCAGGGTCTGGGTGAGTCTGTTCCGGGCTTCGTCTCCATGACGGCACC					
V N L M Q G L G E S V P G F V S M T A P					
1390	1400	1410	1420	1430	1440
TCCACCGGTCAATCAACGGCGTCCGTTGTAALACCACGAAGTGCTCGACGGTCAGCGCCG					
P P V I N G V V V V N H E V L D G Q R R					
1450	1460	1470	1480	1490	1500
CTGGGCTCCGTCGGTGTGATCCGTTGATGCTGAAAGTGGCAAATTCGTATGGCC					
W A P S G V I R G Y D A E S G K F V W A					
1510	1520	1530	1540	1550	1560
CTGGGACGTCAACAATTCGGACGATCACAGCCAGCCTACCGGGTAACCGTCAATTACAGC					
W D V N N S G R S Q P A Y R V T V I T A					
1570	1580	1590	1600	1610	1620
CGTGGAACGCCGAATTCCTGGGCTACCTGACAGGCGACAACGAGGAGGGTCTCGTTTACG					
V E R R I P G L P D R R Q R G G S R L R					
1630	1640	1650	1660	1670	1680
TCCCGACAGGAACCTCTGCTGCTGACTATTACAGCGCCCTGCGTAGTGA.TGCTGAAAACAA					
P D R N S A A D Y Y S A L R S D A E N K					
1690	1700	1710	1720	1730	1740
GGTGTCTCCGCTGTTGTCGCCATTGACGTCAAGACGGGTTCTCCGGCTGGGTCTTCCA					
V S S A V V A I D V K T G S P R W V F Q					
1750	1760	1770	1780	1790	1800
GACGGCTCATAAGCACGCTTGGGATTATGACATCGGTTACAGCGGACCCCTGATCGATAT					
T A H K D V W D Y D I G S Q A T L M D M					

Figura 3b

ES 2 304 792 T3

1810 1820 1830 1840 1850 1860  
 GCCTGGCCCGGATGGCCAGACGGTTCCTCTCATCATGCCGACCAAGCGTGGCCAGAC  
 P G P D G Q T V P A L I M P T K R G Q T  
  
 1870 1880 1890 1900 1910 1920  
 GTTCGTGCTTGACCGTTCGTACCGGCAAGCCAATTCTGCCGGTTGAAGAACGCCAGCTCC  
 F V L D R R T G K P I L P V E E R P A P  
  
 1930 1940 1950 1960 1970 1980  
 GTCCCTGGTGTATTCCGGGTGACCCGGTTCCTCCGACGCAGCCATGGTCCGTCCGGAT  
 S P G V I P G D P R S P T Q P W S V G M  
  
 1990 2000 2010 2020 2030 2040  
 GCCGGCCCTTCGCGTCCCGATCTGAAAGAGACAGACATGTGGGGTATGTCCCCCATCGA  
 P A L R V P D L K E T D M W G M S P I D  
  
 2050 2060 2070 2080 2090 2100  
 TCAGCTCTCTGCCGTATCAAGTTCGCGGTGCGAACTATGTGGGTGAGTTCACACCACC  
 Q L F C R I K F R R A N Y V G E F T P P  
  
 2110 2120 2130 2140 2150 2160  
 GAGCGTTGACAAGCCGTGGATTGAATATCCGGGCTATAACGGTGGCAGTACTGGGGCTC  
 S V D K P W I E Y P G Y N G G S D W G S  
  
 2170 2180 2190 2200 2210 2220  
 CATGTCCTATGATCCGCACTCCGGCATCTGATTGCGAACTGGAACATCACACCGATGTA  
 M S Y D P Q S G I L I A N W N I T P M Y  
  
 2230 2240 2250 2260 2270 2280  
 CGACCAGCTCGTAACCCGCAAGAAGGCAGACTCCCTCCGGCTGATGCCGATCGATGACCC  
 D Q L V T R R K K A D S L G L M P I D D P  
  
 2290 2300 2310 2320 2330 2340  
 CAACTTCAAGCCAGGTGGCGGTGGTCCCGAAGGTAACGGCGCCATGGACGGAACGCCCTTA  
 N F K P G G G G A E G N G A M D G T P Y  
  
 2350 2360 2370 2380 2390 2400  
 CGGTATCGTCTGTGACACCGTCTCTGGGATCAGTACACGGCCATGATGTGCAACCGTCCGCC  
 G I V V T P F W D Q Y T G M M C N R P P  
  
 2410 2420 2430 2440 2450 2460  
 CTACGGTATGATCACAGCCATCGACATGAAGCAGCGCCAGAAGGTTCTGTGGCAGCATCC  
 Y G M I T A I D M K H G Q K V L W Q H P  
 Péptido No. 8  
 2470 2480 2490 2500 2510 2520  
 GCTCGGAACGGCTCGCGCCAACGGTCCATGGGGTCTGCCAACAGGTCGCCATGGGAAT  
 L G T A R A N G P W G L P T G L P W E I  
 -----  
 2530 2540 2550 2560 2570 2580  
 CGGCACTCCGAACAATGGTGGTTCGGTGTGACCGGTGGCGGTCTGATCTTCATCGGTGC  
 G T P N N G G S V V T G G G L I F I G A  
  
 2590 2600 2610 2620 2630 2640  
 GGCAACGGATAACCAGATCCGCGGATGATGAACACACTGGCAAGGTTGTCTGGAGCCG  
 A T D N Q I R A I D E H T G K V V W S A  
  
 2650 2660 2670 2680 2690 2700  
 AGTCTCCCCCGGGCGGTCCAGGCCAATCCGATGACGTATGAAGCCAATGGTCACCAGTA  
 V L P G G G Q A N P M T Y E A N G H Q Y

Figura 3c



3481  
A



: IR1 y IR2 (o 2') son repetitivas invertidas como posibles terminadores de la transcripción para los genes SLDH II y ORF2 respectivamente.

Figura 3d

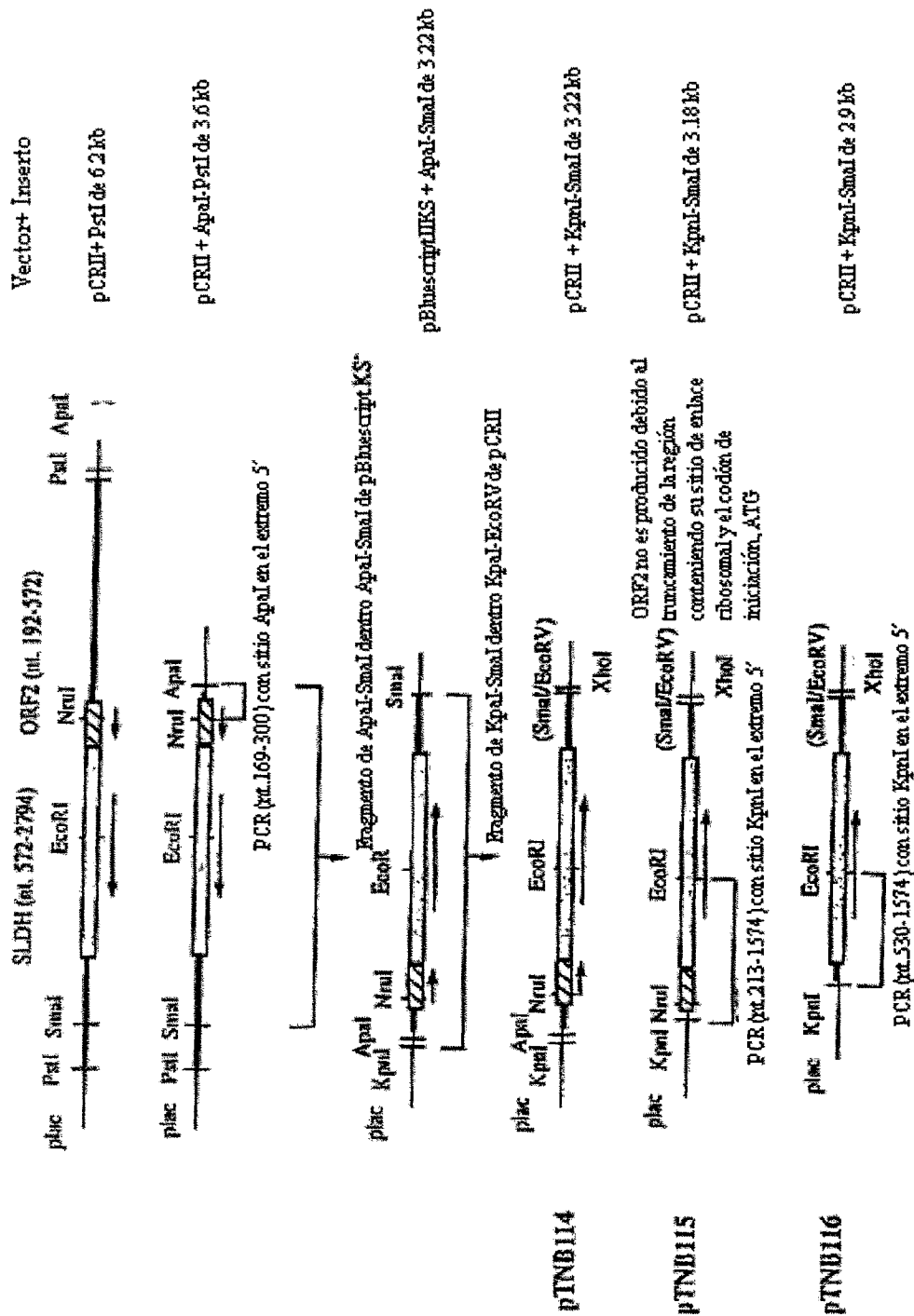


Figura 4



# ES 2 304 792 T3

## LISTA DE SECUENCIAS

### (1) INFORMACIÓN GENERAL

- 5 (i) SOLICITANTE  
NOMBRE: F. HOFFMANN-LA ROCHE AG  
CALLE: Grezacherstrasse 124  
10 CIUDAD: Basle  
PAÍS: Suiza  
CÓDIGO POSTAL: CH-4002  
15 TELÉFONO: 061-688 25 11  
FAX: 061-688 13 95  
TELEX: 962292/965542 hlr c  
20 (ii) TÍTULO DE LA INVENCIÓN:  
Gen de la D-Sorbitol deshidrogenasa  
25 (iii) NÚMERO DE SECUENCIAS: 9  
(iv) FORMA LEGIBLE POR COMPUTADORA:  
(A) TIPO DE MEDIO: Disquete  
30 (B) ORDENADOR: Macintosh  
(C) SISTEMA OPERATIVO:  
(D) SOFTWARE: MS word ver 5.1a  
(v) DATOS DE LA APLICACIÓN EN CURSO:  
35 (A) NÚMERO DE LA APLICACIÓN:  
(B) DATOS CLASIFICADOS:  
(C) CLASIFICACIÓN:

### 40 (2) INFORMACIÓN PARA LA SEQ ID NO:1

- (i) CARACTERÍSTICAS DE SECUENCIA:  
45 (A) LONGITUD: 3481 pares de base  
(B) TIPO: ácido nucleico  
(C) HEBRA: doble  
(D) TOPOLOGÍA: lineal  
50 (ii) TIPO DE MOLÉCULA: ADN (genómico)  
(iii) FUENTE ORIGINAL:  
ORGANISMO: *Gluconobacter suboxydans*  
55 CEPA: IFO 3255  
(iv) RASGOS:  
60 RASGOS CLAVES: CDS  
POSICIÓN: 192..572  
MÉTODO DE SECUENCIACIÓN: E  
65 RASGOS CLAVES: CDS  
POSICIÓN: 572..2794

# ES 2 304 792 T3

## MÉTODO DE SECUENCIACIÓN: E

5 ACAAATCATA CTGGCGGCGC TGTAGTGACA ATTCCGGCGG GTTAAAGAGA ATATTTTTTT 60  
 GGTGACAGGC CACAACAAAT TTTTGTACC TCAAACACAG TTTTGTAGA GCATTTGAAA 120  
 ACGAAGTCCG ATGGACCTGA ACTGAATATG GATTTACCGT CCGGAGGATT CAGTTTGGGA 180  
 10 GGCATTCCGT TATGCCAAAT CTTCAAGGTA ATAGGACTCT GACGGAGTGG CTGACGCTGC 240  
 TTCTCGGGGT CATCGTCCPT CTTGTGGGCC TGTCTTCGT CATTGGGGGT GCTGACCTCG 300  
 15 CGATGCTGGG CGGCTCTACC TACTATGTTT TCTGTGGCAT CCTCCTGGTT GCTAGCGGCG 360  
 TATTCATGCT CATGGGCCGC ACGCTTGGTG CCTTCCGTA TCTGGGTGCC CTGGCCTACA 420  
 CGTGGGTCTG GTCCTTCTGG GAAGTCCGTT TCAGCCCCAT CGATCTTCTG CCCCAGCCTT 480  
 20 TCGGCCGAC CATCCTTGGC ATTCTCGTTG CCTGACCAT TCCGGTCCTG CGCGCATGG 540  
 AAAGCCGTCG TACTCTCAGA GGAGCCGCTT GATGGGCCG CCTTACCTTC TAGCAACAGC 500  
 25 CGCAGGACTC GCCCTTGCCT GTTCGCCGCT CATCGCTCAT GCACAGTTTG CTCCCGCAGG 560  
 GGCTGGCGGC GAACCTTCCT CBTCACTTC TGGGCCAGGA AATGGGAGCG AGCCACCCGA 720  
 AACTCTCCG AAAAGTCAGA GCTACTTCCG AGGACCGTCG CCTATGCCC CGCAGGCTCC 780  
 30 TGGCGTAAAC GCAGCCAACC TGCCGGACAT TGAGTCAATC GATCCCTCGC AGGTCCCGGC 840  
 CATGGCTCCG CAGCAGAGTG CCAATCCGCG ACGTGGAGAC TGGGTTGCTT ACGGACGTGA 900  
 35 CGATCATCAG ACGCGATACT CTCGCTTTC GGAATCAGC CCTGAGAAGC CAAGCAAGCT 960  
 CAAGGTCGCT TTCGTCTACC ACACGGGGAG TTATCCGCT CCGGACAGG TGAACAAATG 1020  
 40 GGCCGCCGAA ACCACGCCGA TCAAGGTTGG TGACGGTCTC TACACATGTT CCGCCATGAA 1080  
 CGACATCATC AAGCTGGATC CGGCTACGGG TAAGCAGATC TGGGTCGGGA ACGTGGATGT 1140  
 CAAATACCAC TCCATTCCCT ATACCGCTGC CTGTAAGGT GTGACGTATT TCACGTCTC 1200  
 45 CGTGGTGCGB GAAGGCCAGC CCTGCCACAA TCGCCTTATC GAAGGCACGC TGGATATGCG 1260  
 TCTGATTGGG GTTACCGCGG AGACAGGGGA TTTCTGCCCT AATTTCGGTC ATGGTGGTCA 1320  
 50 GGTCAACCTG ATGCAGGGTC TGGGTGAGTC TCTTCCGGC TCTGCTCCA TGACGGCAC 1380  
 TCCACCGGTC ATCAACGGCG TCGTGGTTGT AAACCACGAA GTGCTCGACG GTCAGCGCCG 1440  
 CTGGGCTCCG TCCGGTGTGA TCCGTGGTGA CGATGCTGAA AGTGGCAAT TCGTATGGGC 1500  
 55 CTGGGACGTC AACAAATCCG GACGATCACA GCCAGCCTAC CGGGTAACCG TCATTACAGC 1560  
 CGTGGAACGC CGAATTCTG GGTACCTGA CAGGCGACAA CGAGGAGGGT CTCGTTTACG 1620  
 60 TCCCAGACAG AACTCTGCTG CTGACTATTA CAGGCGCCTG CBTAGTGATG CTGAAAACAA 1680

65

# ES 2 304 792 T3

5 GGTGTCTCTCC GCTGTTCTCG CCATTGACGT CAAGACGGGT TCTCCGGCGT GGGTCTTCCA 1740  
 GACGGCTCAT AAGGACGTCT GGGATTATGA CATCGGTCA CAGGCGACCC TGATGSATAT 1800  
 GCCTGGCCCG GATGGCCAGA CCGTTCCTGC TCTCATCATG CCGACCAAGC GTGCCAGAC 1860  
 10 GTTCGTCTTT GACCGTCGTA CCGGCAAGCC AATCTGCCC GTTGAAGAAC GCCCAGCTCC 1920  
 GTCCCTGGT GTTATTCCGG GTGACCCGGG TTCTCCGACG CAGCCATGGT CCGTCCGGAT 1980  
 15 GCCCGCCCTT CGCGTGCCGG ATCTGAAAGA GACAGACATG TGGGGTATGT CCCCATCGA 2040  
 TCAGCTCTTC TGCCGTATCA AGTTCCGGCG TGCGAACTAT GTGGGTGAGT TCACACCACC 2100  
 20 GAGCGTTGAC AAGCCGTGGA TTGAATATCC GGGCTATAAC GGTGGCAGTG ACTGGGGCTC 2160  
 CATGTCTTAT GATCCGCAGT CCGGCATCCT GATTGCGAAC TGGAACATCA CACCGATGTA 2220  
 25 CGACCAGCTC GTAACCCGCA AGAAGGCAGA CTCCTCCGGC CTGATGCCGA TCGATGACCC 2280  
 CAACTTCAAG CCAGGTGGCG GTGGTGCCGA AGGTAACGGC GCCATGGACG GAACGCCTTA 2340  
 30 CGSTATCGTC GTGACACCGT TCTGGGATCA GTACACGGGC ATGATGTGCA ACCGTCCGCC 2400  
 CTACGGTATG ATCACAGCCA TCGACATGAA GCACGGCCAG AAGGTTCTGT GGCAGCATCC 2460  
 35 GCTCGGAACG GCTCGCGCCA ACGGTCCATG GGTCTTGCCA ACAGGTCTGC CATGGGAAAT 2520  
 CGGCACTCCG AACAAATGGT GTTCGGTGTG TACCGGTGGC GGTCTGATCT TCATCGGTGC 2580  
 40 GGCAACGGAT AACCAGATCC GCGCGATTGA TGAACACACT GGCAAGGTTG TCTGGAGCGC 2640  
 AGTCCTCCCC GCGGGCGGTC AGGCCAATCC GATGACGTAT GAAGCCAATG GTCACCAGTA 2700  
 45 CGTTGCCATC ATGGCTGGCG GTCATCACTT CATGATGACG CCAGTGTCTG ACCAGCTTGT 2760  
 GGTTTACGCA CTGCCGGATG CCATCAAGCA GTAATTAAGT CCTGTGGCGG ATGTGTATG 2820  
 50 CATATCCGCC AACTTCCATC GTCAGAAGGA GACTTTCGTG CTAGCCATGC AGGGAAGTCT 2880  
 CCTTTTGACG TTTTGGCTC TTTCCAGCGA GCGGGCAGTC TGAACGGGG CTTCGTCTGG 2940  
 55 CTCGTACTTT CAGAATGGCT GTTCGCACCC TCATGACTGC CCACFCCCC GTPATCTTGC 3000  
 AGGTTCTGCC AGCCCTCAGC ACGGGCGGCC TGGAGCGGGG AGCTATTGAA ATTCCGGCTG 3060  
 60 CCAFCACACA GGCTGGTGGC AAGGCCATTG TCGCTTCGAA GACGGTCCCT CTCTTGTGC 3120

65

ES 2 304 792 T3

AACTCCGCCA CGTGGGAGCA GTGCATGTGC CGCTGGATCT CAAATCGAAA TCGCCGTTTT 3180  
5 CTGTTCCGGC CCGTGCCCGT GAACTCCAGA AACTGATCCG GGAGCAGCAG GTTGATCTGG 3240  
TTCACGCCCG GTCCCGTATT CCGGCATGGG CCGCCTGGCT CGCCTGCCGC CGCGAGAACA 3300  
10 TTCCTTTCGT GACAACGTGG CATGGCGTCC ACGAGGCTGG CTGGTGGGGC AAGAAATTCT 3360  
ACAATTCGGT GCTGSCCCGG GGTGCAAGGG TCATCGCAAT TTCGCACTAC ATTTCCGGGC 3420  
15 GTCTTTCAGG GCAGTACGGC GTTCAGGCAG ATCGICTTGG AACCATTCGG CGTGGTGCCG 3480

A

20 INFORMACIÓN PARA LA SEQ ID NO:2:

- (i) CARACTERÍSTICAS DE SECUENCIA:  
25 (A) LONGITUD: 740 residuos  
(B) TIPO: aminoácidos  
(C) TOPOLOGÍA: lineal (ii)

30 (ii) TIPO DE MOLÉCULA: proteína

(iii) FUENTE ORIGINAL:

ORGANISMO: *Gluconobacter suboxydans*  
35 CEPA: IFO 3255

(iv) RASGOS:

RASGOS CLAVES: péptido sig  
40 POSICIÓN: -24..-1  
MÉTODO DE SECUENCIACIÓN: E  
RASGOS CLAVES: péptido mat  
45 POSICIÓN: 1..716  
MÉTODO DE SECUENCIACIÓN: E

50 Met Arg Arg Pro Tyr Leu Leu Ala Thr Ala Ala Gly Leu Ala Leu

-24 -20 -15 -10

55 Ala Cys Ser Pro Leu Ile Ala His Ala Gln Phe Ala Pro Ala Gly

-5 1 5

60

65

ES 2 304 792 T3

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

ES 2 304 792 T3

	Pro Tyr Thr Ala Ala Cys Lys Gly Val Thr Tyr Phe Thr Ser Ser
5	175 180 185
	Val Val Pro Glu Gly Gln Pro Cys His Asn Arg Leu Ile Glu Gly
10	190 195 200
	Thr Leu Asp Met Arg Leu Ile Ala Val Asp Ala Glu Thr Gly Asp
15	205 210 215
	Phe Cys Pro Asn Phe Gly His Gly Gly Gln Val Asn Leu Met Gln
20	220 225 230
	Gly Leu Gly Glu Ser Val Pro Gly Phe Val Ser Met Thr Ala Pro
25	235 240 245
	Pro Pro Val Ile Asn Gly Val Val Val Val Asn His Glu Val Leu
30	250 255 260
	Asp Gly Gln Arg Arg Trp Ala Pro Ser Gly Val Ile Arg Gly Tyr
35	265 270 275
	Asp Ala Glu Ser Gly Lys Phe Val Trp Ala Trp Asp Val Asn Asn
40	280 285 290
	Ser Gly Arg Ser Gln Pro Ala Tyr Arg Val Thr Val Ile Thr Ala
45	295 300 305
	Val Glu Arg Arg Ile Pro Gly Leu Pro Asp Arg Arg Gln Arg Gly
50	310 315 320
	Gly Ser Arg Leu Arg Pro Asp Arg Asn Ser Ala Ala Asp Tyr Tyr
55	325 330 335

60

65

ES 2 304 792 T3

Ser Ala Leu Arg Ser Asp Ala Glu Asn Lys Val Ser Ser Ala Val  
 5                   340                   345                   350  
 Val Ala Ile Asp Val Lys Thr Gly Ser Pro Arg Trp Val Phe Gln  
 10                   355                   360                   365  
 Thr Ala His Lys Asp Val Trp Asp Tyr Asp Ile Gly Ser Gln Ala  
 15                   370                   375                   380  
 Thr Leu Met Asp Met Pro Gly Pro Asp Gly Gln Thr Val Pro Ala  
 20                   385                   390                   395  
 Leu Ile Met Pro Thr Lys Arg Gly Gln Thr Phe Val Leu Asp Arg  
 25                   400                   405                   410  
 Arg Thr Gly Lys Pro Ile Leu Pro Val Glu Glu Arg Pro Ala Pro  
 30                   415                   420                   425  
 Ser Pro Gly Val Ile Pro Gly Asp Pro Arg Ser Pro Thr Gln Pro  
 35                   430                   435                   440  
 Trp Ser Val Gly Met Pro Ala Leu Arg Val Pro Asp Leu Lys Glu  
 Gen D-Sorbitol deshidrogenasa                   445  
 40                   450                   455  
 Thr Asp Met Trp Gly Met Ser Pro Ile Asp Gln Leu Phe Cys Arg  
 45                   460                   465                   470  
 Ile Lys Phe Arg Arg Ala Asn Tyr Val Gly Glu Phe Thr Pro Pro  
 50                   475                   480                   485  
 Ser Val Asp Lys Pro Trp Ile Glu Tyr Pro Gly Tyr Asn Gly Gly  
 55                   490                   495                   500  
 60  
 65

ES 2 304 792 T3

	Ser Asp Trp Gly Ser Met Ser Tyr Asp Pro Gln Ser Gly Ile Leu	
5	505	510 515
	Ile Ala Asn Trp Asn Ile Thr Pro Met Tyr Asp Gln Leu Val Thr	
10	520	525 530
	Arg Lys Lys Ala Asp Ser Leu Gly Leu Met Pro Ile Asp Asp Pro	
15	535	540 545
	Asn Phe Lys Pro Gly Gly Gly Gly Ala Glu Gly Asn Gly Ala Met	
20	550	555 560
	Asp Gly Thr Pro Tyr Gly Ile Val Val Thr Pro Phe Trp Asp Gln	
25	565	570 575
	Tyr Thr Gly Met Met Cys Asn Arg Pro Pro Tyr Gly Met Ile Thr	
30	580	585 590
	Ala Ile Asp Met Lys His Gly Gln Lys Val Leu Trp Gln His Pro	
35	595	600 605
	Leu Gly Thr Ala Arg Ala Asn Gly Pro Trp Gly Leu Pro Thr Gly	
40	610	615 620
	Leu Pro Trp Glu Ile Gly Thr Pro Asn Asn Gly Gly Ser Val Val	
45	625	630 635
	Thr Gly Gly Gly Leu Ile Phe Ile Gly Ala Ala Thr Asp Asn Gln	
50	640	645 650
	Ile Arg Ala Ile Asp Glu His Thr Gly Lys Val Val Trp Ser Ala	
55	655	660 665

ES 2 304 792 T3

Val Leu Pro Gly Gly Gly Gln Ala Asn Pro Met Thr Tyr Glu Ala  
 5                           670                           675                           680  
 Asn Gly His Gln Tyr Val Ala Ile Met Ala Gly Gly His His Phe  
 10                           685                           690                           695  
 Met Met Thr Pro Val Ser Asp Gln Leu Val Val Tyr Ala Leu Pro  
 15                           700                           705                           710  
 Asp Ala Ile Lys Gln  
 20                           715 716

INFORMACIÓN PARA LA SEQ ID NO:3:

- 25 (i) CARACTERÍSTICAS DE SECUENCIA:
- (A) LONGITUD: 126 residuos
  - (B) TIPO: aminoácidos
  - (C) TOPOLOGÍA: lineal (ii)
- 30 (ii) TIPO DE MOLÉCULA: proteína
- (iii) TIPO DE FRAGMENTO: fragmento interno
- 35 (iv) FUENTE ORIGINAL:
- ORGANISMO: *Gluconobacter suboxydans*
  - CEPA: IFO 3255
- 40 (v) RASGOS:
- RASGOS CLAVES: péptido mat
  - POSICIÓN: 1..126
- 45 MÉTODO DE SECUENCIACIÓN: E

Met Pro Asn Leu Gln Gly Asn Arg Thr Leu Thr Glu Trp Leu Thr  
 50           1                           5                           10                           15  
 Leu Leu Leu Gly Val Ile Val Leu Leu Val Gly Leu Phe Phe Val  
 55                           20                           25                           30  
 Ile Gly Gly Ala Asp Leu Ala Met Leu Gly Gly Ser Thr Tyr Tyr  
 60                           35                           40                           45  
 Val Leu Cys Gly Ile Leu Leu Val Ala Ser Gly Val Phe Met Leu  
 65                           50                           55                           60



## ES 2 304 792 T3

CEPA: IFO 3255

(v) RASGOS:

MÉTODO DE SECUENCIACIÓN: E

**Lys Ser Gln Ser Tyr Phe Ala Gly Pro Ser Gln Tyr Ala Pro Gln**

**1 5 10 15**

**Ala Pro Gly Val Asn Ala Xaa Asn Leu**

**20 24**

INFORMACIÓN PARA LA SEQ ID NO:6:

(i) CARACTERÍSTICAS DE SECUENCIA:

(A) LONGITUD: 16 residuos

(B) TIPO: aminoácidos

(C) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: péptido

(iii) TIPO DE FRAGMENTO: fragmento interno

(iv) FUENTE ORIGINAL:

ORGANISMO: *Gluconobacter suboxydans*

CEPA: IFO 3255

(v) RASGOS:

MÉTODO DE SECUENCIACIÓN: E

**Lys Val Leu Trp Gln His Pro Leu Gly Thr Ala Arg Xaa Asn Gly**

**1 5 10 15**

**Pro**

**16**

INFORMACIÓN PARA LA SEQ ID NO:7:

(i) CARACTERÍSTICAS DE SECUENCIA:

(A) LONGITUD: 30 residuos

(B) TIPO: aminoácidos

(C) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: péptido

(iii) TIPO DE FRAGMENTO: fragmento N-terminal

(iv) FUENTE ORIGINAL:

ORGANISMO: *Gluconobacter suboxydans*

## ES 2 304 792 T3

CEPA: IFO 3255

(v) RASGOS:

5 MÉTODO DE SECUENCIACIÓN: E

**Gln Phe Ala Pro Ala Gly Ala Gly Gly Glu Pro Ser Ser Ser Val**

10 1 5 10 15

**Pro Gly Pro Gly Asn Ala Ser Glu Pro Thr Glu Asn Ser Pro Lys**

15 20 25 30

INFORMACIÓN PARA LA SEQ ID NO:8:

20 (i) CARACTERÍSTICAS DE SECUENCIA:

(A) LONGITUD: 24 residuos

(B) TIPO: aminoácidos

25 (C) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: péptido

(iii) FUENTE ORIGINAL:

30 ORGANISMO: *Gluconobacter suboxydans*

CEPA: IFO 3255

35 (iv) RASGOS:

RASGOS CLAVES: péptido sig

POSICIÓN: 1..24

40 MÉTODO DE SECUENCIACIÓN: E

**Met Arg Arg Pro Tyr Leu Leu Ala Thr Ala Ala Gly Leu Ala Leu**

45 1 5 10 15

**Ala Cys Ser Pro Leu Ile Ala His Ala**

50 20 24

55 INFORMACIÓN PARA LA SEQ ID NO:9:

(i) CARACTERÍSTICAS DE SECUENCIA:

(A) LONGITUD: 6 residuos

60 (B) TIPO: aminoácidos

(C) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: péptido

65 (iii) TIPO DE FRAGMENTO: fragmento C-terminal

(iv) FUENTE ORIGINAL:

# ES 2 304 792 T3

ORGANISMO: *Gluconobacter suboxydans*

CEPA: IFO 3255

5 (v) RASGOS:

MÉTODO DE SECUENCIACIÓN: E

**Pro Asp Ala Ile Lys Gln**

10

**1 5 6**

15

20

25

30

35

40

45

50

55

60

65