

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 684 129**

51 Int. Cl.:

A61K 39/00 (2006.01)
A61K 39/108 (2006.01)
A61K 39/104 (2006.01)
A61K 39/112 (2006.01)
A61K 39/02 (2006.01)
A61K 39/385 (2006.01)
C07K 14/195 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 86 Fecha de presentación y número de la solicitud internacional: **19.02.2009 PCT/IB2009/000287**
 87 Fecha y número de publicación internacional: **27.08.2009 WO09104074**
 96 Fecha de presentación y número de la solicitud europea: **19.02.2009 E 09712440 (8)**
 97 Fecha y número de publicación de la concesión europea: **06.06.2018 EP 2257307**

54 Título: **Bioconjugados preparados a partir de proteínas N-glicosiladas recombinantes procedentes de células procariotas**

30 Prioridad:

20.02.2008 US 64163
05.05.2008 US 71545
30.06.2008 US 129480
24.07.2008 US 129852
25.09.2008 US 136687

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
01.10.2018

73 Titular/es:

GLAXOSMITHKLINE BIOLOGICALS S.A. (100.0%)
Rue de l'Institut, 89
1330 Rixensart, BE

72 Inventor/es:

FERNANDEZ, FABIANA;
WETTER, MICHAEL;
KOWARIK, MICHAEL y
WACKER, MICHAEL

74 Agente/Representante:

ELZABURU, S.L.P

ES 2 684 129 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

Bioconjugados preparados a partir de proteínas N-glicosiladas recombinantes procedentes de células procariotas

Campo de la invención

5 La presente invención se refiere a bioconjugados, específicamente a vacunas de bioconjugados, preparadas a partir de glicoproteínas recombinantes, a saber, proteínas N-glicosiladas. La invención comprende una o varias proteínas N-glicosiladas introducidas con una o unas secuencias de consenso de aminoácidos optimizadas, a ácidos nucleicos que codifican esas proteínas, así como a los vectores y células hospedadoras correspondientes. Además, la presente invención se dirige al uso de dichas proteínas, ácidos nucleicos, vectores y células hospedadoras para la preparación de vacunas de bioconjugados. Además, la presente invención proporciona métodos para producir vacunas de bioconjugados.

Antecedentes de la invención

15 Las glicoproteínas son proteínas que tienen uno o varios polímeros de azúcar fijados covalentemente. Una glicosilación de proteínas ligada a N es un procedimiento esencial y conservado que se produce en el retículo endoplasmático de organismos eucariotas. Es importante para el plegamiento de proteínas, la oligomerización, la estabilidad, el control de calidad, la clasificación y el transporte de proteínas secretoras y de membrana (Helenius, A. y Aebi, M. (2004). Roles of N-linked glycans in the endoplasmic reticulum. *Annu. Rev. Biochem.* 73, 1019-1049).

20 La glicosilación de proteínas tiene una gran influencia sobre la antigenicidad, la estabilidad y la semivida de una proteína. Además, la glicosilación puede ayudar en la purificación de proteínas mediante cromatografía, por ejemplo, la cromatografía por afinidad con ligandos de lectina unidos a una fase sólida que interacciona con restos glicosilados de la proteína. Por lo tanto, es una práctica establecida para producir muchas proteínas glicosiladas de forma recombinante en células eucariotas para proporcionar patrones de glicosilación útiles biológica y farmacéuticamente.

25 Se ha mostrado que una bacteria, el agente patógeno *Campylobacter jejuni* transmitido con los alimentos, también puede N-glicosilar sus proteínas (Szymanski, et al. (1999). Evidence for a system of general protein glycosylation in *Campylobacter jejuni*. *Mol. Microbiol.* 32, 1022-1030). La maquinaria necesaria para la glicosilación está codificada por 12 genes que están agrupados en el locus denominado *pgl*.

30 Una alteración de la N-glicosilación afecta a la invasión y la patogénesis de *C. jejuni*, pero no es letal como en la mayoría de los organismos eucariotas (Burda P. y M. Aebi, (1999). The dolichol pathway of N-linked glycosylation. *Biochim Biophys Acta* 1426(2):239-57). Es posible reconstituir la N-glicosilación de las proteínas de *C. jejuni*, expresando de forma recombinante el locus *pgl* y la glicoproteína aceptora en *E. coli* al mismo tiempo (Wacker et al. (2002). N-linked glycosylation in *Campylobacter jejuni* and its functional transfer into *E. coli*. *Science* 298, 1790-1793, documento WO 06/119987, Feldman et al. (2005), Engineering N-linked protein glycosylation with diverse O antigen lipopolysaccharide structures in *Escherichia coli*. *PNAS* 102; 3016-3021). Woo y Neville (2003), Separation of bivalent anti-T cell immunotoxin from *Pichia pastoris* glycoproteins by borate anion exchange, *BioTechniques* 35; 392-398, documentos WO 05/12495, WO 93/25690 y Johansson et al. (1996) *Journal of Biotechnology* 48; 9-14 describen modos de purificar proteínas.

40 La enfermedad diarreica es un problema importante sanitario, asociado con viajes internacionales en términos de frecuencia y de impacto económico. La diarrea del viajero se refiere a una enfermedad entérica adquirida cuando una persona viaja desde un país desarrollado a un país en desarrollo. Hoy en día, más de 50 millones de personas viajan cada año desde países desarrollados a países en desarrollo y hasta el 50% de estos viajeros informan que han tenido diarrea durante las 2 primeras semanas de su estancia. No ha habido una disminución significativa en la incidencia de la diarrea del viajero desde los años 1970, a pesar de los esfuerzos realizados por la industria del turismo para mejorar la infraestructura local.

45 La diarrea del viajero se adquiere a través de la ingestión de comida contaminada con heces y menos comúnmente agua. Las bacterias son la principal causa de la diarrea del viajero, siendo responsable de hasta un 80% de las infecciones. *E. coli* enterotoxigénica (ETEC) es la bacteria aislada más frecuentemente en todas partes del mundo asociadas con la diarrea del viajero, seguida por *Shigella* spp y *C. jejuni*.

50 La shigelosis sigue siendo una enfermedad grave y común. Además de causar diarrea acuosa, las *Shigellae* son una causa importante de disentería (fiebre, calambres y sangre y/o moco en las heces). El hombre es el único hospedador natural de esta bacteria. El número estimado de infecciones por *Shigella* es más de 200 millones anuales. Aproximadamente 5 millones de estos casos necesitan hospitalización y un millón de personas mueren. Hay tres serogrupos que son los principales responsables de la enfermedad descrita como disentería bacilar: *S. dysenteriae*, *S. flexneri* y *S. sonnei*. El documento WO 05/116063 da a conocer una vacuna oral viva para la protección contra *Shigella dysenteriae* de serotipo 1.

55 *S. dysenteriae* y *S. flexneri* son responsables de la mayoría de las infecciones en los trópicos, con una mortalidad de hasta un 20%. La shigelosis se produce tanto de forma endémica como epidémica. En muchos países tropicales, una infección endémica se debe principalmente a *S. flexneri*, mientras que las epidemias principales de *S. dysente-*

riae se producen en América central, África central y el sudeste asiático. Estas epidemias son riesgos importantes para la salud pública. Las infecciones, principalmente debidas a *S. sonnei* y con menor frecuencia *S. flexneri*, se siguen produciendo en los países industrializados.

5 Las vacunas conjugadas han mostrado unos resultados prometedores contra las infecciones de *Shigella*. Se han empleado polisacáridos específicos de O de *S. dysenteriae* de tipo 1 para sintetizar una vacuna conjugada que ha provocado una respuesta inmune en ratones. Tales vacunas se han sintetizado químicamente y se han conjugado con albúmina de suero humano o se han desarrollado en donde el polisacárido O se ha purificado a partir de *Shigella*. Los polisacáridos específicos de O de *S. sonnei* y *S. flexneri* también se han conjugado químicamente con la exotoxina de *P. aeruginosa* y han provocado una respuesta inmune significativa en ratones. Además, han mostrado ser inmunogénicos y seguros en los seres humanos. Sin embargo, la conjugación química es un proceso costoso y que requiere tiempo, que no siempre proporciona vacunas fiables y reproducibles. Esto conduce a problemas de Normas de Correcta Fabricación (NCF) cuando se intenta desarrollar tales vacunas de bioconjugados a escala comercial.

Compendio de la invención

15 En un aspecto, la presente invención se dirige a un método para producir un bioconjugado de *Shigella* que comprende producir, aislar y purificar una proteína N-glicosilada recombinante, que comprende las etapas de:

a) proporcionar un organismo recombinante, preferiblemente un organismo procarionta que comprende ácidos nucleicos que codifican

i) un operón *pgl* funcional de procedente *Campylobacter* spp., preferiblemente de *C. jejuni*, y

20 ii) al menos una proteína diana recombinante, la exotoxina de *Pseudomonas aeruginosa* (EPA), que comprende una o varias de las siguientes secuencias de consenso de aminoácidos optimizadas N-glicosiladas:



25 en donde X y Z pueden ser cualquier aminoácido natural excepto Pro, y en donde al menos se introduce una de dichas secuencias de aminoácidos de consenso de aminoácidos optimizadas N-glicosiladas, y

(iii) la síntesis del polisacárido de *Shigella*, y

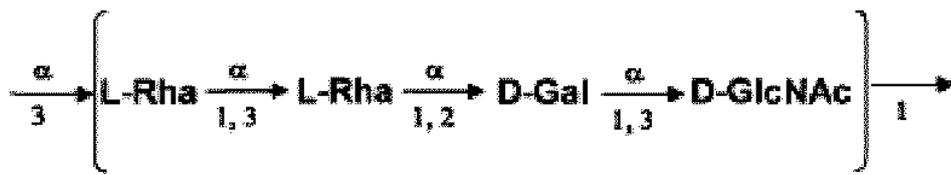
b) cultivar el organismo recombinante de una forma adecuada para la producción y la N-glicosilación de la(s) proteína(s) diana,

c) eliminar la membrana externa de dicha bacteria Gram-negativa recombinante; y

30 d) recuperar dicha proteína recombinante mediante cromatografía de intercambio aniónico y de exclusión por tamaño.

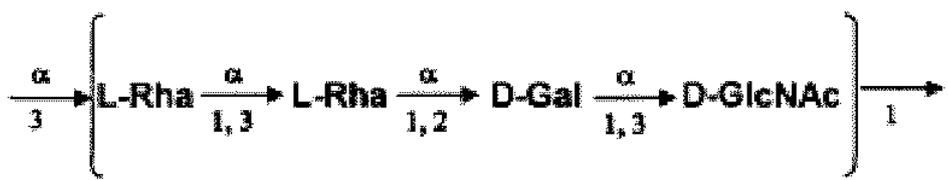
Otro aspecto de la descripción se dirige a una vacuna de bioconjugado que comprende: un vehículo de proteína que comprende una secuencia de consenso insertada, D/E - X - N - Z - S/T, en la que X y Z pueden ser cualquier aminoácido natural excepto prolina; al menos un polisacárido antigénico procedente de al menos una bacteria, ligado al vehículo de proteína, en donde el al menos un polisacárido antigénico es al menos un antígeno O bacteriano procedente de una o varias cepas de *Shigella*, *E. coli* o *Pseudomonas aeruginosa*; y, opcionalmente, un adyuvante.

En otro aspecto, la presente descripción se dirige a una vacuna de bioconjugado de *Shigella* que comprende: un vehículo de proteína que comprende la exotoxina de *Pseudomonas aeruginosa* (EPA) que se ha modificado para contener al menos una secuencia de consenso D/E - X - N - Z - S/T, en la que X y Z pueden ser cualquier aminoácido natural excepto prolina; al menos una cadena de polisacárido ligada al vehículo de proteína y que tiene la siguiente estructura:



y, opcionalmente, un adyuvante.

45 En aún otro aspecto, la presente descripción se refiere a una vacuna de bioconjugado de *Shigella dysenteriae* O1 que comprende: un vehículo de proteína que tiene la secuencia proporcionada en SEQ ID NO: 7; al menos una cadena de polisacárido ligada al vehículo de proteína y que tiene la siguiente estructura:



y un adyuvante.

En aspectos adicionales, la presente descripción se dirige a: un plásmido que comprende SEQ. ID NO. 5; una secuencia genética que comprende SEQ. ID NO. 5; una secuencia de aminoácidos que comprende SEQ. ID NO. 6; una secuencia de aminoácidos que comprende SEQ. ID NO. 7; o un vector pGVXN64.

En otro aspecto, la presente descripción se dirige a un sistema de expresión para producir una vacuna de bioconjugado contra al menos una bacteria que comprende: una secuencia de nucleótidos que codifica una oligosacaril transferasa (OST/OTasa); una secuencia de nucleótidos que codifica un vehículo de proteína; y al menos una agrupación de genes para la síntesis de polisacáridos antigénicos procedente de al menos una bacteria, en donde el polisacárido antigénico es un antígeno O bacteriano.

En todavía otro aspecto, la presente descripción se dirige a un sistema de expresión para producir una vacuna de bioconjugado contra *Shigella dysenteriae* O1 que comprende: una secuencia de nucleótidos que codifica PglB que tiene SEQ. ID NO. 2; una secuencia de nucleótidos que codifica una EPA modificada que tiene SEQ. ID NO. 6; y una agrupación de genes para la síntesis de polisacáridos que comprende SEQ. ID. En aún otro aspecto, la presente invención contempla un método para producir un bioconjugado O1 en un biorreactor que comprende las etapas: expresar en bacterias: EPA modificada que contiene al menos una secuencia de consenso, D/E - X - N - Z - S/T, en la que X y Z pueden ser cualquier aminoácido natural excepto prolina, o AcrA; PglB; y uno o varios de polisacáridos O1; cultivar las bacterias durante un período de tiempo para producir una cantidad del bioconjugado O1 que comprende AcrA o la EPA modificada ligada a uno o varios de los polisacáridos O1; extraer las proteínas periplásmicas; y separar el bioconjugado O1 de las proteínas periplásmicas extraídas.

En un aspecto adicional, la presente descripción contempla un método para producir una vacuna de bioconjugado contra *S. dysenteriae*, comprendiendo dicho método: ensamblar un polisacárido de *S. dysenteriae* en un organismo recombinante mediante el uso de glicosiltransferasas; ligar dicho polisacárido a un residuo de asparagina de una o varias proteínas diana en dicho organismo recombinante, en donde dicha una o varias de dichas proteínas diana contienen uno o varios epítopos de linfocitos T.

En un aspecto adicional, la presente descripción contempla un método para producir una vacuna de bioconjugado contra *S. dysenteriae*, comprendiendo dicho método: introducir una información genética que codifica una maquinaria metabólica que es portadora de una N-glicosilación de una proteína diana dentro de un organismo procarionta para producir un organismo procarionta modificado, en donde la información genética requerida para la expresión de una o varias proteínas diana recombinantes se introduce en dicho organismo procarionta, y en donde la maquinaria metabólica comprende glicosiltransferasas específicas para ensamblar un polisacárido de *S. dysenteriae* sobre un vehículo lipídico y una oligosacariltransferasa, ligando covalentemente la oligosacariltransferasa al polisacárido con un residuo de asparagina de la proteína diana y en donde la proteína diana contiene al menos un epítipo de linfocito T; producir un cultivo del organismo procarionta modificado; y obtener proteínas glicosiladas a partir del medio de cultivo.

Descripción de las Figuras

La **Fig. 1** ilustra la N-glicosilación de derivados de proteínas Lip procedentes de las estructuras artificiales A a C (véase el ejemplo 1). Células *E. coli* Top 10 son portadoras de un operón pgl funcional procedente de *C. jejuni* (Wacker et al., 2002, supra) y un plásmido que codifica las estructuras artificiales A (carril 2), B (carril 1) y C (carril 3) o un mutante de la estructura artificial C con la mutación D121A (carril 4). Las proteínas se expresaron y se purificaron a partir de extractos periplásmicos. Se muestra la SDS-PAGE y una tinción con azul brillante de Coomassie de las fracciones proteicas purificadas.

La **Fig. 2** muestra el análisis de la N-glicosilación de las diferentes proteínas que se analizaron en relación con la N-glicosilación específica de secuencia mediante el operón pgl de *C. jejuni* (Wacker et al., 2002, supra) en células CLM24 (Feldman et al., (2005). Engineering N-linked protein glycosylation with diverse O antigen lipopolysaccharide structures in *Escherichia coli*. Proc. Natl. Acad. Sci. USA 102, 3016-3021) o células Top 10 (carriles 1-6 panel E) o células SCM7 (Alaimo, C, Catrein, L, Morf, L., Marolda, C. L., Callewaert, N., Valvano, M. A., Feldman, M. F., Aebi, M. (2006). Two distinct but interchangeable mechanisms for flipping of lipid-linked oligosaccharides. EMBO Journal 25, 967-976) (panel E, carriles 7, 8) que expresan dichas proteínas a partir de un plásmido. Se muestran extractos periplásmicos separados por SDS-PAGE que se transfirieron a una membrana de nitrocelulosa y se visualizaron con antisueros específicos. En los paneles A-D, los paneles superiores muestran inmunotransferencias analizadas con antisuero anti AcrA (Wacker et al. 2002, supra; Nita-Lazar, M., Wacker, M., Schegg, B., Amber, S. y Aebi, M. (2005). The N-X- S/T consensus sequence is required but not

sufficient for bacterial N-linked protein glycosylation. *Glycobiology* 15, 361-367), mientras que los paneles inferiores muestran inmunotransferencias analizadas con antisuero R12 (Wacker et al., 2002, supra). + y - indican la presencia del operón *pgl* funcional o mutante en las células. El panel A contiene muestras de AcrA soluble de tipo silvestre con la secuencia señal *pelB* y el marcador hexa His (carriles 1, 2), AcrA-N273Q (carril 3, 4), y AcrA-D121A (carril 5). Panel B: AcrA (carriles 1,2), AcrA-T145D (carril 3), AcrA-N123Q-N273Q-T145D (carriles 4, 5). Panel C: AcrA-FI 15D-T145D (carriles 1, 2), AcrA-N123Q-N273Q-N272D (carriles 3, 4). Panel D: AcrA-N273Q (carriles 1, 2), AcrA-N273Q-F122P (carriles 3, 4). Panel E: CtxB (carriles 1, 2), CtxB-W88D (carriles 3, 4), CtxB-Q56/DSNIT (carriles 5, 6), y CtxB-W88D-Q56/DSNIT.

La **Fig. 3** muestra la modificación genética de múltiples sitios de glicosilación en OmpHI. La cepa Δ waalL SCM6 se cotransformó con el plásmido pACYCpgl (que codifica el locus *pgl* completo) y plásmidos que expresan OmpHI de tipo silvestre (carril 1), OmpHI^{N139S}-myc (carril 2), OmpHI^{KGN→NIT-HFGDD→DSNIT}-myc (carril 3), OmpHI^{RGD→NIT-HFGDD→DSNIT}-myc (carril 4), OmpHI^{KGN→NIT, RGD→NIT}-myc (carril 5), OmpHI^{KGN-NIT, RGD→N, T HFGDD→DSNIT}-myc (carril 6) o OmpHI^{RGD→NIT⁻V83T}-myc (carril 7). Las células se cultivaron en condiciones aerobias, se indujeron con 0,5% de arabinosa durante 3 horas antes del análisis. Lisados de células completas precipitaron con TCA después de igualar la densidad óptica de los cultivos como se describe en la sección de materiales y métodos. Las proteínas se separaron mediante SDS-PAGE al 15% y se transfirieron a una membrana de PVDF. El primer panel, inmunotransferencia de lisados de células completas, analizadas con anticuerpos anti-marcador *myc*. Panel inferior, inmunotransferencia de lisados de células completas analizadas con antisuero específico de glicano. Las posiciones de OmpHI sin glicosilar y glicosilada se indican a la derecha.

La **Fig. 4** muestra una microscopía de fluorescencia de células que expresan diferentes variantes de OmpHI. La microscopía de fluorescencia se realizó mediante el uso de un microscopio Axioplan2 (Carl Zeiss). Las imágenes se combinaron mediante el uso de Adobe Photoshop, versión CS2. Células SCM6 que expresaban OmpHI (panel A), OmpHI^{N139S} (panel B), OmpHI^{C20S} (panel C), OmpHI^{KGN→NIT, HFGDD→DSNIT} (panel D), OmpHI^{RGD→NIT-HFGDD-DSNIT} (panel E), OmpHI^{TMN→NIT RGD→NIT} (panel F), OmpHI^{(pand FX V83T KGN→NIT} (panel G), y OmpHI^{KGN→NIT⁻RGD-NIT⁻HFGDD→DSNIT} (panel H). La primera columna es una fusión de las imágenes en las columnas 2, 3 y 4 representada en tonos grises sobre fondo negro. Columna 2: fluorescencia azul en tonos grises procedente de la tinción con DAPI, columna 3: fluorescencia verde procedente de la fluorescencia específica de glicano, columna 4: fluorescencia roja procedente de la tinción anti-*myc*.

La **Fig. 5A** muestra un esquema de los polisacáridos y lipopolisacáridos capsulares en bacterias Gram-positivas y Gram-negativas.

La **Fig. 5B** muestra ADN genómico, con *PgIB* y EPA integrados, y ADN de plásmido, que es intercambiable (es decir, sustituible), que codifica una agrupación de genes para la síntesis de polisacáridos.

La **Fig. 6A** muestra el procedimiento de producción de vacunas conjugadas que utilizan la tecnología de la invención.

La **Fig. 6B** muestra la construcción del plásmido pGVXN64 con expresión del antígeno OI de *Shigella dysenteriae*.

Las **Figs. 7A y 7B** muestran diagramas esquemáticos de la ruta de glicosilación de proteínas utilizada en la presente invención.

Las **Figs. 8A y 8B** son diagramas esquemáticos que representan plataformas de expresión para la producción del bioconjugado de la presente invención.

La **Fig. 9** muestra la producción de bioconjugados de *Shigella*.

La **Fig. 10A** muestra la biosíntesis de polisacáridos del antígeno O de *S. dysenteriae* con serotipo OI sobre undecaprenolpírofosfato (UPP).

La **Fig. 10B** muestra un diagrama esquemático de una proteína de soporte, tal como EPA, sobre la cual se pueden diseñar los sitios de N-glicosilación.

La **Fig. 11** muestra bioconjugados que provocan una respuesta inmune contra el polisacárido OI de *Shigella* en ratones.

La **Fig. 12** muestra los resultados de la producción de un bioconjugado de *Shigella* OI - EPA (por ejemplo, 01-EPA) en un biorreactor.

La **Fig. 13** muestra la purificación de OI - EPA.

La **Fig. 14A** muestra un análisis de transferencia Western de una serie de especímenes de bioconjugados de *Shigella* OI - AcrA, producidos en un matraz de agitación con LB, bajo diversas condiciones.

La **Fig. 14B** proporciona diferentes serotipos de *Shigella* y la estructura del polisacárido que define su antigeni-

cidad (es decir, antígenos O de Shigella).

La **Fig. 15** muestra la expansión de la región anomérica del espectro ¹H de RMN de un ejemplo de bioconjugado de serotipo O1 de *S. dysenteriae* de la invención.

5 La **Fig. 16A** muestra muestras de proteínas de un bioconjugado de Shigella O1 (por ejemplo, EPA-O1) normalizado a una concentración de biomasa (0,1 DO₆₀₀ nm de células/carril).

La **Fig. 16B** muestra el extracto periplásmico procedente de un procedimiento de producción de bioconjugado de Shigella O1 que se cargó en un SDS-PAGE al 7,5%, y se tiñó con Coomassie para identificar EPA y EPA-O1.

La **Fig. 17A** muestra fracciones de proteínas procedentes de 1. la columna Source Q, analizadas por SDS-PAGE y teñidas con Coomassie para identificar el bioconjugado de Shigella O1.

10 La **Fig. 17B** muestra fracciones de proteínas procedentes de 2. la columna Source Q, analizadas por SDS-PAGE y teñidas con Coomassie para identificar el bioconjugado de Shigella O1.

La **Fig. 18A** muestra fracciones de proteínas procedentes de una columna Superdex 200, analizadas por SDS-PAGE y teñidas con Coomassie para identificar el bioconjugado de Shigella O1.

15 La **Fig. 18B** muestra bioconjugados de Shigella procedentes de diferentes etapas de la purificación analizados por SDS-PAGE y teñidos con Coomassie.

Descripción detallada de la invención

Introducción a la invención

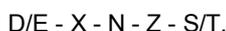
La presente descripción proporciona una plataforma versátil de glicosilación *in vivo*.

20 El documento de Solicitud de Patente Europea nº 03 702 276.1 (patente europea 1 481 057) describe un organismo procarionta en el que se introduce un ácido nucleico que codifica (i) glicosiltransferasas específicas para el ensamblaje de un oligosacárido sobre un vehículo lipídico, (ii) una proteína diana recombinante que comprende una secuencia de consenso "N - X - S/T", en donde X puede ser cualquier aminoácido excepto prolina, e (iii) una oligosacaril transferasa de *C. jejuni* (OTasa) que enlaza covalentemente dicho oligosacárido con la secuencia de consenso de la proteína diana. Dicho organismo procarionta produce N-glicanos con una estructura específica que se define por el tipo de glicosil transferasa específica.

La secuencia de consenso conocida para la N-glicosilación en una proteína permite la N-glicosilación de proteínas diana recombinantes en organismos procariontas que comprenden la oligosacaril transferasa (OTasa) de *C. jejuni*.

30 El objeto de la presente descripción es proporcionar proteínas, así como medios y métodos para producir tales proteínas que tienen una eficacia optimizada para la N-glicosilación, que se pueden producir en organismos procariontas *in vivo*. Otro objeto de la presente descripción se dirige a la introducción más eficaz de N-glicanos en proteínas recombinantes para modificar la antigenicidad, estabilidad, actividad biológica, profiláctica y/o terapéutica de dichas proteínas. Un objeto adicional es proporcionar una célula hospedadora que presenta de manera eficaz proteínas N-glicosiladas recombinantes de la presente descripción en su superficie.

35 En un primer aspecto, la presente descripción proporciona una proteína N-glicosilada recombinante, que comprende una o varias de las siguientes secuencias de aminoácidos optimizadas N-glicosiladas:



(secuencia de consenso optimizada)

en donde X y Z pueden ser cualquier aminoácido natural excepto Pro, y en donde se introduce al menos una de dichas secuencias de aminoácidos parciales N-glicosiladas.

40 Se encontró sorprendentemente que la introducción de la o las secuencias de aminoácidos parciales específicas (secuencia o secuencias de consenso optimizadas) en las proteínas, conduce a proteínas que están N-glicosiladas de forma eficaz a través de la oligosacaril transferasa (OST, OTasa) procedente de *Campylobacter* spp., preferiblemente *C. jejuni*, en esas posiciones introducidas.

45 La expresión "secuencia o secuencias de aminoácidos parciales" tal como se utiliza en el contexto de la presente descripción, también se denomina "secuencia o secuencias de consenso optimizadas" o " secuencia o secuencias de consenso". La secuencia de consenso optimizada está N-glicosilada a través de la oligosacaril transferasa (OST, OTasa) procedente de *Campylobacter* spp, preferiblemente *C. jejuni*, de forma mucho más eficaz que la secuencia de consenso regular, "N - X - S/T", conocida en la técnica anterior.

50 En general, la expresión "proteína N-glicosilada recombinante" se refiere a cualquier polipéptido u oligopéptido heterólogo producido en una célula hospedadora que no comprende de forma natural el ácido nucleico que codifica dicha

proteína. En el contexto de la presente invención, esta expresión se refiere a una proteína producida de forma recombinante en cualquier célula hospedadora, por ejemplo, una célula hospedadora eucariota o procariota, preferiblemente una célula hospedadora procariota, por ejemplo *Escherichia* ssp., *Campylobacter* ssp., *Salmonella* ssp., *Shigella* ssp., *Helicobacter* ssp., *Pseudomonas* ssp., *Bacillus* ssp., más preferiblemente *Escherichia coli*, *Campylobacter jejuni*, *Salmonella typhimurium* etc., en donde el ácido nucleico que codifica dicha proteína ha sido introducido en dicha célula hospedadora y en donde la proteína codificada está N-glicosilada a través de la OTasa procedente de *Campylobacter* spp., preferiblemente *C. jejuni*, estando presente de forma natural dicha enzima transferasa o habiendo sido introducida de forma recombinante en dicha célula hospedadora.

De acuerdo con el código de una letra aceptado internacionalmente para los aminoácidos, las abreviaturas D, E, N, S y T indican ácido aspártico, ácido glutámico, asparagina, serina y treonina, respectivamente. Las proteínas de acuerdo con la invención difieren de las proteínas naturales o de la técnica anterior en que una o varias de las secuencias de consenso optimizadas D/E - X - N - Z - S/T se introducen y se glicosilan en N. Por lo tanto, las proteínas de la presente descripción se diferencian de las proteínas de origen natural de *C. jejuni* en que también contienen la secuencia de consenso optimizada pero no comprenden ninguna secuencia de consenso optimizada adicional (introducida).

La introducción de la secuencia de consenso optimizada se puede lograr mediante la adición, delección y/o sustitución de uno o varios aminoácidos. La adición, delección y/o sustitución de uno o varios aminoácidos con el fin de introducir la secuencia de consenso optimizada, se puede llevar a cabo mediante estrategias de síntesis química bien conocidas por los expertos en la técnica, tales como síntesis química de péptidos asistida en fase sólida. Alternativamente, y de forma preferida para polipéptidos más grandes, las proteínas de la presente descripción se pueden preparar mediante técnicas recombinantes convencionales.

Las proteínas de la presente descripción tienen la ventaja de que se pueden producir con mucha eficiencia y en cualquier hospedador procariota que comprenda un operón pgl funcional procedente de *Campylobacter* spp., preferiblemente *C. jejuni*. Las OTasas alternativas preferidas procedentes de *Campylobacter* spp. para poner en práctica los aspectos y realizaciones de la presente descripción, son *Campylobacter coli* y *Campylobacter lari* (véase, Szymanski, CM. y Wren, B. W. (2005). Protein glycosylation in bacterial mucosal pathogens. *Nat. Rev. Microbiol.* 3:225-237). El operón pgl funcional puede estar presente de forma natural cuando dicho hospedador procariota es *Campylobacter* spp., preferiblemente *C. jejuni*. Sin embargo, como se ha mostrado anteriormente en la técnica y se ha mencionado más arriba, el operón pgl se puede transferir a las células y permanece funcional en dicho nuevo entorno celular.

La expresión "operón pgl funcional procedente de *Campylobacter* spp., preferiblemente *C. jejuni*" se entiende que se refiere a la agrupación de ácidos nucleicos que codifican la oligosacaril transferasa funcional (OTasa) de *Campylobacter* spp., preferiblemente *C. jejuni*, y una o varias glicosiltransferasas específicas, capaces de ensamblar un oligosacárido sobre un vehículo lipídico, y en donde dicho oligosacárido se puede transferir desde el vehículo lipídico a la proteína diana que tiene una o varias secuencias de aminoácidos optimizadas: D/E - X - N - Z - S/T a través de la OTasa. Se debe entender que la expresión "operón pgl funcional procedente de *Campylobacter* spp., preferiblemente *C. jejuni*" en el contexto de esta descripción no se refiere necesariamente a un operón como una unidad transcripcional singular. La expresión solo requiere la presencia de los componentes funcionales para la N-glicosilación de la proteína recombinante en una célula hospedadora. Estos componentes se pueden transcribir como uno o varios ARNm distintos y se pueden regular juntos o por separado. Por ejemplo, la expresión también incluye componentes funcionales posicionados en el ADN genómico y en plásmido(s) en una célula hospedadora. Para fines de eficacia, se prefiere que todos los componentes del operón pgl funcional se regulen y se expresen simultáneamente.

Es importante tener en cuenta que solo una oligosacaril transferasa funcional (OTasa) debe proceder de *Campylobacter* spp., preferiblemente *C. jejuni*, y que una o varias glicosiltransferasas específicas capaces de ensamblar un oligosacárido sobre un vehículo lipídico, pueden proceder de la célula hospedadora o se pueden introducir de forma recombinante en dicha célula hospedadora, en donde la única limitación funcional es que el oligosacárido ensamblado a través de dichas glicosiltransferasas se pueda transferir desde el vehículo lipídico a la proteína diana, que tiene una o varias secuencias de consenso optimizadas por la OTasa. Por lo tanto, la selección de la célula hospedadora que comprende glicosiltransferasas específicas de forma natural o glicosiltransferasas específicas que incapacitan presentes de forma natural en dicho hospedador, así como la introducción de glicosiltransferasas específicas heterólogas, permitirá a los expertos en la técnica variar los N-glicanos unidos al sitio de consenso de N-glicosilación optimizado en las proteínas de la presente descripción.

Como resultado de lo anterior, la presente descripción proporciona el diseño individual de patrones de N-glicanos sobre las proteínas de la presente descripción. Por tanto, las proteínas se pueden individualizar en su patrón de N-glicanos para satisfacer las necesidades biológicas, farmacéuticas y de purificación.

En una realización preferida, las proteínas de la presente descripción pueden comprender una, pero también más de una, preferiblemente al menos dos, preferiblemente al menos 3, más preferiblemente al menos 5 de dichas secuencias de aminoácidos optimizadas N-glicosiladas.

La presencia de una o varias secuencias de aminoácidos optimizadas N-glicosiladas en las proteínas de la presente

descripción puede ser una ventaja para aumentar su antigenicidad, incrementando su estabilidad, que afecta a su actividad biológica, prolongando su semivida biológica y/o simplificando su purificación.

5 La secuencia de consenso optimizada puede incluir cualquier aminoácido excepto prolina en la o las posiciones X y Z. La expresión "cualquier aminoácido" se entiende que incluye aminoácidos naturales comunes y raros, así como derivados de aminoácidos sintéticos y análogos que todavía permitirán que la secuencia de consenso optimizada sea N-glicosilada a través de la OTasa. Los aminoácidos de origen natural, comunes y raros son los preferidos para X y Z. X y Z pueden ser iguales o diferentes.

Se observa que X y Z pueden ser diferentes para cada secuencia de consenso optimizada en una proteína de acuerdo con la presente descripción.

10 El N-glicano unido a la secuencia de consenso optimizada se determinará por las glicosiltransferasas específicas y su interacción cuando se ensambla el oligosacárido sobre un vehículo lipídico para la transferencia a través de la OTasa. Los expertos en la técnica pueden diseñar el N-glicano variando el o los tipos y la cantidad de glicosiltransferasas específicas presentes en la célula hospedadora deseada.

15 Los N-glicanos se definen en esta memoria como monosacáridos, oligosacáridos o polisacáridos de composiciones variables que están unidos a un nitrógeno de ϵ -amida de un residuo de asparagina en una proteína, mediante un enlace N-glicosídico. Preferiblemente, los N-glicanos transferidos por la OTasa se ensamblan sobre un anclaje lipídico de undecaprenol-pirofosfato que está presente en la membrana citoplasmática de las bacterias Gram-negativas o Gram-positivas. Están implicados en la síntesis del antígeno O, polisacárido O y peptidoglicano (Bugg, T. D., y Brandish, P. E. (1994). From peptidoglycan to glycoproteins: common features of lipid-linked oligosaccharide biosynthesis. FEMS Microbiol Lett 119, 255-262; Valvano, M. A. (2003). Export of O-specific lipopolysaccharide. Front Biosci 8, S452-471).

25 En una realización preferida, la proteína recombinante de la presente descripción comprende uno o varios N-glicanos seleccionados a partir del grupo de N-glicanos procedentes de *Campylobacter* spp., preferiblemente *C. jejuni*, los N-glicanos obtenidos a partir de oligosacáridos y polisacáridos transferidos a un polisacárido O que forma un antígeno O en bacterias Gram-negativas o polisacáridos capsulares procedentes de bacterias Gram-positivas, preferentemente: *P. aeruginosa* 09, 011; *E. coli* 07, 09, 016, 0157 y *Shigella dysenteriae* O1, y variantes modificadas genéticamente de los mismos, obtenidos mediante la inserción o la eliminación de glicosiltransferasas y epimerasas que afectan a la estructura del polisacárido.

30 En una realización preferida adicional, la proteína recombinante de la presente descripción comprende dos o más N-glicanos diferentes.

35 Por ejemplo, unos N-glicanos diferentes en la misma proteína se pueden preparar mediante el control del momento de la expresión de glicosiltransferasas específicas usando promotores tempranos o tardíos, o la introducción de factores para el arranque, silenciamiento, mejora y/o reducción de la actividad promotora de glicosiltransferasas específicas de un individuo. Los promotores adecuados y los factores que regulan su actividad están disponibles rutinariamente para los expertos en la técnica.

40 No existe ninguna limitación sobre el origen de la proteína recombinante de la descripción. Preferiblemente, dicha proteína se obtiene a partir de proteínas de mamífero, bacterianas, víricas, fúngicas o vegetales. Más preferiblemente, la proteína se obtiene a partir de proteínas de mamífero, lo más preferiblemente de seres humanos. Para la preparación de proteínas recombinantes antigénicas de acuerdo con la invención, preferiblemente para uso como componentes activos en vacunas, se prefiere que la proteína recombinante se obtenga a partir de una proteína bacteriana, vírica o fúngica.

45 En una realización preferida adicional, la presente descripción proporciona proteínas recombinantes, en donde o bien la proteína y/o el o los N-glicanos son activos terapéutica y/o profilácticamente. La introducción de al menos una secuencia de consenso optimizada y N-glicosilada puede modificar o incluso introducir una actividad terapéutica y/o profiláctica en una proteína. En una realización más preferida, la proteína y/o el o los N-glicanos son inmunogénicamente activos. En ese caso, la o las N-glicosilaciones introducidas pueden tener un efecto modificador sobre la actividad biológica de las proteínas y/o introducir nuevos sitios antigénicos y/o puede enmascarar a la proteína para evitar etapas degradantes y/o aumentar la semivida.

50 Las proteínas recombinantes de la presente descripción se pueden dirigir de manera eficaz a la membrana externa y/o a la superficie de células hospedadoras, preferiblemente bacterias, más preferiblemente bacterias Gram-negativas. Para facilitar la presentación en la superficie y/o la localización en la membrana externa, se prefiere que la proteína recombinante de la descripción comprenda además al menos una secuencia de polipéptido capaz de dirigir dicha proteína recombinante hacia la membrana externa y/o la superficie celular de una bacteria, preferiblemente una bacteria Gram-negativa.

55 En una realización preferida, la proteína recombinante de la descripción es una, en donde dicha secuencia de polipéptido dirigida se selecciona a partir del grupo que consiste en péptidos señal de tipo II (Paetzl, M., Karla, A., Strynadka, N.C., y Dalbey, R. E. 2002. Signal peptidases. Chem Rev 102: 4549-4580) o proteínas de la membrana

externa (revisado en Wemerus, H. y Stahl, S. 2004. Biotechnological applications for surface-engineered bacteria. *Biotechnol Appl Biochem* 40: 209-228), se selecciona preferiblemente a partir del grupo que consiste en la proteína de longitud completa o los péptidos señal de OmpH1 procedente de *C. jejuni*, Jipa procedente de *C. jejuni*, proteínas de la membrana externa procedente de *E. coli*, preferiblemente OmpS, OmpC, OmpA, OprF, PhoE, LamB, Lpp'OmpA (una proteína de fusión para la tecnología de presentación en superficie, véase Francisco, JAi Earhart, C. F., y Georgiou, G. 1992. Transport and anchoring of beta- lactamase to the external surface of *Escherichia coli*. *Proc Natl Acad Sci U S A* 89: 2713- 2717), y la proteína Inp procedente de *Pseudomonas aeruginosa*.

En un aspecto diferente, la presente descripción se refiere a un ácido nucleico que codifica una proteína recombinante de acuerdo con la descripción. Preferiblemente, dicho ácido nucleico es un ARNm, un ADN o un PNA, más preferiblemente un ARNm o un ADN, lo más preferiblemente un ADN. El ácido nucleico puede comprender la secuencia que codifica dicha proteína y, además, otras secuencias, tales como secuencias reguladoras, por ejemplo, promotores, potenciadores, codones de parada, codones de inicio y los genes necesarios para regular la expresión de la proteína recombinante a través de las secuencias reguladoras mencionadas, etc. La expresión "ácido nucleico que codifica una proteína recombinante de acuerdo con la descripción" se dirige a un ácido nucleico que comprende dicha secuencia codificante y, opcionalmente, cualquier secuencia de ácido nucleico adicional, independientemente de la información de la secuencia, siempre y cuando el ácido nucleico sea capaz de producir la proteína recombinante de la descripción en una célula hospedadora que contiene un operón #g7 funcional procedente de *Campylobacter* spp., preferiblemente *C. jejuni*. Más preferiblemente, la presente descripción proporciona ácidos nucleicos aislados y purificados ligados funcionalmente a un promotor, preferiblemente ligados a un promotor seleccionado a partir del grupo que consiste en promotores procariontes inducibles y constitutivos conocidos, más preferiblemente el promotor de tetraciclina, el promotor de arabinosa, el promotor de salicilato, promotores lac, trc, tac (Baneyx, F. (1999). Recombinant protein expression in *Escherichia coli*. *Curr Opin Biotechnol* 10, 411-421; Billman-Jacobe, H. (1996). Expression in bacteria other than *Escherichia coli*. *Curr Opin Biotechnol* 7, 500-504). Dichos ácidos nucleicos ligados funcionalmente se pueden usar para, por ejemplo, una vacunación.

Además, otro aspecto de la presente descripción se refiere a una célula hospedadora que comprende un ácido nucleico y/o un vector de acuerdo con la presente descripción. El tipo de célula hospedadora no es limitativo, siempre y cuando se acomode un operón pgl funcional procedente de *C. jejuni* y uno o varios ácidos nucleicos que codifican la o las proteínas diana recombinantes de la presente descripción. Las células hospedadoras preferidas son células hospedadoras procariontes, más preferiblemente bacterias, más preferiblemente aquellas seleccionadas a partir del grupo que consiste en *Escherichia* spp., *Campylobacter* spp., *Salmonella* spp., *Shigella* spp., *Helicobacter* spp., *Pseudomonas* spp., *Bacillus* spp., preferiblemente *Escherichia coli*, más preferiblemente las cepas de *E. coli* Top 10, W3110, CLM24, BL21, SCM6 y SCM7 (Feldman et al., (2005). Engineering N-linked protein glycosylation with diverse O antigen lipopolysaccharide structures in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* 102, 3016-3021; Alaimo, C, Catrein, I., Morf, L., Marolda, C. L., Callewaert, N., Valvano, M. A., Feldman, M. F., Aebi, M. (2006). Two distinct but interchangeable mechanisms for flipping of lipid-linked oligosaccharides. *EMBO Journal* 25, 967-976) y las cepas de *S. enterica* SL3261 {*Salmonella enterica* sv. Typhimurium LT2 (delta) aroA, véase Hoiseth, S. K., y Stocker, B.A. 1981, Aromatic-dependent *Salmonella typhimurium* are non-virulent and effective as live vaccines. *Nature* 291:238-239), SL3749 {*Salmonella enterica* sv. Typhimurium LT2 waaL, véase Kaniuk et al., *J. Biol. Chem.* 279: 36470-36480) y SL3261 ΔwaaL.

En una realización más preferida, la célula hospedadora de acuerdo con la descripción es una que es útil para dirigir hacia la membrana exterior y/o presentar en la superficie proteínas recombinantes de acuerdo con la invención, preferiblemente una, en donde dicha célula hospedadora es una bacteria Gram-negativa recombinante que tiene:

i) un genotipo que comprende secuencias de nucleótidos que codifican

a) al menos una glicosiltransferasa específica natural o recombinante para ensamblar un oligosacárido sobre un vehículo lipídico,

b) al menos una oligosacaril transferasa (OTasa) procarionte natural o recombinante procedente de *Campylobacter* spp., preferiblemente *C. jejuni*,

c) al menos una proteína recombinante de acuerdo con la descripción, preferiblemente una proteína que comprende además un polipéptido que dirige hacia una diana, e ii) un fenotipo que comprende una proteína N-glicosilada recombinante de acuerdo con la invención que se encuentra en y/o sobre la membrana externa de la bacteria Gram-negativa.

La célula hospedadora para la realización anterior se selecciona preferiblemente a partir del grupo que consiste en *Escherichia* spp., *Campylobacter* spp., *Shigella* spp., *Helicobacter* spp. y *Pseudomonas* spp., *Salmonella* spp., preferentemente *E. coli*, más preferiblemente cepas de *E. coli* Top 10, W3110, CLM24, BL21, SCM6 y SCM7, y cepas de *S. enterica* SL3261, SL3749 y SL3261ΔwaaL (véase Hoiseth, S. K., y Stocker, B. A. 1981. Aromatic- dependent *Salmonella typhimurium* are non-virulent and effective as live vaccines. *Nature* 291: 238-239), SL3749 (Kaniuk, N.A., Vinogradov, E., y Whitfield, C. 2004. Investigation of the structural requirements in the lipopolysaccharide core acceptor for ligation of O antigens in the genus *Salmonella*: WaaL "ligase" is not the sole determinant of acceptor specificity. *J Biol Chem* 279: 36470-36480).

Dado que las proteínas preferidas de la presente descripción pueden tener una actividad terapéutica o profiláctica por sí mismas y/o debido a los sitios de N-glicosilación introducidos, se pueden utilizar para la preparación de un medicamento. El tipo de proteína para la puesta en práctica de la invención no está limitada y, por lo tanto, las proteínas de la invención, tales como EPO, INF-alfa, TNF-alfa, IgG, IgM, IgA, interleucinas, citocinas, proteínas víricas y bacterianas para la vacunación, como proteínas de *C. jejuni* tales como HisJ (CjO734c), AcrA (CjO367c), OmpH1 (CjO982c), toxina diftérica (CRM 197), toxina del cólera, exoproteína de *P. aeruginosa*, por nombrar solo unas pocas, y a las que se ha introducido la secuencia de consenso N-glicosilada optimizada, son útiles para preparar un medicamento (Wyszynska, A., Raczko, A., Lis, M., y Jagusztyń-Krynicka, E. K. (2004). Oral immunization of chickens with avirulent *Salmonella* vaccine strain carrying *C. jejuni* 72Dz/92 *cjaA* gene elicits specific humoral immune response associated with protection against challenge with wild-type *Campylobacter*. *Vaccine* 22, 1379-1389).

Además, los ácidos nucleicos y/o los vectores de acuerdo con la descripción también son útiles para la preparación de un medicamento, preferiblemente para uso en terapia génica.

Por otra parte, una célula hospedadora de acuerdo con la descripción, preferiblemente una que tiene un fenotipo que comprende una proteína recombinante N-glicosilada de la descripción, que se encuentra en y/o sobre la membrana externa de una bacteria, preferiblemente una bacteria Gram-negativa, más preferiblemente una de las bacterias Gram-negativas mencionadas anteriormente, es particularmente útil para la preparación de un medicamento.

Más preferiblemente, una proteína de la descripción se utiliza para la preparación de un medicamento para la vacunación terapéutica y/o profiláctica de un sujeto que lo requiere.

En una realización más preferida, la presente descripción se refiere al uso de un ácido nucleico y/o un vector de acuerdo con la descripción para la preparación de un medicamento para la vacunación terapéutica y/o profiláctica de un sujeto que lo requiere, preferiblemente mediante terapia génica.

Las células hospedadoras de la descripción que presentan dichas proteínas recombinantes N-glicosiladas, son particularmente útiles para la preparación de vacunas, debido a que las proteínas N-glicosiladas presentadas están presentes de forma abundante en la superficie de la célula hospedadora y son de acceso fácil para las células inmunes, en particular sus N-glicanos hidrófilos, y debido a que las células hospedadoras tienen el efecto añadido de que un adyuvante, si está vivo, incluso se puede replicar en cierta medida y amplificar sus efectos de la vacunación.

Preferiblemente, la célula hospedadora para la puesta en práctica de los aspectos médicos de esta descripción, es una célula hospedadora atenuada o muerta.

Otra ventaja de la utilización de las células hospedadoras de acuerdo con la invención para la preparación de medicamentos, preferiblemente vacunas, es que inducen anticuerpos IgA debido al componente celular.

Preferiblemente, dichas células hospedadoras se usan de acuerdo con la descripción para la inducción de anticuerpos IgA en un animal, preferiblemente un mamífero, roedor, ovino, equino, canino, bovino o un ser humano. Se prefiere que dicho sujeto que tiene necesidad de una vacunación sea aviar, mamífero o un pez, preferiblemente un mamífero, más preferiblemente un mamífero seleccionado a partir del grupo que consiste en ganado vacuno, bovino, equino, perros, gatos y seres humanos, lo más preferiblemente seres humanos. También se prefieren las aves.

Un aspecto adicional de la presente descripción se refiere a una composición farmacéutica, que comprende al menos una proteína, al menos un ácido nucleico, al menos un vector y/o al menos una célula hospedadora de acuerdo con la descripción. La preparación de medicamentos que comprenden proteínas o células hospedadoras, preferiblemente células hospedadoras atenuadas o muertas, y la preparación de medicamentos que comprenden ácidos nucleicos y/o vectores para la terapia génica, es bien conocida en la técnica. El esquema de preparación de la composición farmacéutica final y el modo y los detalles de su administración, dependerán de la proteína, la célula hospedadora, el ácido nucleico y/o el vector utilizado.

En una realización preferida, la composición farmacéutica de la descripción comprende un excipiente, diluyente y/o adyuvante farmacéuticamente aceptable.

La presente descripción proporciona una composición farmacéutica que comprende al menos uno de los siguientes, (i) una proteína recombinante, una célula hospedadora, un ácido nucleico y/o un vector recombinante que es/está codificado por/expresa una proteína recombinante de acuerdo con la presente descripción, y (ii) un excipiente, diluyente y/o adyuvante farmacéuticamente aceptable.

Los excipientes, diluyentes y/o adyuvantes adecuados son bien conocidos en la técnica. Un excipiente o un diluyente puede ser un material sólido, semisólido o líquido que puede servir como vehículo o medio para el ingrediente activo. Un experto ordinario en la técnica del campo de preparación de composiciones, puede seleccionar fácilmente la forma y el modo de administración adecuados, dependiendo de las características particulares del producto seleccionado, la enfermedad o afección que se va a tratar, el estadio de la enfermedad o afección, y otras circunstancias relevantes (Remington's Pharmaceutical Sciences, Mack Publishing Co. (1990)). La proporción y la naturaleza del diluyente o excipiente farmacéuticamente aceptable están determinadas por las propiedades de solubilidad y químicas del compuesto farmacéuticamente activo seleccionado, la ruta de administración elegida, y la práctica farmacéu-

5 tica convencional. La preparación farmacéutica se puede adaptar para uso oral, parenteral o tópico y se puede administrar al paciente en forma de comprimidos, cápsulas, supositorios, solución, suspensiones o similares. Los compuestos farmacéuticamente activos de la presente invención, aunque son eficaces por sí mismos, se pueden formular y administrar en forma de sus sales farmacéuticamente aceptables, tales como sales de adición de ácido o sales de adición de base, para fines de estabilidad, conveniencia de la cristalización, solubilidad incrementada y similares.

Un aspecto adicional de la presente descripción se dirige a un método para producir proteínas glicosiladas ligadas a N, que comprende las etapas de:

a) proporcionar un organismo recombinante, preferiblemente un organismo procariota que comprende ácidos nucleicos que codifican

10 i) un operón pgl funcional procedente de *Campylobacter* spp, preferiblemente *C. jejuni*, e ii) al menos una proteína diana recombinante que comprende una o varias de la o las siguientes secuencias de consenso de aminoácidos optimizadas N-glicosiladas:

D/E - X - N - Z - S/T,

15 en donde X y Z pueden ser cualquier aminoácido natural excepto Pro, y en donde se introduce al menos una de dicha o dichas secuencias de consenso de aminoácidos optimizadas N-glicosiladas, y

b) cultivar el organismo recombinante de una forma adecuada para la producción y N-glicosilación de la o las proteínas diana.

Preferiblemente, la proteína diana es una de las proteínas recombinantes descritas anteriormente de acuerdo con la descripción.

20 En un método preferido de la invención, el operón pgl funcional procedente de *Campylobacter* spp., preferiblemente *C. jejuni*, comprende ácidos nucleicos que codifican

i) OTasa recombinante procedente de *Campylobacter* spp., preferiblemente *C. jejuni*, y

ii) glicosiltransferasas específicas recombinantes y/o naturales procedentes de *Campylobacter* spp., preferiblemente *C. jejuni*, y/o

25 iii) glicosiltransferasas específicas recombinantes y/o naturales procedentes de especies distintas de *Campylobacter* spp.,

para ensamblar un oligosacárido sobre un vehículo lipídico que se va a transferir a la proteína diana mediante la OTasa.

30 Además, en una realización preferida, la presente descripción se refiere a un método para preparar una célula hospedadora de acuerdo con la descripción que comprende las etapas de:

i) proporcionar una bacteria Gram-negativa,

ii) introducir en dicha bacteria al menos una secuencia de nucleótidos que codifica

35 a) al menos una glicosiltransferasa específica recombinante para ensamblar un oligosacárido sobre un vehículo lipídico, y/o b) al menos una oligosacaril transferasa recombinante (OTasa) procedente de *Campylobacter* spp., preferiblemente *C. jejuni*, y/o

c) al menos una proteína recombinante que comprende una o varias de las siguientes secuencias de consenso de aminoácidos optimizadas N-glicosiladas:

D/E - X - N - Z - S/T,

40 en donde X y Z pueden ser cualquier aminoácido natural excepto Pro, y en donde se introduce al menos una de dichas secuencias de consenso de aminoácidos optimizadas N-glicosiladas, y

iii) cultivar dicha bacteria hasta que al menos una proteína N-glicosilada recombinante codificada por la secuencia de nucleótidos de c) esté situada en y/o sobre la membrana externa de la bacteria Gram-negativa.

45 Para la puesta en práctica de los métodos preferidos anteriores, el organismo procariota recombinante o la célula hospedadora se selecciona preferiblemente a partir del grupo de bacterias que consiste en *Escherichia* spp., *Campylobacter* spp., *Salmonella* spp., *Shigella* spp., *Helicobacter* spp., *Pseudomonas* spp., *Bacillus* spp., preferiblemente *Escherichia coli*, preferiblemente las cepas de *E. coli* Top 10, W3110, W31 10Δ waal, BL21, SCM6 y SCM7, y las cepas de *S. enterica* SL3261, SL3749 y SL3261 Δwaal.

Otro método preferido para producir, aislar y/o purificar una proteína recombinante de acuerdo con la descripción

comprende las etapas de:

- a) cultivar una célula hospedadora,
- b) eliminar la membrana externa de dicha bacteria Gram-negativa recombinante; y
- c) recuperar dicha proteína recombinante.

- 5 Los métodos a modo de ejemplo para la eliminación de la membrana externa de una célula, preferiblemente una célula procariota, más preferiblemente una célula bacteriana Gram-negativa, son métodos de tratamiento enzimático adecuado, solubilización con detergente y choque osmótico y el método de prensa francesa.

De forma más preferida, la presente descripción se refiere a un método, en el que las glicosiltransferasas específicas recombinantes o naturales procedentes de especies distintas a *Campylobacter* spp., preferiblemente *C. jejuni*, se seleccionan a partir del grupo de glicosiltransferasas y epimerasas procedentes de bacterias, archea y/o eucariotas que se pueden expresar funcionalmente en dicha célula hospedadora.

Vacunas de bioconjugados

15 Una realización de la descripción implica nuevas vacunas de bioconjugados. Una realización adicional de la invención implica un enfoque novedoso para la producción de tales vacunas de bioconjugados que emplea células bacterianas recombinantes que producen directamente bioconjugados inmunogénicos o antigénicos. En una realización, las vacunas de bioconjugados se pueden usar para tratar o prevenir enfermedades bacterianas, tales como diarrea, infecciones nosocomiales y meningitis. En realizaciones adicionales, las vacunas de bioconjugados pueden tener un potencial terapéutico y/o profiláctico contra el cáncer u otras enfermedades.

20 Las vacunas conjugadas se pueden administrar a los niños para protegerles contra infecciones bacterianas y pueden proporcionar una respuesta inmune de larga duración en los adultos. Se ha encontrado que las estructuras artificiales de la invención generan una respuesta de IgG en los animales. Se ha encontrado que una respuesta de IgG frente a una vacuna conjugada de polisacárido específico de O de *Shigella*-proteína en seres humanos se correlaciona con una protección inmune en los seres humanos. (Passwell, J.H. et al., "Safety and Immunogenicity of Improved *Shigella* O-Specific Polysaccharide-Protein Conjugate Vaccines in Adults in Israel" *Infection and Immunity*, 69(3):1351-1357 (Mar. 2001)). Se cree que el polisacárido (es decir, el residuo de azúcar) desencadena una respuesta inmunitaria a corto plazo que es específica del azúcar. En efecto, el sistema inmune humano genera una fuerte respuesta frente a estructuras específicas de la superficie de polisacáridos de bacterias, tales como los antígenos O y polisacáridos capsulares. Sin embargo, ya que la respuesta inmune frente a los polisacáridos es dependiente de IgM, el sistema inmune no genera memoria. El vehículo de proteína que transporta el polisacárido desencadena una respuesta de IgG que es dependiente de los linfocitos T y que proporciona una protección de larga duración, ya que el sistema inmune genera memoria.

Una dosificación típica de vacunación para seres humanos es de aproximadamente 1 a 25 µg, preferiblemente de aproximadamente 1 µg hasta aproximadamente 10 µg, lo más preferiblemente de aproximadamente 10 µg. Opcionalmente, una vacuna, tal como una vacuna de bioconjugado de la presente descripción, incluye un adyuvante.

35 Complejos de polisacáridos sintetizados (es decir, residuos de azúcar) y proteínas (es decir, vehículos de proteínas) se pueden utilizar como vacunas conjugadas para proteger contra una serie de infecciones bacterianas. En un aspecto, la presente invención se dirige a un nuevo enfoque de bioingeniería para la producción de vacunas conjugadas inmunogénicas que proporcionan ventajas sobre los métodos clásicos de conjugación química. En una realización, el enfoque implica la producción *in vivo* de glicoproteínas en células bacterianas, por ejemplo, células Gram-negativas tales como *E. coli*.

45 La biosíntesis de diferentes polisacáridos se conserva en las células bacterianas. Los polisacáridos se ensamblan sobre los lípidos de soporte a partir de precursores comunes (nucleótidos de azúcar activados) en la membrana citoplasmática a través de diferentes glicosiltransferasas con especificidad definida. Los lipopolisacáridos (LPS) se proporcionan solo en las bacterias Gram-negativas, por ejemplo, *Shigella* spp., *Pseudomonas* spp. y *E. coli* (ExPEC, EHEC).

La síntesis de lipopolisacáridos (LPS) se inicia con la adición de un monosacárido al fosfato de undecaprenilo del lípido de soporte en el lado citoplásmico de la membrana. El antígeno se construye mediante una adición secuencial de monosacáridos procedentes de nucleótidos de azúcar activados por diferentes glicosiltransferasas y el polisacárido ligado al lípido se da la vuelta a través de la membrana con una flipasa. La unidad repetitiva del antígeno se polimeriza mediante una reacción enzimática. El polisacárido se transfiere entonces al lípido A a través de la ligasa WaaL que forma el LPS que se exporta a la superficie, mientras que el polisacárido capsular se libera desde el lípido de soporte después de la polimerización y se exporta a la superficie. La ruta biosintética de estos polisacáridos permite la producción de bioconjugados de LPS *in vivo*, capturando los polisacáridos en el periplasma en un vehículo de proteína. Los bioconjugados, tales como los bioconjugados de LPS, se prefieren en la presente descripción.

55 Como se muestra en la Figura 5A, tanto las bacterias Gram-positivas como las Gram-negativas tienen una membra-

na celular que está rodeada por polisacáridos capsulares. Además, las bacterias Gram-negativas tienen una membrana externa sobre la membrana celular, con un espacio periplásmico que separa las dos membranas. Además, contienen lipopolisacáridos en la superficie. Cuando las bacterias Gram-negativas, tales como *E. coli*, se utilizan para producir una vacuna conjugada de la presente descripción, la glicoproteína empleada en la vacuna conjugada se ensambla en el espacio periplásmico.

Las vacunas conjugadas se han utilizado con éxito para proteger contra infecciones bacterianas. Se requiere la conjugación de un polisacárido antigénico con un vehículo de proteína para obtener una respuesta de memoria protectora, ya que los polisacáridos son antígenos independientes de los linfocitos T. Los polisacáridos se han conjugado con vehículos de proteínas mediante diferentes métodos químicos, utilizando grupos reactivos de activación en el polisacárido, así como el vehículo de proteína.

La Figura 6A muestra el procedimiento de producción de vacunas conjugadas que utiliza una tecnología de la invención, en comparación con el procedimiento utilizado actualmente. Actualmente, las vacunas conjugadas se producen usando dos rondas de fermentación y después de varias etapas de purificación y escisión química, como se muestra esquemáticamente en el panel superior. Este enfoque actual tiene una serie de problemas. En primer lugar, se requiere un cultivo a gran escala de organismos patógenos. En segundo lugar, el enfoque de la conjugación depende del polisacárido. En tercer lugar, el enfoque tiene una baja reproducibilidad. En cuarto lugar, el enfoque tiene una homogeneidad baja debido a una conjugación inespecífica. En quinto lugar, el enfoque también tiene una pureza baja debido a un exceso de polisacárido en la conjugación. Por último, el enfoque actual proporciona rendimientos menores del 20%.

Como se muestra en el panel inferior de la Figura 6A, en una realización, la tecnología innovadora de la invención se puede utilizar para desarrollar vacunas conjugadas (por ejemplo, vacunas de bioconjugados) completamente *in vivo* con células no patógenas, evitando las reacciones químicas y proporcionando una pureza elevada después de pocas etapas de purificación. Este nuevo método también permite la producción de vacunas de bioconjugados que no es factible con los métodos actuales. Además, el procedimiento de conjugación y purificación es independiente del antígeno de polisacárido que se utiliza. Como resultado, las vacunas de bioconjugados se pueden diseñar más rápido usando estructuras de glicano novedosas. El aumento de la homogeneidad de los conjugados resultantes y la reproducibilidad mejorada (es decir, sin variabilidad de lote a lote) de tales conjugados, hacen que sea un procedimiento muy atractivo por el control de calidad y las perspectivas de regulación. Además, el nuevo método proporciona un buen rendimiento (30-60 mg/L y hasta 200 mg/L).

La presente descripción se dirige a un procedimiento de conjugación novedoso que implica células bacterianas modificadas genéticamente para producir las vacunas de bioconjugados finales. Una realización de la descripción permite la producción de vacunas de bioconjugados *in vivo*, evitando la conjugación química y, por lo tanto, simplificando el proceso de producción. La tecnología incluye un nuevo mecanismo genético/enzimático para la síntesis *in vivo* de nuevos bioconjugados que consisten en sacáridos ligados a proteínas.

La base de un aspecto de la descripción incluye el descubrimiento de que *Campylobacter jejuni* contiene un sistema de glicosilación de proteínas ligadas a N general, una característica inusual para los organismos procariontas. Se ha mostrado que varias proteínas de *C. jejuni* se han modificado a través de un heptasacárido. Este heptasacárido se ensambla sobre pirofosfato de undecaprenilo, el lípido de soporte, en el lado citoplásmico de la membrana interna, mediante una adición gradual de monosacáridos activados con nucleótidos, catalizada por glicosiltransferasas específicas. El oligosacárido ligado a lípido se da la vuelta entonces (difunde transversalmente) en el espacio periplásmico a través de una flipasa, por ejemplo, PglK. En la etapa final de la glicosilación de proteínas ligada a N, la oligosacaryltransferasa (por ejemplo, PglB) cataliza la transferencia del oligosacárido desde el lípido de soporte a residuos de Asn dentro de la secuencia de consenso Asp/Glu-Xaa-Asn-Zaa-Ser/Thr (es decir, D/E - X - N - Z - S/T), en donde Xaa y Zaa pueden ser cualquier aminoácido excepto Pro (Figura 7A). Se ha transferido con éxito la agrupación génica para la glicosilación para el heptasacárido en *E. coli* y se han podido producir glicoproteínas ligadas a N de *Campylobacter*.

Se ha podido demostrar que PglB no tiene una especificidad estricta hacia el sustrato de azúcar ligado a lípido. Los polisacáridos antigénicos ensamblados sobre pirofosfato de undecaprenilo son capturados por PglB en el periplasma y se transfieren a un vehículo de proteína (Feldman, 2005; Wacker, M., et al., Substrate specificity of bacterial oligosaccharyltransferase suggests a common transfer mechanism for the bacterial and eukaryotic systems. Proc Natl Acad Sci USA, 2006. 103(18): p. 7088-93). La enzima también transferirá una diversa gama de oligosacáridos ligados a pirofosfato de undecaprenilo (UPP) si contienen una hexosamina N-acetilada en el extremo terminal reductor. La secuencia de nucleótidos de pglB se proporciona en SEQ. ID NO. 1, mientras que la secuencia de aminoácidos de PglB se proporciona en SEQ. ID. NO. 2.

Las Figuras 7A y 7B muestran diagramas esquemáticos de la ruta de glicosilación de proteínas (es decir, el sistema de N-glicosilación) de la presente invención. En una realización, la ruta de glicosilación de proteínas de *C. jejuni* (por ejemplo, incluyendo el operón *pgl*) se puede introducir en *E. coli*. En la Figura 7A, un oligosacárido, específicamente un heptasacárido formado a base de cinco unidades de N-acetil-D-galactosamina, una unidad de glucosa y una unidad de 2,4-diacetamido-2,4,6-trideoxi-D-glucosa, se ensambla sobre un vehículo lipídico, undecaprenilpirofosfato (UDP), utilizando glicosiltransferasas (por ejemplo, *pglA*, *pglC*, *pglH*, J; I) en el lado citoplásmico de la membrana

interna y se transfiere al espacio periplásmico por medio de una flipasa denominada PglK. Por otro lado, una proteína de soporte representada como una espiral y que contiene la secuencia de consenso D/E - X - N - Z - S/T (es decir, Asp/Glu-Xaa-Asn-Zaa-Ser/Thr) se traduce en el citoplasma y se secreta en el espacio periplásmico. En la etapa final, una oligosacaril transferasa (OST o OTasa) (por ejemplo, PglB) transfiere el heptasacárido a residuos de Asn dentro de una secuencia de consenso de la proteína de soporte para producir una glicoproteína.

La Figura 7B también muestra la biosíntesis de un polisacárido (es decir, un polisacárido antigénico o un antígeno) mediante la acción gradual de las glicosiltransferasas, y la transferencia del antígeno O al periplasma por medio de la flipasa, seguido de la polimerización en un polisacárido usando una polimerasa (por ejemplo, wzy). Por otro lado, una proteína de soporte, tal como EPA, se produce y se secreta en el periplasma. Un oligosacaril transferasa (OST u OTasa), tal como PglB, tiene una especificidad de sustrato relajada y transfiere el polisacárido desde un vehículo lipídico a Asn en la secuencia de consenso dentro de EPA.

La Figura 8A muestra un diagrama esquemático que representa una realización de la plataforma de expresión para la producción de un bioconjugado de la presente descripción. La tecnología de la descripción es versátil porque se pueden emplear diversas proteínas de soporte existentes, siempre que la proteína de soporte contenga o se modifique para contener la secuencia de consenso, como se ha descrito anteriormente. En particular, la Figura 8A ilustra la construcción de un hospedador para la expresión, tal como una bacteria *E. coli* modificada genéticamente, en una realización de la invención. Una *E. coli* de este tipo contiene los componentes generales de un sistema de glicosilación (es decir, OST/OTasa, por ejemplo, PglB, y un vehículo de proteína, por ejemplo, EPA). Tales componentes se pueden integrar en el genoma de una cepa de *E. coli*. Además, la ligasa WaaL, así como WecG se eliminan. Adicionalmente, componentes específicos para la expresión del antígeno polisacárido (es decir, una agrupación de genes para la síntesis de un polisacárido que contiene, por ejemplo, glicosil transferasa, polimerasa, flipasa y enzimas para la biosíntesis de azúcar) se pueden proporcionar mediante la adición de un plásmido intercambiable. Esta construcción permite una glicosilación específica del vehículo de proteína con un polisacárido de elección *in vivo*.

En una realización del sistema de expresión para un bioconjugado bacteriano que es compatible con las Normas de Correcta Fabricación (NCF), el ADN que codifica la oligosacariltransferasa inducible y la proteína de soporte se puede integrar de forma estable en un genoma bacteriano (por ejemplo, *E. coli*) de tal manera que se pueden omitir los genes para la selección con antibióticos. Por ejemplo, como se muestra en la Figura 5B, PglB y EPA se integran en el ADN genómico, mientras que el ADN del plásmido, que es intercambiable (es decir, sustituible), codifica una agrupación de genes para la síntesis de polisacáridos.

En otra realización, la Figura 8B muestra un sistema de expresión para un bioconjugado bacteriano que incluye tres plásmidos. Un primer plásmido codifica la proteína de soporte, por ejemplo, AcrA procedente de *Campylobacter jejuni*, que tiene dos sitios de N-glicosilación y se dirige al periplasma a través de un péptido señal PelB. Un segundo plásmido codifica la OST/OTasa, por ejemplo, PglB procedente de *C. jejuni*, que está ligada a la membrana. Un tercer plásmido es un plásmido natural que codifica, por ejemplo, una agrupación génica para la síntesis de un polisacárido (antígeno O), tal como el de *Shigella dysenteriae* 01.

En una realización, un plásmido de expresión para un antígeno O bacteriano, tal como el antígeno O1 de *Shigella dysenteriae*, se puede construir como en pGVXN64 mostrado en la Figura 6B. Este plásmido codifica todas las enzimas necesarias para sintetizar los polisacáridos en la cepa de *Shigella dysenteriae* que componen el serotipo O1. Estas enzimas se enumeran en la columna de la izquierda de la Figura 6B. El vector pGVXN64 que expresa el antígeno O1 de *Shigella dysenteriae* se construyó mediante una digestión de pLARFRI (Vanbleu, E. et al., "Genetic and physical map of the pLARFRI vector" DNA Seq. 15(3):225-227 (2004)) con EcoRI, seguido de la inserción de un cassette de oligonucleótidos (5' -AATTCTGC AGGATCCTCTAGAAGCTTGG (SEQ. ID NO. 3) y 5' -AATTCCAAGCTTCTAGAGGATCCTGCAG (SEQ. ID NO. 4). El fragmento BamHI de pSDM7 (Fait, I. et al., "Construction of recombinant *aroA* salmonellae stably producing the *Shigella Dysenteriae* serotype 1 O-antigen and structural characterization of the *Salmonella/Shigella* hybrid LPS" Microb. Pathog. 20(1): 11-30 (1996)) que contiene la agrupación *rfb* y *rfp* de O1 de *Shigella dysenteriae* se clonó a continuación a través del sitio BamHI en el cassette de oligonucleótidos que contenía pLARFRI. La secuencia completa de nucleótidos que codifica el antígeno O1 de *Shigella dysenteriae* en el plásmido pGVXN64 se indica como SEQ ID NO: 5 en la Lista de Secuencias proporcionada a continuación.

El organismo hospedador de un sistema de expresión de la descripción puede ser, por ejemplo, una cepa de *Escherichia coli*, tal como *Escherichia coli* W311 IOΔwaaL. La delección de WaaL impide la transferencia de cualquier polisacárido al lípido A del núcleo. La copia cromosómica de WaaL también se puede sustituir por PglB. La cepa contiene también una mutación en *wbbL*, por lo que no produce ningún polisacárido de *E. coli* 016. Para aumentar aún más la producción de polisacárido ligado a lípidos de soporte, *wecG* ha sido eliminado para evitar la formación de ECA (antígeno "Entero Common Antigen").

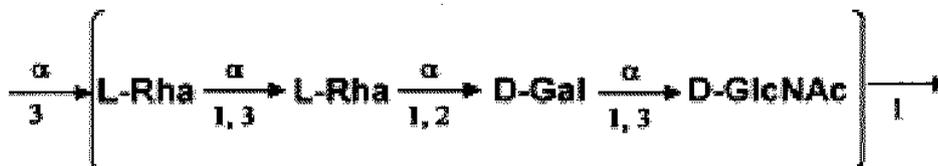
En un aspecto, la presente descripción se dirige además al desarrollo de vacunas de bioconjugados, preferiblemente vacunas de bioconjugados LPS, contra una o varias especies de *Shigella*, que son bacterias invasoras, Gram-negativas. Las especies de *Shigella* causan shigelosis, una inflamación grave del colon. Existen 165 millones de casos en el mundo cada año, en donde un 70% de estos casos, son niños menores de 5 años de edad. En los países en desarrollo, la shigelosis provoca 1,1 millones de muertes por año. Esta es una enfermedad grave que se

transmite por la ruta fecal-oral y es altamente transmisible. Los grupos potenciales que podrían beneficiarse de una inmunización frente a las especies de *Shigella* incluyen, por ejemplo, niños, viajeros y personas en campos de refugiados.

5 Existen cuatro serogrupos diferentes de *Shigella*, a saber, *S. dysenteriae*, *S. flexneri*, *S. sonnei* y *S. boydii*. En realizaciones de la presente invención, los bioconjugados inmunogénicos se pueden preparar contra cada uno de estos diferentes serogrupos de *Shigella*. Por ejemplo, la Figura 14B proporciona diferentes serotipos de *Shigella* y la estructura del polisacárido que define su antigenicidad (es decir, antígenos O de *Shigella*).

10 En realizaciones adicionales de la presente descripción, los bioconjugados de LPS inmunogénicos se podrían preparar contra otras bacterias utilizando las enseñanzas de esta memoria descriptiva, incluyendo bacterias: (1) que causan infecciones nosocomiales, tales como *Pseudomonas aeruginosa*; y (2) que causan una infección del tracto urinario, tales como *E. coli* extraintestinal (ExPEC).

15 En una realización, los inventores han desarrollado una vacuna de bioconjugado de LPS y *Shigella dysenteriae* O1 (también denominada bioconjugado de *S. dysenteriae*), utilizando *E. coli* modificada genéticamente con métodos simples de fermentación y purificación. La Figura 9 muestra la producción de bioconjugados de *Shigella*. El panel superior muestra la síntesis de bioconjugados en *E. coli*. En una realización, la unidad que se repite del antígeno O de *Shigella dysenteriae* O1, se ensambla sobre el lípido de soporte pirofosfato de undecaprenilo (UPP), que se ha dado la vuelta hacia el espacio periplásmico y se polimeriza. La estructura de *S. dysenteriae* O1 es la siguiente y también se proporciona en la parte central derecha de la Figura 9:



20 PglB transfiere el polisacárido activado a residuos de Asn de vehículos de proteínas, formando los bioconjugados de *Shigella*. El vehículo de proteína puede ser, por ejemplo, AcrA o un vehículo de proteína que ha sido modificado para contener la secuencia de consenso para la glicosilación de proteínas, es decir, D/E - X - N - Z - S/T, en donde X y Z pueden ser cualquier aminoácido excepto prolina (por ejemplo, una exotoxina modificada de *Pseudomonas aeruginosa* (EPA)). La EPA se ha utilizado con éxito en las vacunas conjugadas.

25 En una realización ilustrada en la Figura 9, las proteínas periplásmicas de células de *E. coli* que expresan la EPA modificada en presencia de PglB y la agrupación del polisacárido O1, se separaron mediante SDS-PAGE y, después se transfirieron a nitrocelulosa, la EPA se inmunodetectó con un antisuero que se había obtenido contra EPA (carril 2). En el carril 1, las proteínas periplásmicas de células de *E. coli* que expresaban la proteína AcrA de *Campylobacter* en presencia de PglB y la agrupación de polisacárido O1, se separaron y se inmunodetectaron con un antisuero que se había obtenido contra AcrA. Ambas proteínas se glicosilaron con la agrupación de polisacárido O1. En el panel más inferior de la Figura 9, los LPS de *E. coli* se separaron mediante SDS-PAGE y se visualizaron con tinción con plata. El carril de la izquierda representa el LPS extraído a partir de una cepa que no expresa WaaL, mientras que el carril de la derecha muestra el patrón típico del LPS O1. Ambas cepas expresan la agrupación para la biosíntesis del polisacárido de *S. dysenteriae* O1.

35 La producción de un bioconjugado bacteriano, tal como un bioconjugado de *Shigella*, se describe en una realización más detallada con referencia a las Figuras 10A y 10B. Antes de ensamblar el bioconjugado bacteriano utilizando un sistema bacteriano, tal como *E. coli*, es necesario introducir en el sistema bacteriano ciertas secuencias genéticas que codifican las diversas enzimas y proteínas que se deben utilizar, como se ha descrito anteriormente con referencia a las Figuras 8A y 8B. Por ejemplo, esto incluye una OST/OTasa, preferiblemente de procedente *C. jejuni* (por ejemplo, PglB), un vehículo de proteína (por ejemplo, EPA) y una agrupación de genes dirigida a la síntesis de polisacáridos antigénicos (por ejemplo, la agrupación de genes para la síntesis del polisacárido de *S. dysenteriae* O1).

45 La Figura 10A muestra la Etapa 1 en el desarrollo de un bioconjugado bacteriano, a saber, la biosíntesis de un polisacárido, tal como el antígeno O de *S. dysenteriae* serotipo O1. En esta etapa, el antígeno se sintetiza sobre el lípido de soporte, undecaprenil pirofosfato (UPP), y después se transfiere al periplasma usando una flipasa. El antígeno se polimeriza con la polimerasa Wzy y se transfiere al núcleo del lípido A a través de la ligasa WaaL. Para transferir el polisacárido a una proteína de soporte, la ligasa se sustituye por la oligosacariltransferasa; PglB.

50 La Etapa 2 en la producción de un bioconjugado bacteriano implica modificar genéticamente un vehículo de proteína adecuado. Los vehículos de proteínas que son útiles, preferentemente deben tener ciertas características inmunológicas y farmacológicas. Desde una perspectiva inmunológica, preferiblemente, un vehículo de proteína debe: (1) tener epítomos de linfocitos T; (2) ser capaz de entregar un antígeno a células presentadoras de antígeno (APCs) en el sistema inmune; (3) ser potente y duradero; y (4) ser capaz de generar una respuesta de IgG sistémica, específica de antígeno. Desde una perspectiva farmacológica, preferiblemente, un vehículo de proteína debe: (1) no ser tóxico; y (2) ser capaz de entregar antígenos de manera eficaz a través de barreras epiteliales intactas. Más preferiblemente

te, además de estas características inmunológicas y farmacológicas, un vehículo de proteína adecuado para la producción de un bioconjugado bacteriano debe: (1) ser secretado fácilmente en el espacio periplásmico; y (2) ser capaz de tener epítomos de antígenos que se introducen fácilmente como bucles o secuencias lineales en el mismo.

5 Los inventores han encontrado que una exotoxina de *Pseudomonas aeruginosa* (EPA) destoxicada genéticamente y la proteína AcrA de *Campylobacter* son vehículos adecuados de proteínas, la más preferible EPA. AcrA contiene sitios de glicosilación naturales mientras que EPA se tiene que modificar para codificar sitios de glicosilación. Preferiblemente, EPA se modifica con la introducción de dos sitios de glicosilación dirigidos al antígeno OI de *Shigella*. Más preferiblemente, se introducen dos secuencias de consenso como se describe en el Ejemplo 10.

10 La secuencia de aminoácidos de EPA, tal y como se modifica en una realización de esta invención para contener dos sitios de glicosilación, se proporciona como SEQ ID NO: 6 (con secuencia señal) y SEQ ID NO: 7 (sin secuencia señal) en la Lista de Secuencias proporcionada a continuación. Los sitios de glicosilación en cada una de SEQ ID NO: 6 y SEQ ID NO: 7 se indican con un subrayado.

15 La Figura 10B muestra un diagrama esquemático de una proteína de soporte, tal como EPA, sobre la que se pueden diseñar sitios de N-glicosilación como Etapa 3 en la producción de un bioconjugado bacteriano. Los sitios de N-glicosilación requieren la introducción de las secuencias de consenso descritas anteriormente, a saber, la inserción de los sequones D/E - X - N - Z - S/T, en donde X y Z pueden ser cualquier aminoácido natural excepto prolina. Se ha encontrado que tales secuencias de consenso se introducen preferiblemente a través de bucles de la superficie, mediante inserción en lugar de una mutación, y teniendo en cuenta el uso de residuos flanqueantes para optimizar el funcionamiento del sitio de N-glicosilación.

20 La Figura 11 muestra bioconjugados que provocan una respuesta inmune contra el polisacárido OI de *Shigella* en ratones. OI-AcrA y OI-EPA se purificaron mediante cromatografía en columna de afinidad y por intercambio aniónico. El bioconjugado puro se inyectó en ratones ($n = 10$). El suero de los ratones que se habían inmunizado con OI-AcrA (parte superior) u OI-EPA (parte inferior) tres veces (días 1, 21, 60) se reunió y se analizó mediante ELISA el día 70 para estudiar una respuesta de IgG específica de azúcar. Las placas (Nunc, Polysorb) se recubrieron con LPS aislado a partir de *S. dysenteriae* OI y se incubaron con el suero y anti-polivalente de ratón-HRP. Los ratones que recibieron algún conjugado desarrollaron una respuesta de IgG contra el polisacárido, lo que confirma la presencia de epítomos de linfocitos T en los dos vehículos de proteína.

25

En consecuencia, los bioconjugados bacterianos de la presente descripción mostraron una inmunogenicidad *in vivo*. En una realización, los bioconjugados bacterianos son capaces de mostrar: (1) una respuesta específica de hidratos de carbono; y (2) una respuesta específica del vehículo o una respuesta similar independientemente de la proteína de soporte. Por otra parte, una respuesta específica de IgG muestra la dependencia de los linfocitos T de la respuesta inmune, de modo que se espera que genere memoria de la respuesta.

30

La Figura 12 refleja la producción de un bioconjugado OI de *Shigella*, por ejemplo, OI-EPA, en un biorreactor. Células de *E. coli* que expresaban EPA, PglB y el polisacárido OI se cultivaron en un biorreactor hasta $DO_{600} = 40$ mediante dos pulsos de nutrientes. Se indujo la expresión de PglB y EPA y las células se cultivaron durante una noche mediante alimentación lineal de nutrientes. La curva de crecimiento se representa en el panel superior. Los extractos de células completas se separaron mediante SDS-PAGE y la expresión y la glicosilación de EPA se analizaron mediante inmunodetección, utilizando un antisuero policlonal que se obtuvo contra EPA (parte inferior). Las células glicosilaban EPA de manera eficaz con una densidad celular elevada. El procedimiento es reproducible y conduce a una densidad óptica total (DO) de 90, que es un incremento de 45 veces en comparación con el cultivo en matraz de agitación. En consecuencia, es posible una ampliación usando el procedimiento de alimentación discontinua.

35

40

La Figura 13 muestra un ejemplo de purificación de OI-EPA. Más específicamente, la Figura 13 muestra el fraccionamiento y la purificación cromatográfica del bioconjugado OI de *S. dysenteriae*. Las células de *E. coli* que expresaban OI-EPA se cultivaron en el biorreactor hasta una densidad celular elevada (véase la Figura 12). Las células se sedimentaron por centrifugación y las proteínas periplásmicas se extrajeron por choque osmótico. Las proteínas periplásmicas se separaron mediante cromatografía de intercambio aniónico (Source Q). Las fracciones enriquecidas para OI-EPA se purificaron adicionalmente mediante una segunda columna (fluorapatita). Las diferentes fracciones se separaron mediante SDS-PAGE y las proteínas se visualizaron con azul de Coomassie. El carril 1 muestra los extractos de células completas, el carril 2 proteínas periplásmicas después del choque osmótico, el carril 3 proteínas periplásmicas cargadas sobre una columna de intercambio aniónico, los carriles 4 y 5 material eluido desde un intercambio aniónico y el carril 6 material eluido de OI-EPA después de la segunda purificación en columna. Este procedimiento permite la purificación de OI-EPA a gran escala. En esta realización, el procedimiento de purificación es: (1) eficaz (>10 mg/L de cultivo); (2) posible a gran escala; y (3) compatible con las Normas de Correcta Fabricación (NCF). Después de dicha purificación, el rendimiento de EPA-OI para el lote alimentado con LB-glicerol era de hasta 200 mg/L, que es sustancialmente mayor que el rendimiento en matraz de agitación con LB, que era de 0,6 mg/L.

45

50

55

La Figura 14A muestra un análisis de transferencia Western de una serie de especímenes de bioconjugados de AcrA-*Shigella* OI producidos en un matraz de agitación con LB en diversas condiciones, incluyendo una inducción previa, 4 horas después de la inducción, y 4 horas y 19 horas después de la inducción en circunstancias con limita-

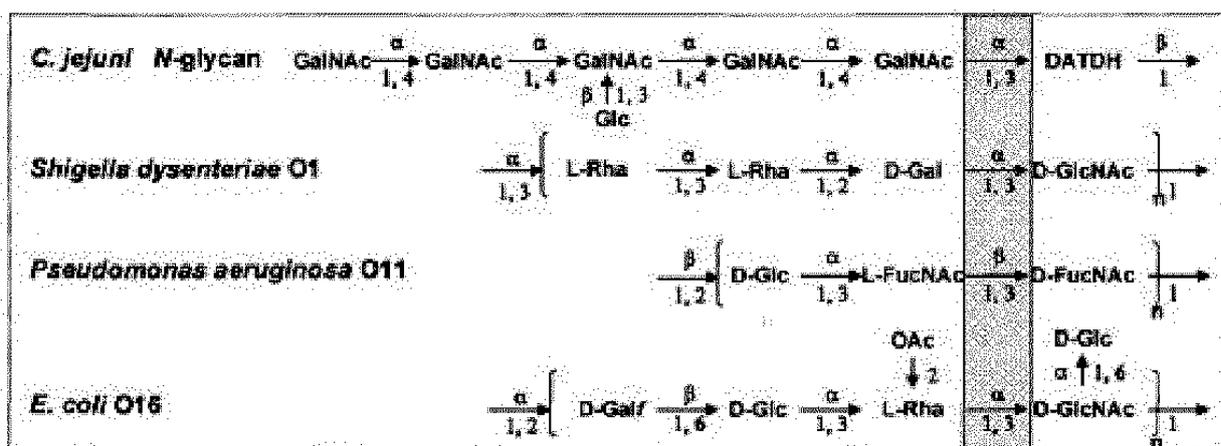
ción de oxígeno. Después de la extracción y la purificación, las proteínas periplásmicas se separaron mediante SDS-PAGE. AcrA y los bioconjugados de AcrA-Shigella O1 se detectaron utilizando un anticuerpo anti-AcrA y una detección quimioluminiscente por medio de un anticuerpo secundario. Las muestras cargadas se normalizaron a una DO₆₀₀ del cultivo en el momento del muestreo.

5 En resumen, en un aspecto, la tecnología de la presente descripción se ha utilizado para desarrollar una vacuna contra una infección con *S. dysenteriae* O1. Por ejemplo, el polisacárido de *S. dysenteriae* O1 se puede conjugar con EPA en *E. coli*. Esto es muy beneficioso ya que EPA ya se había utilizado con éxito en ensayos clínicos con diferentes vacunas conjugadas. En la presente invención, el bioconjugado de *S. dysenteriae* O1 se produjo en un biorreactor a escala 31. Las células se cultivaron con una DO elevada y el bioconjugado se extrajo por choque osmótico. Los bioconjugados se purificaron hasta tener un 98% de pureza mediante cromatografía de intercambio aniónico y de exclusión por tamaño. Los bioconjugados se inyectaron en diferentes cepas de ratones. Después de dos inyecciones, así como tres, se detectó una respuesta de IgG específica de azúcar contra el polisacárido, usando LPS procedente de *Shigella dysenteriae* O1 para el análisis (Figura 11). Como se esperaba, la respuesta de IgM específica se obtuvo cuando se inyectó el LPS. Los bioconjugados produjeron una respuesta de IgG específica contra el polisacárido aislado a partir de *S. dysenteriae*. Se ha mostrado que la respuesta de IgG contra el antígeno de azúcar correspondiente, que se había acoplado químicamente a una proteína de soporte, se correlaciona con la protección en seres humanos.

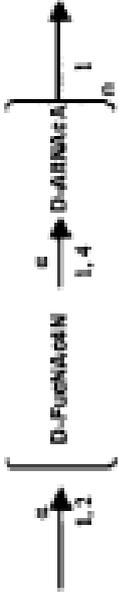
Estos resultados sugieren fuertemente que nuestra cepa de *E. coli* de la invención es adecuada para la producción potencial de una vacuna bacteriana antigénica, como una vacuna contra *Shigella* antigénica. En una realización, el bioconjugado EPA-Shigella se caracterizó intensamente por diferentes métodos, como RMN, HPLC y MS. La Figura 15 muestra la expansión de la región anomérica del espectro 1H RMN de un ejemplo de un bioconjugado con serotipo O1 de *S. dysenteriae* de la descripción. El bioconjugado contiene una relación azúcar/proteína de 0,15, con 13,2 unidades de repetición del antígeno que se unen a la proteína, y 1-2 sitios que se glicosilan. Dos secuencias de consenso para la glicosilación se introdujeron en EPA y aproximadamente 20% de la proteína se glicosiló completamente. El polisacárido se une a través del extremo reductor a la proteína de soporte; por lo tanto, los epítomos del antígeno del polisacárido no están modificados. Además, el método de conjugación *in vivo* fija solo las unidades de repetición del antígeno O a la proteína, pero los monosacáridos del núcleo del lípido A no están fijados.

Utilizando esta tecnología, se pueden producir bioconjugados bacterianos que son inmunogénicos. Se pueden realizar modificaciones genéticas que permiten una conjugación *in vivo* de los polisacáridos bacterianos en las proteínas deseadas y en las posiciones deseadas. Por ejemplo, en una realización y como se ha descrito anteriormente, el polisacárido antigénico de *S. dysenteriae* O1 se puede expresar en *E. coli* y conjugar con dos vehículos de proteínas diferentes *in vivo* (es decir, EPA y AcrA). Ambos bioconjugados provocan una respuesta de IgG específica contra el polisacárido en ratones. Como otro ejemplo, la Tabla 1 a continuación describe diferentes sustratos de polisacáridos para OST/OTasas bacterianas, tales como PgIB, que se pueden utilizar en el método *in vivo* de la presente invención para conjugar un vehículo de proteína con el polisacárido.

Tabla 1



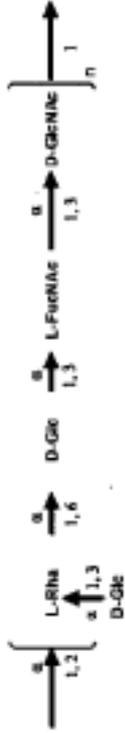
La Tabla 2 siguiente describe los diferentes sustratos adicionales de polisacáridos LPS que se podrían utilizar en la presente invención, con respecto a varias cepas de *Shigella* y *E. coli*., así como de *Pseudomonas aeruginosa* O-11 y *Francisella tularensis*.



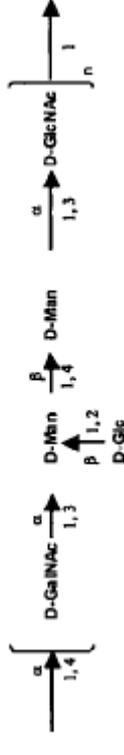
E. coli O4:K52 (ExPEC)



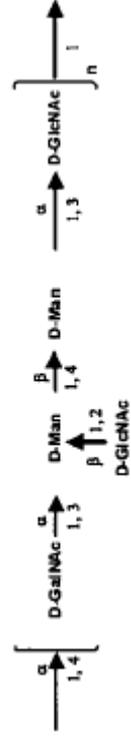
E. coli O4:K6 (ExPEC)



E. coli O6:K2 (ExPEC)

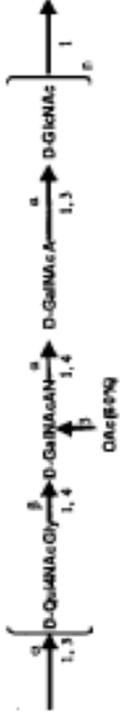


E. coli O6:K54 (ExPEC)



E. coli O22 (ExPEC)

E. coli O121



E. coli O157 (EHEC)



Pseudomonas aeruginosa O11



Francisella tularensis



Qui4NFm, 4,6-didesoxi-4-formamido-D-glucosa
 GalNAcAN, 2-acetamido-2-desoxi-D-galacturonamida
 QuiNAc, 2-acetamido-2,6-didesoxi-D-glucosa

Por ejemplo, en una realización adicional de la descripción, también se pueden desarrollar vacunas de bioconjugados contra E. coli. E. coli es una especie bacteriana bien conocida. Desde un punto de vista genético y clínico, las cepas de E. coli de importancia biológica para los seres humanos se pueden clasificar ampliamente como cepas comensales, cepas patógenas intestinales y E. coli patógena extraintestinal (ExPEC). Las cepas ExPEC pueden formar parte de la flora intestinal normal y se aíslan en el 11% de los individuos sanos. No causan gastroenteritis en humanos, pero su principal característica es su capacidad para colonizar sitios extraintestinales e inducir infecciones en diversos órganos o sitios anatómicos. Son la causa principal de infecciones del tracto urinario (ITU), están involucradas en la septicemia, diversas infecciones abdominales y meningitis. La bacteriemia puede surgir con un riesgo de sepsis grave. La sepsis grave debida a ExPEC se asoció con 41.000 muertes estimadas en 2001. Las cepas de ExPEC han sido susceptibles a los antibióticos; sin embargo, se han desarrollado más y más cepas resistentes a los antibióticos, tanto en hospitales como de forma extrahospitalaria. Esta resistencia antimicrobiana está haciendo que el tratamiento de las infecciones por ExPEC sea más difícil; por lo tanto, las nuevas vacunas serían una estrategia alternativa para evitar esas infecciones.

En modelos animales, la inmunización pasiva o activa contra antígenos capsulares, específicos de O y diferentes proteínas de la membrana externa han proporcionado una protección contra las infecciones sistémicas y la inmunización con estos antígenos diferentes protege contra las infecciones del tracto urinario procedentes de cepas de ExPEC que expresan estos factores de virulencia. Los serotipos 04, 06, 014, 022, 075 y 083 causan el 75% de las ITU. En una realización, la nueva tecnología de la presente invención se puede utilizar para desarrollar un bioconjugado monovalente de LPS que incluye un antígeno (por ejemplo, el serotipo 06, uno de los principales serotipos) e incluso un bioconjugado multivalente de LPS incluyendo estos 6 antígenos. Por ejemplo, la agrupación de genes que codifica las enzimas que sintetizan el antígeno O para ExPEC, se podría amplificar y luego expresar en la cepa de producción de Shigella.

La presente invención implica un procedimiento de producción altamente eficaz, con rendimientos potenciales elevados que se puede utilizar para preparaciones a escala industrial en un procedimiento rentable. Este nuevo enfoque rentable de bioingeniería para la producción de bioconjugados, se puede aplicar a otros conjugados y a aplicaciones. Una característica adicional de la invención implica una simplificación considerable de la producción de vacunas bacterianas con reproducibilidad elevada y un riesgo potencial reducido de aparición de lotes defectuosos.

Procedimiento para la preparación de la vacuna conjugada

Ahora es posible modificar genéticamente sistemas de expresión bacterianos de modo que se producen bioconjugados específicos que son biológicamente activos. Por ejemplo, el polisacárido específico de O de S. dysenteriae se ha conjugado con diferentes vehículos de proteínas y el bioconjugado resultante ha provocado una respuesta de IgG específica contra el polisacárido en ratones. En una realización, la tecnología de la invención hace uso de una oligosacariiltransferasa, por ejemplo, PglB de Campylobacter jejuni para acoplar polisacáridos bacterianos (antígenos O) *in vivo* para expresar de forma simultánea proteínas de soporte recombinantes, produciendo vacunas de bioconjugados altamente inmunogénicas.

Se ha establecido un procedimiento de producción que se puede utilizar a escala industrial. Esto abre la posibilidad de que una multitud de diferentes vacunas conjugadas se pueda desarrollar y fabricar utilizando una simple fermentación bacteriana. El procedimiento tiene varias ventajas en comparación con el método de conjugación *in vitro* que se muestra en el panel superior de la Figura 6A. Como es un procedimiento completamente *in vivo*, el coste y el riesgo de fallos se reducen significativamente y el procedimiento es más reproducible. Además, la secuencia de consenso de captura permite la conjugación de polisacáridos con proteínas definidas en sitios específicos incorporados, lo que facilita una aceptación reglamentaria y el control de calidad. Por último, el desarrollo de vacunas conjugadas es mucho más rápido ya que el procedimiento se simplifica y solamente requiere herramientas biotecnológicas. Además, el procedimiento de conjugación *in vivo* es adecuado para una aplicación en la que las composiciones de polisacáridos evitan una reticulación química.

En una realización, la presente descripción se refiere a la producción a mayor escala de proteínas glicosiladas recombinantes en bacterias y factores determinantes de la eficacia de la glicosilación. Por ejemplo, las proteínas glicosiladas recombinantes de la presente invención se pueden preparar usando el procedimiento de matraz de agitación. Los bioconjugados previamente se producían principalmente en cultivos en matraz de agitación con LB. Más preferiblemente, en un aspecto de la descripción, un primer procedimiento de alimentación discontinua se puede utilizar para la producción de proteínas glicosiladas recombinantes en bacterias. En un procedimiento de preparación preferido, el objetivo es lograr concentraciones finales de biomasa marcadamente crecientes, manteniendo al mismo tiempo una eficacia de la glicosilación y un rendimiento en proteína recombinante por célula y manteniendo al mismo tiempo la simplicidad y la reproducibilidad en el procedimiento.

En una realización, los bioconjugados bacterianos de la presente descripción se pueden preparar a escala comercial mediante el desarrollo de un método de preparación optimizado, empleando procedimientos de producción de E. coli típicos. En primer lugar, se pueden utilizar varios tipos de estrategias de alimentación, tales como por lotes, quimios-tato y alimentación discontinua. En segundo lugar, se puede utilizar un procedimiento que requiere el suministro de oxígeno y uno que no requiere un suministro de oxígeno. En tercer lugar, se puede variar la forma en la que se produce la inducción en el sistema para permitir un rendimiento máximo de producto.

Se ha encontrado que, en contraste con la expresión de la proteína de soporte, el grado de glicosilación ligada a N reacciona fuertemente frente a los cambios en la disponibilidad de nutrientes, tipo de fuente de carbono y energía, suministro de oxígeno y momento de la inducción. Por ejemplo, en un procedimiento de alimentación discontinua, la adición de inductores a los cultivos por lotes y alimentación discontinua conduce inmediatamente a una disminución de 3 veces de la tasa de crecimiento específica, lo que indica una carga metabólica elevada y/o un estrés debido a la síntesis de la proteína de soporte y la oligosacariltransferasa unida a la membrana. Basándose en el descubrimiento de los inventores de un retraso recurrente en la aparición de una proteína de soporte glicosilada, en comparación con la forma no glicosilada después de la inducción, se concluye que la glicosilación parece ser la etapa limitante de la tasa, en la biosíntesis de bioconjugados.

Basándose en esos resultados, en un ejemplo de una realización de la descripción, se ha desarrollado el siguiente diseño de procedimiento para el cultivo: el modo de cultivo de alimentación discontinua para la consecución de altas densidades celulares; la incubación prolongada después de la inducción para facilitar una glicosilación máxima; el exceso de oferta de nutrientes (por ejemplo, los componentes de LB, extracto de levadura y triptona) durante la acumulación de biomasa hasta la inducción para proporcionar un suministro suficiente de elementos estructurales para el proceso de producción; y glicerol como fuente principal de carbono y energía para evitar la represión por catabolito y la formación de acetato. Este bioprocedimiento permite un aumento del rendimiento de 50 veces en comparación con un cultivo por lotes con LB, allanando el camino hacia una producción rentable de vacunas conjugadas en *Escherichia coli* recombinante. En este ejemplo, se pueden tener condiciones óxicas en todo el proceso de producción, por ejemplo, logradas mediante una aireación enriquecida con oxígeno; sin embargo, un contenido bajo de oxígeno también es factible. El Ejemplo 9 describe este ejemplo de un procedimiento de alimentación discontinua con mayor detalle. Debe reconocerse, sin embargo, que se pueden emplear otros procedimientos para producir los bioconjugados de LPS bacterianos de la presente descripción.

Por consiguiente, en una realización de la descripción, *E. coli* se puede utilizar para la producción *in vivo* de proteínas glicosiladas y es adecuada para una producción industrial de proteínas glicosiladas.

Ejemplos

Ejemplo 1: Selección de AcrA como proteína modelo para la optimización de la N-glicosilación

Para optimizar los requerimientos de proteína aceptora para la N-glicosilación, se realizaron estudios detallados en la glicoproteína AcrA de *C. jejuni* (CjO367c). AcrA es una lipoproteína periplásmica de 350 residuos de aminoácidos. Se ha mostrado que la secreción al periplasma, pero no el anclaje sobre lípidos, es un requisito previo para la glicosilación (Nita-Lazar et al., 2005, supra). La señal para la exportación puede ser la secuencia señal natural de AcrA o la señal heteróloga de PelB cuando se expresa en *E. coli*. De los cinco sequones potenciales para la glicosilación λ Minked (NI 17, N123, N147, N273, N274) se utilizan dos iguales en *C. jejuni* y *E. coli* (N 123 y N273 (Nita-Lazar et al., 2005, supra)). AcrA fue elegida como modelo, ya que es la única N-glicoproteína periplásmica de *C. jejuni* para la cual está disponible una información estructural detallada. Se ha publicado recientemente, la estructura cristalina de un homólogo de AcrA, la proteína MexA procedente de la bacteria Gram-negativa *P. aeruginosa* (Higgins et al., (2004). Structure of the periplasmic component of a bacterial drug efflux pump. Proc. Natl. Acad. Sci. USA 70fi 9994-9999). Ambas proteínas son miembros de las denominadas proteínas de bomba de eflujo periplásmicas (PEP, Johnson, J. M. y Church, G. M. (1999). Alignment and structure prediction of divergent protein families: periplasmic and outer membrane proteins of bacterial efflux pumps. J. Mol. Biol. 287, 695-715)). La molécula alargada contiene tres subdominios dispuestos linealmente: una hélice superenrollada α -helicoidal, anti-paralela que se mantiene unida en la base por un dominio lipoilo, que viene seguido por un dominio de β -barril de seis hebras. Los 23-28 residuos en el extremo N-terminal y 95-101 residuos en el extremo C-terminal no están estructurados en los cristales. Las secuencias de las proteínas MexA y AcrA son idénticas en un 29,3% y similares en un 50%. Por lo tanto, las dos proteínas muestran probablemente un plegamiento global similar.

Ejemplo 2: Elucidación de la secuencia peptídica primaria que desencadena la glicosilación

Se sabe que los dominios lipoilo similares a MexA de *P. aeruginosa* y, en consecuencia, también en AcrA de *C. jejuni*, forman una proteína compacta que se puede expresar individualmente en *E. coli* (revisado por Berg, A., y de Kok, A. (1997). 2-Oxo acid dehydrogenase multienzyme complexes. The central role of the lipoyl domain. Biol. Chem. 378, 617-634). Para comprobar qué secuencia aceptora del péptido era necesaria para la N-glicosilación con la maquinaria de pgl en *E. coli*, se tomó el dominio lipoilo de AcrA. Se utilizó como una plataforma molecular para el transporte de péptidos de diferentes longitudes hasta el periplasma y para presentarlos a la maquinaria de pgl *in vivo*.

Por lo tanto, un plásmido que codificaba el dominio lipoilo (Lip) se construyó y se fusionó el extremo N-terminal con la secuencia señal de OmpA (Choi, J. H., y Lee, S. Y. (2004). Secretory and extracellular production of recombinant proteins using *Escherichia coli*. Appl Microbiol Biotechnol 64, 625-635) y el extremo C-terminal con un marcador hexa His. Se realizó una clonación para que la expresión del gen estuviera bajo el control del promotor de arabinosa. Para los límites del dominio Lip, se eligieron posiciones de aminoácidos que aparecían en las mismas posiciones que los límites de la parte del dominio lipoilo en MexA. Para someter a ensayo diferentes péptidos para estudiar su capacidad para aceptar un N-glicano, se insertaron tramos de la secuencia entre las dos partes de tipo cabeza de

martillo del dominio Lip. Los tramos consistían en secuencias que comprendían el sitio de N-glicosilación N 123 de AcrA de *C. jejuni*. Los marcos de lectura abiertos resultantes consistían en las secuencias que codificaban la secuencia señal OmpA, la parte de AcrA de tipo cabeza de martillo como extremo N-terminal (D60-D95, la numeración de los aminoácidos se refiere a la numeración de la secuencia polipeptídica madura de AcrA), en donde los diferentes tramos contenían el sitio de glicosilación N 123 de AcrA (véase más abajo), la parte de AcrA-Lip de tipo cabeza de martillo del extremo C-terminal (L167-D210) y el marcador His C-terminal.

5

La construcción de los plásmidos se logró mediante técnicas convencionales de biología molecular. Se insertaron tres tramos que contenían el sitio de glicosilación N 123 natural de AcrA de longitudes diferentes, entre las dos mitades de Lip, dando como resultado tres ORFs diferentes:

10 La estructura artificial A contiene AI 18-SI 30 que genera una secuencia proteica de:

MKKTAIAIAVALAGFATVAQADVIIKPOVSGVIVNKLKAGDKVKKGQTLFHEODOAS
KDFNRSKALFSQLDHTEIKAPFDGTIGDALVNIGDYVSASTTELVRVTNLNPIYADGS
HHH HHH (SEQ. ID NO. 8).

La estructura artificial B contiene FI 22-EI 38 que genera una secuencia proteica de:

MKKTAIAIAVALAGFATVAOADVIIKPOVSGVIVNKLKAGDKVKKGOTLFIIEQDOFN
RSKALFSQSAISQKELDHTEIKAPFDGTIGDALVNIGDYVSASTTELVRVTNLNPIYA
DGS HHHHHH (SEQ. ID NO. 9).

La estructura artificial C contiene D 121 -A 127 que genera una secuencia proteica de:

MKKTAIAIAVALAGFATVAOADVIIKPOVSGVIVNKLKAGDKVKKGQTLFIIEQDQDF
NRSKALDHTEIKAPFDGTIGDALVNIGDYVSASTTELVRVTNLNPIYADGSHHHHHH
(SEQ. ID. NO. 10).

15 Los tramos subrayados de la secuencia indican el péptido señal OmpA, se introdujeron residuos subrayados aislados para la clonación o para hacer que la proteína fuera resistente a la degradación. Negrita: sitio de glicosilación que corresponde a N 123 de AcrA. Cursiva: marcador hexa-His. Los genes correspondientes se expresaban bajo el control del promotor de arabinosa en la cadena principal del plásmido pEC415 (Schulz, H., Hennecke, H., y Thony-Meyer, L. (1998). Prototype of a heme chaperone essential for cytochrome c maturation. *Science* 281, 1197-1200).

20 Para comprobar cuál de los tres tramos desencadenaba la glicosilación de las proteínas Lip, se realizaron experimentos de expresión de proteínas. Células *E. coli* Top 10 (Invitrogen, Carlsbad, CA, EE.UU.) portadoras de pACYCpgl o pACYC/?g/mut (Wacker et al., 2002, supra) y un plásmido que codificaba las estructuras artificiales A, B o C, se cultivaron en medio LB que contenía ampicilina y cloranfenicol hasta una DO de 0,5 a 37°C. Para la inducción se añadió una solución de 1/1000 volúmenes de 20% de arabinosa (p/v) y las células se cultivaron durante otras 2 horas. Después, las células se recogieron por centrifugación y se resuspendieron en Tris/HCl 20 mM, pH 8,5, 20% de sacarosa (p/v), EDTA 1 mM, PMSF 1 mM, y 1 g/l (p/v) de lisozima y se incubaron a 4°C durante 1 h. Los extractos periplásmicos se obtuvieron después de la sedimentación de los esferoblastos y se diluyeron con 1/9 volúmenes (v/v) de 10x tampón A (NaCl 3 M, Tris/HCl 0,5 M, pH 8,0 e imidazol 0,1 M) y se añadió MgSO₄ hasta 2,5 mM. La purificación con afinidad hacia Ni se realizó en columnas de 1 ml de Ni-Sefarosa de Amersham Pharmacia Biotech (Uppsala, Suecia) en tampón A. Las proteínas eluyeron en tampón A que contenía imidazol 0,25 M.

30 La Fig. 1 muestra un gel de SDS-PAGE teñido con azul brillante de Coomassie de las fracciones pico de la elución a partir de extractos periplásmicos purificados con Ni. El análisis de la expresión mostró que la estructura artificial B producía una especie proteica aislada prominente (Fig. 1, carril 1). Las estructuras artificiales A y C ambas conducían, además de a la proteína prominente, a una segunda banda de proteína con movilidad electroforética más lenta (Fig. 1, carriles 2 y 3). Que la especie proteica más pesada estaba de hecho glicosilada, se demostró por MALDI-TOF/TOF (no se muestra). El único aminoácido que faltaba en la estructura artificial B pero que estaba presente en A y C era D121, un residuo aspartato, 2 posiciones en dirección N-terminal de N 123 glicosilada. Esto demuestra que D 121 tiene un papel importante en la glicosilación a través de la OTasa. Para verificar que D 121 es esencial para la glicosilación, se mutó a alanina en la estructura artificial C. El análisis de la expresión dio como resultado una sola banda de proteína (Figura 1, carril 4), mostrando de este modo que D 121 es importante para la glicosilación. Además, el hecho de que una proteína artificial que muestra el péptido pueda estar glicosilada, muestra que un péptido corto del tipo D/E-X-N-Y-S/T contiene toda la información para que se produzca la N-glicosilación transportada por C-*jejuni*.

45

Ejemplo 3: Verificación del Ejemplo 2: AcrA-D121A no está glicosilada en N 123

Para confirmar los hallazgos del enfoque de presentación de péptidos, se insertó una mutación de aspartato a alanina en la posición 121 (D121A, es decir, 2 residuos antes de N 123 glicosilada) en la versión soluble de longitud completa de la proteína AcrA y se sometió a ensayo si el sitio N 123 todavía se podía glicosilar en *E. coli*. Con el fin de comprobarlo, AcrA-D121A se expresó y se analizó su estado de glicosilación. Para el análisis se utilizó una AcrA modificada genéticamente. Se diferenciaba del gen original de *C. jejuni* en que contenía la secuencia señal PelB (Choi y Lee, 2004, *supra*) para la secreción en el periplasma y un marcador hexa His C-terminal para la purificación. Se ha mostrado que esta variante de AcrA se secreta, se escinde el péptido señal y se glicosila como la proteína natural anclada al lípido (Nita-Lazar et al., 2005, *supra*). Lo siguiente es la secuencia de aminoácidos de la proteína AcrA soluble:

MKYLLLPTAAAGLLLLAAOPAMAMHMSKEEAPKIQMPPQPVTTMSAKSEDLPLS/TY
 PAKLVSDYDVIKPKQVSGVIVNKLKFKAGDKVKKGQTLFIIIEQDKFKASVDSAYGQALM
 AKATFENASKDFNRSKALFSKSAISQKEYDSSLATFNNSKASLASARAQLANARIDL
 DHTEIKAPFDGTIGDALVNIGDYVSASTTELVRVTNLNPIYADFFISDTDKLNLVRNTQ
 SGKWDLDSIHANLNLNGETVQGKLYFIDSVIDANSGTVKAKAVFDNNNSTLLPGAFA
 TITSEGFQKNGFKVPQIGVKQDQNDVYVLLVKNKVEKSSVHISYQNNEYAIIIDKGL
 QNGDKIILDNF KKIQVGSEVKEIGAQLEHHHHHH (SEQ. ID NO. 11)

Los residuos subrayados son el péptido señal PelB, en cursiva el marcador hexa-His y en negrita los dos sitios de glicosilación naturales, en N 123 y N273. Un plásmido que contenía el ORF para la proteína anterior en el plásmido pEC415 (Schulz et al., 1998) se construyó para producir pAcrAper.

El ensayo para analizar el estado de glicosilación de AcrA y de sus mutantes (véase más adelante) fue del modo siguiente: la expresión de AcrA fue inducida con 0,02% de arabinosa en células de *E. coli* CLM24 en crecimiento exponencial (Feldman et al, 2005, *supra*) que contenían el operón *lpg* transportado por el plásmido en su forma activa o inactiva (pACYCpg/ o pACYCpg/mut, véase (Wacker et al., 2002, *supra*)) y un plásmido que codificaba AcrA (pAcrAper). Después de cuatro horas de inducción, se prepararon extractos periplásmicos como se ha descrito anteriormente y se analizaron por SDS-PAGE, electrotransferencia e inmunodetección con antisuero anti-AcrA o antisuero R12. Este último es específico de proteínas que contienen N-glicano de *C. jejuni* (Wacker et al., 2002, *supra*).

Los dos primeros carriles de la Fig. 2A muestran AcrA en ausencia y en presencia de un operón *pgl* funcional. En ausencia del operón *pgl* funcional solo aparece una banda, pero en presencia aparecen tres (Fig. 2A, panel superior). Estas se corresponden con AcrA sin glicosilar (carril 1) y a AcrA sin glicosilar, monoglicosilada y diglicosilada (carril 2). Que las dos proteínas más pesadas en el carril 2 estuvieran glicosiladas se confirmó con la transferencia Western RI 2 (carril 2, panel inferior). Cuando el mutante AcrA-N273Q se expresaba de la misma manera, solo se detectó la AcrA monoglicosilada en presencia del operón *pgl* funcional para la glicosilación (carril 3). AcrA sin glicosilar se detectó en ausencia del locus *pgl* funcional (carril 4). El análisis del mutante AcrA-D121A producía solamente dos bandas, una de ellas glicosilada (carril 5) como se observó con AcrA-N273Q en el carril 3. Esto significa que D121 es esencial para una glicosilación eficaz en la posición 123-125.

Ejemplo 4: Introducción de sitios de glicosilación artificiales en AcrA

Para comprobar si la introducción de un residuo de aspartato podía generar un sitio de glicosilación, se generaron mutantes de AcrA en los que el residuo en la posición -2 de los sitios de glicosilación no utilizados en las posiciones NI 17 y N 147 de AcrA soluble, se intercambiaron por aspartato (FI 15D, T 145D). A continuación, se sometió a ensayo si los sitios de glicosilación modificados se podían glicosilar, con el mismo ensayo que el que se ha descrito en el Ejemplo 2. Ambas mutaciones se insertaron individualmente en la secuencia de tipo silvestre de la versión soluble de AcrA o en el doble mutante en el que se habían delecionado ambos sitios de glicosilación empleados (N123Q y N273Q). Se prepararon extractos de periplasma de cultivos inducidos durante 4 horas, se separaron mediante SDS-PAGE y se analizaron por transferencia Western (Fig. 2B). Como controles, las muestras de AcrA de tipo silvestre glicosilada y no glicosilada se hicieron migrar sobre el mismo gel (carriles 1 y 2). La mutación T145D afectaba a la posición -2 del sequón de glicosilación no utilizado de forma natural NI 47-SI 49. Después de la expresión de AcrA-T145D, una transferencia Western con antisuero anti AcrA dio lugar a cuatro bandas, la más superior de ellas con movilidad electroforética más lenta que la proteína doblemente glicosilada en el carril 2 (carril 3 en la figura 2B). La transferencia RI 2 confirmó que la cuarta banda era una AcrA glicosilada de forma triple. A pesar de la baja intensidad frente a anti AcrA, la banda más pesada proporcionó la señal más fuerte con el antisuero R12 específico de la glicosilación. Cuando el mismo mutante AcrA-T145D se expresaba en ausencia de la secuencia de N-glicosilación natural (AcrA-N123Q-N273Q-T145D), solamente se detectaba AcrA monoglicosilada en presencia de un operón *lpg* funcional (Fig. 2B, carril 4), que faltaba en ausencia de un operón *pgl* funcional (carril 5). Esto demuestra que la banda más pesada en el carril 4 estaba glicosilada. Por lo tanto, mediante una simple introducción de la mutación T145D, se generó un sitio de glicosilación optimizado (DFNNS).

Para confirmar adicionalmente que es posible introducir un sitio de glicosilación mediante la inserción de un residuo de aspartato en la posición -2, los sitios no empleados de forma natural NI 17-SI 19 y N274-T276 se cambiaron para optimizar la N-glicosilación. Para este fin se generaron mutantes adicionales (Fig. 2C). La expresión de AcrA-FI 15D-T145D en el sistema descrito anteriormente dio lugar a cinco especies de proteínas detectadas con el antisuero anti AcrA (carril 2). Esto es indicativo de que tienen lugar cuatro glicosilaciones en la misma molécula de AcrA. Cuando la detección se realizó con el antisuero R12 específico de N-glicano de *C. jejuni*, se detectó una escala de cinco bandas. La banda débil más inferior es AcrA no glicosilada, ya que también está presente en ausencia de glicosilación (carril 1), la más superior da lugar a una señal fuerte, probablemente debida a los cinco determinantes antigénicos en una AcrA glicosilada cuatro veces. Por lo tanto, los dos sitios introducidos (en NI 17 y N 147) y los dos sitios utilizados de forma natural (N 123 y N273) son utilizados y glicosilados por la maquinaria de pgl. La expresión de AcrA-N123Q-N273Q-N272D con y sin el operón pgl, demostró que un tercer sitio de glicosilación introducido artificialmente, N274 (DNNST), también era reconocido por el operón pgl (Fig. 2C, carriles 3 y 4).

Los experimentos anteriores confirman el hallazgo de que el sitio de N-glicosilación bacteriano reconocido por la OTasa de *C. jejuni* consiste en parte en la misma secuencia de consenso que el sitio eucariota (N - X - S/T, con X ≠ P) pero, además, se requiere un aspartato en la posición -2 para aumentar la eficacia. Además, se demuestra que es posible glicosilar una proteína en un sitio deseado mediante la introducción de manera recombinante de una secuencia de consenso optimizada de este tipo.

Ejemplo 5: Verificación de la posición -1 en una secuencia de N-glicosilación optimizada

Un experimento adicional se realizó para analizar si la posición -1 en el sitio de glicosilación bacteriana muestra las mismas restricciones que la posición +1 en eucariotas (Imperiali, B., y Shannon, K. L. (1991). Differences between Asn-Xaa-Thr-containing peptides: a comparison of solution conformation and substrate behaviour with oligosaccharyl-transferase. *Biochemistry* 30, 4374-4380; Rudd, P. M., y Dwek, R. A. (1997). Glycosylation: heterogeneity and the 3D structure of proteins. *Crit. Rev. Biochem. Mol. Biol.* 32, 1-100). Se cree que un residuo de prolina en +1 restringe el péptido de tal manera que la glicosilación es inhibida. Para someter a ensayo si un efecto similar también se podía observar en la posición -1, un residuo de prolina se introdujo en esa posición del primer sitio utilizado de forma natural en un mutante puntual que tenía el segundo sitio natural desactivado (AcrA-N273Q-F122P). La expresión de control de AcrA-N273Q mostraba una proteína monoglicosilada en presencia de un operón funcional /?g/ (Fig. 2D, carril 1 y 2). Sin embargo, AcrA-N273Q-F122P no estaba glicosilada (Fig. 2D, carriles 3 y 4). Esto indica que la prolina inhibe la N-glicosilación bacteriana cuando constituye el residuo entre la asparagina y el residuo cargado negativamente de la posición -2.

Las alineaciones de secuencias de todos los sitios que se sabe que están glicosilados por la maquinaria de pgl de *C. jejuni*, indican que todos ellos comprenden un D o E en la posición -2 (Nita-Lazar et al., 2005, supra; Wacker et al., 2002, supra; Young et al., (2002). Structure of the N-linked glycan present on multiple glycoproteins in the Gram-negative bacterium, *Campylobacter jejuni*. *J. Biol. Chem.* 277, 42530-42539). Por lo tanto, se estableció que la secuencia de consenso de glicosilación para las bacterias se puede optimizar con un aminoácido cargado negativamente en la posición -2, dando lugar a D/E - X - N - Z - S/T, en donde X y Z ≠ P.

Ejemplo 6: N-glicosilación de una proteína que no es de *C. jejuni*

Para demostrar que el requisito de una secuencia primaria (secuencia de consenso optimizada) es suficiente para la N-glicosilación en las bacterias, se sometió a ensayo si una proteína que no era de *C. jejuni* se podía glicosilar mediante la aplicación de la estrategia anterior. Se empleó la subunidad B de la toxina del cólera (CtxB) como diana de la glicosilación. El gen correspondiente se amplificó a partir de *Vibrio cholerae* de una manera tal que contenía la secuencia codificante de la secuencia señal OmpA en el extremo N-terminal y un marcador hexaHis en el extremo C-terminal, igual que en las estructuras artificiales A a C en el ejemplo 1. El ADN resultante se clonó para reemplazar la estructura artificial A en los plásmidos empleados en el ejemplo 1. Una mutación puntual de W88 a D o una inserción de D después de W88 generaba un sitio de glicosilación optimizada (DNNKT). Las proteínas de CtxB de tipo silvestre y W88D que contenían la secuencia señal y el marcador His, se expresaron en *E. coli* Top 10 y otros tipos de células, en presencia y ausencia del locus pgl funcional procedente de *C. jejuni*. Cuando se analizaron los extractos periplásmicos de la células Top 10 mediante SDS-PAGE, electrotransferencia e inmunotransferencia consecutiva con un antisuero de CtxB, solamente CtxB W88D producía una banda superior, y por tanto glicosilada, en la dotación del locus pgl (Fig. 2E, comparar los carriles 3 y 4). Una secuencia de consenso (DSNIT) también se insertó mediante la sustitución de G54 o Q56 de CtxB (este último se denomina CtxB-Q56/DSNIT), es decir, en uno de los bucles de los que se había informado que contribuían a la actividad de unión del gangliósido GM 1 de CtxB. Los carriles 5 y 6 de la Figura 2E muestran que la proteína modificada genéticamente (ejemplificada por la estructura artificial que contiene la secuencia peptídica DSNIT en lugar de Q56 expresada en células Top 10) producía una menor movilidad y, por tanto, glicosilaba la banda en células que eran competentes para la glicosilación pero no carecían de glicosilación, cuando se analizaban de la misma manera como se ha descrito anteriormente. También se mostró que una CtxB que contenía dos manipulaciones, es decir, la inserción de D después de W88, así como DSNIT sustituyendo a Q56, estaba doblemente glicosilada en células SCM7 (Alaimo et al., *EMBO Journal* 25: 967-976 (2006)) (panel E, carriles 7 y 8). La proteína CtxB doblemente glicosilada mostrada en el carril 7 se purificó por afinidad hacia Ni²⁺ y se analizó mediante ESI- MS/MS después de una tripsinización en gel de acuerdo con protocolos convencionales. Los glicopéptidos esperados se detectaron confirmando que la N-glicosilación bacteriana tam-

bién se puede dirigir a una proteína que no es de *C. jejuni*, mediante una mutación o inserción de la secuencia de consenso optimizada de acuerdo con la invención, para la N-glicosilación bacteriana (no se muestra). Ejemplos de otras cepas de *E. coli* ejemplares, adecuadas para la puesta en práctica de la presente invención son W3110, CLM24, BL21 (Stratagene, La Jolla, CA, EE.UU.), SCM6 y SCM7.

- 5 La secuencia de aminoácidos de la proteína de CtxB utilizada en este documento se indica a continuación (secuencia señal OmpA recombinante subrayada, marcador hexa-His en cursiva, W88 en negrita):

MKKTAIAIAVALAGFATVAOATPONITDLCAEYHNTIOIHTLNDKIFSYPESLAGKREM
 AIITFKNGATFQVEVPGSQHIDSQKKAIERMKDTRLRIAYLTEAKVEKLCVWNNKTPHA
 IAAI SMANGSHHHHHH (SEQ. ID NO. 12)

Ejemplo 7: Introducción de sitios de N-glicosilación artificiales en la proteína de la membrana externa de *C. jejuni*. OmpHI

10 Una aplicación potencial de la N-glicosilación en bacterias es la presentación del glicano en la superficie de una célula hospedadora bacteriana con el fin de ligar el fenotipo al genotipo y de ese modo seleccionar mutaciones genéticas específicas. Para demostrar que los N-glicanos se pueden presentar sobre proteínas de la membrana externa, la proteína OmpHI se modificó genéticamente de manera que contenía múltiples sitios de consenso optimizados de acuerdo con la invención. Los sitios se modificaron genéticamente en regiones de bucle de la proteína como se
 15 había deducido por la estructura cristalina conocida (Muller, A., Thomas, G. H., Horler, R., Brannigan, J.A., Blagova, E., Levdikov, V.M., Fogg, M.J., Wilson, K.S., y Wilkinson, A.J. 2005. An ATP-binding cassette-type cysteine transporter in *Campylobacter jejuni* inferred from the structure of an extracytoplasmic solute receptor protein. *Mol. Microbiol.* 57: 143-155). Experimentos previos habían mostrado que los mejores sequones de glicosilación se generaban con las mutaciones V83T, K59N-G601-N61T, R190N-G191I-D192T y H263D-F264S-G265N-D266I-D267T. Para la
 20 presentación en la superficie se deseaba evaluar diferentes combinaciones de esos sitios introducidos, con el fin de establecer la muestra más específica de N-glicano. Las combinaciones se generaron en una estructura artificial de plásmido que codificaba OmpHI de tipo silvestre y se sometieron a ensayo de una manera similar a la descrita para AcrA. La Figura 3 muestra el análisis de diversas variantes de OmpHI que albergaban múltiples sequones de glicosilación, además del sequón existente de tipo silvestre. Las variantes de OmpHI se generaron con tres (carril 3, 4, 5 y
 25 7) y cuatro sequones de glicosilación (carril 6). Una OmpHI de tipo silvestre con un solo sequón de glicosilación y un mutante que carecía de la asparagina decisiva para la glicosilación, también se incluyeron en el experimento. Todas las variantes sometidas a ensayo en este documento no solo demostraron un alto nivel de eficacia de la glicosilación, sino que también se utilizaba cada sequón de glicosilación. Los resultados se confirmaron con inmunosuero específico de N-glicano de *Campylobacter* (Fig. 3 panel inferior).

- 30 A continuación se muestra la secuencia de aminoácidos de la proteína OmpHI de *Campylobacter jejuni* (cepa 81-176) con el marcador myc fijado en cursiva:

MKKILLSVLTTFFVAVVLAACGGNSDSKTLNSLDKIKONGWRIGVFGDKPPFGYVDEK
 GNNQGYDIALAKRIAKELFGDENKVQFVLVEAANRVEFLKSNKVDIILANFTQTPERA
 EQVDFCLPYMKVALGVAVPKDSNITSVEDLKDKTLLLNGGTTADAYFTQDYPNIKTL
 KYDQNTETFAALMDKRGDALSHDNTLLFAWVKDHPDFKMGIKELGNKDVIAPAVKK
 GDKELKEFIDNLIKLGQEQFFHKAYDETLKAHFGDDVKADDWIEGGKILEQKLZSEE
 DL (SEQ. ID NO. 13)

El sitio de glicosilación natural en la proteína está en negrita, la secuencia señal está subrayada.

- 35 Ejemplo 8: Presentación en la superficie de N-glicanos procedentes de *C. jejuni* sobre OmpHI en la membrana externa de células *E. coli*

Para responder a la cuestión de si múltiples variantes de OmpHI glicosiladas se pueden presentar en la superficie de células bacterianas, se realizó una inmunofluorescencia sobre células bacterianas CLM24 o SCM6 (que es SCM7 Δ*waal*) que expresaban diversas variantes de OmpHI. Una OmpHI de tipo silvestre y un mutante que carecía de la asparagina decisiva para la glicosilación, se incluyeron en el experimento. Además, un mutante C20S se construyó
 40 con el fin de retener la proteína en el periplasma, sirviendo de este modo como un control en el experimento. La inmunotinción se llevó a cabo en las células tratadas con paraformaldehído. El paraformaldehído fija las células sin destruir la estructura celular o la compartimentación. Se empleó inmunosuero específico de c-Myc y N-glicano en combinación con los correspondientes anticuerpos secundarios conjugados con FITC y Cy3 para detectar la proteína (fluorescencia roja) y el N-glicano (verde) en la superficie celular bacteriana, respectivamente. Además, se empleó
 45 4,6-diamino-2-fenilindol (DAPI, azul) para teñir el ADN bacteriano para diferenciar sin ambigüedad entre las células bacterianas y los restos celulares. Cuando se teñieron las células que expresaban OmpHI de tipo silvestre, se detectó una inmunofluorescencia específica de la proteína así como del N-glicano (Fig. 4A). Cuando un mutante que carecía de la asparagina decisiva N139S se teñía tanto con inmunosuero específico anti-Myc y anti-N-glicano, solamente se

obtuvieron señales específicas de la proteína pero no del glicano (panel 4 B) lo que indica una especificidad del inmunosuero específico de N-glicano. Cuando la proteína se retenía dentro del periplasma como en el mutante C20S, no se detectó una inmunofluorescencia roja específica de la proteína, lo que indicaba que los anticuerpos no eran capaces de difundir dentro de la célula y eran lo suficientemente competentes para detectar cualquier fenómeno de la superficie (panel 4 C). A continuación, se tiñeron las células que expresaban múltiples variantes de OmpHI con diferentes glicosilaciones: OmpHI^{KGN→NIT, HFGDD→DSNIT} (panel 4 D)

OmpHI^{RGD→NIT, HFGDD→DSNIT} (panel 4 E), OmpHI^{KGN→NIT, RGD→NIT} (panel 4 F), OmpHI^{V83T, KGN→NIT} (panel 4 G) y OmpHI^{ANIT, RGD→NIT, HFGDD→DSNIT} (panel, 4 H) AU le

Las variantes de OmpHI se tiñeron de forma doble, indicando la presencia de proteína glicosilada en la superficie bacteriana. La Figura 4 se representa en escala de grises, la primera columna es una imagen de combinación de las otras imágenes de la misma fila.

La Figura 4 muestra la microscopía de fluorescencia de células que expresaban diferentes variantes de OmpHI. Los cultivos de cepas de *E. coli* CLM24 o SCM6 que contenían el plásmido de expresión para OmpHI de tipo silvestre y sus variantes, se igualaron a una DO₆₀₀ de 0,25/ml. Las células se lavaron dos veces con solución salina tamponada con fosfato (PBS), pH 7,4 y suspensiones de células de 100 µl se dejaron gotear sobre portaobjetos de vidrio gelatinizados y se incubaron a temperatura ambiente (TA) durante 30 min dentro de una cámara humidificada. Todas las etapas posteriores para marcar con inmunofluorescencia las células completas, se realizaron a temperatura ambiente en una cámara humidificada. Las células no unidas se eliminaron y el resto se fijó con paraformaldehído al 4% que contenía PBS durante 30 min a TA. Es importante destacar que se considera que el paraformaldehído no permeabiliza las células, pero conserva la compartimentación mediante membranas intactas. Las células fijadas se lavaron dos veces con PBS y se resuspendieron en tampón de bloqueo que contenía 5% de BSA en PBS. Después del bloqueo, las células se incubaron con IgG de ratón monoclonal anti-myc (1:50, Calbiochem) y/o antisuero anti-glicano (1:4000) durante 1 h en 100 µl de PBS que contenía 5% de BSA. Las células se lavaron tres veces con 100 µl de PBS durante 5 min cada vez y se incubaron con anticuerpo secundario anti-conejo conjugado con FITC (1:250, Jackson ImmunoResearch Laboratories) y/o anticuerpo anti-ratón conjugado con Cy3 (1:250, Jackson ImmunoResearch Laboratories) durante 1 h en 100 µl de PBS que contenía 5% de BSA. Si era necesario, se añadió 4,6-diamino-2-fenilindol (DAPI) (Sigma) (0,5 µg/ml) en el momento de la incubación con el anticuerpo secundario para teñir el ADN bacteriano. El anticuerpo secundario se aclaró de las células con PBS y los cubreobjetos se montaron sobre portaobjetos usando medio para montar Vectashield (Vector Laboratories) y se sellaron con esmalte de uñas. La microscopía de fluorescencia se realizó empleando un microscopio Axioplan2 (Carl Zeiss). Las imágenes se combinaron mediante el uso de Adobe Photoshop, versión CS2. Las células SCM6 expresaban OmpHI (panel A), OmpHI^{N139S} (panel B), OmpHI^{C20S} (panel C), OmpHI^{KGN→NIT, HFGDD→DSNIT} (panel D), OmpHI^{RGD→NIT, HFGDD→DSNIT} (panel E), OmpHI^{KGN→NIT, RGD→NIT} (panel F), OmpHI^{V83T, KGN→NIT} (panel G), y OmpHI^{KGN, Nrr, RGD, Nrr, HFGDD→DSNrr} (panel H). La primera columna es una fusión de las imágenes en las columnas 2, 3 y 4 representadas en tonos grises sobre fondo negro. Columna 2: fluorescencia azul en tonos grises de la tinción con DAPI, columna 3: fluorescencia verde procedente de la fluorescencia específica de glicano, columna 4: fluorescencia roja procedente de la tinción anti-myc.

Ejemplo 9: Un ejemplo de un procedimiento de producción de bioconjugado de Shigella OI LPS

Este es un ejemplo de un procedimiento de producción; sin embargo, diferentes condiciones también conducen a la formación de un producto similar.

A. Procedimiento de producción

La cepa de *E. coli* W31 IOΔwaaL que contenía tres plásmidos que expresaban PglB, EPA y las enzimas para la biosíntesis del polisacárido OI de *Shigella* fue utilizada para la producción del bioconjugado LPS. Una sola colonia se inoculó en 50 ml de medio LB y se cultivó a 37°C O/N. El cultivo se utilizó para inocular un cultivo de 1 l en un biorreactor de 2 l. El biorreactor se agitó con 500 rpm, el pH se mantuvo a 7,0 mediante la adición controlada automáticamente de KOH 2 M o 20% de H₃PO₄ y la temperatura de cultivo se ajustó a 37°C. El nivel de oxígeno disuelto (pO₂) se mantuvo entre 0 y 10% de oxígeno. Las células se cultivaron en un medio de glicerol semidefinido que contenía kanamicina a una DO₆₀₀= 15. El medio contenía los siguientes ingredientes: glicerol 330 mM, 10 g de extracto de levadura, 20 g de triptona, K₂HPO₄ 34 mM, KH₂PO₄ 22 mM, (NH₄)₂SO₄ 38 mM, MgSO₄·7 H₂O 2 mM y ácido cítrico 5 mM. Después de una fase inicial por lotes de aproximadamente 5 h, se añadió un primer pulso de nutrientes para mantener la acumulación rápida de biomasa (glicerol, triptona y extracto de levadura). Después de 1,5 h adicionales, el cultivo alcanzó una DO₆₀₀= 30. En ese momento, se añadió un segundo dosis-pulso de nutrientes de glicerol y triptona junto con los inductores requeridos 1% de L-arabinosa e IPTG 1 mM. Con el fin de mantener la inducción en los niveles máximos y suministrar otros aminoácidos para la síntesis de proteína recombinante, una alimentación lineal de nutrientes/inductor (28,8 ml/h) se inició con la adición de este pulso. La alimentación se mantuvo hasta el final del procedimiento. El cultivo del biorreactor se recogió después de un total de ~ 24 h de cultivo, cuando debería haber llegado a una DO₆₀₀ de ± 80.

El procedimiento de producción se analizó por transferencia Western como se ha descrito previamente (Wacker, M., et al., N-linked glycosylation in *Campylobacter jejuni* and its functional transfer into *E. coli*. *Science*, 2002. 298(5599):

p. 1790-3). Después de la transferencia a la membrana de nitrocelulosa, la muestra se inmunotizó con el anticuerpo anti-EPA específico (Wacker, M., et al., N-linked glycosylation in *Campylobacter jejuni* and its functional transfer into *E. coli*. *Science*, 2002. 298(5599): p. 1790-3). Anti-IgG de conejo-HRP (Biorad) se utilizó como anticuerpo secundario. La detección se realizó con reactivos de detección de transferencia Western ECL™ (Amersham Biosciences, Little Chalfont Buckinghamshire).

La Figura 16A muestra proteínas extraídas del bioconjugado de LPS y *Shigella* OI (es decir, EPA-OI) a partir de un procedimiento de alimentación discontinua que se normalizó a una concentración de biomasa (DO_{600nm} 0,1 células/carril). Las proteínas se separaron mediante SDS-PAGE, se transfirieron a una membrana de nitrocelulosa y se visualizaron mediante el anticuerpo anti EPA de conejo. El tiempo de inducción para la expresión de PglB y EPA era de 1 h y O/N.

B. Extracción de proteínas periplásmicas

Las células se recogieron por centrifugación durante 20 min a 10000 g y se resuspendieron en 1 volumen de 0,9% de NaCl. Las células se sedimentaron mediante centrifugación durante 25-30 min a 7.000 g. Las células se resuspendieron en tampón de suspensión (25% de sacarosa, EDTA 100 mM, Tris HCl 200 mM, pH 8,5, DO 250/ml) y la suspensión se incubó con agitación a 4-8°C durante 30 min. La suspensión se centrifugó a 4-8°C durante 30 min a 7.000-10.000 g. El material sobrenadante se desechó, las células se resuspendieron en el mismo volumen de Tris HCl 20 mM enfriado sobre hielo, pH 8,5 y se incubaron con agitación a 4-8°C durante 30 min. Los esferoblastos se centrifugaron a 4-8°C durante 25-30 min a 10.000 g, se recogió el material sobrenadante y se pasó a través de una membrana de 0,2 μ .

Como se muestra en la Figura 16B, el extracto periplásmico se cargó en un SDS-PAGE al 7,5%, y se tiñó con Coomassie para identificar EPA y EPA-OI. EPA es una banda gruesa que migra por encima del marcador de 70 kDa. OI-EPA (es decir, EPA-OI) migra como un líder entre 100 y 170 kDa.

C. Purificación de bioconjugados

El material sobrenadante que contenía las proteínas periplásmicas obtenidas a partir de una DO 100.000 de las células, se cargó en una columna de intercambio aniónico Source Q (XK 26/40 ~ 180 ml de material de lecho) equilibrada con tampón A (Tris HCl 20 mM pH 8,0). Después de lavar con 5 volúmenes de columna (VC) de tampón A, las proteínas eluyeron con un gradiente lineal de 15 VC hasta 50% de tampón B (Tris HCl 20 mM + NaCl 1 M pH 8,0) y luego 2 VC hasta 100% de tampón B. Las proteínas se analizaron mediante SDS-PAGE y se tiñeron con Coomassie. Se reunieron las fracciones que contenían OI-EPA. Normalmente, el bioconjugado eluía con una conductividad entre 6-17 mS. La muestra se concentró 10 veces y el tampón se intercambió a Tris HCl 20 mM pH 8,0.

Como se muestra en la Figura 17A, las fracciones de proteínas procedentes de 1. Source Q se analizaron por SDS-PAGE y se tiñeron con Coomassie. Las fracciones C1 a G9 contenían bioconjugado OI y se agruparon.

El bioconjugado OI se cargó una segunda vez en una columna Source Q (XK 16/20 ~ 28 ml de material de lecho) que se había equilibrado con tampón A: Tris HCl 20 mM pH 8,0. Se utilizó un gradiente idéntico al que se había empleado anteriormente para eluir el bioconjugado. Las proteínas se analizaron por SDS-PAGE y se tiñeron con Coomassie. Se reunieron las fracciones que contenían OI-EPA. Normalmente, el bioconjugado eluía con una conductividad entre 6-17 mS. La muestra se concentró 10 veces y el tampón se intercambió a Tris HCl 20 mM pH 8,0.

Como se muestra en la Figura 17B, las fracciones proteicas de la columna 2. Source Q se analizaron con SDS-PAGE y se tiñeron con Coomassie. Las fracciones A1 a B3 que contenían el bioconjugado OI, se agruparon.

El bioconjugado OI se cargó en una columna Superdex 200 (Hi Load 26/60, grado prep) que se equilibró con Tris HCl 20 mM pH 8,0.

Como se muestra en la Figura 18A, las fracciones proteicas procedentes de la columna Superdex 200 se analizaron con SDS-PAGE y se tiñeron con Coomassie. Las fracciones F1 a F1 se agruparon.

Como se muestra en la Figura 18B, el bioconjugado de *Shigella* procedente de diferentes etapas de purificación se analizó por SDS-PAGE y se tiñó con Coomassie. OI-EPA se purificó hasta más de un 98% de pureza mediante el procedimiento, mostrando que el bioconjugado OI-EPA se puede producir con éxito utilizando esta tecnología.

Ejemplo 10: Modificación genética de la exotoxina A de *Pseudomonas aeruginosa* para la glicosilación con carbohidratos antigénicos

La exotoxina A de *Pseudomonas aeruginosa* (EPA) es una proteína secretada extracelularmente de 67 kDa que codifica 613 aminoácidos en su forma madura y que contiene cuatro puentes disulfuro (C1-C15, C197-C214, C265-C287, C372-C379). Para permitir su glicosilación en *E. coli*, la proteína se debe localizar en el espacio periplásmico para que se produzca la glicosilación. Por lo tanto, un péptido señal de la proteína DsbA de *E. coli* se fusionó genéticamente con el extremo N-terminal de la secuencia madura de EPA. Un plásmido obtenido a partir de pEC415 [Schulz, H., Hennecke, H., y Thony-Meyer, L., Prototype of a heme chaperone essential for cytochrome c maturation,

Science, 281, 1197-1200, 1998] que contenía el código del péptido señal DsbA seguido por una secuencia de AR-Nasa se digirió (NdeI para EcoRI) para conservar la señal DsbA y eliminar el inserto de ARNasa. EPA se amplificó usando PCR (el oligo directo era 5'-AAGCTAGCGCCGCGGAGGAAGCCTTCGACC (SEQ. ID NO. 14) y el oligo inverso era 5'-AAGAATTCTCAGTGGTGGTGGTGGTGGTGGTTCAGGTCCTCGCGCGGCGG (SEQ. ID NO. 15)) y se digirió con NheI/EcoRI y se ligó para reemplazar la secuencia de la ARNasa eliminada previamente. La estructura artificial resultante (pGVXN69) codificaba un producto proteico con un péptido señal DsbA, la secuencia madura de EPA y un marcador hexa-His. La destoxificación se logró mediante la mutación/eliminación de los residuos catalíticamente esenciales L552VΔE553 de acuerdo con [Lukac, M., Pier, G.B., y Collier, R.J., Toxoid of Pseudomonas aeruginosa exotoxin A generated by deletion of an active-site residue, Infect Immun, 56, 3095-3098, 1988] y [Ho, M.M., et al., Preclinical laboratory evaluation of a bivalent Staphylococcus aureus saccharide-exotoxin A protein conjugate vaccine, Hum Vaccin, 2, 89-98, 2006] usando mutagénesis Quick Change (Stratagene) y los oligonucleótidos fosforilados 5'-GAAGGCGGGCGCGTGACCATTCTCGGC (SEQ. ID NO. 16) and 5'-GCCGAGAATGGTCACGCGCCCGCCTTC (SEQ. ID NO. 17), dando lugar a la estructura artificial pGVXN70.

Se sabe que la inserción de una secuencia de pentapéptido del tipo D/E-Z-N-X-S/T en una posición adecuada da lugar a una glicosilación. Para glicosilar EPA en células de E. coli, se insertaron dos sitios de glicosilación diferentes en las estructuras artificiales descritas anteriormente de acuerdo con la siguiente descripción.

Para insertar un sitio en la posición 375, se realizaron dos etapas. Primero, mutagénesis Quick Change usando los oligos 5'-CCTGACCTGCCCGGGGAATGCGCGG (SEQ. ID NO. 18) y 5'-CCGCGCATTCGCCGGGCAGGT-CAGG (SEQ. ID NO. 19) con pGVXN70 como molde, dio lugar a una estructura artificial que contenía un único sitio SmaI en la posición aminoacídica 375 de la secuencia de la proteína EPA, mediante la delección de tres residuos, pero por lo demás conservando intacta la secuencia proteica de partida. En una segunda etapa, un inserto compuesto por dos oligonucleótidos complementarios, fosforilados que codificaban (i) los residuos previamente delecionados (cuando se insertaba el sitio SmaI), (ii) el sequeón de glicosilación pentapéptido y (iii) residuos adicionales de lisina que flanqueaban el consenso para mejorar la eficacia de la glicosilación (como se observó con experimentos adicionales), se ligó en el sitio SmaI (5'-GTCGCCAAAGATCAAAATAGAACTAAA (SEQ. ID NO. 20) y 5'-TTTAGTTCTATTTTGATCTTTGGCGAC (SEQ. ID NO. 21). La estructura artificial resultante era pGVXN137. Para insertar un sitio de glicosilación adicional en la estructura artificial en el aminoácido 240, se aplicó un procedimiento de una etapa empleando la mutagénesis Quick Change con los oligonucleótidos 5'-CATGACCTGGACATCAAGGATAAT AATAATTCTACTCCACGGTCATCAGTCATC (SEQ. ID NO. 22) y 5'-GATGACTGATGACCGTGGGAGTAGA ATTATTATTATCCTTGATGTCCAGGTCATG (SEQ. ID NO. 23) sobre la estructura artificial pGVXN137. La estructura artificial resultante contenía, por lo tanto, varios cambios en comparación con la proteína EPA de tipo silvestre: dos sitios de glicosilación, un péptido señal DsbA, mutación de destoxificación.

Por otra parte, en los casos en la memoria descriptiva en los que se observaron secuencias específicas de nucleótidos o de aminoácidos, se entenderá que la presente descripción incluye secuencias homólogas que todavía tienen la misma funcionalidad que las secuencias indicadas. Preferiblemente, tales secuencias son al menos 85% homólogas. Más preferiblemente, tales secuencias son al menos 90% homólogas. Lo más preferiblemente, tales secuencias son al menos 95% homólogas.

Lista de Secuencias

SEQ. ID NO. 1: SECUENCIA DE NUCLEÓTIDOS PARA PglB
 LOCUS pglb 2169 pb de ADN SYN 17-FEB-2009
 DEFINICIÓN
 ACCESO pglb
 PALABRAS CLAVE.
 FUENTE.
 ORGANISMO C. jejuni 81.116
 REFERENCIA 1 (bases 1 hasta 2169)
 AUTORES Los mismos
 REVISTA Sin publicar.
 RECUENTO DE BASES 650 a 264 c 386 g 869 t
 ORIGEN

1 ATGTTGAAAA AAGAGTATTT AAAAAACCCT TATTTAGTTT TGTTTGCGAT GATTATATTA
 61 GCTTATGTTT TTAGTGTATT TTGCAGGTTT TATTGGGTTT GGTGGGCAAG TGAGTTTAAAT
 121 GAGTATTTTT TCAATAATCA GTTAATGATC ATTTCAAATG ATGGCTATGC TTTTGCTGAG
 181 GGCGCAAGAG ATATGATAGC AGGTTTTTCAT CAGCCTAATG ATTTGAGTTA TTATGGATCT
 241 TCTTTATCCG CGCTTACTTA TTGGCTTTAT AAAATCACAC CTTTTTCTTT TGAAAGTATC
 301 ATTTTATATA TGAGTACTTT TTTATCTTCT TTGGTGGTGA TTCCTACTAT TTTGCTAGCT
 361 AACGAATACA AACGTCCTTT AATGGGCTTT GTAGCTGCTC TTTTAGCAAG TATAGCAAAC
 421 AGTTATTATA ATCGCACTAT GAGTGGGTAT TATGATACGG ATATGCTGGT AATTGTTTTG
 481 CCTATGTTTA TTTTATTTTT TATGGTAAGA ATGATTTTAA AAAAAAGACTT TTTTTCATTG
 541 ATTGCCCTGC CGTTATTTAT AGGAATTTAT CTTTGGTGGT ATCCTTCAAG TTATACTTTA
 601 AATGTAGCTT TAATTGGACT TTTTFTAATT TATACACTTA TTTTTCATAG AAAAGAAAAG
 661 ATTTTTTATA TAGCTGTGAT TTTGTCTTCT CTTACTCTTT CAAATATAGC ATGGTTTTAT
 721 CAAAGTGCCA TTATAGTAAT ACTTTTTGCT TTATTCGCCT TAGAGCAAAA ACGCTTAAAT
 781 TTTATGATTA TAGGAATTTT AGGTAGTGCA ACTTTGATAT TTTTGATTTT AAGTGGTGGG
 841 GTTGATCCTA TACTTTATCA GCTTAAATTT TATATTTTAA GAAGTGATGA AAGTGCGAAT
 901 TTAACGCAGG GCTTTATGTA TTTTAATGTC AATCAAACCA TACAAGAAGT TGAAAATGTA
 961 GATCTTAGCG AATTTATGCG AAGAATTAGT GGTAGTGAAG TTGTTTTTTT GTTTTCTTTG
 1021 TTTGGTTTTG TATGGCTTTT GAGAAAACAT AAAAGTATGA TTATGGCTTT ACCTATATTG
 1081 GTGCTTGGGT TTTTAGCCTT AAAAGGGGGG CTTAGATTTA CCATTTATTC TGTACCTGTA
 1141 ATGGCCTTAG GATTTGGTTT TTTATTGAGC GAGTTTAAAG CTATAATGGT TAAAAAATAT
 1201 AGCCAATTA CTTCAAATGT TTGTATTGTT TTTGCAACTA TTTTGACTTT AGCTCCAGTA
 1261 TTTATCCATA TTTACAATA TAAAGCGCCA ACAGTTTTTT CTCAAATGA AGCATCATT
 1321 TTAAATCAAT TAAAAAATAT AGCCAATAGA GAAGATTATG TGGTAACTTG GGCGGCTTAT
 1381 GGTTATCCTG TCGGTTATTA TAGCGATGTG AAAACTTTAG TAGATGGTGG AAAGCATTTA
 1441 GGTAAGGATA ATTTTTTCCC TTCTTTTGCT TTAAGCAAAG ATGAACAAGC TGCAGCTAAT
 1501 ATGGCAAGAC TTAGTGTAGA ATATACAGAA AAAAGCTTTT ATGCTCCGCA AAATGATATT
 1561 TAAAAACAG ACATTTTGCA AGCCATGATG AAAGATTATA ATCAAAGCAA TGTGGATTTG
 1621 TTTCTAGCTT CATTATCAA ACCTGATTTT AAAATCGATA CGCCAAAAAC TCGTGATATT
 1681 TATCTTTATA TGCCCGCTAG AATGTCTTTG ATTTTTTCTA CGGTGGCTAG TTTTCTTTT
 1741 ATTAATTTAG ATACAGGAGT TTTGGATAAA CCTTTTACCT TTAGCACAGC TTATCCACTT
 1801 GATGTTAAAA ATGGAGAAAT TTATCTTAGC AACGGAGTGG TTTAAGCGA TGATTTTAGA
 1861 AGTTTTAAAA TAGGTGATAA TGTGGTTTCT GTAAATAGTA TCGTAGAGAT TAATTCTATT
 1921 AAACAAGGTG AATACAAAAT CACTCCAATT GATGATAAGG CTCAGTTTTA TATTTTTTAT
 1981 TTAAAGGATA GTGCTATTCC TTACGCACAA TTTATTTTAA TGGATAAAAC CATGTTAAT
 2041 AGTGCTTATG TGCAAATGTT TTTTTTAGGA AATTATGATA AGAATTTATT TGACTTGGTG
 2101 ATTAATTCTA GAGATGCTAA GGTTTTTAAA CTAAAATTT ACCCATACGA TGTTCCAGAT
 2161 TACGCTTAA

SEQ. ID NO. 2: SECUENCIA DE AMINOÁCIDOS PARA PgIB
 722 AA

MLKKEYLKNPYLVLFAMIILAYVFSVFCRFYVWWWASEFNEYFFNNQLMIISNDGYAF
 GARDMIAGFHQPNDLSYYGSSLSALTYWLYKITPFSFESIILYMSTFLSSLVVIPTILLA
 NEYKRPLMGFVAALLASIANSYNRTMSGYYDMDLVIVLPMFILFFMVRMILKKDFFSL
 IALPLFIGIYLWWYPSSYTLNVALIGLFLIYTLIFHRKEKIFYIAVILSSLTSLNIAWFY
 QSIIIVILFALFALEQKRLNFMIIIGILGSATLIFLILSGGVDPIYQLKFYIFRSDESAN
 LTQGFMYFNVNQTIQEVENVDLSEFMRRISGSEIVFLFSLFGFVWLLRKHKSMIMALPIL
 VLGFLALKGGLRFTIYSVPVMALGFGFLSEFKAIMVKKYQLTSNVCIVFATILTLAPV
 FIIHINYKAPT VFSQNEASLLNQLKNIANREDYVVTWAAAYGYPVRYSDVKTLVDGGKHL
 GKDNFFPSFALS KDEQAAANMARLSVEYTEKSFYAPQNDILKTDILQAMMKDYNQSNVDL
 FLASLSKPDFKIDTPKTRDIYLYMPARMSLIFSTVASFSFINLDTGVLDKPFTFSTAYPL
 DVKNGEIYLSNGVVLSDDFRSFKIGDNVSVNSIVEINSIKQGEYKITPIDDKAQFYIFY
 LKDSAIPYAQFILMDKTMFNSAYVQMFFLGNYDKNLFDLVINSRDAKVFLLKIYPYDVPD
 YA

5

SEQ. ID NO. 3: CASETE DE OLIGONUCLEÓTIDOS EN EL PLÁMIDO pGVXN64
 5'-AATTCTGCAGGATCCTCTAGAAGCTTGG

SEQ. ID NO. 4: CASETE DE OLIGONUCLEÓTIDOS EN EL PLÁMIDO pGVXN64
5' AATTCCAAGCTTCTAGAGGATCCTGCAG

5 **SEQ. ID NO 5: SECUENCIA COMPLETA DE NUCLEÓTIDOS QUE CODIFICA EL ANTÍGENO O1 DE SHIGELLA DYSENTERIAE EN EL PLÁSMIDO PGVXN64**

LOCUS i-O1Cluste 10174 pb de ADN SYN 23-ENE-2009

DEFINICIÓN Agrupación O1 invertida

ACCESO i-O1Cluste

PALABRAS CLAVE.

10 FUENTE Desconocida.

ORGANISMO Desconocido

Sin clasificar.

REFERENCIA 1 (bases 1 a 10174)

AUTORES Los mismos

15 REVISTA Sin publicar.

RECUENTO DE BASES 3223 a 1581 c 2074 g 3296 t

ORIGEN

1 ATGAAAAACC ATAAGGTTAG TATTATCATC CCCTGTTTTA ATAACGGAAA AACTATAGAA
61 AGATGCGTTC TTTCCGCATT AAAACAAACC CATAGAAATA TAGAGATTAT TTGTATAAAT
121 GATGGACCGT CAGATAATAG CTGGAATATT TAAAGAAAT TAAGCTCATT GTATGGAAAT
181 GTTTTTGCTT TTAATAATGA GGACAATTCA GGGCCGTCAT TTTCCAGGAT TAAAGGGGTT
241 TCACTATCTA CGGGGCATTT TTTGTCATTT TTGGATGCAG ATGATTACTG GCATCCAAAA
301 AAACTAGAAT TACAACATC ATTTATTAAT GATGAAAAC TGGATTTTTT AGGTTCAACG
361 TGTTCATTG GTGAGAAAAA TAACCAAGAA ATTAACCAAG GAATTAAAAA AGAACATTTA

421 AAATTA AAAA TAATTT CATT TAACATGATG TTGTTTAAGA ATTATTTCCA GACTCCAGCT
 481 GTCATTATGA AAAGAGATAT TTTTATTCCA TTTAATGAGA ATCAGCGTTT TTCAGAGGAC
 541 TACATGTCAT GGCTTGTTAT CGTTTATAAT AAAAAACAA ATGTGGATTT AATATATGGA
 601 AGGGATTTGG TTTTCTCGA TAAATTTAAC TTTGGAGTGT CAGGGTTGAG TGGAATTTA
 661 TGGTTGATGG AGAAGTGGGA GTTAAAAAAT ATATTTAACT TCTTGTTGAA AGGTAAAATA
 721 ATGGCAGTGC CTGCGATCTT GTTTTCTTTG ATAAAATATG AAAGAAGATG CGCTTTAACA
 781 AAGAAAAATA AAGGTAAGGG TAATAAATAA TGAAGATCTC AATAATAGGG AACACAGCAA
 841 ATGCTATGAT TTTGTTTAGA TTGGATTTAA TAAAAACT AACCAACAAA GGGATTTTCA
 901 TCTATGCTTT TGCTACTGAC TATAATGATT CATCCAAGGA AATAATAAAA AAAGCAGGGC
 961 CCATTCCTGT TGATTATAAT TTAAGTCGCA GTGGTATTA CTTGCTGGT GATTTATGGA
 1021 ATACTTACTT ATTAAGTAAA AAATAAAGA AGATAAAACC AGATGCTATT TTATCTTTTT
 1081 TTTCAAAGCC CTCTATCTTT GGATCGTTGG CTGGTATTTT TTCAGGCGTT AAAAATAATA
 1141 ACGCTATGCT TGAGGGGTTA GTTTTTTTAT TTACAGAGCA GCCACATGGA ACTCCGTTAA
 1201 AAACAAAGTT ACTTAAAAAT ATCCAGGTTT CCTGTATAA AATAATATTT CCACATATCA
 1261 ACTCATAAT ACTCCTAAC AAGGATGATT ATCATGATTT GATAGATAAA TACAAAATAA
 1321 AATTA AAAATC TGCCATATT CTTGGTGGCA TTGGTTTAGA TATGAATAAT TACTGTAAAA
 1381 GCACGCCACC AACAAATGAA ATATCATTCA TTTTATAGC TCGTTTGCTA GCAGAAAAAG
 1441 GAGTCAATGA GTTTGTGCT GCCGAAAAA AAATAAAAAA AACACATCCC AATGTTGAAT
 1501 TTATTATACT TGGCGCTATA GATAAGGAAA ACCCGGAGG GTTATCTGAA TCTGACGTAG
 1561 ATACTTTAAT TAAATCAGGA GTTATTTCTT ATCCCGGATT TGTTTCTAAT GTGGCTGATT
 1621 GGATTGAAAA ATCAAGCGTA TTTGTTCTT CTTCTATTA TCGAGAGGGA GTTCTCGTA
 1681 GTACACAAGA AGCGATGGCT ATGGGGAGGC CGATTTTAACTACTAATTTA CCAGGCTGCA
 1741 AAGAAACAAT TATTGATGGT GTGAATGGAT ATGTTGTA AAAATGGTCA CATGAAGATC
 1801 TTGCAGAAAA AATGCTGAAG TTAATTAATA ATCCTGAAAA AATAATCAGT ATGGGAGAAG
 1861 AAAGTTATAA GTTAGCAAGA GAAAGATTCG ATGCAAATGT AAATAATGTA AAGTTATTA
 1921 AAATACTAGG GATTCCTGAT TAATAAACGA AAAGCGGCTC TGATTCATTC GGAACATAA
 1981 ACCTATCTCA ATAGGAGCTA AATTCATGAC CTTACCCAGC CATATCGACC TGCAGCCTGA
 2041 CTGATGCCAT TGCCGAACCTG GCGAAAAAAC AGTCCGTTGA TGCCATGCTG ATGACTGGAG
 2101 ACAGCTACGA CTGCGGAAAA AAAATGGGCT ATATGCAGGC GTTTGTGAAG TATGGGCTGC
 2161 GCAACCTCAA AGAAGGGGCG AAGTTCCGTA AAGGGATTGA GAAGCTGTTA AGCGAATAAT
 2221 GAAAACTGA CCGAATGTAA CGGTTGATAA GAAAATTATA ACGCAGTGA AGATTCATGG
 2281 CGAAAGTAAT TTGTTGCGAA TATTCCTGCC GTTGTTTTAT ATAAACAATC AGAATAACAA
 2341 AGAGTTAGCA ATAGGATTTT CGTCAAAGTT TTCCAGGATT TTCCTTGTTT CCAGAGCGGA
 2401 TTGGTAAGAC AATTAGTGTT TGAATTTTTT GGGTTTAGCG CGAGTGGGTA ACGCTCGTCA
 2461 CATCGTGGAC ATGTATGCAG TGCTCTGGTA GCTGTAAAGC CAGGGGCGGT AGCGTGCATT
 2521 AATACCTCTA TTAATCAAAC TGAGAGCCGC TTATTTTACA GCATGCTCTG AAGTAATATG
 2581 GAATAATAAA GTGAAGATAC TTGTTACTGG TGGCGCAGGA TTTATTGGTT CTGCTGTAGT
 2641 TCGTCACATT ATAAATAATA CGCAGGATAG TGTTGTTAAT GTCGATAAAT TAACGTACGC
 2701 CGGAAACCTG GAGTCACTTG CTGATGTTT TGACTCTAAA CGCTATGTTT TTGAACATGC
 2761 GGATATTTGC GATGCTGCTG CAATGGCGCG GATTTTTGCT CAGCATCAGC CGGATGCAGT
 2821 GATGCACCTG GCTGCTGAAA GCCATGTGGA TCGTTCAATT ACAGGCCCTG CGGCATTTAT
 2881 TGAAACCAAT ATTGTTGGTA CTTATGTCCT TTTGGAAGCG GCTCGCAATT ACTGGTCTGC
 2941 TCTTGATGGC GACAAGAAAA ATAGCTCCG TTTTCATCAT ATTTCTACTG ACGAAGTCTA
 3001 TGGTGATTTG CCTCATCCTG ACGAAGTAAA TAATAAAGAA CAATTACCCC TCTTTACTGA
 3061 GACGACAGCT TACGCGCCTA GTAGTCCTTA TTCCGCATCA AAAGCATCCA GCGATCATT
 3121 AGTCCGTGCG TGGAACGTA CCTATGGTTT ACCGACCATT GTGACTAAT GTTCGAATAA
 3181 CTACGGTCTT TATCACTTTT CGGAAAAATT GATTCCACTA GTAATTCTTA ATGCTCTGGA
 3241 AAGTAAGGCA TTACCTATT ATGGCAAAGG GATCAAATT CGTGACTGGC TGTATGTTGA
 3301 AGATCATGCG CGTGCGTTAT ATATCGTCGT AACCGAAGGT AAAGCGGGTG AAAC TTATAA
 3361 CATTGGTGG AACAACGAAA AGAAAAACAT CGATGTAGTG CTCACTATTT GTGATTTGTT
 3421 GGATGAGATT GTACCGAAAG AGAAATCTTA CCGCGAGCAA ATTACTTATG TTGCCGATCG
 3481 CCCGGGACAC GATCGCCGTT ATGCGATTGA TGCAGAGAAG ATTAGCCGCG AATTGGGCTG
 3541 GAAACCGCAG GAAACGTTT AGAGCGGGAT TCGTAAAACG GTGGGATGGT ACCTCTCAA
 3601 TACAAAATGG GTTGATAATG TAAAAAGTGG TGCCTATCAA TCGTGGATTG AACAGAATA
 3661 TGAGGGCCGC CAGTAATGAA TATCCTCTT TCGGCAAAA CAGGGCAGGT AGGTTGGGAA
 3721 CTACAGCGTG CTCTGGCACC TCTGGGTAAT TTGATTGCTC TTGATGTTCA CTCCACTGAT

3781 TACTGTGGTG ATTTTAGTAA TCCTGAAGGT GTAGCTGAAA CCGTAAGAAG CATTCCGGCCT
 3841 GATATTATTG TCAACGCAGC CGCTCACACC GCAGTAGACA AAGCAGAATC AGAACCGGAG
 3901 TTTGCACAAT TACTTAACGC GACGAGTGTC GAAGCGATCG CGAAAGCAGC CAATGAAGTC
 3961 GCGCCTGGG TTATTACTA CTCTACTGAC TACGTATTTT CGGGGACCGG TGAAATACCA
 4021 TGGCAGGAGG CGGATGCAAC CGCACCCTA AATGTTTACG GTGAAACCAA GTTAGCTGGA
 4081 GAAAAAGCAT TACAAGAGCA TTGTGCGAAG CACCTAATTT TCCGTACAAG CTGGGTCTAT
 4141 GCAGGTAAAG GAAATAACTT CGCCAAAACG ATGTTGCGTC TGGGAAAAGA GCGTGAAGAA
 4201 TTAGCCGTTA TTAATGATCA GTTTGGTGCG CCAACAGGTG CTGAACTGCT GGCTGATTGT
 4261 ACGGCACATG CAATTCGTGT GGCCTGAAT AAACCAGAAG TCGCAGGCTT GTACCATCTG
 4321 GTAGCCACTG GTACCACAAC CTGGCACGAT TATGCTGCGC TGTTTTTTGA AGAGGCACGA
 4381 AAAGCAGGTA TTCCCTTGC ACTCAACAAG CTCAACGCAG TACCAACAAC AGCTTATCCT
 4441 ACACCAGCTC GTCGTCCACA TAACTCTCGC CTTAATACAG AAAAATTTCA GCAAAATTTT
 4501 GCGCTTGT TTGCTGACTG GCAGGTTGGC GTGAAACGAA TGCTCAACGA ATTATTTACG
 4561 ACTACAGCAA TTTAATAGTT TTTGCATCTT GTTCGTGATG ATGGAGCAAG ATGAATTTAA
 4621 AGGAATGATG TAATGAAAAC GCGTAAAGGT ATTATTTTAG CGGGTGGCTC TGGTACTCGT
 4681 CTTTATCCTG TGACTATGGC TGTCAGTAAA CAGCTATTAC CTATTTATGA TAAGCCGATG
 4741 ATCTATTACC CGCTCTCTAC ACTGATGTTG GCGGGTATT CCGATATTCT GATTATTAGT
 4801 ACGCCACAGG ATACTCCTCG TTTTCAACAA CTCCTGGGTG ATGGTAGCCA GTGGGGTTA
 4861 AATCTTCAGT ACAAAGTGCA ACGAGTCCA GATGGTCTTG CGCAGGCATT TATCATCGGT
 4921 GAAGAGTTTA TCGGTGGTGA TGATTGTGCT CTGGTTCTCG GTGATAATAT CTTCTACGGT
 4981 CATGATCTGC CGAAGTTAAT GGATGTCGCT GTCAACAAAG AAAGTGGTGC AACGGTATTT
 5041 GCCTATCACG TTAATGATCC TGAACGCTAC GGTGTTGTTG AGTTTGATAA AAACGGTACG
 5101 GCAATCAGCC TGGAAGAAAA ACCGCTACAA CCAAAAAGTA ATTATGCGGT AACCGGGCTT
 5161 TATTTCTATG ATAACGACGT TGTCGAAATG GCGAAAAACC TTAAGCCTTC TGCCCGTGGT
 5221 GAACTGGAAA TTACCGATAT TAACCGTATT TATATGGAGC AGGGGCGTTT ATCCGTTGCC
 5281 ATGATGGGAC GTGGTTATGC ATGGCTGGAC ACGGGGACAC ATCAAAGTCT TATTGAAGCA
 5341 AGCAACTTCA TTGCAACAAT TGAAGAGCGC CAAGGGTTAA AGGTATCTTG CCTGGAAGAG
 5401 ATTGCTTATC GTAAAGGCTT TATTGACGCA GAGCAGGTTA ATGTATTAGC CGAACCGCTA
 5461 AAGAAAAATG CTTATGGTCA GTATCTGTTG AAAATGATTA AAGGTTATTA AAAATGAATG
 5521 TAATTTAAAC TGAAATTTCA GATGTATTA TTTTCGAGCC GAAAGTTTTT GGTGATGAAC
 5581 GTGGTTTTTT TATGGAAAGC TTAACCAGA AAGTTTTCGA AGAGGCTGTA GGGCGGAAGG
 5641 TTGAATTTGT TCAGGATAAC CATTCTAAAT CAACTAAGGG TGTGTTACGC GGACTGCACT
 5701 ATCAGTTGGA ACCTTATGCT CAAGGTAAAT TAGTTCGTTG TGTTGTCGGT GAAGTTTTTG
 5761 ATGTAGCAGT TGATATTCGT AAATCGTCAC CTACATTTGG GAAATGGATT GGGGTGAATT
 5821 TGTCTGCTGA GAATAAGCGT CAGTTGTGGA TACCTGAAGG ATTTGCGCAT GGATTTTTGG
 5881 TGCTGAGTGA AACGGCTGAG TTTGTTTATA AAACAACAAA CTATTACAAT CCAAGTTTTG
 5941 AAAAAAGTAT TTCATACTCA GATCCTACCA TAAAATTCA GTGGCCCAAT TTACAGGATA
 6001 TGCATTTTAA ATTATCAAAT AAGGATTTGA ATGCTAAGAA CTTTTTAAAT AACAAATGTT
 6061 TAATGCAATG AAGAAAAATA TATTGCTCTT GTTCTTAGTA CATGGGACCA ATTATTTGTT
 6121 CCCGTTTATA GTTCTCCAT ATCAAACCTG AATATTAAGC ATCGAGACAT TCGCAGATGT
 6181 AGCAAAAATT CAAGCCGCTG TGATGCTTTT ATCTTTAATC GTAAATTATG GATATAACTT
 6241 ATCAAGTACA AGAGCTATAG CTAGGGCCGT ATCTCAAGCA GAAATAAATA AGATCTATAG
 6301 TGAGACTCTT ATTGTA AAAAT TATTATTGGC AACCATTTGT CTTGCACTTG GTTGCGTACA
 6361 TTTGATGTAT GTCAAAGAGT ACTCATTGAT ATATCCTTTT ATAATCAGTT CGATATATCT
 6421 TTATGGTAGT GCATTATTTG CTACTTGGTT ATTCCAAGGA CTTGAGAAAA TGAAAGCGGT
 6481 CGTTATAGCA ACAACAATCG CTAAACTGAC TGGTGTGATA CTTACTTTTA TTTTAGTTAA
 6541 GTCTCCAAAT GATATAGTTG CAGCTCTTTT TACACAAAAC ATTGGGATGT TTATAAGTGG
 6601 TATAATATCT ATTTATTTGG TAAGGAAAAA CAAATATGCA ACCGTAATAT GTTTTCGACT
 6661 TAAAAATATT ATTGTAAGCT TAAAAGAAGC GTGGCCGTTT TTTTATCAT TAGCTGCAAC
 6721 AAGTGTATAT ACATATTTTA ATGTGATTTT ATTATCTTTT TATGCTGGCG ACTATGTTGT
 6781 GGCAAATTTT AATGCTGCTG ATAAATTAAG AATGGCTGCT CAAGGGTTAC TTATTCCAAT
 6841 AGGACAGGCT GTTTTCCAC GATTATCTAA ACTAGAGGGC TATGAATATA GTTCTAAACT
 6901 TAAAATTTAT GCAATAAGGT ATGCTATTTT TGGTGTGTTG ATTAGTGCGG GACTTGTATT
 6961 TTTAGGTTCC ATGTTAACTA CTATTTATTT AGGCAAAGAA TATTCGTTGT CAGGAGAATA
 7021 TCTTCAAAGT ATGTTTTTAC TACCTGCCAC TATTTCAATA TCGACTATAC TGAGTCAATG
 7081 GATGTTGATA CCTCAAGGCA AAGAAAAAAT ATTAAGCAGA ATCTATATTC TAGGCGCCAT

7141 TGTCCATTTA TTATATGCAT TTCCTTTAGT TTACTIONTAT GGGGCTTGGG GCATGGTAAT
 7201 ATCAATTTTA TTTACTGAAG TCTTAATTGT ATTATTTATG CTTAAGGCTG TGAAATGACT
 7261 TACTTTACTG GTTTTATTTT AATATTGTTT GCTATTATAA TTAAAAGATT AACTCCAAGT
 7321 CAAAGCAAGA AAAATATTGT CTTAATAGCT AATGCGTTTT GGGGAATATT GTTGGTAGGT
 7381 TATGCTTTCA ATGAACAATA TTTCTGACCA TTAAGTGCAA CAACCTTGTT TTTTATACTT
 7441 GCATTCTTAT TTTTCTTTAG TATGACTTAT ATTTTAATTG CTAGGAGTGG AAGGGTTGTT
 7501 TTTTCTTTTCG GTACTGGTTT TATAGAAAGC AAATATATTT ACTGGTTTGC TGGGATGATT
 7561 AATATTATTA GTATCTGCTT TGGCATTATC CTTTTATATA ATAATCATT TCTTTAAAA
 7621 GTAATGAGAG AAGGAATTTT AGATGGTTCT ATTAGTGGGT TTGGATTGGG GATAAGTTTG
 7681 CCACTTTCCT TCTGCTGTAT GTATTTAGCA AGACATGAGA ATAAAAAAA TTATTTCTAT
 7741 TGTTTTACAC TACTTTCATT CTGCTTGGC GGTATTCAA CTTCAAAGAT CTTCTTAATA
 7801 TTATTCCTTG TATATATTGT TGGAATAAAT AGTTATGTAA GCAAAAAGAA ATTGCTTATT
 7861 TATGGAGTGT TTGTATTTGG ACTGTTGCTT TATCAAGTA TTATCTTGGG TAAGTTCTCT
 7921 TCAGACCCTG AAGGCAAGAT TATTTTCAGCA ATATTTGATA CGTTAAGGGT TTATCTTTTC
 7981 CTTGGATTGG CAGCCTTTAA TCTTTATGTT GAAAAGAATG CCACGCTCCC CGAAAATTTA
 8041 CTTTTGTATC CATTAAAGGA GTTTGGGGG ACGACAAAAG ATATTCCCAA AACTGATATT
 8101 TTGCCTTGGG TCAACATTGG TGTATGGGAC ACGAATGTAT ATACAGCTTT TGCACCATGG
 8161 TATCAGTCAT TGGGATTATA TGCAGCTATA ATTATTGGTA TTCTCTTAGG GTTTTATTAC
 8221 GGGATATGGT TTAGCTTTTCG TCAAAATTTA GCTGTGGGT TTTATCAAAC ATTTTTGTGT
 8281 TTTCTCTTT TAATGTTGTT TTTCCAGGAG CATTATTTGT TGTCATGGAA AATGCATTTT
 8341 ATTTATTTTT TATGTGCAAT TTTATTAGCG ATGAGAAAAG CATTAGAGTA TGAATAAATA
 8401 TTGTATCTTA GTACTATTTA ATCCAGATAT AAGTGTTTTT ATTGATAATG TCAAAAAGAT
 8461 TTTATCTTTG GATGTAAGTT TATTTGTATA TGACAATTCA GCAAATAAAC ATGCATTCTT
 8521 TGCTCTATCC TCACAAGAGC AAACAAAGAT AAATTACTTT TCGATATGTG AAAATATCGG
 8581 ATTGTCGAAA GCTTATAATG AGACACTAAG GCATATTCTT GAATTTAATA AGAATGTGAA
 8641 AAATAAAAGC ATTAATGATA GTGTGCTTTT TCTCGACCAA GACTCTGAAG TTGATTTAAA
 8701 TTCCATCAAT ATTTTGTTTG AAACATATC AGCAGCAGAG TCTAATGTGA TGATAGTCGC
 8761 GGGGAATCCC ATAAGGAGAG ATGGACTACC GTATATAGAT TACCCCAACA CTGTAACAA
 8821 TGTAATAATTT GTAATTAGTA GTTATGCTGT GTATCGCTTA GACGCATTTA GAAACATCGG
 8881 CTTGTTTCAA GAAGATTTTT TTATAGATCA TATCGATAGT GATTTTTGTT CAAGGCTGAT
 8941 AAAAAGCAAT TACCAAATC TCCTTAGAAA AGATGCCTTT TTTTACAAC CAATAGGAAT
 9001 AAAACCATC AATCTCTGTG GTAGATATTT ATTCCCTATC CCATCACAAC ACCGAACATA
 9061 TTTTCAAATT AGAAATGCTT TTTAAGTTA CAGGCGCAAT GGTGTTACAT TTAATTTTTT
 9121 ATTTAGGGAA ATTGTAATA GATTGATTAT GAGTATATTC TCAGGCCTTA ACGAGAAAGA
 9181 CTTATTGAAA CGATTGCATT TATATTTAAA AGGAATAAAA GATGGTCTTA AAATGTAATT
 9241 CTTGGCTAGA AGTGGGGGCG TTGTGATTAA AAAAAAGTG GCGGCGATAA TTATAACATA
 9301 TAATCCAGAT CTAACAATC TCGGAGAAAG TTATACGAGT CTATATAAGC AAGTCGATAA
 9361 AATAATTCTT ATTGATAACA ACTCTACAAA CTATCAAGAA CTTAAGAAGT TATTCGAAAA
 9421 AAAAGAAAAA ATAAAAATAG TGCCCTTGAG TGATAATATA GGACTAGCAG CAGCTCAAAA
 9481 TTTAGGTTTG AACTTAGCTA TAAAAATAA CTATACTTAT GCTATTTTAT TCGATCAGGA
 9541 TAGCGTCTTA CAAGACAATG GAATTAACAG TTTCTTTTTT GAATTTGAGA AATTAGTTAG
 9601 TGAAGAAAAA TTAAATATAG TTGCCATTGG GCCAAGTTTT TTTGACGAAA AGACAGGAAG
 9661 ACGCTTTCGG CCTACAAAAT TTATCGGTCC CTTTTTATAT CCCTTTCGTA AAATAACCAC
 9721 AAAAAATCCT CTAACAGAAG TTGACTTCTT GATTGCTTCT GGTTGTTTCA TAAAATTGGA
 9781 GTGTATTAAA TCAGCCGGA TGTGACTGA ATCGTTATC ATCGATTATA TTGATGTTGA
 9841 ATGGTCATAT CGTATGCGTT CGTATGGCTA TAAGCTATAT ATTCATAATG ATATTCACAT
 9901 GAGTCATTTA GTGGGAGAAT CTCGAGTTAA TTTAGGATTG AAAACTATTT CTTTACATGG
 9961 GCCGCTAAGA CGATATTACT TATTTAGGAA TTATATTTCA ATTTTAAAAG TGAGATATAT
 10021 ACCGTTAGGA TATAAAATAC GTGAGGGTTT TTTAATATC GGAAGATTTT TGGTAAGTAT
 10081 GATTATAACT AAAAATAGAA AAACTTTAA TTTATACACT ATAAAAGCAA TTAAGGACGG
 10141 AATAAATAAT GAAATGGGGA AATATAAAGG CTA

SEQ. ID NO. 6: SECUENCIA DE AA DE EPA (643 AA), QUE CONTIENE DOS SITIOS DE GLICOSILACIÓN, CON SECUENCIA SEÑAL

MKKIWLALAGLVLAFSASAAEEAFDLWNECAKACVLDLKDGVSRSSRMSVDPADTN
 GQGV LHYSMVLEGGNDALKLAIDNALSITS DGLTIRLEGGVEPNKPVRYSYTRQARGSW
 SLNWL VPIGHEKPSNIKVFIHEL NAGNQLSHMSPIYTIEMGDELLAKLARDATFFVRAHE
 SNEMQPTLAISHAGVSVVMAQAQPRREKRWSEWASGKVLCLLDPLDGVYNYLAQQRC
 NLDDT WEGKIYRVLAGNPAKHDLDIK DNNNSTPTVISHRLHFPEGGSLAALTAHQACHL
 PLEAFTRHRQPRGWEQLEQCGYPVQRLVALYLAARLSWNQVDQVIRNALASPGSGGDL
 GEAIREQPEQARLALTLAAAESERFVRQGTGNDEAGAASADVSLTCPVAK DQNRTKG
 ECAGPADSGDALLERNYPTGAEFLGDGGDV SFSTRGTQNWTVRLLQHRQLEERGYV
 FVG YHGT FLEAAQSIVFGGVRARSQDLDAIWRGFYIAGDPALAYGYAQDQEPDARGRIR
 NGALLRVYVPRWSLPGFYRTGLTLAAPEAAGEVERLIGHPLPLRLDAITGPEEEGGRVTI
 LGWPLAERTVVIPSAIPTDPRNVGGDLDPSSIPDKEQAISALPDYASQPGKPPREDLK

SEQ. ID NO. 7: SECUENCIA DE AA DE EPA PROCESADA (624 AA), QUE CONTIENE DOS SITIOS DE GLICOSILACIÓN

AEEAFDLWNECAKACVLDLKDGVSRSSRMSVDPADTN GQGV LHYSMVLEGGNDALK
 LAIDNALSITS DGLTIRLEGGVEPNKPVRYSYTRQARGSWSLNWL VPIGHEKPSNIKVFIH
 ELNAGNQLSHMSPIYTIEMGDELLAKLARDATFFVRAHESNEMQPTLAISHAGVSVVMA
 QAQPRREKRWSEWASGKVLCLLDPLDGVYNYLAQQRCNLDDT WEGKIYRVLAGNPAK
 HDL DIK DNNNSTPTVISHRLHFPEGGSLAALTAHQACHL PLEAFTRHRQPRGWEQLEQC
 GYPVQRLVALYLAARLSWNQVDQVIRNALASPGSGGDLGEAIREQPEQARLALTLAA
 ESERFVRQGTGNDEAGAASADVSLTCPVAK DQNRTKGECAGPADSGDALLERNYPTG
 AEFLGDGGDV SFSTRGTQNWTVRLLQHRQLEERGYV FVG YHGT FLEAAQSIVFGG
 RARSQDLDAIWRGFYIAGDPALAYGYAQDQEPDARGRIRNGALLRVYVPRWSLPGFYR
 TGLTLAAPEAAGEVERLIGHPLPLRLDAITGPEEEGGRVTILGWPLAERTVVIPSAIPTDPR
 NVGGDLDPSSIPDKEQAISALPDYASQPGKPPREDLK

5

SEQ. ID NO. 8: ESTRUCTURA ARTIFICIAL QUE CONTIENE A118-S130 PROCEDENTE DEL EJEMPLO 2

MKKTAIAIAVALAGFATVAQADVIIKPQVSGVIVNKLKFKAGDKVKKGQTLFIIEQDQASK
 DFNRSKALFSQLDHTEIKAPFDGTIGDALVNIGDYVSASTTELVRVTNLNPIYADGSHHH
 HHH

SEQ. ID NO. 9: ESTRUCTURA ARTIFICIAL B QUE CONTIENE F122-E138 PROCEDENTE DEL EJEMPLO 2

MKKTAIAIAVALAGFATVAQADVIIKPQVSGVIVNKLKFKAGDKVKKGQTLFIIEQDQFN
 RSKALFSQSAISQKELDHTEIKAPFDGTIGDALVNIGDYVSASTTELVRVTNLNPIYADGS
 HHHHHH

10

SEQ. ID NO. 10: ESTRUCTURA ARTIFICIAL C QUE CONTIENE D121-A127 PROCEDENTE DEL EJEMPLO 2

MKKTAIAIAVALAGFATVAQADVIIKPQVSGVIVNKLKFKAGDKVKKGQTLFIIEQDQDF
 NRSKALDHTEIKAPFDGTIGDALVNIGDYVSASTTELVRVTNLNPIYADGSHHHHHH

15

SEQ. ID NO. 11: SECUENCIA DE AMINOÁCIDOS DE LA PROTEÍNA AcrA SOLUBLE PROCEDENTE DEL EJEMPLO 3

MKYLLPTAAAGLLLLAAQPAMAMHMSKEEAPKIQMPPQVTTMSAKSEDLPLS/TYPA
 KLVSDYDVIKPQVSGVIVNKLKFKAGDKVKKGQTLFIIEQDKFKASVDSAYGQALMAKA
 TFENASKDFNRSKALFSKSAISQKEYDSSLATFNNSKASLASARAQLANARIDLHTEIK
 APFDGTIGDALVNIGDYVSASTTELVRVTNLNPIYADFFISD TDKLNLV RNTQSGKWDLD
 SIHANLNLNGETVQGKLYFIDSVIDANSGTVKAKAVFDNNNSTLLPGAFATITSEGFQK
 NGFKVPQIGVKQDQNDVYVLLVKNKGVKSSVHISYQNNEYAIIDKGLQNGDKIILDNF
 KKIQVGSEVKEIGAQLHHHHHHH

20

SEQ. ID NO. 12: SECUENCIA DE AMINOÁCIDOS DE LA PROTEÍNA CTXB PROCEDENTE DEL EJEMPLO 6

REIVINDICACIONES

1. Un método para producir un bioconjugado de *Shigella* que comprende producir, aislar y purificar una proteína N-glicosilada recombinante, que comprende las etapas de:

a) proporcionar una bacteria gram-negativa recombinante que comprende ácidos nucleicos que codifican

- 5 i) un operón pgl funcional procedente de *Campylobacter* spp., preferiblemente de *C. jejuni*, y
 ii) al menos una proteína diana recombinante, la exotoxina de *Pseudomonas aeruginosa* (EPA), que comprende una o varias de las siguientes secuencias de consenso de aminoácidos N-glicosiladas optimizadas:



10 en donde X y Z pueden ser cualquier aminoácido natural excepto Pro, y en donde se introduce al menos una de dicha o dichas secuencias de consenso de aminoácidos N-glicosiladas optimizadas, y

(iii) la síntesis del polisacárido de *Shigella*, y

b) cultivar el organismo recombinante de una forma adecuada para la producción y la N-glicosilación de la(s) proteína(s) diana,

15 c) eliminar la membrana externa de dicha bacteria gram-negativa recombinante; y

d) recuperar dicha proteína recombinante mediante cromatografía de intercambio aniónico y de exclusión por tamaño.

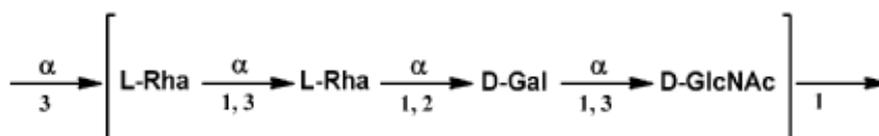
2. El método según la reivindicación 1, en el que dicha EPA modificada tiene una secuencia de aminoácidos que comprende SEQ ID NO: 6.

20 3. El método según la reivindicación 1, en el que dicha EPA modificada tiene una secuencia de aminoácidos que comprende SEQ ID NO: 7.

4. El método según una cualquiera de las reivindicaciones precedentes, en el que dicho polisacárido de *Shigella* se selecciona a partir de *Shigella dysenteriae* O1, *S. flexneri* 2a, *S. flexneri* 3a, *S. flexneri* 3 b, *S. flexneri* 6 y *S. sonnei*.

25 5. El método según una cualquiera de las reivindicaciones precedentes, en el que dicho polisacárido de *Shigella* procede de *Shigella dysenteriae* O1.

6. El método según una cualquiera de las reivindicaciones precedentes, en el que dicho polisacárido de *Shigella* comprende la siguiente estructura:



30 7. El método según una cualquiera de las reivindicaciones precedentes, en el que el organismo comprende una secuencia de nucleótidos que comprende SEQ ID NO: 5.

8. El método según una cualquiera de las reivindicaciones precedentes, en el que el operón pgl funcional procedente de *Campylobacter* spp. codifica una oligosacaril transferasa funcional procedente de *Campylobacter jejuni*.

9. El método según una cualquiera de las reivindicaciones 1-8, en el que dicho organismo es *E. coli*.

35 10. Uso del método según una cualquiera de las reivindicaciones precedentes para la producción de una vacuna de bioconjugado.

FIGURA 1

1 2 3 4

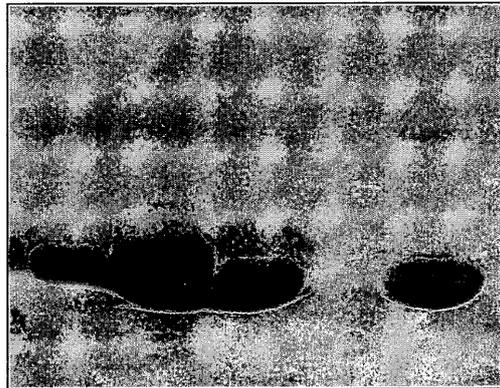


FIGURA 2

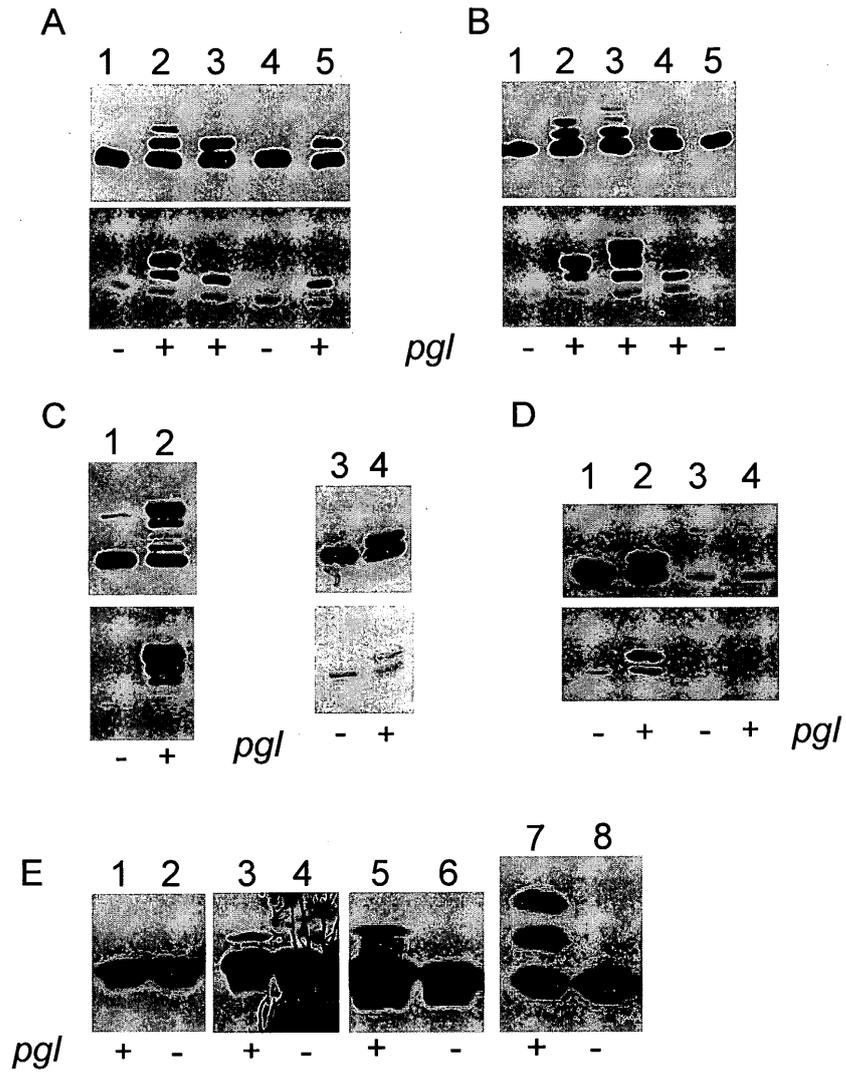


FIGURA 3

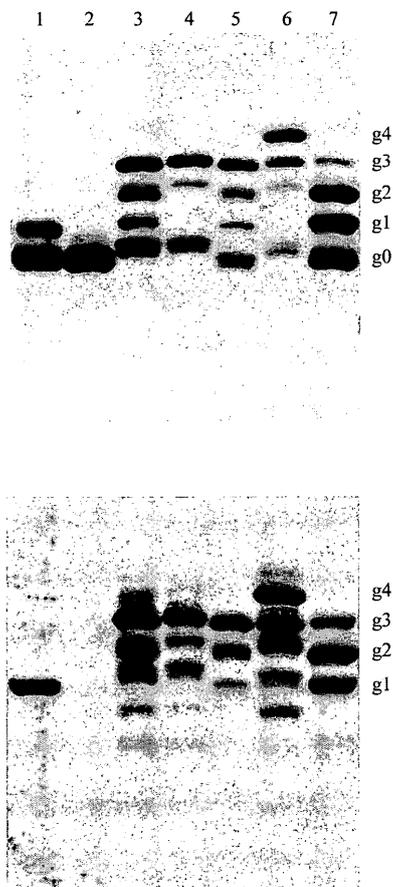


FIGURA 4

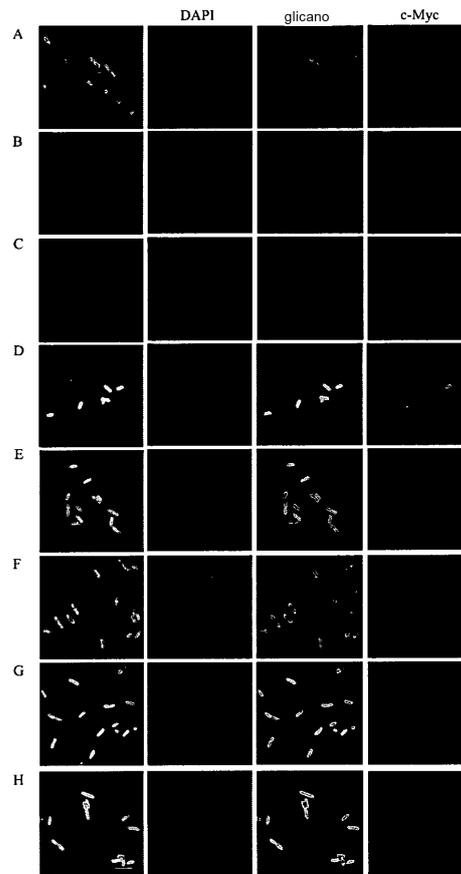


FIGURA 5A

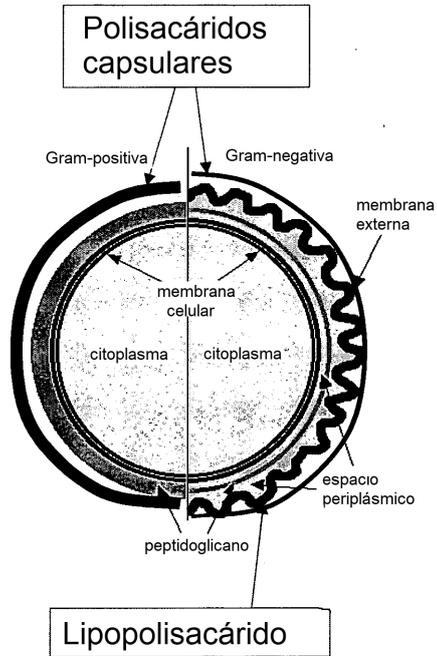
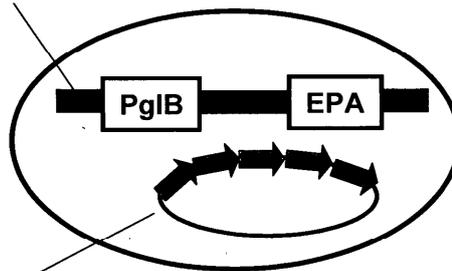


FIGURA 5B

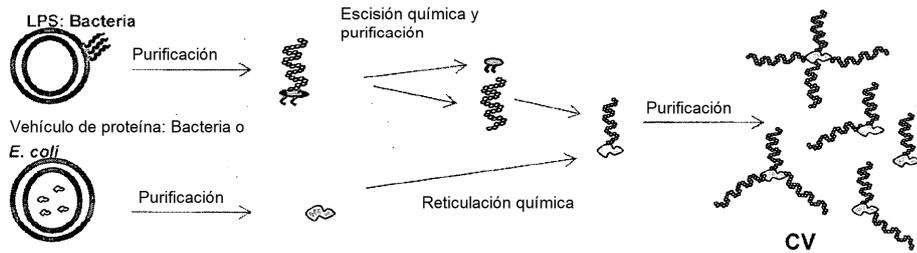
ADN genómico, con PglB y EPA integradas



ADN plasmídico, intercambiable, codifica la agrupación génica de síntesis de polisacáridos

FIGURA 6A

2 rondas de fermentación, diversas etapas de purificación



1 ronda de fermentación, pocas etapas de purificación

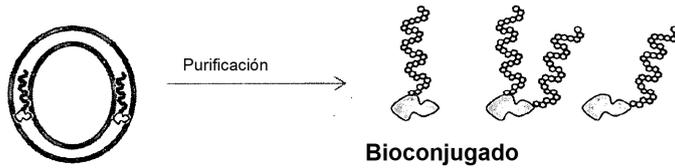


FIGURA 6B

- galF: subunidad de glucosa-1-fosfato uridiltransferasa
- rmIB: dTDP-glucosa 4,6-deshidratasa
- rmID: dTDP-4-deshidroramnosia reductasa
- rmIA: dTDP-glucosa pirofosforilasa
- rmIC: dTDP-4-deshidroramnosia 3,5-epimerasa
- wzx: flipasa de antígeno O
- wzy: polimerasa de antígeno O
- wbbR: ramnosil transferasa II
- wbbQ: ramnosil transferasa I
- rfaA: no necesaria
- wbbP: galactosil transferasa

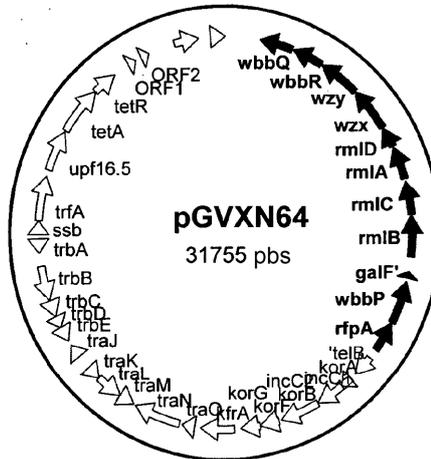


FIGURA 7A

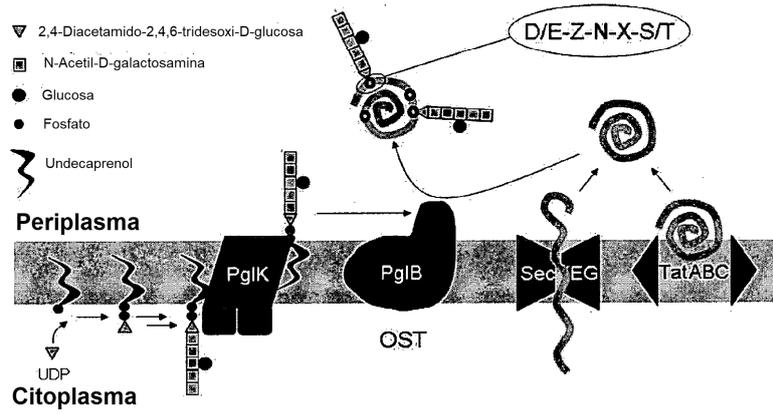


FIGURA 7B

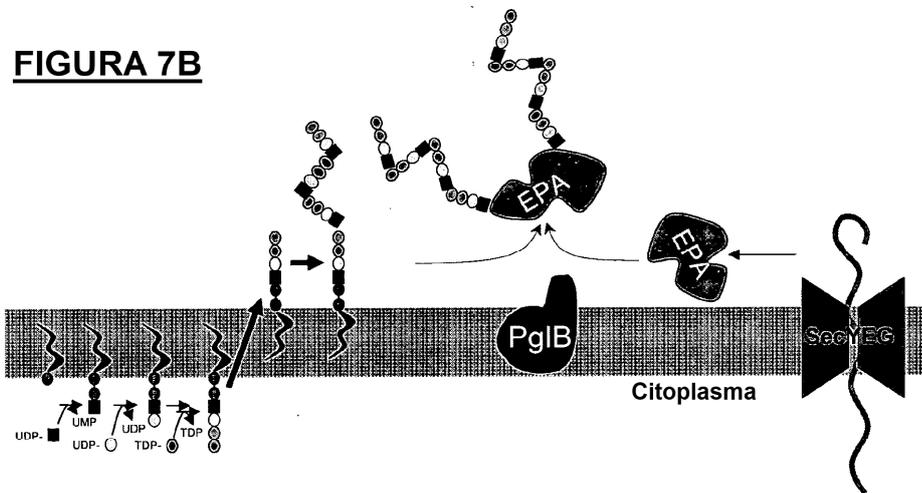


FIGURA 8A

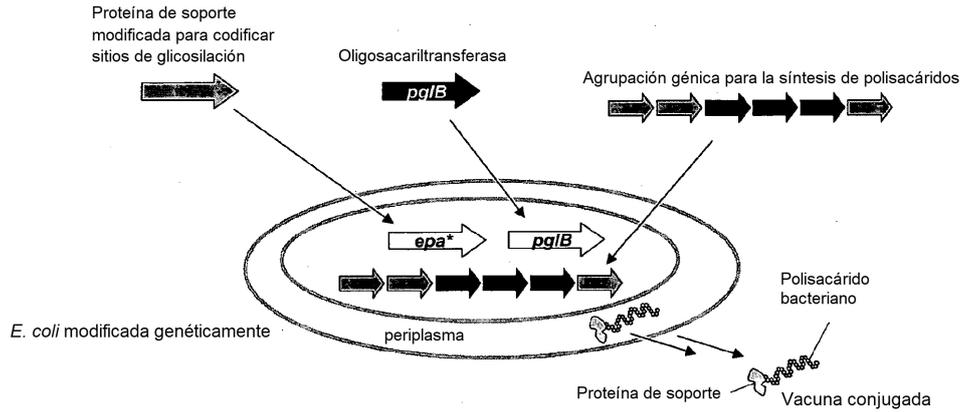


FIGURA 8B

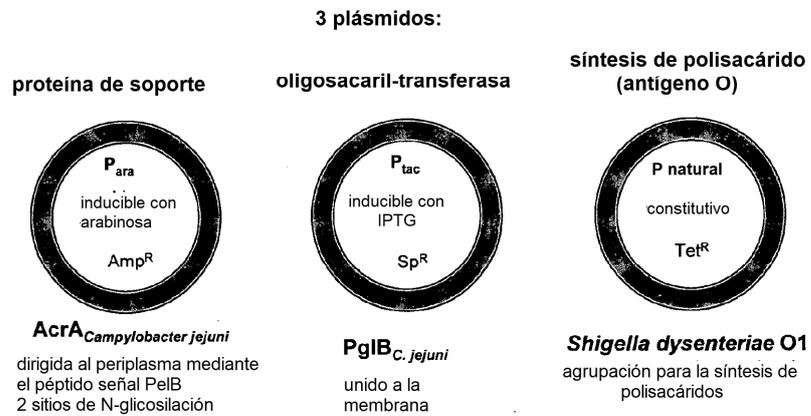


FIGURA 9

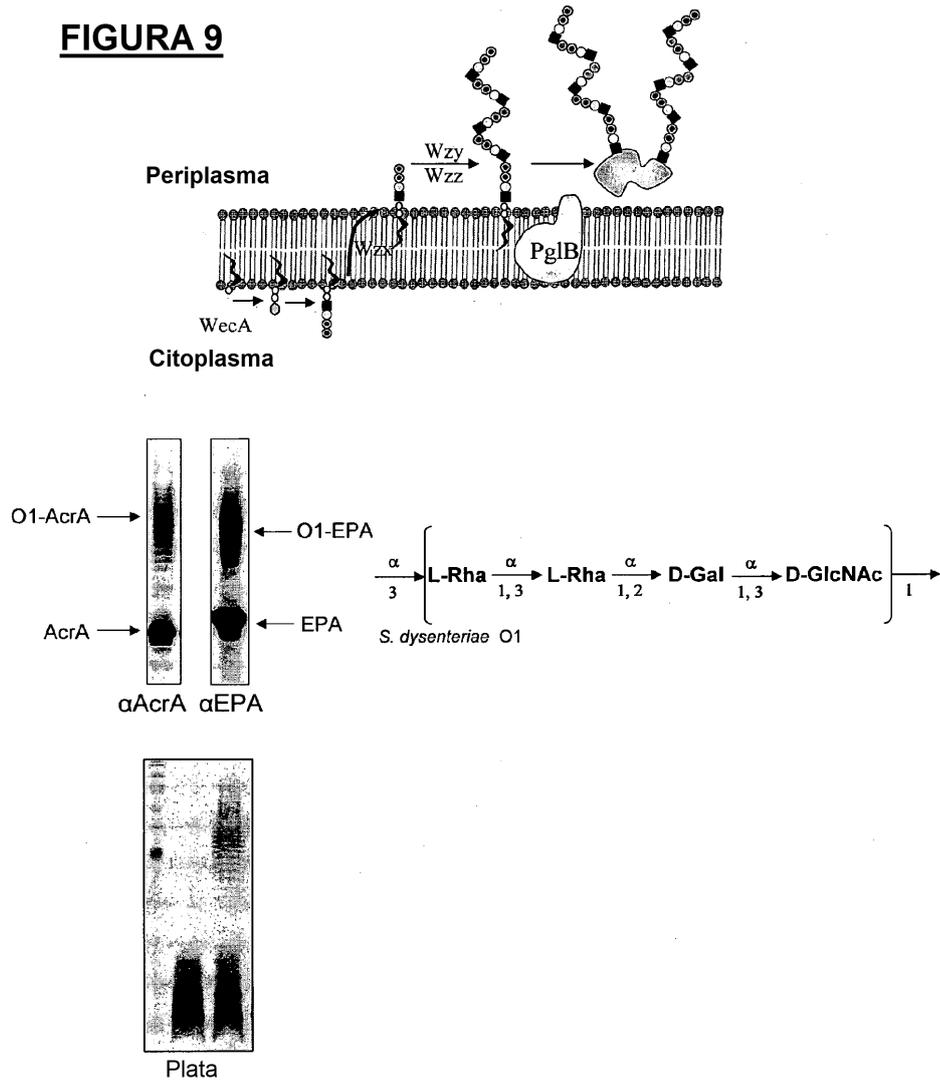


FIGURA 10A

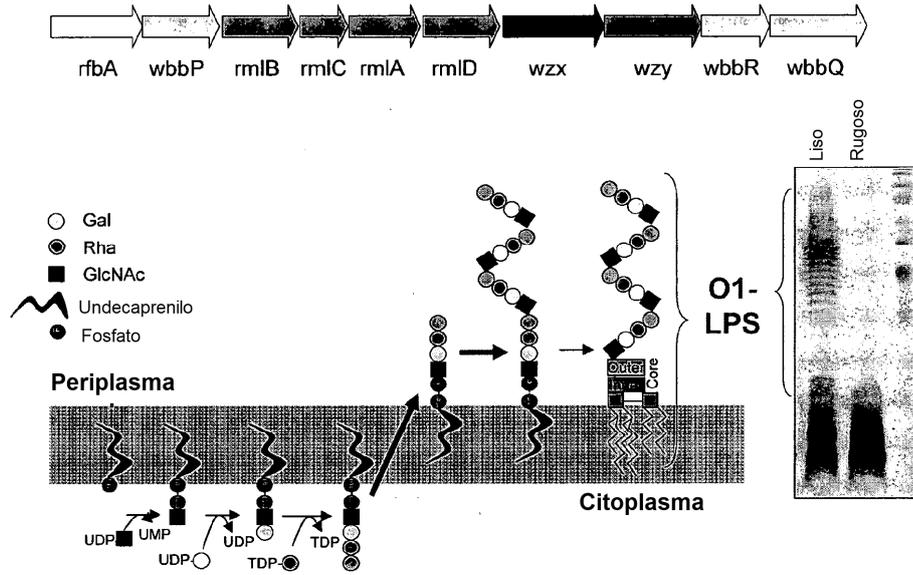


FIGURA 10B

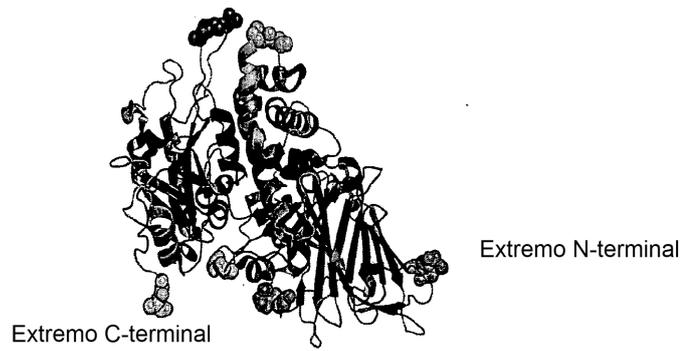


FIGURA 11

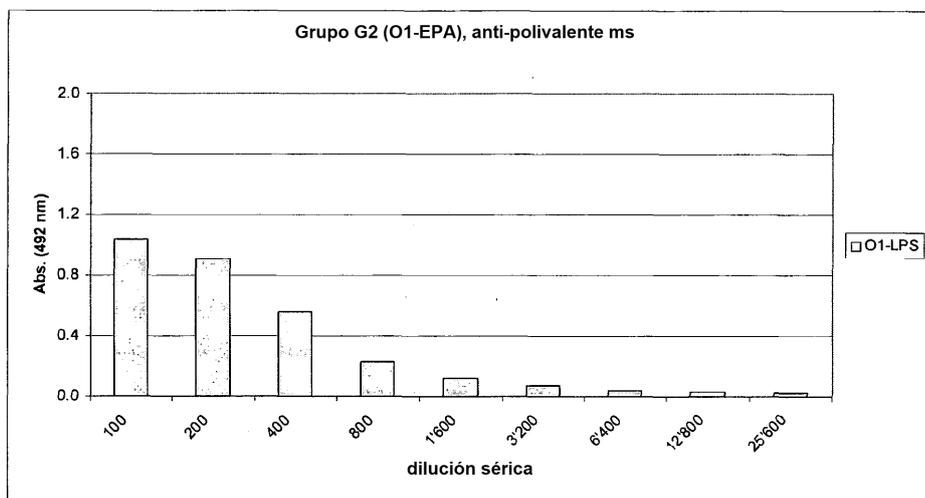
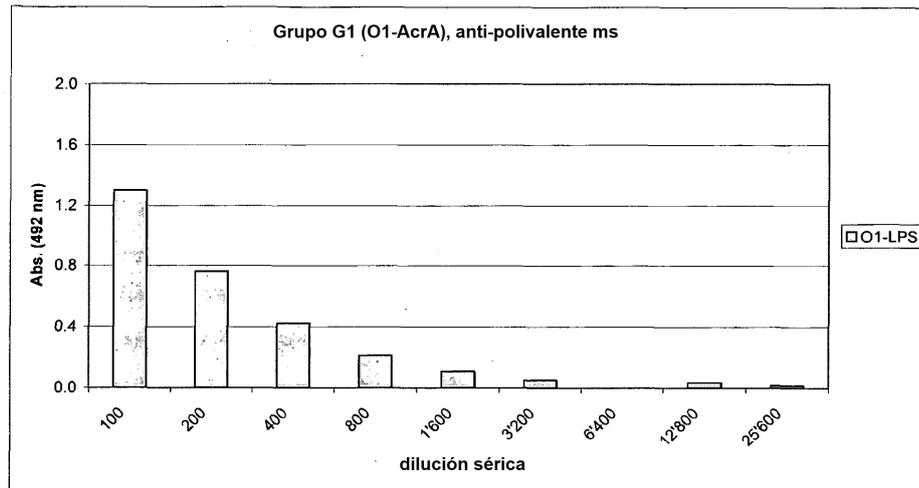


FIGURA 12

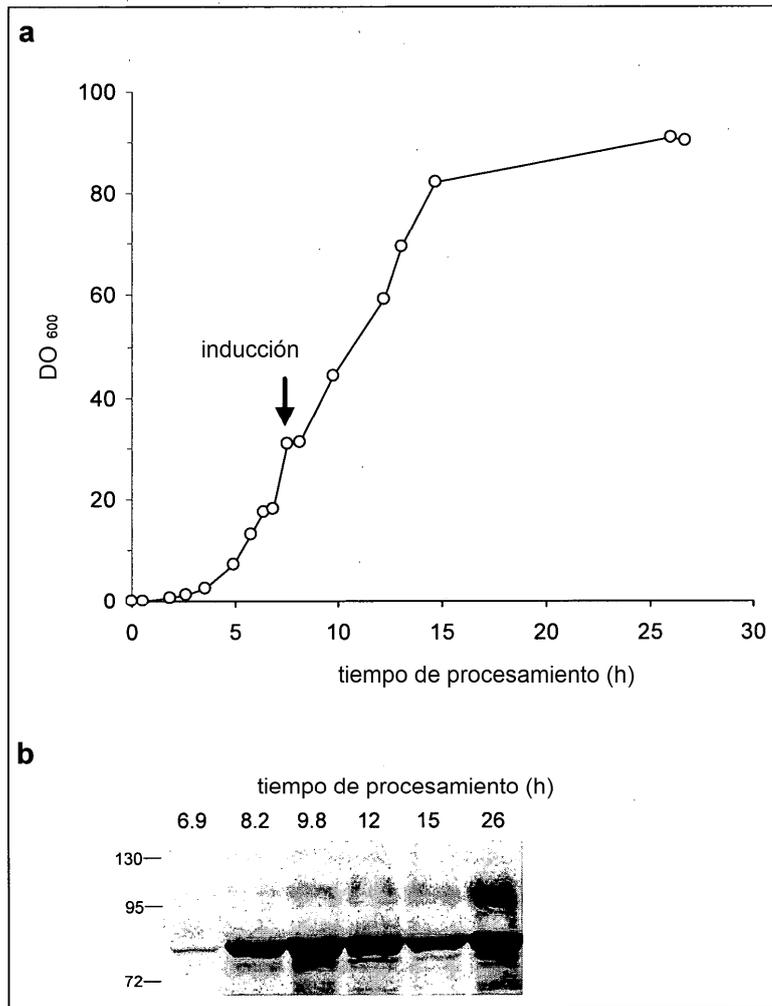
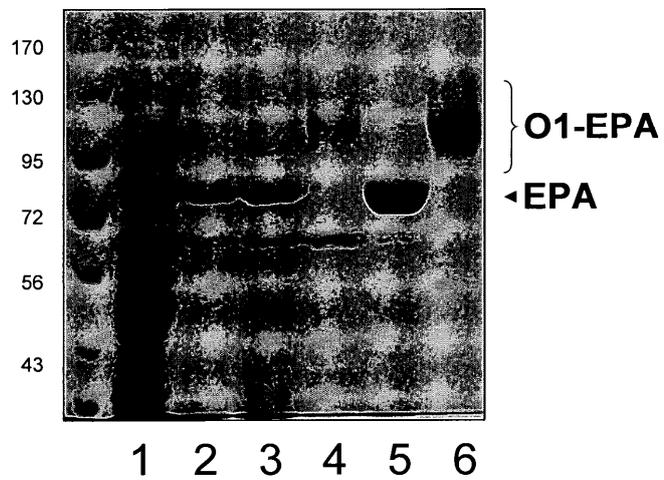


FIGURA 13



1. Células completas
2. Extracto periplásmico
3. Carga IEX (Source Q)
4. Material eluido 1 de IEX
5. Material eluido 2 de IEX
6. Material eluido 1 (Fluoroapatita)

FIGURA 14A

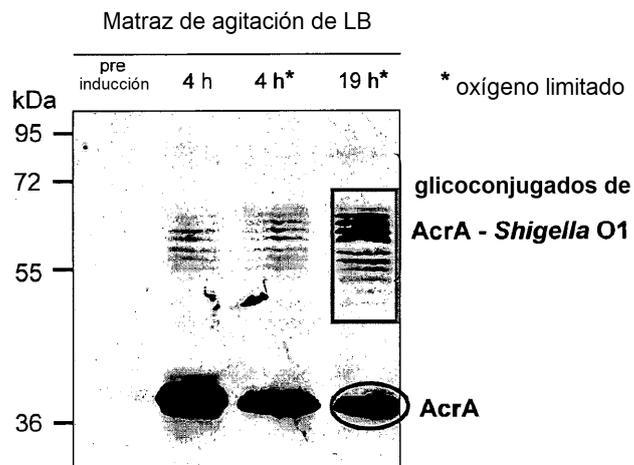


FIGURA 14B

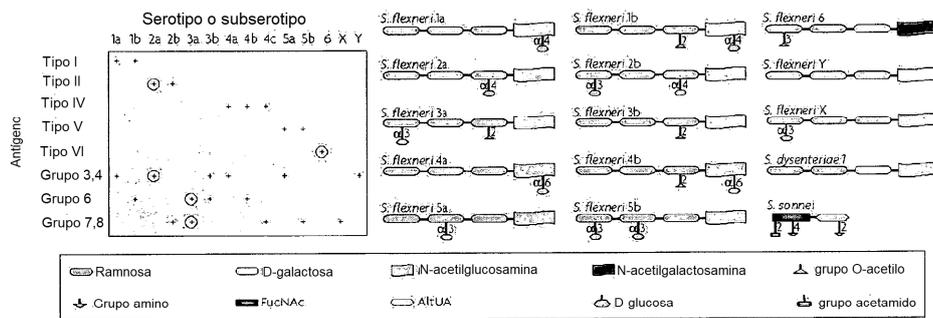


FIGURA 15

Expansión de la región anomérica del espectro ^1H RMN de O1-EPA

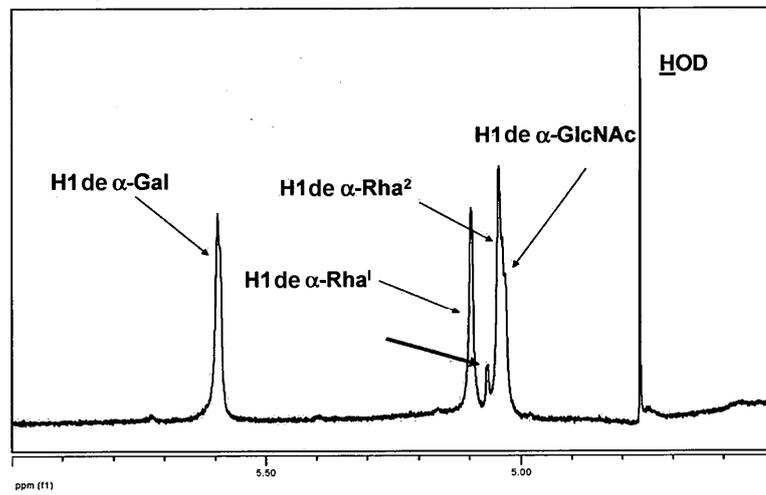


FIGURA 16A

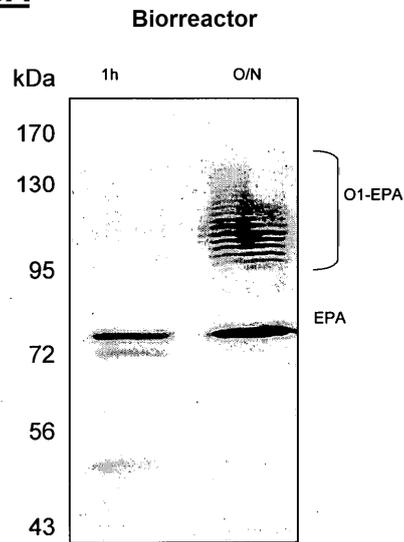


FIGURA 16B

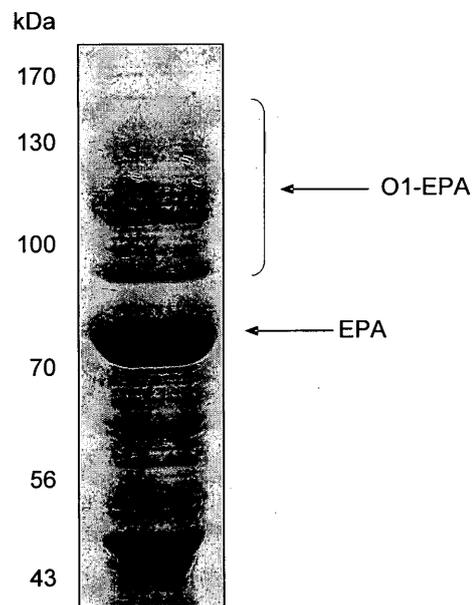


FIGURA 17A

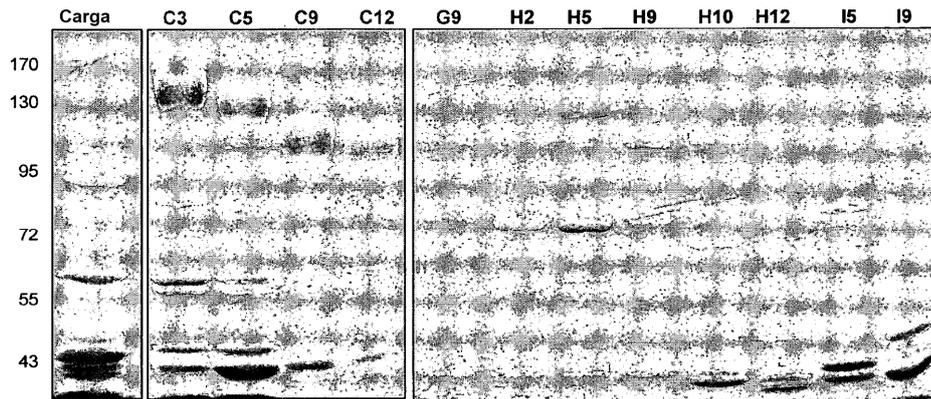


FIGURA 17B

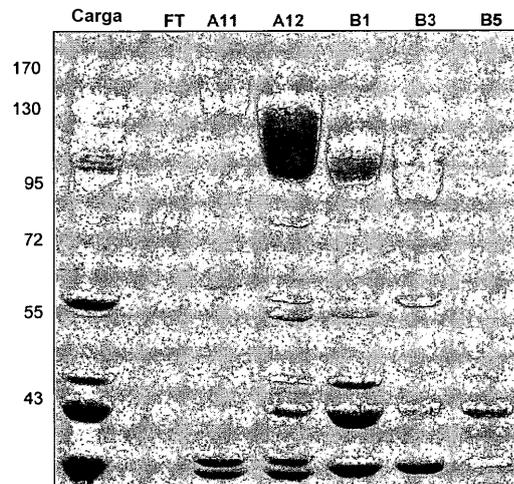


FIGURA 18A

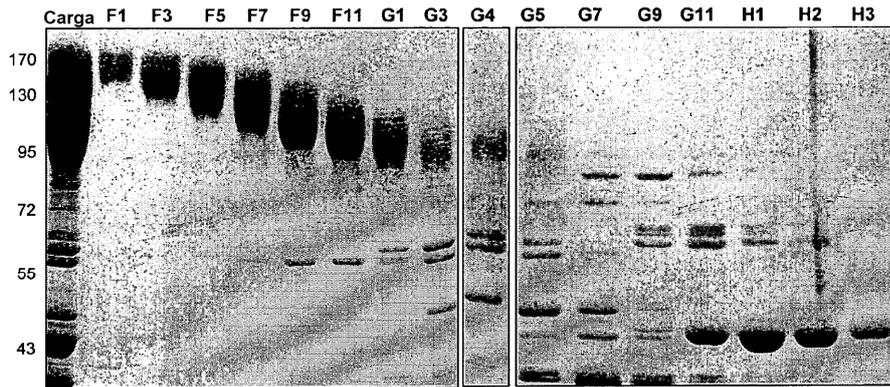


FIGURA 18B

