

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 535 084**

51 Int. Cl.:

C07K 14/25 (2006.01)

C07K 14/205 (2006.01)

C12P 21/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **10.05.2006 E 10179208 (3)**

97 Fecha y número de publicación de la concesión europea: **21.01.2015 EP 2311972**

54 Título: **Proteínas N-glicosiladas recombinantes procedentes de células procariotas**

30 Prioridad:

11.05.2005 EP 05010276

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

05.05.2015

73 Titular/es:

**ETH ZURICH (100.0%)
Raemistrasse 101/ETH Transfer
8092 Zurich , CH**

72 Inventor/es:

**AEBI, MARKUS;
KOWARIK, MICHAEL y
AHUJA, UMESH**

74 Agente/Representante:

ISERN JARA, Jorge

ES 2 535 084 T3

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

DESCRIPCIÓN

Proteínas N-glicosiladas recombinantes procedentes de células procariotas.

La presente invención se refiere al uso de una o varias secuencias de consenso de aminoácidos N-glicosilados optimizados, en la modificación de una proteína de modo que la proteína es capaz de ser N-glicosilada por una oligosacaril transferasa de *Campylobacter* spp.

Antecedentes de la invención

La glicosilación de proteínas ligada a N es un proceso esencial y conservado que tiene lugar en el retículo endoplásmico de organismos eucariotas. Es importante para el plegamiento de las 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).

La glicosilación de proteínas influye extremadamente sobre la antigenicidad, la estabilidad y la semivida de una proteína. Además, la glicosilación puede ayudar a la purificación de proteínas mediante cromatografía, por ejemplo, cromatografía de afinidad con ligandos de lectina unidos a una fase sólida que interaccionan con restos glicosilados de la proteína. Por lo tanto, producir de forma recombinante muchas proteínas glicosiladas en células eucariotas para obtener patrones de glicosilación de utilidad biológica y farmacéutica es una práctica establecida.

Solo en los últimos años se ha demostrado que una bacteria, el agente patógeno *Campylobacter jejuni* transmitido por 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 se agrupan en el locus denominado *pgl*. La 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* mediante una expresión recombinante del locus *pgl* y de 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).

En el documento de solicitud de Patente Europea nº 03 702 276.1 (Patente Europea 1 481 057), una invención anterior de los presentes inventores, se describe un organismo procariota en el que se introduce un ácido nucleico que codifica (i) glicosiltransferasas específicas para el acoplamiento 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, y (iii) una oligosacaril transferasa de *C. jejuni* (OTasa) que une covalentemente dicho oligosacárido con la secuencia de consenso sobre la proteína diana. Dicho organismo procariota produce N-glicanos con una estructura específica que se define por el tipo de las glicosiltransferasas específicas.

Aun cuando la presencia de la secuencia de consenso de N-glicosilación conocida en una proteína permite la N-glicosilación de proteínas diana recombinantes en organismos procariotas que comprenden la oligosacaril transferasa (OTasa) de *C. jejuni*, la N-glicosilación de algunas proteínas diana es frecuentemente ineficaz.

El documento WO 03074687 describe un sistema para la producción de glicoproteínas recombinantes en *E. coli* utilizando la proteína AcrA como una proteína modelo.

Wacker et al., Science, American Association for the Advancement of Science, EE.UU., Vol 298, 29 de noviembre de 2002, págs.1790-1793, describen el sistema de glicosilación ligado a N del locus *pgl* obtenido a partir de la bacteria *Campylobacter jejuni* y transferido funcionalmente a *E. coli*.

Nita-Lazar et al, Glycobiology, vol 15, nº 4, abril de 2005, págs. 361-367, describen los requisitos de la secuencia de polipéptidos del receptor para la N-glicosilación en la N-glicosilación reconstituida de la proteína modelo AcrA en *E. coli*.

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

En un primer aspecto, la presente invención proporciona el uso de una o varias secuencias de aminoácidos que comprenden la secuencia D/E- X- N- Z-S/T, en donde X y Z pueden ser cualquier aminoácido natural excepto prolina, para modificar una proteína de tal manera que la proteína sea capaz de ser N-glicosilada por una oligosacaril transferasa de *Campylobacter* spp.

Se encontró sorprendentemente que la introducción de una(s) secuencia(s) de aminoácidos específica(s) parcial(es) (secuencia(s) de consenso optimizada(s)) en las proteínas conduce a proteínas que se N-glicosilan de manera eficaz con la oligosacaril transferasa (OST, OTasa) de *Campylobacter* spp., preferiblemente *C. jejuni*, en estas posiciones introducidas.

- 5 La expresión "secuencia(s) parcial(es) de aminoácidos" tal como se utiliza en el contexto de la presente invención, también se puede denominar "secuencia(s) de consenso optimizada(s)". La secuencia de consenso optimizada se N-glicosila con la oligosacaril transferasa (OST,_[MW1] OTasa) de *Campylobacter* spp., preferiblemente *C. jejuni*, mucho más eficazmente que la secuencia de consenso normal "N - X - S/T", conocida en la técnica anterior.

10 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, p. ej., *Escherichia* spp., *Campylobacter* spp., *Salmonella* spp., *Shigella* spp., *Helicobacter* spp., *Pseudomonas* spp., *Bacillus* spp., 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 por la OTasa de *Campylobacter* spp., preferiblemente *C. jejuni*, siendo dicha enzima transferasa de origen natural o habiéndose introducido de forma recombinante en dicha célula hospedadora.

20 De acuerdo con el código de una letra aceptado internacionalmente para los aminoácidos, las abreviaturas D, E, N, S y T significan 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 N-glicosilan. Por lo tanto, las proteínas de la presente invención difieren de las proteínas de origen natural de *C. jejuni* que también contienen la secuencia de consenso optimizada pero no comprenden ninguna secuencia de consenso optimizada adicional (introducida).

25 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 lograr mediante estrategias de síntesis química bien conocidas por los expertos en la técnica, tales como la síntesis química de péptidos facilitada por una fase sólida. Alternativamente y de forma preferida para polipéptidos más largos, las proteínas de la presente invención se pueden preparar mediante técnicas recombinantes convencionales.

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

35 La expresión "operón *pgl* funcional procedente de *Campylobacter* spp., preferiblemente de *C. jejuni*" se entiende que se refiere a la agrupación de ácidos nucleicos que codifican la oligosacaril transferasa funcional (OTasa) de *Campylobacter* spp., preferiblemente de *C. jejuni*, y una o varias glicosil transferasas más 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, mediante la OTasa. Se debe entender que la expresión "operón *pgl* funcional de *Campylobacter* spp., preferiblemente de *C. jejuni*" en el contexto de esta invención no se refiere necesariamente a un operón como una unidad transcripcional singular. La expresión meramente 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 pueden estar regulados juntos o por separado. Por ejemplo, la expresión también incluye componentes funcionales situados en el ADN genómico y plásmido(s) en una célula hospedadora. Para fines de eficacia, se prefiere que todos los componentes del operón *pgl* funcional estén regulados y se expresen simultáneamente.

40 Es importante saber que solo la oligosacaril transferasa funcional (OTasa) debe proceder de *Campylobacter* spp., preferiblemente de *C. jejuni*, y que la 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, siendo la única limitación funcional que el oligosacárido ensamblado por dicha 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 naturales y/o glicosiltransferasas específicas incapacitantes naturales en dicho hospedador, así como la introducción de glicosiltransferasas específicas heterólogas, permitirá que los

expertos en la técnica varíen los N-glicanos unidos al sitio de consenso de N-glicosilación optimizado en las proteínas de la presente invención.

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

En una realización preferida, las proteínas modificadas de acuerdo con la presente invención 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 N-glicosiladas optimizadas.

10 La presencia de una o varias secuencias de aminoácidos N-glicosiladas optimizadas en las proteínas puede ser una ventaja para aumentar su antigenicidad, incrementar su estabilidad, afectar a su actividad biológica, prolongar su semivida biológica y/o simplificar su purificación.

15 La secuencia de consenso optimizada puede incluir cualquier aminoácido excepto prolina en la(s) posición(es) X y Z. La expresión "cualquier aminoácido" se entiende que incluye aminoácidos naturales comunes y raros, así como derivados sintéticos y análogos de aminoácidos que seguirán permitiendo que la secuencia de consenso optimizada se N-glicosile con la OTasa. Aminoácidos de origen natural, comunes y raros son los preferidos para X y Z. X y Z pueden ser iguales o diferentes.

Cabe señalar que X y Z pueden diferir en una proteína para cada secuencia de consenso optimizada de acuerdo con la presente invención.

20 El N-glicano unido a la secuencia de consenso optimizada será determinado por las glicosiltransferasas específicas y su interacción cuando se ensambla el oligosacárido sobre un vehículo lipídico para la transferencia mediante la OTasa. Los expertos en la técnica pueden diseñar el N-glicano variando el(los) tipo(s) y la cantidad de glicosiltransferasas específicas presentes en la célula hospedadora deseada.

25 Los N-glicanos se definen en esta memoria como mono, oligo o polisacáridos de composiciones variables que están ligados a un nitrógeno ϵ -amídico de un residuo de asparagina en una proteína, a través de un enlace N-glicosídico. Preferiblemente, los N-glicanos transferidos por la OTasa están ensamblados sobre un anclaje lipídico de undecaprenol-pirofosfato que está presente en la membrana citoplásmica de las bacterias Gram negativas o positivas. Están involucrados en la síntesis de 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).

30 Preferiblemente, la proteína recombinante comprende uno o varios N-glicanos seleccionados entre el grupo de N-glicanos procedentes de *Campylobacter* spp., preferiblemente de *C. jejuni*, los N-glicanos obtenidos a partir de oligosacáridos y polisacáridos transferidos al polisacárido O que forma el antígeno O en bacterias Gram negativas o polisacáridos capsulares de bacterias Gram positivas, preferiblemente: *P. aeruginosa* O9, O11; *E. coli* O7, O9, O16, O157 y *Shigella dysenteriae* O1 y variantes modificadas genéticamente de los mismos, obtenidas a través de la inserción o delección de glicosiltransferasas y epimerasas que afectan a la estructura del polisacárido.

Más preferiblemente, la proteína recombinante comprende dos o más N-glicanos distintos.

40 Por ejemplo, se pueden preparar N-glicanos diferentes sobre la misma proteína mediante el control de la regulación temporal de la expresión de glicosiltransferasas específicas, usando promotores tempranos o tardíos, o la introducción de factores para iniciar, silenciar, potenciar y/o reducir la actividad promotora de glicosiltransferasas específicas individuales. Los promotores y factores adecuados que regulan su actividad están a disposición de los expertos en la técnica rutinaria y no se explicarán adicionalmente.

45 No hay ninguna limitación para el origen de la proteína recombinante de la invención. 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 mamíferos, más preferiblemente seres humanos. Para la preparación de las proteínas recombinantes antigénicas de acuerdo con la invención, preferiblemente para un uso como componentes activos en vacunas, se prefiere que la proteína recombinante proceda de una proteína bacteriana, vírica o fúngica.

50 En una realización preferida adicional, la proteína modificada y/o el(los) N-glicano(s) son terapéutica y/o profilácticamente activos. 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, es la proteína y/o el(los) N-glicano(s) los que son inmunogénicamente activos. En este caso, la(las) N-glicosilación(es) introducida(s) puede(n) tener un efecto modificador sobre la actividad biológica de las proteínas y/o introducir nuevos sitios antigénicos y/o puede(n) enmascarar la proteína para evitar etapas de degradación y/o

55 aumentar la semivida.

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

En una realización preferida, la proteína recombinante es una en la que dicha secuencia polipeptídica dirigida se selecciona a partir del grupo que consiste en péptidos señal de tipo II (Paetzel, 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 Wernerus, H. y Stahl, S. 2004. Biotechnological applications for surface-engineered bacteria. Biotechnol Appl Biochem 40: 209-228.^[mk2]), seleccionadas preferiblemente a partir del grupo que consiste en la proteína de longitud completa o los péptidos señal OmpH1 de *C. jejuni*, JlpA de *C. jejuni*, proteínas de la membrana externa de *E. coli*, preferentemente OmpS, OmpC, OmpA, OprF, PhoE, LamB, Lpp/OmpA (una proteína de fusión para tecnología de presentación en superficie; véase Francisco, J.A., 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 de *Pseudomonas aeruginosa*.

Los ácidos nucleicos que codifican una proteína recombinante también se describen en el presente documento. Preferiblemente, dicho ácido nucleico es ARNm, ADN o APN, 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 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 invención" se refiere a un ácido nucleico que comprende dicha secuencia codificadora 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 invención en una célula hospedadora que contiene un operón *pgl* funcional de *Campylobacter* spp., preferiblemente de *C. jejuni*. Más preferiblemente, los ácidos nucleicos aislados y purificados están ligados funcionalmente a un promotor, preferiblemente ligados a un promotor seleccionado entre el grupo que consiste en promotores procariotas inducibles y constitutivos conocidos, más preferiblemente el promotor de tetraciclina, el promotor de arabinosa, el promotor de salicilato, los promotores *lac*-, *trc*- y *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, por ejemplo, para la vacunación.

Además, las células hospedadoras que comprenden el ácido nucleico y/o un vector también se describen en el presente documento.

El tipo de célula hospedadora no es limitativo, siempre que se adapte a un operón *pgl* funcional procedente de *C. jejuni* y a uno o varios ácidos nucleicos que codifican la(s) proteína(s) diana recombinante(s).

Las células hospedadoras preferidas son células hospedadoras procariotas, más preferiblemente bacterias, lo más preferiblemente las seleccionadas entre el grupo que consiste en *Escherichia coli* ssp., *Campylobacter* ssp., *Salmonella* ssp., *Shigella* ssp., *Helicobacter* ssp., *Pseudomonas* ssp., *Bacillus* ssp., preferiblemente *Escherichia coli*, más preferentemente las cepas de *E. coli* Top10, 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. U S A 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 Δ *waal*.

En una realización más preferida, la célula hospedadora es una que es útil para dirigir proteínas recombinantes de acuerdo con la invención, preferiblemente una, a la membrana externa y/o para la presentación en la superficie, 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 el acoplamiento de un oligosacárido sobre un vehículo lipídico,

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

c) al menos una proteína recombinante de acuerdo con la invención, preferiblemente una proteína que comprende adicionalmente un polipéptido dirigido, y

ii) un fenotipo que comprende una proteína N-glicosilada recombinante de acuerdo con la invención que se encuentra en la membrana externa y/o sobre la membrana externa de la bacteria Gram negativa.

5 La célula hospedadora para la realización anterior se selecciona preferiblemente a partir del grupo que consiste en *Escherichia* ssp., *Campylobacter* spp., *Shigella* spp., *Helicobacter* ssp. y *Pseudomonas* spp., *Salmonella*, ssp., preferiblemente *E. coli*, más preferiblemente las cepas de *E. coli* Top10, W3110, CLM24, BL21, SCM6 y SCM7 y las cepas de *S. enterica* SL3261, SL3749 y SL3261 Δ waaL. (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).^[mk4]

10 Debido a que las proteínas preferidas descritas en este documento 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 limitado y, por lo tanto, proteínas tales como EPO, IFN-alfa, TNFalfa, IgG, IgM, IgA, interleucinas, citocinas, proteínas víricas y bacterianas para la vacunación, tales como proteínas de *C. jejuni* tales como HisJ (Cj0734c), AcrA (Cj0367c), OmpH1 (Cj0982c), toxina diftérica (CRM197), toxina colérica, exoproteína de *P. aeruginosa*, por nombrar solo unas pocas, y en las que se ha introducido la secuencia de consenso N-glicosilada optimizada, son útiles para la preparación de un medicamento (Wyszynska, A., Raczko, A., Lis, M. y Jagusztyn-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 descritos en este documento también son útiles para la preparación de un medicamento, preferiblemente para uso en terapia génica.

25 Por otra parte, una célula hospedadora descrita en la presente memoria, preferiblemente una que tiene un fenotipo que comprende una proteína recombinante N-glicosilada tal y como se describe en el presente documento, que está ubicada en la membrana externa 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.

30 Más preferiblemente, una proteína descrita en este documento se utiliza para la preparación de un medicamento para la vacunación terapéutica y/o profiláctica de un sujeto que lo requiera.

En una realización más preferida, el ácido nucleico y/o un vector descrito en la presente memoria se utiliza en la preparación de un medicamento para la vacunación terapéutica y/o profiláctica de un sujeto que lo requiera, preferiblemente mediante terapia génica.

35 Las células hospedadoras que presentan dichas proteínas recombinantes N-glicosiladas son particularmente útiles para la preparación de vacunas, porque las proteínas N-glicosiladas presentadas están profusamente presentes en la superficie de la célula hospedadora y son muy accesibles para las células inmunes, en particular, sus N-glicanos hidrófilos, y porque las células hospedadoras tienen el efecto añadido de un adyuvante que, si están vivas, incluso se pueden replicar en cierta medida y amplificar sus efectos de vacunación.

40 Preferiblemente, la célula hospedadora para la puesta en práctica de los aspectos médicos descritos en el presente documento, es una célula hospedadora atenuada o muerta.

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

45 Preferiblemente, dichas células hospedadoras se utilizan para la inducción de anticuerpos IgA en un animal, preferiblemente un mamífero, roedor, ovino, equino, canino, bovino o humano. Se prefiere que dicho sujeto que requiere la vacunación sea un ave, un mamífero o un pez, preferiblemente un mamífero, más preferiblemente un mamífero seleccionado entre el grupo que consiste en ganado bovino, ovino, equino, perros, gatos y seres humanos, lo más preferiblemente seres humanos. También se prefieren las aves.

50 Se dan a conocer composiciones farmacéuticas que comprenden al menos una proteína, al menos un ácido nucleico, al menos un vector y/o al menos una célula hospedadora, tal y como se describen en el presente documento. 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, son bien conocidas 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.

55 Preferiblemente, la composición farmacéutica comprende un excipiente, un diluyente y/o un adyuvante farmacéuticamente aceptable.

Por lo tanto, se da a conocer 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/codifica/expresa una proteína recombinante de acuerdo con la presente invención, y (ii) un excipiente, un diluyente y/o un adyuvante farmacéuticamente aceptable.

- 5 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 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 la afección a tratar, el estadio de la enfermedad o la afección y otras circunstancias relevantes (Remington's Pharmaceutical Sciences, Mack Publishing Co. 1990)). La proporción y la naturaleza del diluyente o del excipiente farmacéuticamente aceptable vienen determinadas por la solubilidad y las propiedades químicas del compuesto farmacéuticamente activo seleccionado, la ruta de administración elegida y la práctica farmacéutica 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.
- 10 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 ácidos o sales de adición de bases, para fines de estabilidad, conveniencia de cristalización, aumento de la solubilidad y similares.

En este documento también se describen métodos para producir proteínas glicosiladas ligadas a N, que comprenden las etapas de:

- a) proporcionar un organismo recombinante, preferiblemente un organismo procariota que comprende ácidos nucleicos que codifican
- i) un operón *pgl* funcional procedente de *Campylobacter* spp., preferiblemente de *C. jejuni*, y
- ii) al menos una proteína diana recombinante que comprende una o varias de las siguientes secuencias de consenso de aminoácidos N-glicosilados optimizados:
- 25 D/E - X - N - Z - S/T,
- en donde X y Z pueden ser cualquier aminoácido natural ácido excepto Pro, y en donde se introduce al menos una de dichas secuencias de consenso de aminoácidos N-glicosilados optimizados, y
- b) cultivar el organismo recombinante en una forma adecuada para la producción y la N-glicosilación de la(s) proteína(s) diana.
- 30

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

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

- i) OTasa recombinante procedente de *Campylobacter* spp., preferiblemente de *C. jejuni*, y
- 35 ii) glicosiltransferasas específicas recombinantes y/o naturales procedentes de *Campylobacter* spp., preferiblemente de *C. jejuni*, y/o
- iii) glicosiltransferasas específicas recombinantes y/o naturales procedentes de especies distintas de *Campylobacter* spp.,

40 para el acoplamiento de un oligosacárido sobre un vehículo lipídico para ser transferido a la proteína diana a través de la OTasa.

Por otra parte, un método preferido comprende las etapas de:

- i) proporcionar una bacteria Gram negativa,
- ii) introducir en dicha bacteria al menos una secuencia nucleotídica que codifica
- 45 a) al menos una glicosiltransferasa específica recombinante para el acoplamiento de un oligosacárido sobre un vehículo lipídico, y/o
- b) al menos una oligosacárid transferasa (OTasa) recombinante procedente de *Campylobacter* spp., preferiblemente de *C. jejuni*, y/o
- c) al menos una proteína recombinante que comprende una o varias de las siguientes secuencias de consenso de aminoácidos N-glicosilados optimizados:

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

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 N-glicosilados optimizados, y

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

Para la puesta en práctica de los métodos preferidos anteriores, el organismo procarionte recombinante o la célula hospedadora se selecciona preferentemente a partir del grupo de bacterias que consisten en *Escherichia* spp., *Campylobacter* spp., *Salmonella* spp., *Shigella* spp., *Helicobacter* spp., *Pseudomonas* spp., *Bacillus* spp., 10 preferiblemente *Escherichia coli*, preferentemente las cepas de *E. coli* Top10, W3110, CLM24, BL21, SCM6 y SCM7 y las cepas de *S. enterica* SL3261, SL3749 y SL3261Δ*waalL*.

Otro método preferido para producir, aislar y/o purificar una proteína recombinante tal y como se describe en el presente documento comprende las etapas de:

- 15 a) cultivar una célula hospedadora según las reivindicaciones 15 o 16,
 b) eliminar la membrana externa de dicha bacteria Gram negativa recombinante y
 c) recuperar dicha proteína recombinante.

Los métodos ejemplares para la eliminación de la membrana externa de una célula, preferiblemente una célula procarionte, más preferiblemente una célula bacteriana Gram negativa, son métodos de tratamiento enzimático adecuados, solubilización con detergente y choque osmótico y método de prensa francesa.

- 20 Lo más preferiblemente se emplea un método, en el que se seleccionan glicotransferasas específicas recombinantes o naturales procedentes de especies distintas de *Campylobacter* spp., preferiblemente de *C. jejuni*, a partir del grupo de glicotransferasas y epimerasas procedentes de bacterias, arqueobacterias y/o eucariotas que se pueden expresar funcionalmente en dicha célula hospedadora.

Figuras

- 25 La Fig. 1 ilustra la N-glicosilación de proteínas Lip obtenidas a partir de las estructuras artificiales A a C (véase el ejemplo 1). Células de *E. coli* Top 10 que 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 el SDS-PAGE y la tinción con azul brillante de Coomassie de las 30 fracciones proteicas purificadas.

- La Fig. 2 muestra el análisis de la N-glicosilación de las proteínas diferentes en las que se analizó 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. U S A 102, 3016-3021) o células Top10 (carriles 1-6 del panel E) o células SCM7 (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) (carriles 7, 8 del panel E) 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 35 antisueros específicos. En los paneles A-D, los paneles superiores muestran inmunotransferencias investigadas 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 investigadas 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 de tipo silvestre soluble con la secuencia señal pelB y el marcador hexa His (carriles 1, 2), AcrA-N273Q (carriles 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-F115D-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 sitios múltiples de glicosilación en OmpH1. La cepa Δ*waalL* de SCM6 se cotransformó con el plásmido pACYC*pgl* (que codificaba el locus *pgl* completo) y plásmidos que expresaban 50 OmpH1 de tipo silvestre (carril 1), OmpH1^{N139S}-myc (carril 2), OmpH1^{KGN → NIT, HFGDD → DSNIT}-myc (carril 3), OmpH1^{RGD → NIT, HFGDD-DSNIT}-myc (carril 4), OmpH1^{KGN → NIT, RGD → NIT}-myc (carril 5), OmpH1^{KGN → NIT, RGD → NIT, HFGDD → DSNIT}-myc (carril 6) o OmpH1^{RGD → NIT, V83T}-myc (carril 7). Las células se cultivaron en condiciones aerobias, inducidas con arabinosa al 0,5% durante 3 horas antes del análisis. Los lisados de células completas se precipitaron con TCA después de igualar la densidad óptica de los cultivos, tal y como se describe en la sección de materiales y métodos. 55

Las proteínas se separaron mediante SDS-PAGE al 15% y se transfirieron a una membrana de PVDF. *Primer panel*, inmunotransferencia de lisados de células completas examinadas con anticuerpos anti-marcador myc. *Panel inferior*, inmunotransferencia de lisados de células completas examinadas con antisuero específico de glicano. Las posiciones de OmpH1 sin glicosilar y glicosilada se indican a la derecha.

5 La Fig. 4. Microscopía de fluorescencia de células que expresaban diversas variantes de OmpH1. Los cultivos de cepas de *E. coli* CLM24 o SCM6 que contenían el plásmido de expresión para OmpH1 de tipo silvestre y sus variantes, se igualaron a una DO_{600} de 0,25/ml. Las células se lavaron dos veces con solución salina tamponada con fosfato (PBS), pH 7,4 y 100 μ l de las suspensiones celulares se pipetearon sobre portaobjetos de vidrio gelatinizados y se incubaron a temperatura ambiente (TA) durante 30 min dentro de una cámara humidificada. Todas las etapas
10 posteriores de la marcación con inmunofluorescencia de células completas se realizaron a temperatura ambiente dentro de una cámara humidificada. Las células no unidas se retiraron y el resto se fijó con paraformaldehído al 4% que contenía PBS, durante 30 min a TA. Es importante destacar que el paraformaldehído no se considera que permeabilice las células, sino que conserva intacta la compartimentación mediante membranas. Las células fijadas se lavaron dos veces con PBS y se resuspendieron en tampón de bloqueo que contenía 5% de BSA en PBS.
15 Después del bloqueo, las células se incubaron con IgG monoclonal de ratón anti-myc (1:50, Calbiochem) y/o antisuero anti-glicano (1:4000) durante 1 h en 100 μ l de PBS que contenía 5% de BSA. Las células se lavaron tres veces con 100 μ l 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 μ l de PBS que contenía 5% de BSA. Si era necesario, se añadió 4,6-diamino-2-fenilindol (DAPI) (Sigma) (0,5 μ g/ml) en el momento de la incubación con el anticuerpo secundario para teñir el ADN bacteriano. El anticuerpo secundario se eliminó de las células por lavado con PBS, y se montaron cubreobjetos sobre los portaobjetos usando medio de montaje 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 utilizando Adobe Photoshop, versión CS2. Las células SCM6 expresaban
20 OmpH1 (panel A), OmpH1^{N139S} (panel B), OmpH1^{C20S} (panel C), OmpH1^{KGN \rightarrow NIT, HFGDD \rightarrow DSNIT} (panel D), OmpH1^{RGD \rightarrow NIT, HFGDD-DSNIT} (panel E), OmpH1^{KGN \rightarrow NIT, RGD \rightarrow NIT} (panel F), OmpH1^{V83T, KGN \rightarrow NIT} (panel G) y OmpH1^{KGN \rightarrow NIT, RGD \rightarrow NIT, HFGDD \rightarrow DSNIT} (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 DAPI, columna 3: fluorescencia verde procedente de la fluorescencia específica de glicanos, columna 4: fluorescencia roja procedente de la tinción anti-myc.
30

Los siguientes ejemplos sirven para ilustrar adicionalmente la presente invención y no se pretende de ningún modo que limiten su alcance.

Ejemplos

Selección de AcrA como proteína modelo para optimizar la N-glicosilación

35 Para optimizar los requerimientos de la proteína aceptora para la N-glicosilación, se realizaron estudios detallados sobre la glicoproteína AcrA de *C. jejuni* (Cj0367c). AcrA es una lipoproteína periplásmica de 350 residuos de aminoácidos. Se ha mostrado que la secreción en el periplasma, pero no el anclaje lipídico, 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 PelB heteróloga cuando se expresa en *E. coli*. De los cinco secuencias potenciales para la
40 glicosilación ligada a N (N117, N123, N147, N273, N274) se utilizan dos iguales en *C. jejuni* y *E. coli* (N123 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 que está disponible una información estructural detallada. Recientemente, se ha publicado 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. U S A 101, 9994-9999). Ambas proteínas son miembros de las proteínas denominadas bomba de flujo de salida periplásmico (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 α -hélice, un bucle enrollado de forma antiparalela que se mantiene unido a la base por un dominio de lipoilo, que va seguido de un dominio en barril β de
50 seis cadenas. Los residuos 23-28 en el extremo N-terminal y los residuos 95-101 en el extremo C-terminal no están estructurados en los cristales. Las secuencias proteicas de MexA y AcrA tienen una identidad del 29,3% y son similares en un 50%. Por lo tanto, las dos proteínas probablemente presentan un plegamiento global similar.

Ejemplo 1 Clarificación de la secuencia peptídica primaria que desencadena la glicosilación

55 Se sabe que los dominios de lipoilo similares a los de 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 determinar qué secuencia peptídica aceptora era necesaria para la N-glicosilación por la maquinaria de *pgl* en *E. coli*, se tomó el dominio lipóilico de AcrA. Se usó como un andamio molecular para transportar péptidos de diferentes longitudes al periplasma y presentarlos *in vivo* a la maquinaria de *pgl*.

Por lo tanto, se construyó un plásmido que codificaba el dominio lipófilo (Lip) y se fusionó N-terminalmente 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 C-terminalmente con un marcador hexa His. Se llevó a cabo la clonación para poner la expresión génica bajo el control del promotor de arabinosa. Para los límites del dominio Lip, se escogieron posiciones de aminoácidos que aparecían en las mismas posiciones que las de los límites de dominio de la parte del dominio lipófilo en MexA. Para examinar diferentes péptidos en cuanto a 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 N123 de AcrA de *C. jejuni*. Los marcos de lectura abierta resultantes consistían en las secuencias que codificaban la secuencia señal de OmpA, la parte N-terminal de tipo cabeza de martillo de AcrA (D60-D95 (la numeración de los aminoácidos se refiere a la numeración de la secuencia del polipéptido AcrA maduro)), los diferentes tramos que contenían el sitio de glicosilación N123 natural de AcrA (véase más adelante), la parte C-terminal de tipo cabeza de martillo de AcrA-Lip (L167-D210) y el marcador C-terminal his.

La construcción de los plásmidos se llevó a cabo mediante técnicas convencionales de biología molecular. Se insertaron tres tramos que contenían el sitio de glicosilación N123 natural de AcrA de diferentes longitudes, entre las dos mitades de Lip para dar lugar a tres ORFs diferentes:

La estructura artificial A contiene A118-S130 que da como resultado una secuencia proteica:

MKKTAIAlAVALAGFATVAQADVlIKPQVSGVlVnKlFKAGDKVKKGQTLFllEQDQASKDF
NRSKALFSQLDHTElKAPFDGTlGDALVNlGDYVSASTTELVRVTNLNPIYADGSHHHHH
***H* (secuencia 1).**

La estructura artificial B contiene F122-E138 que da como resultado una secuencia proteica:

MKKTAIAlAVALAGFATVAQADVlIKPQVSGVlVnKlFKAGDKVKKGQTLFllEQDQFNRSK
ALFSQSAISQKELDHTElKAPFDGTlGDALVNlGDYVSASTTELVRVTNLNPIYADGSHHH
***HHH* (secuencia 2).**

La estructura artificial C contiene D121-A127 que da como resultado una secuencia proteica:

MKKTAIAlAVALAGFATVAQADVlIKPQVSGVlVnKlFKAGDKVKKGQTLFllEQDQDFNR
SKALDHTElKAPFDGTlGDALVNlGDYVSASTTELVRVTNLNPIYADGSHHHHHH
(secuencia 3).

Los tramos subrayados de la secuencia indican el péptido señal de OmpA; se introdujeron residuos subrayados individualmente por razones de clonación o para hacer que la proteína fuera resistente a la degradación. Letra negrita: sitio de glicosilación que corresponde a N123 de AcrA. Letra cursiva: marcador hexa-His. Los genes correspondientes se expresaron 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).

Para comprobar cuál de los tres tramos desencadenaba la glicosilación de las proteínas Lip, se llevaron a cabo experimentos de expresión proteica. Se cultivaron células Top10 de *E. coli* (Invitrogen, Carlsbad, California, EE.UU.) que eran portadoras de pACYC*pgl* o pACYC*pglmut* (Wacker et al., 2002, *supra*) y un plásmido que codificaba las estructuras artificiales A, B o C, en medio LB que contenía ampicilina y cloranfenicol, a 37°C, hasta una DO de 0,5. Para la inducción, se añadió una dilución volúmica 1/1000 de una solución de arabinosa al 20% (p/v) y las células se cultivaron durante otras 2 horas. Las células fueron recogidas a continuación por centrifugación, resuspendidas en Tris/HCl 20 mM, pH 8,5, sacarosa al 20% (p/v), EDTA 1 mM, PMSF 1 mM y 1 g/l (p/v) de lisozima, y se incubaron a 4°C durante 1 hora. Se obtuvieron extractos periplásmicos después de recoger los esferoblastos como sedimento de centrifugación y se diluyeron los extractos con una dilución volúmica 1/9 (v/v) de tampón A 10x (NaCl 3 M, Tris/HCl 0,5 M, pH 8,0 e imidazol 0,1 M), y se añadió MgSO₄ hasta 2,5 mM. Se llevó a cabo una purificación por afinidad hacia Ni en columnas de Ni-Sefarosa de Amersham Pharmacia Biotech (Uppsala, Suecia) de 1 ml, en tampón A. Las proteínas se eluyeron en tampón A que contenía imidazol 0,25 M.

En la Figura 1 se muestra el gel SDS-PAGE teñido con azul brillante de Coomassie de las fracciones de elución de picos procedentes de los extractos periplásmicos purificados con Ni. El análisis de la expresión mostraba que la estructura artificial B producía una sola especie proteica destacada (Figura 1, carril 1). Tanto la estructura artificial A como la C conducían, además de a la proteína destacada, a una segunda banda proteica con menor movilidad

electroforética (Figura 1, carriles 2 y 3). El que la especie proteica más pesada estuviera realmente 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, el residuo de aspartato situado 2 posiciones N-terminalmente con respecto al N123 glicosilado. Esto demuestra que D121 desempeña un importante papel en la glicosilación a través de la OTasa. Para verificar que D121 es esencial para la glicosilación, fue mutado a alanina en la estructura artificial C. El análisis de la expresión dio como resultado una sola banda proteica (Figura 1, carril 4), lo que mostraba que D121 es importante para la glicosilación. Además, el hecho de que una proteína artificial de presentación peptídica pueda ser glicosilada muestra que un péptido corto del tipo D/E-X-N-Y-S/T contiene toda la información para que tenga lugar la N-glicosilación ejercida por *C. jejuni*.

10 Ejemplo 2 Verificación del Ejemplo 1; AcrA-D121A no está glicosilada en N123

Para confirmar los hallazgos de la metodología de presentación peptídica, se insertó una mutación de aspartato a alanina en la posición 121 (D121A; es decir, dos residuos antes del N123 glicosilado) en la versión soluble de longitud completa de la proteína AcrA y se analizó si el sitio N123 aún se podía glicosilar en *E. coli*. Con objeto de someter esto a ensayo, se expresó esta AcrA-D121A y se analizó su estado de glicosilación. Para el análisis, se usó una AcrA modificada genéticamente. Difería del gen original de *C. jejuni* en que contenía la secuencia señal PelB (Choi y Lee, 2004, *supra*) para la secreción al periplasma y un marcador C-terminal hexa His para la purificación. Se ha observado que esta variante de AcrA se secreta, es escindida por el péptido señal y se glicosila como la proteína natural anclada a lípido (Nita-Lazar et al., 2005, *supra*). A continuación se muestra la secuencia proteica de la proteína AcrA soluble:

MKYLPTAAAGLLLLAAQPAMAMHMSKEEAPKIQMPPQPVTTMSAKSEDLPLSFTYPAK
LVSDYDVIIKPQVSGVIVNKLKFKAGDKVKKGQTLFIEQDKFKASVDSAYGQALMAKATFE
NASKDFNRSKALFSKSAISQKEYDSSLATFNNSKASLASARAQLANARIDLHTEIKAPF
DGTIGDALVNIGDYVSASTTELVRVTNLNPIYADFFISDTDKLNLVRNTQSGKWDLDSIHA
NLNLNGETVQGKLYFIDSVIDANSQTVKAKAVFDNNNSTLLPGAFATITSEGFIQKNGFK
VPQIGVKQDQNDVYVLLVKNGKVEKSSVHISYQNNEYAIIKGLQNGDKIILDNFKKIQVG
SEVKEIGAQLEHHHHH (secuencia 4)

Los residuos subrayados son el péptido señal PelB, los residuos en letra cursiva el marcador hexa His y los residuos en letra negrita los dos sitios de glicosilación naturales en N123 y N273. Se construyó un plásmido que contenía el ORF para la proteína anterior en el plásmido pEC415 (Schulz et al., 1998), para producir pAcrAper.

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

Los dos primeros carriles de la Figura 2A muestran AcrA en ausencia y presencia de un operón *pgl* funcional. Solo aparece una banda en ausencia del operón *pgl* funcional pero tres en presencia del mismo (Figura 2A, panel superior). Éstas se corresponden a AcrA no glicosilada (carril 1) y a AcrA no glicosilada, monoglicosilada y diglicosilada (carril 2). El que las dos proteínas más pesadas del carril 2 estuvieran glicosiladas fue confirmado mediante transferencia Western con R12 (carril 2, panel inferior). Cuando se expresaba el mutante AcrA-N273Q del mismo modo, solo se detectaba AcrA monoglicosilada en presencia del operón *pgl* funcional para la glicosilación (carril 3). Se detectó AcrA no glicosilada en ausencia del locus *pgl* funcional (carril 4). El análisis del mutante AcrA-D121A solo produjo 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 3 Introducción de sitios de glicosilación artificiales en AcrA

Para someter a ensayo 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 N117 y N147 de AcrA soluble, se había intercambiado por aspartato (F115D, T145D). Luego se examinó si los sitios de glicosilación modificados se podían glicosilar mediante un ensayo igual al 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 mutante doble en el que se habían delecionado ambos sitios de glicosilación utilizados (N123Q y N273Q). Se

prepararon extractos periplásmicos de cultivos inducidos durante 4 horas, se separaron mediante SDS-PAGE y se analizaron por transferencia Western (Fig. 2B). Como controles, se desarrollaron las muestras de AcrA glicosilada y no glicosilada de tipo silvestre en el mismo gel (carriles 1 y 2). La mutación T145D afectaba a la posición -2 del secunón de glicosilación N147-S149 no utilizado naturalmente. Después de la expresión de AcrA-T145D, la transferencia Western con antisuero anti-AcrA dio lugar a cuatro bandas, la mayor de ellas con menor movilidad electroforética que la proteína doblemente glicosilada en el carril 2 (carril 3 en la figura 2B). La transferencia con R12 confirmó que la cuarta banda era una AcrA triplemente glicosilada. A pesar de la baja intensidad con respecto al anti-AcrA, la banda más pesada proporcionó la señal más intensa con el antisuero R12 específico de la glicosilación. Cuando se expresaba el mismo mutante AcrA-T145D en ausencia de la secuencia de N-glicosilación natural (AcrA-N123Q-N273Q-T145D), solo se detectó AcrA monoglicosilada en presencia de un operón *pgl* funcional (Fig. 2B, carril 4), que no aparecía en ausencia de un operón *pgl* funcional (carril 5). Esto demuestra que la banda más pesada en el carril 4 estaba glicosilada. De esta manera, se generó un sitio de glicosilación optimizado (DFNNS) introduciendo simplemente la mutación T145D.

Para confirmar adicionalmente que es posible introducir un sitio de glicosilación insertando un residuo de aspartato en la posición -2, se cambiaron los sitios N117-S119 y N274-T276 no usados naturalmente para optimizar la N-glicosilación. Con este fin, se generaron otros mutantes (Fig. 2C). La expresión de AcrA-F115D-T145D en el sistema descrito anteriormente dio lugar a cinco especies proteínicas detectadas con el antisuero anti-AcrA (carril 2). Esto es indicativo de que tienen lugar cuatro glicosilaciones en la misma molécula de AcrA. Cuando se llevó a cabo la detección con el antisuero R12 específico de N-glicano de *C. jejuni*, se detectó una escalera de cinco bandas. La banda inferior apenas visible es AcrA no glicosilada porque también está presente en ausencia de glicosilación (carril 1); la superior da lugar a una señal intensa debida probablemente a los cinco determinantes antigénicos de una AcrA glicosilada con cuatro plegamientos. De este modo, los dos sitios introducidos (en N117 y N147) y los dos sitios usados naturalmente (N123 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 parcialmente en el mismo sitio de consenso que el eucariótico (N - X - S/T, con X≠P) pero, además, se requiere un aspartato en la posición -2 para aumentar la eficacia. Además, demostraron que es posible glicosilar una proteína en un sitio deseado al introducir mediante recombinación una secuencia de consenso optimizada de este tipo.

Ejemplo 4 Verificación de la posición -1 en la secuencia de N-glicosilación optimizada

Se llevó a cabo un experimento adicional para analizar si la posición -1 del sitio de glicosilación bacteriano presenta 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 modo que se inhibe la glicosilación. Para analizar si también se podría observar un efecto similar en la posición -1, se introdujo un residuo de prolina en esa posición del primer sitio usado naturalmente en un mutante puntual que tenía inoperante el segundo sitio natural (AcrA-N273Q-F122P). La expresión control de AcrA-N273Q mostró una proteína monoglicosilada en presencia de un operón *pgl* funcional (Fig. 2D; carriles 1 y 2). Sin embargo, AcrA-N273Q-F122P no estaba glicosilada (Fig. 2D; carriles 3 y 4). Esto indica que la prolina inhibía la N-glicosilación bacteriana cuando constituye el residuo entre la asparagina y el residuo cargado negativamente de la posición -2.

Alineaciones de secuencia de todos los sitios conocidos por estar 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). De este modo, se estableció que la secuencia de consenso de glicosilación para bacterias se puede optimizar a través de un aminoácido cargado negativamente en la posición -2, para dar lugar a D/E - X - N - Z - S/T, en donde X y Z ≠ P.

Ejemplo 5 N-glicosilación de una proteína que no es de *C. jejuni*

Para demostrar que el requisito primario de la secuencia (secuencia de consenso optimizada) es suficiente para la N-glicosilación en bacterias, se examinó si una proteína que no era de *C. jejuni* se podía glicosilar aplicando la estrategia anterior. Como diana de la glicosilación se empleó la subunidad B de la toxina colérica (CtxB). El gen correspondiente se amplificó a partir de *Vibrio cholerae* de tal modo que contenía la secuencia que codificaba la secuencia señal de OmpA en el extremo N-terminal y un marcador hexa His en el extremo C-terminal, justo igual que 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 detrás de W88 generó un sitio de glicosilación optimizado (DNNKT). Las proteínas CtxB de tipo silvestre y W88D que contenían la secuencia señal y el marcador His se expresaron en *E. coli* Top10 y otros tipos celulares en presencia y ausencia del locus *pgl* funcional de *C. jejuni*. Cuando se analizaron extractos periplásmicos de células

Top10 mediante SDS-PAGE, electrotransferencia e inmunotransferencia consecutiva con un antisuero para CtxB, solo CtxB W88D produjo una banda superior y, por lo tanto, glicosilada en la base del locus *pgl* (Fig. 2E; compárense los carriles 3 y 4). También se insertó una secuencia de consenso (DSNIT) al reemplazar G54 o Q56 de CtxB (la última se denomina CtxB-Q56/DSNIT), es decir, en uno de los bucles de los que se había comunicado que contribuyen a la actividad ligante de CtxB relativa al gangliósido GM1. 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 Top10) producía una banda de menor movilidad y, por lo tanto, glicosilada en células competentes para glicosilación pero no en células carentes de glicosilación, cuando se analizó de un modo igual al descrito anteriormente. También se mostró que una CtxB que contenía dos manipulaciones, es decir, la inserción de D detrás de W88 así como DSNIT reemplazando a Q56, se glicosilaba doblemente en células SCM7 (Alaimo et al., EMBO Journal 25, 967-976 (2006)) (panel E, carriles 7 y 8). La proteína CtxB glicosilada de forma doble mostrada en el carril 7, se purificó por afinidad hacia Ni²⁺ y se analizó por ESI-MS/MS después de una tripsinización en gel de acuerdo con protocolos convencionales. Se detectaron los glicopéptidos esperados, lo que confirma que la N-glicosilación bacteriana también se puede dirigir a una proteína que no sea de *C. jejuni*, al mutar o insertar la secuencia de consenso optimizada de acuerdo con la invención para una N-glicosilación bacteriana (no se muestra). W3110, CLM24, BL21 (Stratagene, La Jolla, California, EE.UU.), SCM6 y SCM7 son ejemplos de otras cepas adecuadas a modo de ejemplo de *E. coli* para poner en práctica la presente invención.

A continuación se indica la secuencia de aminoácidos de la proteína CtxB utilizada en esta memoria (secuencia señal de OmpA recombinante subrayada; marcador hexa His en letra cursiva; W88 en letra negrita).

MKKTAI^{W88}IAVALAGFATVAQATPQNITDLCAEYHNTQIH^{His}TLN^{His}DKIFSYTESLAGKREMAIT
FKNGATFQVEVPGSQHIDSQKKAIERMKD^{W88}TLRIAYL^{His}TEAKVEKLCVWNNKTPHAI^{His}AAISM
ANGSHHHHHH (secuencia 5)

Ejemplo 6 Introducción de sitios de N-glicosilación artificiales en la proteína de la membrana externa de *C. jejuni*, OmpH1

Una posible aplicación de la N-glicosilación en las bacterias es la presentación del glicano en la superficie de una célula hospedadora bacteriana con objeto de enlazar el fenotipo con el genotipo y, de este modo, seleccionar mutaciones genéticas específicas. Para mostrar que los N-glicanos se pueden presentar en proteínas de la membrana externa, la proteína OmpH1 se modificó genéticamente de tal modo que contuviera múltiples sitios de consenso optimizados de acuerdo con la invención. Los sitios se modificaron genéticamente en regiones bucle de la proteína, como se dedujo a partir de 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 mostraron que los mejores secuencias para la glicosilación se generaban mediante las mutaciones V83T, K59N-G60I-N61T, R190N-G191I-D192T y H263D-F264S-G265N-D266I-D267T. Para la presentación superficial, se quiso evaluar diferentes combinaciones de esos sitios introducidos con objeto de establecer la muestra más específica para N-glicanos. Las combinaciones se generaron en una estructura artificial plasmídica que codificaba OmpH1 de tipo silvestre y se examinaron de una manera similar a la descrita para AcrA. En la Figura 3 se muestra el análisis de diversas variantes de OmpH1 que contienen múltiples secuencias con glicosilación además del secución existente de tipo silvestre. Se generaron las variantes de OmpH1 con tres (carriles 3, 4, 5 y 7) y cuatro secuencias de glicosilación (carril 6). También se incluyeron en el experimento una OmpH1 de tipo silvestre con un único secución de glicosilación, y un mutante que carecía de la asparagina decisiva para la glicosilación. Todas las variantes sometidas a ensayo en esta memoria no solo mostraron un elevado nivel de eficacia para la glicosilación sino que también se utilizó cada secución de glicosilación. Los resultados se confirmaron con un inmunosuero específico de N-glicanos de *Campylobacter* (Fig. 3, panel inferior).

A continuación se muestra la secuencia proteica de la proteína OmpH1 de *Campylobacter jejuni* (cepa 81-176), con el marcador myc fijado en letra cursiva:

MKKILLSVLTTFVAVVLAACGGNSDSKTLNSLDKIKQNGVVRIGVFGDKPPFGYVDEKG
 NNQGYDIALAKRIAKELFGDENKVQFVLVEAANRVEFLKSNKVDIILANFTQTPERA^{myc}EQV
 DFCLP^{myc}YMKVALGVAVPKDSNITSVEDLKDKTLLL^{myc}NKGTTADAYFTQDYP^{myc}NIKTLKYDQNT
 ETFAALMDKRGD^{myc}ALSHDNTLLFAWVKDHPDFKMGIKELGNKDVIAPAVKKGDKELKEFI
 DNLI^{myc}IKLGQE^{myc}QFFHKAYDETLKAHFGDDVKADDV^{myc}IEGGKILEQKLISEEDL (secuencia 6)

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

Ejemplo 7: Presentación superficial de N-glicanos procedentes de *C. jejuni* en OmpH1 sobre la membrana externa de células *E. coli*

5 Con objeto de responder a la pregunta de si se pueden presentar múltiples variantes de OmpH1 glicosiladas en la superficie de células bacterianas, se llevó a cabo una inmunofluorescencia sobre células bacterianas CLM24 o SCM6 (que es SCM7 $\Delta waaL$) que expresaban diversas variantes de OmpH1. En el experimento se incluyó una OmpH1 de tipo silvestre y un mutante que carecía de la asparagina decisiva para la glicosilación. Además, se construyó un mutante C20S con objeto de retener la proteína en el periplasma, sirviendo así como control en el experimento. La inmunotinción se llevó a cabo sobre las células tratadas con paraformaldehído. El paraformaldehído fija las células sin destruir la estructura celular ni la compartimentación. Para detectar la proteína (fluorescencia roja) y el N-glicano (verde) en la superficie de las células bacterianas, se utilizaron respectivamente inmunosueros específicos para c-Myc y N-glicano, en combinación con los correspondientes anticuerpos secundarios conjugados con FITC y Cy3. Además, se empleó 4,6-diamino-2-fenilindol (DAPI; azul) para teñir el ADN bacteriano con el fin de diferenciar inequívocamente las células bacterianas de los desechos celulares. Cuando se tiñeron las células que expresaban OmpH1 de tipo silvestre, se detectó una inmunofluorescencia específica para la proteína y también para el N-glicano (Fig. 4A). Cuando se tiñó un mutante que carecía de la asparagina decisiva N139S con el inmunosuero específico, tanto anti-Myc como anti-N-glicano, solo se obtuvieron señales específicas de la proteína, pero no del glicano (panel 4B), lo que indicaba la especificidad del suero inmune específico de N-glicano. Cuando la proteína quedó retenida en el periplasma, como en el mutante C20S, no se detectó inmunofluorescencia roja, específica de la proteína, lo que indicaba que los anticuerpos eran incapaces de difundir dentro de la célula y eran suficientemente competentes para detectar cualquier fenómeno superficial (panel 4C). A continuación, se tiñeron células que expresaban múltiples variantes de OmpH1 diferentes en cuanto a la glicosilación: OmpH1^{KGN → NIT, HFGDD → DSNIT} (panel 4 D), OmpH1^{RGD → NIT, HFGDD → DSNIT} (panel 4 E), OmpH1^{KGN → NIT, RGD → NIT} (panel 4 F), OmpH1^{V83T, KGN → NIT} (panel 4 G) y OmpH1^{KGN → NIT, RGD → NIT, HFGDD → DSNIT} (panel 4 H). Todas las variantes de OmpH1 resultaron doblemente teñidas, lo que indicaba la presencia de proteína glicosilada en la superficie bacteriana. La Figura 4 se representa en escala de grises; la primera columna es una fotografía de fusión de las otras fotografías de la misma fila.

LISTADO DE SECUENCIAS

<110> Eidgenoessische Technische Hochschule zuerich

30 <120> Proteínas N-glicosiladas recombinantes procedentes de células procariotas

<130> P44148EP1/PWC

<140> 10179208.3

< 141> 2006-05-10

<150> EP 05010276.3

35 < 151> 2005-05-11

<160> 6

<170> PatentIn versión 3.3

<210> 1

< 211> 123

40 < 212> PRT

< 213> Artificial

<220>

< 223> Lip con sitio de glicosilación N123 de AcrA y marcador hexa His

ES 2 535 084 T3

<400> 1

Met Lys Lys Thr Ala Ile Ala Ile Ala Val Ala Leu Ala Gly Phe Ala
1 5 10 15

Thr Val Ala Gln Ala Asp Val Ile Ile Lys Pro Gln Val Ser Gly Val
20 25 30

Ile Val Asn Lys Leu Phe Lys Ala Gly Asp Lys Val Lys Lys Gly Gln
35 40 45

Thr Leu Phe Ile Ile Glu Gln Asp Gln Ala Ser Lys Asp Phe Asn Arg
50 55 60

Ser Lys Ala Leu Phe Ser Gln Leu Asp His Thr Glu Ile Lys Ala Pro
65 70 75 80

Phe Asp Gly Thr Ile Gly Asp Ala Leu Val Asn Ile Gly Asp Tyr Val
85 90 95

Ser Ala Ser Thr Thr Glu Leu Val Arg Val Thr Asn Leu Asn Pro Ile
100 105 110

Tyr Ala Asp Gly Ser His His His His His His
115 120

<210> 2

<211> 126

<212> PRT

<213> Artificial

5

<220>

<223> Lip con sitio de glicosilación N123 de AcrA y marcador hexa His

<400> 2

Met Lys Lys Thr Ala Ile Ala Ile Ala Val Ala Leu Ala Gly Phe Ala
1 5 10 15

Thr Val Ala Gln Ala Asp Val Ile Ile Lys Pro Gln Val Ser Gly Val
20 25 30

Ile Val Asn Lys Leu Phe Lys Ala Gly Asp Lys Val Lys Lys Gly Gln
35 40 45

Thr Leu Phe Ile Ile Glu Gln Asp Gln Phe Asn Arg Ser Lys Ala Leu
50 55 60

Phe Ser Gln Ser Ala Ile Ser Gln Lys Glu Leu Asp His Thr Glu Ile
65 70 75 80

Lys Ala Pro Phe Asp Gly Thr Ile Gly Asp Ala Leu Val Asn Ile Gly
85 90 95

Asp Tyr Val Ser Ala Ser Thr Thr Glu Leu Val Arg Val Thr Asn Leu
100 105 110

Asn Pro Ile Tyr Ala Asp Gly Ser His His His His His His His
115 120 125

10

<210> 3

<211> 116

ES 2 535 084 T3

< 212> PRT
 < 213> Artificial

<220>
 < 223> Lip con sitio de glicosilación N123 de AcrA y marcador hexa His

5

<400> 3
 Met Lys Lys Thr Ala Ile Ala Ile Ala Val Ala Leu Ala Gly Phe Ala
 1 5 10 15
 Thr Val Ala Gln Ala Asp Val Ile Ile Lys Pro Gln Val Ser Gly Val
 20 25 30
 Ile Val Asn Lys Leu Phe Lys Ala Gly Asp Lys Val Lys Lys Gly Gln
 35 40 45
 Thr Leu Phe Ile Ile Glu Gln Asp Gln Asp Phe Asn Arg Ser Lys Ala
 50 55 60
 Leu Asp His Thr Glu Ile Lys Ala Pro Phe Asp Gly Thr Ile Gly Asp
 65 70 75 80
 Ala Leu Val Asn Ile Gly Asp Tyr Val Ser Ala Ser Thr Thr Glu Leu
 85 90 95
 Val Arg Val Thr Asn Leu Asn Pro Ile Tyr Ala Asp Gly Ser His His
 100 105 110
 His His His His
 115

10

<210> 4
 < 211> 379
 < 212> PRT
 < 213> Artificial

<220>
 < 223> Proteína AcrA con secuencia señal PelB y marcador hexa His

ES 2 535 084 T3

<400> 4

Met Lys Tyr Leu Leu Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala
 1 5 10 15

Ala Gln Pro Ala Met Ala Met His Met Ser Lys Glu Glu Ala Pro Lys
 20 25 30

Ile Gln Met Pro Pro Gln Pro Val Thr Thr Met Ser Ala Lys Ser Glu
 35 40 45

Asp Leu Pro Leu Ser Phe Thr Tyr Pro Ala Lys Leu Val Ser Asp Tyr
 50 55 60

Asp Val Ile Ile Lys Pro Gln Val Ser Gly Val Ile Val Asn Lys Leu
 65 70 75 80

Phe Lys Ala Gly Asp Lys Val Lys Lys Gly Gln Thr Leu Phe Ile Ile
 85 90 95

Glu Gln Asp Lys Phe Lys Ala Ser Val Asp Ser Ala Tyr Gly Gln Ala
 100 105 110

Leu Met Ala Lys Ala Thr Phe Glu Asn Ala Ser Lys Asp Phe Asn Arg
 115 120 125

Ser Lys Ala Leu Phe Ser Lys Ser Ala Ile Ser Gln Lys Glu Tyr Asp
 130 135 140

Ser Ser Leu Ala Thr Phe Asn Asn Ser Lys Ala Ser Leu Ala Ser Ala
 145 150 155 160

Arg Ala Gln Leu Ala Asn Ala Arg Ile Asp Leu Asp His Thr Glu Ile
 165 170 175

Lys Ala Pro Phe Asp Gly Thr Ile Gly Asp Ala Leu Val Asn Ile Gly
 180 185 190

Asp Tyr Val Ser Ala Ser Thr Thr Glu Leu Val Arg Val Thr Asn Leu
 195 200 205

ES 2 535 084 T3

Asn Pro Ile Tyr Ala Asp Phe Phe Ile Ser Asp Thr Asp Lys Leu Asn
 210 215 220

Leu Val Arg Asn Thr Gln Ser Gly Lys Trp Asp Leu Asp Ser Ile His
 225 230 235 240

Ala Asn Leu Asn Leu Asn Gly Glu Thr Val Gln Gly Lys Leu Tyr Phe
 245 250 255

Ile Asp Ser Val Ile Asp Ala Asn Ser Gly Thr Val Lys Ala Lys Ala
 260 265 270

Val Phe Asp Asn Asn Asn Ser Thr Leu Leu Pro Gly Ala Phe Ala Thr
 275 280 285

Ile Thr Ser Glu Gly Phe Ile Gln Lys Asn Gly Phe Lys Val Pro Gln
 290 295 300

Ile Gly Val Lys Gln Asp Gln Asn Asp Val Tyr Val Leu Leu Val Lys
 305 310 315 320

Asn Gly Lys Val Glu Lys Ser Ser Val His Ile Ser Tyr Gln Asn Asn
 325 330 335

Glu Tyr Ala Ile Ile Asp Lys Gly Leu Gln Asn Gly Asp Lys Ile Ile
 340 345 350

Leu Asp Asn Phe Lys Lys Ile Gln Val Gly Ser Glu Val Lys Glu Ile
 355 360 365

Gly Ala Gln Leu Glu His His His His His His
 370 375

5

<210> 5
 < 211> 132
 < 212> PRT
 < 213> Artificial

<220>
 < 223> Proteína CtxB con marcador hexa His y secuencia señal de OmpA

<400> 5
 Met Lys Lys Thr Ala Ile Ala Ile Ala Val Ala Leu Ala Gly Phe Ala
 1 5 10 15

Thr Val Ala Gln Ala Thr Pro Gln Asn Ile Thr Asp Leu Cys Ala Glu
 20 25 30

10

Tyr His Asn Thr Gln Ile His Thr Leu Asn Asp Lys Ile Phe Ser Tyr
 35 40 45

ES 2 535 084 T3

Thr Glu Ser Leu Ala Gly Lys Arg Glu Met Ala Ile Ile Thr Phe Lys
 50 55 60

Asn Gly Ala Thr Phe Gln Val Glu Val Pro Gly Ser Gln His Ile Asp
 65 70 75 80

Ser Gln Lys Lys Ala Ile Glu Arg Met Lys Asp Thr Leu Arg Ile Ala
 85 90 95

Tyr Leu Thr Glu Ala Lys Val Glu Lys Leu Cys Val Trp Asn Asn Lys
 100 105 110

Thr Pro His Ala Ile Ala Ala Ile Ser Met Ala Asn Gly Ser His His
 115 120 125

His His His His
 130

<210> 6
 <211> 290
 <212> PRT
 <213> Artificial

5

<220>
 <223> OmpH1 con marcador myc

Met Lys Lys Ile Leu Leu Ser Val Leu Thr Thr Phe Val Ala Val Val
 1 5 10 15

Leu Ala Ala Cys Gly Gly Asn Ser Asp Ser Lys Thr Leu Asn Ser Leu
 20 25 30

Asp Lys Ile Lys Gln Asn Gly Val Val Arg Ile Gly Val Phe Gly Asp
 35 40 45

Lys Pro Pro Phe Gly Tyr Val Asp Glu Lys Gly Asn Asn Gln Gly Tyr
 50 55 60

Asp Ile Ala Leu Ala Lys Arg Ile Ala Lys Glu Leu Phe Gly Asp Glu
 65 70 75 80

Asn Lys Val Gln Phe Val Leu Val Glu Ala Ala Asn Arg Val Glu Phe
 85 90 95

Leu Lys Ser Asn Lys Val Asp Ile Ile Leu Ala Asn Phe Thr Gln Thr
 100 105 110

Pro Glu Arg Ala Glu Gln Val Asp Phe Cys Leu Pro Tyr Met Lys Val
 115 120 125

Ala Leu Gly Val Ala Val Pro Lys Asp Ser Asn Ile Thr Ser Val Glu
 130 135 140 <400> 6

10

ES 2 535 084 T3

Asp Leu Lys Asp Lys Thr Leu Leu Leu Asn Lys Gly Thr Thr Ala Asp
 145 150 155 160
 Ala Tyr Phe Thr Gln Asp Tyr Pro Asn Ile Lys Thr Leu Lys Tyr Asp
 165 170 175
 Gln Asn Thr Glu Thr Phe Ala Ala Leu Met Asp Lys Arg Gly Asp Ala
 180 185 190
 Leu Ser His Asp Asn Thr Leu Leu Phe Ala Trp Val Lys Asp His Pro
 195 200 205
 Asp Phe Lys Met Gly Ile Lys Glu Leu Gly Asn Lys Asp Val Ile Ala
 210 215 220
 Pro Ala Val Lys Lys Gly Asp Lys Glu Leu Lys Glu Phe Ile Asp Asn
 225 230 235 240
 Leu Ile Ile Lys Leu Gly Gln Glu Gln Phe Phe His Lys Ala Tyr Asp
 245 250 255
 Glu Thr Leu Lys Ala His Phe Gly Asp Asp Val Lys Ala Asp Asp Val
 260 265 270
 Val Ile Glu Gly Gly Lys Ile Leu Glu Gln Lys Leu Ile Ser Glu Glu
 275 280 285
 Asp Leu
 290

REIVINDICACIONES

1. Uso de una o varias secuencias de aminoácidos que comprenden la secuencia D/E- X- N- Z-S/T, en donde X y Z pueden ser cualquier aminoácido natural excepto prolina, para modificar una proteína de tal manera que la proteína es susceptible de ser N-glicosilada por una oligosacaril transferasa de *Campylobacter* spp.
- 5 2. Uso según la reivindicación 1, en donde la proteína es modificada mediante la introducción en dicha proteína dichas una o varias secuencias de aminoácidos.
3. Uso según la reivindicación 1, en donde la proteína es modificada mediante adición, delección y/o sustitución de uno o varios aminoácidos en dicha proteína.
4. Uso según una cualquiera de las reivindicaciones 1 a 3, en donde *Campylobacter* spp. es *C. jejuni*.
- 10 5. El uso según una cualquiera de las reivindicaciones 1 a 4, en donde la proteína comprende adicionalmente un glicano enlazado a cada una de dichas una o varias secuencias de consenso introducidas mediante un enlace N-glicosídico.
6. El uso según una cualquiera de las reivindicaciones 1 a 5, en donde dicha proteína es una exoproteína de *P. aeruginosa*.
- 15 7. El uso según la reivindicación 7, en donde dicha exoproteína de *P. aeruginosa* comprende dos de dichas secuencias de consenso introducidas.
8. El uso según la reivindicación 5, en donde dicho glicano comprende un oligosacárido o un polisacárido procedente de una bacteria Gram negativa.
- 20 9. El uso según la reivindicación 8, en donde dicho oligosacárido o polisacárido procede de *Shigella* spp, *Pseudomonas* spp o *E. coli*.
10. El uso según la reivindicación 9, en donde dicha *Shigella* spp es *Shigella dysenteriae* O1.
11. El uso según la reivindicación 9, en donde dicha *Pseudomonas* spp es *P. aeruginosa*.
12. El uso según la reivindicación 11, en donde dicha *P. aeruginosa* es *P. aeruginosa* O11.
- 25 13. El uso según la reivindicación 5, en donde dicho glicano comprende un oligosacárido o un polisacárido procedente de una bacteria Gram negativa.
14. El uso según la reivindicación 1, en donde dicha proteína comprende al menos dos de dichas secuencias de consenso introducidas.
15. El uso según la reivindicación 14, en donde al menos uno de dichos glicanos es diferente de otro de dichos glicanos.
- 30 16. El uso según la reivindicación 1, en donde dicha proteína es CRM.
17. El uso según la reivindicación 1, en donde dicha proteína es la toxina colérica.

Fig. 1

1 2 3 4



Fig. 2

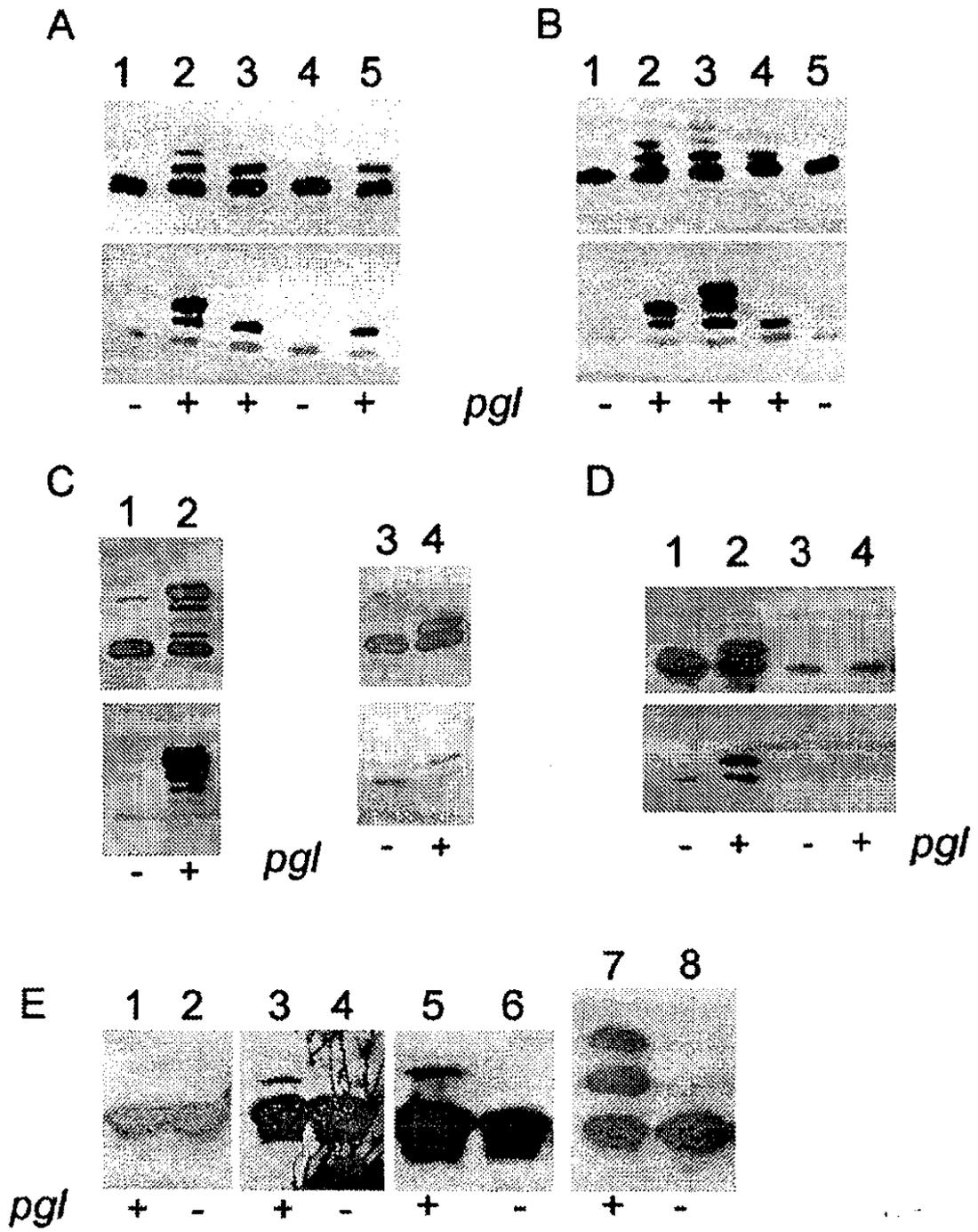


Fig 3

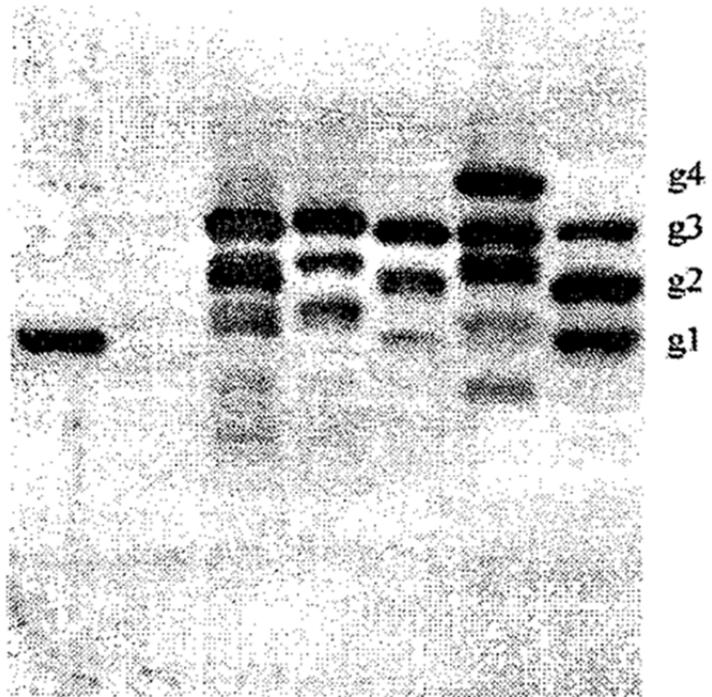
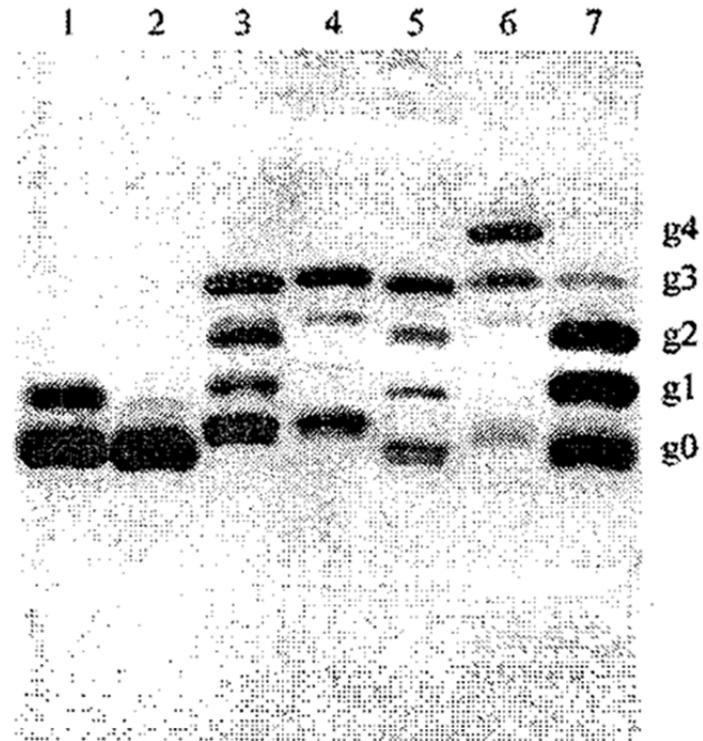


Fig. 4

