

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 765 484**

51 Int. Cl.:

C07K 14/005 (2006.01)

A61K 39/245 (2006.01)

A61K 39/12 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **07.12.2015 PCT/IB2015/059420**

87 Fecha y número de publicación internacional: **16.06.2016 WO16092460**

96 Fecha de presentación y número de la solicitud europea: **07.12.2015 E 15830990 (6)**

97 Fecha y número de publicación de la concesión europea: **06.11.2019 EP 3230305**

54 Título: **Antígenos de citomegalovirus**

30 Prioridad:

08.12.2014 EP 14196854

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

09.06.2020

73 Titular/es:

**GLAXOSMITHKLINE BIOLOGICALS SA (100.0%)
Rue de l'Institut 89
1330 Rixensart, BE**

72 Inventor/es:

**CARFI, ANDREA;
CHANDRAMOULI, SUMANA y
SETTEMBRE, ETHAN C.**

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 765 484 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

Antígenos de citomegalovirus

Campo de la invención

La presente invención se refiere a proteínas de citomegalovirus (CMV) adecuadas para usos vacunales.

5 Antecedentes de la invención

El citomegalovirus humano (HCMV) produce infecciones humanas persistentes ampliamente diseminadas, que varían con la edad y la inmunocompetencia del huésped. Puede permanecer latente toda la vida del huésped con esporádicos eventos de reactivación. La infección primaria de los huéspedes con un sistema inmunitario funcional se asocia con síntomas leves, aunque pueden progresar con fiebre, hepatitis, esplenomegalia y enfermedad tipo mononucleosis. Por el contrario, cuando se produce la infección primaria o la reactivación en huéspedes inmunocomprometidos o inmunodeficientes, a menudo experimentan enfermedades que ponen en peligro su vida, incluyendo neumonía, hepatitis, retinitis y encefalitis (Sinclair y Sissons, J. Gen. Virol. 87:1763-1779, 2006). La infección por HCMV ha sido reconocida por su asociación con tres poblaciones diferentes: los neonatos con sistemas inmunitarios inmaduros; receptores de trasplantes con sistemas inmunitarios discapacitados debido al uso de fármacos y los pacientes infectados con el VIH con sistemas inmunitarios comprometidos debido a la disminución de linfocitos T CD4+.

El HCMV es particularmente devastador en neonatos, produciendo defectos en el desarrollo neurológico. En los países industrializados, la infección intrauterina es la más común. Las estimaciones sugieren que entre un 0,6 % y un 0,7 % (dependiendo de la seroprevalencia de la población examinada) de todos los nuevos neonatos se infectan en el útero (Dollard y col., Rev. Med. Virol., 17(5):355-363, 2007). Solo en los Estados Unidos, esto corresponde aproximadamente a 40.000 nuevas infecciones cada año. Alrededor del 1,4 % de infecciones de CMV intrauterinas se produce por transmisión de mujeres con infección establecida. Una nueva infección materna se produce en el 0,7 a 4,1 % de las gestaciones y se transmite al feto aproximadamente en el 32 % de los casos. Alrededor del 90 % de los bebés infectados son asintomáticos al nacer y la mayoría desarrollará graves consecuencias de la infección en el curso de varios años, incluyendo el retraso mental y la pérdida de audición. Otros niños infectados muestran la enfermedad por HCMV sintomática con síntomas de implicación irreversible del sistema nervioso central en forma de microencefalia, encefalitis, convulsiones, ceguera, trastornos de la neurona motora superior y retraso psicomotor (Kenneson y col., Rev. Med. Virol., 17(4):253-276, 2007). En suma, aproximadamente 8.000 niños en los Estados Unidos desarrollan una enfermedad neurológica relacionada con virus cada año. La infección congénita es la fuerza motriz principal tras los esfuerzos para desarrollar una vacuna contra HCMV.

Las glicoproteínas de la envoltura del CMV gB, gH, gL, gM y gN representan candidatos vacunales atractivos ya que se expresan en la superficie del virus y pueden desencadenar respuestas inmunitarias humorales neutralizantes del virus protectoras.

Algunas estrategias vacunales con CMV se han dirigido a la glicoproteína B (gB) de superficie principal, que puede inducir una respuesta de anticuerpos dominante (Go y Pollard, J Infect Dis. 2008; 197:1631-1633). La glicoproteína B (gB) es una proteína trimérica que está altamente conservada entre las diferentes cepas de HCMV, así como entre otros herpesvirus, tales como el virus del herpes simple (HSV) y el virus de Epstein Barr (EBV). Pertenece a las proteínas de fusión vírica Clase III y tiene un papel crítico en el ciclo de replicación vírica fusionando la membrana vírica con la de la célula diana, facilitando el suministro del genoma vírico en el citoplasma. Los ensayos clínicos están progresando para evaluar la eficacia de la subunidad, así como candidatos vacunales de partículas tipo virus que incorporan distintas formas de gB de HCMV. La glicoproteína gB de CMV puede inducir una respuesta de anticuerpos neutralizantes, y los sueros de pacientes positivos a CMV está compuesto sobre todo por anticuerpos dirigidos contra gB (Britt, Journal of Virology 64:1079-1085, 1990).

El documento WO/2012/049317 desvela un polipéptido de la gB de CMV que comprende un dominio de bucle de fusión 1 (FL1) y un dominio de bucle de fusión 2 (FL2), en el que al menos uno de los dominios FL1 y FL2 comprende al menos una eliminación o sustitución de aminoácidos. Los ejemplos muestran que el porcentaje de trímeros de gB era alrededor del 70 %.

Existe la necesidad de una vacuna eficaz que se dirija a la glicoproteína gB de CMV y de procedimientos de inmunización que produzca mejores respuestas inmunitarias.

Sumario de la invención

Como se desvela y ejemplifica en el presente documento, se pueden introducir ciertas mutaciones en la proteína gB de citomegalovirus (CMV) (o un fragmento inmunogénico de la misma), en particular, para facilitar la producción recombinante y purificación de proteína gB recombinantes para usos vacunales. Los inventores reconocieron que un ectodominio de la proteína gB de tipo silvestre, que forma naturalmente un trímero monomérico, tiene una superficie expuesta hidrofóbica que puede hacer muy difícil la recombinación de la expresión de la proteína, y la secreción posterior por la célula huésped. Las mutaciones particularmente útiles son las que (i) al menos enmascaran parcialmente dicha superficie hidrofóbica, (ii) reducen la hidrofobicidad total de dicha superficie hidrofóbica, o (iii)

ambas.

En consecuencia, en un aspecto, la invención proporciona una proteína gB de citomegalovirus (CMV), en el que (i) dicha proteína gB no comprende un dominio transmembrana (TM); y (ii) dicha proteína gB comprende una superficie hidrofóbica 1, región formada por los aminoácidos 154-160 y 236-243 de la secuencia que se expone en la SEQ ID NO: 1, o por restos de aminoácido correspondientes en otras proteínas gB de CMV que se identifican mediante alineamiento de la secuencia en cuestión contra la SEQ ID NO: 1, que comprende una mutación que da como resultado un sitio de glicosilación dentro de dicha región. Preferentemente, dicho sitio de glicosilación es un sitio de N-glicosilación que comprende un motivo N-X-S/T/C, en el que X es cualquier resto de aminoácido (excepto prolina).

En otro aspecto, la invención proporciona una proteína gB recombinante de citomegalovirus (CMV), en el que (i) dicha proteína gB no comprende un dominio transmembrana (TM); y (ii) dicha proteína gB comprende una mutación que da como resultado un sitio de glicosilación, en el que dicho sitio de glicosilación está (1) dentro de una superficie hidrofóbica 2 formada por los restos de aminoácido 145-167 y 230-252 de la secuencia que se expone en la SEQ ID NO: 1, o por restos de aminoácido correspondientes en otras proteínas gB de CMV que se identifican mediante alineamiento de la secuencia en cuestión contra la SEQ ID NO: 1; o (2) en un resto que esté dentro de los restos de aminoácido 696-698 de la secuencia que se expone en la SEQ ID NO: 1.

La presente divulgación también describe una proteína gB de citomegalovirus (CMV), o un fragmento inmunogénico de la misma, en el que (i) dicha proteína gB, o fragmento inmunogénico de la misma no comprende un dominio transmembrana (TM); y (ii) dicha proteína gB, o fragmento inmunogénico de la misma, comprende una mutación dentro de la superficie hidrofóbica 1 (restos de aminoácido 154-160 y 236-243), en la que dicha mutación da como resultado la reducción de índice global de hidrofobicidad de dicha superficie hidrofóbica 1; en la que dicha mutación no es una eliminación o sustitución de un aminoácido del bucle de fusión 1 (FL1) o el bucle de fusión 2 (FL2).

La presente divulgación describe adicionalmente una proteína gB de citomegalovirus (MCV), o un fragmento inmunogénico de la misma, en la que (i) dicha proteína gB, o fragmento inmunogénico de la misma, no comprende un dominio transmembrana (TM); (ii) dicha proteína gB, o fragmento inmunogénico de la misma, comprende un ectodominio; y (iii) dicha proteína gB, o fragmento inmunogénico de la misma, comprende una secuencia heteróloga que tiene al menos 12 restos de longitud en el extremo C. En algunos aspectos, la proteína gB puede ser una proteína de fusión en la que la secuencia heteróloga se fusiona en el extremo C del ectodominio. En algunos aspectos, la secuencia heteróloga puede ser un péptido anfipático.

También se describe en el presente documento composiciones inmunogénicas que comprenden proteínas gB de CMV y fragmentos inmunogénicos de las mismas, como se describe en el presente documento. Las composiciones inmunogénicas pueden comprender un adyuvante inmunológico, y/u otro antígeno de CMV.

También se describen en el presente documento ácidos nucleicos que codifican la gB de CMV y fragmentos inmunogénicos de la misma, como se describe en el presente documento. El ácido nucleico puede utilizar como una vacuna basada en ácido nucleico (por ejemplo, una molécula de ARN autorreplicante que codifique la gB o un fragmento inmunogénico de la misma). El ácido nucleico también se puede utilizar para la producción recombinante de proteína gB.

La presente divulgación describe adicionalmente una célula huésped que comprende los ácidos nucleicos descritos en el presente documento. Los ácidos nucleicos pueden expresar una proteína gB (o un fragmento inmunogénico de la misma), y preferentemente forma un trímero monomérico. Preferentemente, el trímero monomérico se puede secretar por la célula huésped. Las células huésped preferidas son células huésped de mamífero, tal como células CHO o células HEK-293.

La presente divulgación describe adicionalmente un cultivo celular que comprende la célula huésped que se describe en el presente documento. Preferentemente, el cultivo es al menos de 20 litros de tamaño, y/o el rendimiento de proteína gB (o un fragmento inmunogénico de la misma) es al menos de 0,1 g/l.

La presente divulgación también describe un procedimiento de inducción de una respuesta inmunitaria contra el citomegalovirus (CMV), que comprende la administración a un sujeto que lo necesite de una cantidad inmunológicamente eficaz de la proteína gB (o un fragmento inmunogénico de la misma) que se describe en el presente documento. La presente divulgación, también describe un procedimiento de inhibición de la entrada de un citomegalovirus (CMV) en una célula, que comprende el contacto de la célula con la proteína gB (o un fragmento inmunogénico de la misma) que se describe en el presente documento.

También se describe en el presente documento el uso de las composiciones descritas en el presente documento para la inducción de una respuesta inmunitaria contra el citomegalovirus (CMV) y el uso de las composiciones descritas en el presente documento en la fabricación de un medicamento para la inducción de una respuesta inmunitaria contra el citomegalovirus (CMV).

55 **Breve descripción de las figuras**

La Figura 1 es una representación esquemática de la proteína gB de CMV de la cepa Merlin. TM: dominio

transmembrana; Cito: dominio citoplasmático; SP: péptido de señal; MPR: región proximal de membrana; I: Dominio I; II: Dominio II; III: Dominio III; IV: Dominio IV; V: Dominio V. Los números indicados se refieren a la posición de los restos de aminoácido de la gB de CMV de la cepa Merlin y se exponen en la SEQ ID NO: 1.

5 La Figura 2 muestra los resultados de la cromatografía de exclusión por tamaño. La figura muestra que la gB-698, con las mutaciones del bucle de fusión, pero sin glicosilación, formaban trímeros diméricos, mientras que la mutante gB-698glyc no formaba trímeros diméricos incluso a alta concentración proteica.

La Figura 3 muestra la estructura cristalina de Δ NgB.

10 La Figura 4 muestra la transferencia de Western del sobrenadante del cultivo celular utilizando un anticuerpo anti-His. Calles: 1: tipo silvestre (WT); 2: R236N; 3: G237N; 4: T158N/Y160T; 5: Y160E; 6: R236E/S238E; 7: R236E/S238E/T239E; 8: NGT insertado antes de W240; 9: I156H/H157R/W240N/Y242T. El panel de la izquierda muestra las muestras reducidas con 50 mM de DTT y cocidas a 95 °C durante 5 minutos. El panel de la derecha muestra muestras en condiciones no reductoras sin cocer. Todas las construcciones excepto el ectodominio de gB de tipo silvestre y R236E/S238E tenían una expresión detectable en condiciones de cocción y reducción (panel de la izquierda).

15 Descripción detallada de la invención

1. SUMARIO

Como se describe y ejemplifica en el presente documento, los inventores han descubierto que se pueden introducir ciertas mutaciones en la proteína gB de citomegalovirus (CMV) (o un fragmento inmunogénico de la misma, tal como el ectodominio) para facilitar la producción recombinante de esta proteína.

20 En general la proteína gB de CMV forma un trímero monomérico (que comprende tres proteínas gB, a las que también se hace referencia como subunidades) que se pueden utilizar como antígeno contra el CMV. Sin embargo, el trímero monomérico comprende una superficie hidrofóbica expuesta, que puede producir problemas significativos tanto en la producción como en la purificación de antígeno. Por ejemplo, la superficie hidrofóbica puede producir agregación de la proteína gB producida de manera recombinante (por ejemplo, dos trímeros monoméricos pueden formar un trímero
25 dimérico mediante la superficie hidrofóbica, lo que puede producir problemas de producción). La superficie hidrofóbica también produce que el trímero monomérico de gB se adhiera a la célula huésped (por ejemplo, a la membrana celular, a la membrana del ER, a otras proteínas hidrofóbicas, proteínas agregadas, etc.). Los inventores han descubierto que modificando esta superficie hidrofóbica se puede facilitar mucho la producción y purificación posteriores del antígeno gB.

30 Como se desvela en el presente documento, los inventores resolvieron la estructura cristalina de la gB de CMV, en forma de trímero monomérico, formando un complejo con un anticuerpo anti-gB (Fab). Basándose en la estructura cristalina, los inventores identificaron varias categorías de mutaciones que pueden reducir la agregación entre los trímeros monoméricos, y/o la adhesión del trímero monomérico a la célula huésped (por ejemplo, a la membrana celular, la membrana del ER, otras proteínas hidrofóbicas, proteínas agregadas, etc.). En particular, las mutaciones
35 en general también permiten que el trímero monomérico se secrete de una célula huésped, mejorando de esta manera significativamente la eficacia de los procedimientos de producción.

Basándose en la estructura cristalina, los inventores descubrieron que se pueden introducir varias categorías de mutaciones.

40 Primero, se puede introducir un sitio de glicosilación en una superficie hidrofóbica más estrecha, a la que se hace referencia en el presente documento como "superficie hidrofóbica 1" (que comprende los restos de aminoácido 154-160 y 236-243). Sin el deseo de quedar ligados por teoría alguna, se cree que la unión de un resto de glicano puede crear una barrera física (así como una superficie más hidrofílica), que reduzca la agregación no deseada de los trímeros monoméricos y la adhesión no deseada de un trímero monomérico a la célula huésped (por ejemplo, a la membrana celular, la membrana del ER, otras proteínas hidrofóbicas, proteínas agregadas, etc.) mediante la superficie
45 hidrofóbica. La mutación permite que la proteína gB (o un fragmento inmunogénico de la misma) se glicosile mediante el enlazamiento covalente con un resto de azúcar (glicano) en la superficie hidrofóbica 1. Preferentemente, el sitio de glicosilación es un sitio de N-glicosilación que comprende un motivo N-X-S/T/C, en el que X es cualquier resto de aminoácido (excepto prolina). Dicho sitio de glicosilación se puede crear, por ejemplo, sustituyendo un resto con un N, o insertando un resto de N.

50 Segundo, se puede introducir un sitio de glicosilación (i) en una superficie hidrofóbica más amplia, a la que se hace referencia en el presente documento como "superficie hidrofóbica 2" (que comprende los restos de aminoácido 145-167 y 230-252), o (ii) en un resto que está dentro de 20 angstroms de uno de los dos bucles de fusión altamente hidrofóbicos: el bucle de fusión 1 (FL1) (restos de aminoácido 155-157), o el bucle de fusión 2 (FL2) (restos de aminoácido 240-242), o ambos. Por ejemplo, basándose en la estructura cristalina, se descubrió que la región del extremo C del ectodominio está en proximidad conformacional con el FL1 o FL2 altamente hidrofóbico. Por ejemplo,
55 los restos 696-698 (Y, E y E, respectivamente, en la gB de la cepa Merlin) se cree que están dentro de 20 angstroms de FL1 y/o FL2. Por lo tanto, la introducción de un resto de glicano en la región del extremo C del ectodominio (por

ejemplo, creando un sitio de glicosilación en los restos 696-698, por ejemplo, reemplazando un resto con N, o insertando una secuencia N-X-S/T/C) también puede crearse una barrera física para reducir la agregación y/o la adhesión de los trímeros monoméricos.

5 Tercero, se puede introducir una mutación en la superficie hidrofóbica 1 (que comprende los restos 154-160 y 236-243), en el que la mutación da como resultado la reducción del índice global de hidrofobicidad de dicha superficie hidrofóbica 1. La creación de una superficie más hidrofílica puede reducir la agregación y/o adhesión de los trímeros monoméricos.

10 Cuarto, debido a que la región del extremo C del ectodominio está en proximidad conformacional con la superficie hidrofóbica 1, se puede añadir una secuencia heteróloga a la región del extremo C del ectodominio para "enmascarar" la superficie hidrofóbica. Se cree que la secuencia heteróloga puede servir como un lípido para cubrir la superficie hidrofóbica, reduciendo de esta manera la agregación y/o adhesión de los trímeros monoméricos. Preferentemente, la secuencia heteróloga comprende un péptido anfipático. Un péptido anfipático comprende una parte hidrofóbica que puede interactuar con la superficie hidrofóbica del trímero monomérico, y el péptido también comprende una superficie hidrofílica que se puede exponer a la solución acuosa.

15 Estos cuatro tipos de mutaciones se pueden utilizar de manera singular, o en combinación, para producir una proteína gB recombinante. Por ejemplo, la gB puede comprender dos mutaciones, una que crea un sitio de glicosilación en dicha superficie hidrofóbica 1 o la superficie hidrofóbica 2, y la otra sustituyendo un resto hidrofóbico con un resto más hidrofílico.

20 En consecuencia, la presente divulgación describe proteínas gB modificadas y fragmentos inmunogénicos de las mismas que comprenden un sitio de glicosilación y/o una o más mutaciones que reducen la agregación y/o adhesión de los trímeros de gB monoméricos.

En general, las proteínas gB y los fragmentos inmunogénicos descritos en el presente documento no comprenden dominio transmembrana (TM) (es decir, se ha eliminado el dominio TM de gB).

2. DEFINICIONES

25 La gB es una glicoproteína B de la envoltura que tiene numerosos papeles, uno de los cuales es la implicación en la fusión del citomegalovirus con las células huésped. Está codificada por el gen UL55 del genoma del HCMV. El tamaño de la forma nativa de gB depende del tamaño de la fase de lectura abierta (ORF) que puede variar un poco según la cepa. Por ejemplo, la ORF de la cepa AD169, que tiene 2717 pb de largo, codifica una gB de longitud completa de 906 aminoácidos, mientras que la ORF de la cepa Towne codifica una gB de longitud completa de 907 aminoácidos.

30 Aunque la presente invención es aplicable a las proteínas gB que se originan de cualquier cepa de CMV, con el fin de facilitar su entendimiento, cuando se hace referencia a las posiciones de aminoácidos en la presente memoria descriptiva, la numeración se da en relación con la secuencia de aminoácidos de la proteína gB de SEQ ID NO: 1 que se origina de la cepa Merlin que es un aislado clínico, a menos de que se establezca otra cosa. La presente invención, sin embargo, no se limita a la cepa Merlin. Las posiciones de aminoácidos comparables en una proteína gB de cualquier cepa de CMV pueden ser determinada por los expertos habituales en la técnica mediante el alineamiento de las secuencias de aminoácidos utilizando algoritmos de alineamiento bien conocidos y fácilmente disponibles (tales como BLAST, utilizando los ajustes por defecto, ClustalW2, utilizando los ajustes por defecto; o algoritmos desvelados por Corpet, Nucleic Acids Research, 1998, 16(22):10881-10890, utilizando los parámetros por defecto). En consecuencia, cuando se hace referencia a una "proteína gB de CMV", se entiende como una proteína gB de CMV de cualquier cepa (además de la cepa Merlin). El número actual se puede que tenga que ajustarse para las proteínas gB de otras cepas dependiendo del alineamiento de secuencia actual.

40 Por ejemplo, el bucle de fusión 1 (FL1) se define que consiste en los restos de aminoácido 155-157, y el bucle de fusión 2 (FL2) se define que consiste en los restos de aminoácido 240-242. Estos números son en relación a la secuencia de aminoácidos de la proteína gB de la SEQ ID NO: 1. Las secuencias/posiciones de FL1 y FL2 de las proteínas gB de otras cepas de CMV, u otras mutantes o variantes de gB, o fragmentos de gB se pueden concebir utilizando los programas de alineamiento de secuencia convencionales que alinean una secuencia cuestión con la SEQ ID NO: 1, e identifica restos que coinciden con los 155-157 y 240-242 de la SEQ ID NO: 1.

45 Las posiciones de restos de aminoácido específicas se numeran también de acuerdo con la SEQ ID NO: 1. Por ejemplo, "Y160" se refiere a la posición 160 de la SEQ ID NO: 1 (que es una Y), así como a los restos correspondientes de otras secuencias de gB (o variantes o fragmentos) que coinciden con Y160 de la SEQ ID NO: 1, cuando la secuencia se alinea con la SEQ ID NO: 1. Por simplicidad, se hace referencia a cualquier resto de la secuencia de gB (o variante o fragmento) que se corresponde con la Y160 de la SEQ ID NO: 1 como Y160, aunque la posición actual de ese resto puede ser el 160 o no, y el resto actual puede ser Y o no. Por ejemplo, una sustitución conservadora (por ejemplo, F) puede alinearse con Y160 de la SEQ ID NO: 1. Una sustitución conservadora se identifica normalmente como "positiva" o "+" por el BLAST 2.

55 De manera similar, las mutaciones también se identifican de acuerdo con la numeración de la SEQ ID NO: 1. Por ejemplo, Y160T significa que cualquier resto de la secuencia de gB (o variante o fragmento) que se corresponde con la Y160 de la SEQ ID NO: 1 ha mutado en T.

Un resto de aminoácido de una secuencia de consulta "se corresponde con" una posición designada de una secuencia de referencia (por ejemplo, Y160 de la SEQ ID NO: 1) cuando, alineando la secuencia de aminoácidos de consulta con la secuencia de referencia, la posición del resto coincide con la posición designada. Dichos alineamientos se pueden hacer a mano o utilizando programas de alineamiento de secuencias bien conocidos tales como ClustalW2 o "secuencias BLAST 2" utilizando los parámetros por defecto.

La forma nativa de gB de Merlin contiene en la dirección del extremo N al extremo C de la proteína (véase, la FIG. 1) (i) un péptido de señal, que se conoce por estar implicado en el tráfico intracelular polipeptídico incluyendo el direccionamiento del polipéptido hacia la secreción, seguido por (ii) una región llamada la secuencia líder, (iii) un dominio extracelular que también comprende un sitio de escisión endoproteolítico de tipo furina entre los restos de aminoácido 456 y 460, (iv) un dominio transmembrana y (v) un dominio citoplásmico en el extremo C. El ectodominio comprende dos bucles de fusión hidrofóbicos: el bucle de fusión 1 (FL1) (restos de aminoácido 155-157) o el bucle de fusión 2 (FL2) (restos de aminoácido 240-242).

La "superficie hidrofóbica 1" consiste en los restos de aminoácido 154-160 y 236-243. La "superficie hidrofóbica 2" consiste en los restos de aminoácido 145-167 y 230-252. De nuevo, los restos de aminoácido se identifican por la posición de acuerdo con la proteína gB de la cepa Merlin (SEQ ID NO: 1). Los restos de aminoácido correspondientes de otras secuencias o fragmentos de gB se pueden concebir alineando la secuencia de consulta contra la SEQ ID NO: 1.

Ectodominio de gB de CMV se refiere a un fragmento de gB de CMV que comprende sustancialmente la parte extracelular de la proteína gB de CMV madura, con o sin el péptido de señal y que carece del dominio transmembrana y el dominio del extremo C de la proteína gB de CMV de origen natural. Por ejemplo, el ectodominio puede comprender los restos de aminoácido 69 a 698.

El dominio transmembrana (TM) se refiere a la región que abarca la membrana celular. La región mínima consiste en los restos de aminoácido 750-766.

Se forma un trímero monomérico mediante tres proteínas gB (a las que también se hace referencia como subunidades). Se forma un trímero dimérico por dimerización de dos trímeros monoméricos. Por lo tanto, un trímero monomérico comprende seis subunidades de gB.

Un fragmento inmunogénico de gB se refiere a un fragmento que mantiene al menos un epítipo inmunogénico predominante de gB de longitud completa. Se han descrito varios dominios antigénicos (AD) de gB. Véase, por ejemplo, Wieggers y col., *J Virol.* 15 de oct de 2014. pii: JVI.02393-14; Epub antes de impresión; Pöttsch y col., *PLoS Pathog* 7(8): e1002172. doi:10.1371/journal.ppat.1002172 (Ago 2011); Spindler y col., *J Virol.* Ago 2013; 87(16):8927-39. doi: 10.1128/JVI.00434-13; Ohlin y col., *J. Virol.* 67(2): 703-710, 1993.

El AD-1 tiene aproximadamente 80 restos entre las posiciones 560 y 640 de la gB de la cepa AD169. Es la región inmunodominante de gB ya que casi todos los sueros de individuos infectados con HCMV reconocen el AD-1. El AD-2, se localiza en el extremo amino de la proteína, comprende al menos dos sitios distintos entre el a.a. 50 y 78 de la gB de AD169. Sitio I (a.a. 50-54 de AD169) se diferencia entre las cepas y es reconocido por los anticuerpos específicos de cepa algunos de los cuales neutralizan de una manera dependiente del complemento. Sitio II (a.a. 69-78 de AD169) es común a todas las cepas de HCMV e induce anticuerpos ampliamente neutralizantes. Una secuencia de a.a. Lineal, AD-3 (a.a. 783-906), reconocida por los anticuerpos específicos de gB en el suero humano, incluye los epítopos de la parte intraluminal/intravírica de la molécula. El AD-4 está formado por una secuencia discontinua que comprende los aminoácidos 121 a 132 y 344 a 438 de la GB de la cepa AD169. El AD-5 está formado por una secuencia continua que comprende los a.a. 133-343 de la cepa AD169. El fragmento inmunogénico descrito en el presente documento puede comprender un dominio antigénico seleccionado de entre el grupo que consiste en AD-1, AD-2, AD-3, AD-4, AD-5, y una combinación de los mismos.

Una secuencia heterogénea se refiere a un aminoácido o secuencia de nucleótidos que no se encuentra en la proteína gB de CMV o un ácido nucleico que codifica una proteína gB de CMV de origen natural.

Un péptido anfipático se refiere a péptidos que contiene tanto restos de aminoácido hidrofílicos como hidrofóbicos, donde la separación espacial de estos restos, tales como, por ejemplo, mediante la estructura secundaria del péptido, da como resultado su capacidad para la partición en una interfaz entre un medio polar y no polar tal como una interfaz lipídica, una interfaz aire/agua, una interfaz disolvente hidrofílico/disolvente hidrofóbico e interfaz aire/material de empaquetamiento. Normalmente, los péptidos anfipáticos presentan una anfipaticidad definida por una media del momento hidrofóbico entre aproximadamente 0 y aproximadamente 0,9, de acuerdo con el gráfico de Eisenberg (Eisenberg y col., *J. Mol. Biol.* 179:125-142, 1984).

3. PROTEÍNAS GB RECOMBINANTES

La proteína gB de CMV comprende un dominio extracelular en el extremo N de aproximadamente 725 aminoácidos, seguidos por una región transmembrana y un dominio del extremo C. Los epítopos neutralizantes más conocidos del mapa de la gB están hacia el ectodominio de gB, como son las dos regiones hidrofóbicas, a las que se hace referencia como bucles de fusión 1 (FL1) y 2 (FL2). Se cree que la gB inserta estos bucles de fusión en la membrana celular de

la diana y, mediante un cambio conformacional de su estructura, pone la membrana vírica y celular en yuxtaposición para facilitar la fusión de la membrana vírica/celular.

Como se desvela en el presente documento, los inventores resolvieron la estructura cristalina de la gB de CMV, en forma de trímero monomérico, formando un complejo con un anticuerpo anti-gB (fragmento Fab). Basándose en la estructura cristalina, los inventores han identificado una superficie hidrofóbica expuesta que puede producir agregación de los trímeros gB monoméricos (por ejemplo, la formación de trímeros diméricos, que comprenden seis subunidades de gB), así como la adhesión no deseada de trímero monomérico a la célula huésped (por ejemplo, a la membrana celular, la membrana del ER, otras proteínas hidrofóbicas, proteínas agregadas, etc.). Una superficie hidrofóbica más estrecha (superficie hidrofóbica 1) que consiste en los restos 154-160 y 236-243. Una superficie hidrofóbica más ancha (superficie hidrofóbica 2) que consiste en los restos 145-167 y 230-252.

Basándose en esta información, los inventores identificaron varias categorías de mutaciones que pueden reducir la agregación entre los trímeros monoméricos, y/o la adhesión del trímero monomérico a la célula huésped (por ejemplo, a la membrana celular, la membrana del ER, otras proteínas hidrofóbicas, proteínas agregadas, etc.). En particular, las mutaciones en general permiten que el trímero monomérico se secrete de una célula huésped (por ejemplo, cuando se expresan de manera heteróloga en células de mamífero), mejorando de esta manera significativamente la eficacia del procedimiento de producción.

En particular se han identificado cuatro estrategias. La primera es introducir un sitio de glicosilación dentro de la superficie hidrofóbica 1 (restos 154-160 y 236-243). La segunda es introducir un sitio de glicosilación en la superficie hidrofóbica 2 (restos 145-167 y 230-252), y/o un resto que está dentro de 20 angstroms del bucle de fusión 1 (FL1) (restos 155-157) y/o el bucle de fusión 2 (FL2) (restos 240-242). Por ejemplo, los restos 696, 697, y 698 se localizan dentro de 20 angstroms de FL1 y/o FL2 y se puede utilizar para introducir un sitio de glicosilación. La tercera es reducir el índice global de hidrofobicidad de la superficie hidrofóbica 1 o 2. La cuarta es incluir una secuencia heteróloga en la región del extremo C del ectodominio para enmascarar la superficie hidrofóbica 1 o 2 al menos parcialmente.

Estas mutaciones se pueden utilizar individualmente y en combinación para superar la hidrofobicidad de la superficie hidrofóbica expuesta. Por ejemplo, se pueden evaluar los mutantes mediante cromatografía de exclusión por tamaño y potencialmente mediante microscopía electrónica para verificar la formación de trímeros de gB monoméricos solubles.

En general, la proteína gB de CMV humana recombinante (o un fragmento inmunogénico de la misma) que se describe en el presente documento no comprende un dominio transmembrana (TM). El dominio TM consiste mínimamente en los restos de aminoácido 750-766. Sin embargo, debido a que los restos adyacentes de esta mínima región de TM son generalmente hidrofóbicos, la proteína gB (o un fragmento inmunogénico de la misma) puede no comprender (i) los restos de aminoácido 700-766, (ii) los restos de aminoácido 700-776, (iii) los restos de aminoácido 698-766, o (iv) los restos de aminoácido 698-776. Se cree que la eliminación de estos restos de aminoácido facilita la producción recombinante de proteínas gB que se desvela en el presente documento. De manera alternativa, la proteína gB (o un fragmento inmunogénico de la misma) puede no comprender los dos tramos de restos hidrofóbicos, siendo el primero los restos 725 a 744, y siendo el segundo del 751 al 773.

Las proteínas gB descritas en el presente documento pueden ser variantes gB que tienen distintos grados de identidad con la SEQ ID NO: 1 tal como al menos un 60 %, 70 %, 80 %, 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % idénticas a la secuencia mencionada en la SEQ ID NO: 1. Se desvelan en el presente documento variantes proteicas de gB que: (i) pueden no formar una cantidad sustancial de trímeros diméricos; (ii) pueden comprender al menos un epítipo de la SEQ ID NO: 1; y/o (iii) puede desencadenar anticuerpos (preferentemente anticuerpos neutralizantes) *in vivo* que reaccionan de manera cruzada con un virión de CMV. Una cantidad sustancial de trímero dimérico significa que, por ejemplo, al menos un 35 %, al menos un 30 %, al menos un 25 %, al menos un 20 %, al menos un 15 %, al menos un 10 %, o al menos un 5 %, de las subunidades totales de gB están en forma de trímero dimérico.

A. Glicosilación

En un aspecto, se desvela en el presente documento proteínas gB, o fragmentos inmunogénicos de la misma, que comprende un sitio de glicosilación dentro de la superficie hidrofóbica 1. En particular la divulgación describe adicionalmente una proteína gB de citomegalovirus (MCV), o un fragmento inmunogénico de la misma, en la que (i) dicha proteína gB, o fragmento inmunogénico de la misma, no comprende un dominio transmembrana (TM); y (ii) dicha proteína gB, o fragmento inmunogénico de la misma, comprende una mutación que da como resultado un sitio de glicosilación dentro de la superficie hidrofóbica 1. Preferentemente, dicho sitio de glicosilación es un sitio de N-glicosilación que comprende un motivo N-X-S/T/C, en el que X es cualquier resto de aminoácido (preferentemente que no sea prolina). Más preferentemente, dicho sitio de glicosilación es un sitio de N-glicosilación que comprende un motivo N-X-S/T, en el que X es cualquier resto de aminoácido (preferentemente que no sea prolina). En consecuencia, cuando se produce de manera recombinante en una célula huésped adecuada (por ejemplo, una célula huésped que comprende una enzima de glicosilación), la proteína gB o un fragmento inmunogénico de la misma comprende un resto de glicano unido a un resto dentro de la superficie hidrofóbica 1.

En otro aspecto, se desvela en el presente documento proteínas gB, o fragmentos inmunogénicos de la misma, que comprende un sitio de glicosilación dentro de la superficie hidrofóbica 2, o en un resto que está en proximidad conformacional con FL1 y/o FL2. En particular, la divulgación describe una proteína gB de citomegalovirus (CMV) recombinante, o un fragmento inmunogénico de la misma, en la que (i) dicha proteína gB o un fragmento inmunogénico de la misma, no comprende un dominio transmembrana (TM); y (ii) dicha proteína gB o fragmento inmunogénico de la misma, comprende una mutación que da como resultado un sitio de glicosilación, en el que dicho sitio de glicosilación está (1) dentro de la superficie hidrofóbica 2 (restos de aminoácido 145-167 y 230-252); o (2) en un resto que está dentro de 20 angstroms del bucle de fusión 1 (FL1) (restos de aminoácido 155-157) y/o el bucle de fusión 2 (FL2) (restos de aminoácido 240-242). En consecuencia, cuando se produce de manera recombinante en una célula huésped adecuada (por ejemplo, una célula huésped que comprende una enzima de glicosilación), la proteína gB o un fragmento inmunogénico de la misma comprende un resto de glicano unido a un resto dentro de la superficie hidrofóbica 2, o un resto que está dentro de 20 angstroms del bucle de fusión 1 (FL1) y/o el bucle de fusión 2 (FL2).

Como se desvela en el presente documento, la unión de un glicano crea una barrera física (así como una superficie más hidrofílica) para reducir la agregación/adhesión mediante la superficie hidrofóbica. El sitio de glicosilación puede estar dentro de la superficie hidrofóbica 1 más estrecha o la superficie hidrofóbica 2 más ancha que se desvelan en el presente documento. De manera alternativa, el sitio de glicosilación puede estar en un resto que esté en proximidad conformacional con el FL1 y/o FL2 altamente hidrofóbicos (por ejemplo, dentro de 30 angstroms, o dentro de 25 angstroms, o dentro de 20 angstroms, o dentro de 15 angstroms, o dentro de 14 angstroms, o dentro de 13 angstroms, o dentro de 12 angstroms, o dentro de 11 angstroms, o dentro de 10 angstroms, o dentro de 9 angstroms, o dentro de 8 angstroms, o dentro de 7 angstroms, o dentro de 6 angstroms, o dentro de 5 angstroms, de uno de los átomos del FL1 y/o el FL2). Por ejemplo, basándose en la estructura cristalina, la región del extremo C del ectodominio está en proximidad conformacional con el FL1 y/o FL2. Por ejemplo, los restos 696, 697, y 698 se localizan todos dentro de 20 angstroms de FL1 y/o FL2 y se puede utilizar para introducir un sitio de glicosilación.

Los sitios de glicosilación se pueden introducir en localizaciones deseadas mediante una modificación adecuada de secuencias de aminoácidos de la proteína gB. Preferentemente, se introducen los sitios de glicosilación unidos en N, que comprenden el motivo N-X-S/T/C. Preferentemente, el motivo es N-X-S/T. Preferentemente, X no es prolina.

Por ejemplo, la glicosilación unida en N se puede introducir en la superficie hidrofóbica cambiando la secuencia de aminoácidos de la proteína gB para incluir el motivo N-X-S/T/C para la glicosilación unida a N. Esto se puede conseguir insertando el motivo N-X-S/T/C en la secuencia de gB, o sustituyendo uno o más aminoácidos para producir el sitio de glicosilación, o cualquier combinación de adición y mutación que dé como resultado el motivo N-X-S/T/C. Por ejemplo, se puede añadir N, aunque la posición S/T/C puede mutarse; o N puede mutarse con otro resto, mientras que se añade la posición S/T/C. Cuando la proteína se expresa en células adecuadas, por ejemplo, células de mamífero, los glicanos unidos a N se unirán al resto N para crear una gB N-glicosilada.

De manera similar, también se pueden añadir sitios para la glicosilación unida a O. En la glicosilación unida a O, el resto de carbohidrato está unido al oxígeno del hidroxilo de una serina y treonina. Además, la glicosilación unida a O también se produce en la tirosina, 5-hidroxiserina, y 4-hidroxiprolina.

En ciertas realizaciones, la mutación comprende una inserción de una secuencia N-X-S/T/C (tal como N-X-S, N-X-T, N-X-C, N-G-S, N-G-T, N-A-S, N-A-T, etc., donde X es cualquiera aminoácido, pero preferentemente no es prolina). En ciertas realizaciones, la mutación comprende una inserción de una secuencia N-X-S/T/C (tal como N-X-S, N-X-T, N-X-C, N-G-S, N-G-T, N-A-S, N-A-T, etc., donde X es cualquiera aminoácido, pero preferentemente no es prolina), en el bucle de fusión 1 (FL1) (restos de aminoácido 155-157), bucle de fusión 2 (FL2) (restos de aminoácido 240-242), o ambos. La mutación puede comprender una inserción de una secuencia N-X-S/T/C (tal como N-X-S, N-X-T, N-X-C, N-G-S, N-G-T, N-A-S, N-A-T, etc., donde X es cualquiera aminoácido, pero preferentemente no es prolina) sin la mutación de otros restos de FL1 y FL2.

La mutación puede comprender la mutación ²³⁶RGSTW (SEQ ID NO: 12) a ²³⁶RGSTNGTW (SEQ ID NO: 13); ²⁴⁰WLYR (SEQ ID NO: 14) a ²⁴⁰WLYNGTR (SEQ ID NO: 15), o una combinación de las mismas.

La proteína gB o fragmento inmunogénico de la misma puede comprender una mutación que se selecciona de entre el grupo que consiste en (i) R236N, (ii) G237N, (iii) T158N, (iv) W240N y Y242T, (v) W240N y Y242S, (vi) W240N y Y242C, y una combinación de las mismas. La proteína gB o fragmento inmunogénico de la misma puede comprender una mutación que se selecciona de entre el grupo que consiste en (i) R236N, (ii) G237N, (iii) T158N, y una combinación de las mismas. En ciertas realizaciones, la mutación comprende la mutación T158N.

La mutación puede comprender las mutaciones (i) W240N; y (ii) Y242T, Y242S, o Y242C. La combinación de las dos mutaciones crea un sitio de glicosilación.

Además de la glicosilación, la proteína gB y los fragmentos inmunogénicos de la misma que se describen en el presente documento también pueden comprender una o más mutaciones que reducen el índice global de hidrofobicidad de la superficie hidrofóbica, como se describe posteriormente, y/o comprende una secuencia heteróloga en el extremo C, como se describe posteriormente.

B. Reducción del índice global de hidrofobicidad

En otro aspecto, la divulgación describe una proteína gB recombinante de citomegalovirus (CMV), o un fragmento inmunogénico de la misma, en el que (i) dicha proteína gB, o fragmento inmunogénico de la misma no comprende un dominio transmembrana (TM); y (ii) dicha proteína gB, o fragmento inmunogénico de la misma, comprende una mutación en la superficie hidrofóbica 1 (restos de aminoácido 154-160 y 236-243) o la superficie hidrofóbica 2 (restos de aminoácido 145-167 y 230-252); en la que dicha mutación da como resultado la reducción de índice global de hidrofobicidad de dicha superficie hidrofóbica 1 o 2.

La mutación puede no ser una eliminación o sustitución de un aminoácido en el bucle de fusión 1 (FL1) (restos de aminoácido 155-157) o el bucle de fusión 2 (FL2) (restos de aminoácido 240-242). Por lo tanto, por ejemplo, se pueden excluir las proteínas gB que comprenden una eliminación o sustitución de un resto de aminoácido seleccionado de entre el grupo que consiste en los restos de aminoácido de 155, 156, 157, 240, 241 y 242.

La hidrofobicidad de una secuencia de aminoácidos particular se puede determinar utilizando una escala de hidrofobicidad, tal como la escala de Kyte y Doolittle (Kyte y col. 1982. J. Mol. Bio. 157: 105-132). La hidrofobicidad de una secuencia de aminoácidos o un fragmento de la misma se dicta por el tipo de aminoácidos que componen esta secuencia o un fragmento de la misma. Los aminoácidos se clasifican comúnmente en grupos distintos de acuerdo con sus cadenas laterales. Por ejemplo, algunas cadenas laterales se consideran no polares, es decir, hidrofóbicas, mientras que algunas otras se consideran polares. En el sentido de la presente invención, la alanina (A), glicina (G), valina (V), leucina (L), isoleucina (I), metionina (M), prolina (P), fenilalanina (F) y triptófano (W) se consideran parte de los aminoácidos hidrofóbicos, mientras que la serina (S), treonina (T), asparagina (N), glutamina (Q), tirosina (Y), cisteína (C), lisina (K), arginina (R), histidina (H), ácido aspártico (D) y ácido glutámico (E) se consideran parte de los aminoácidos polares. Independientemente de su hidrofobicidad, los aminoácidos también se clasifican en grupos basándose en propiedades comunes compartidas por sus cadenas laterales. Por ejemplo, la fenilalanina, el triptófano y la tirosina se clasifican juntos como aminoácidos aromáticos y se considerarán como aminoácidos aromáticos dentro del significado de la presente invención. El aspartato (D) y glutamato (E) son parte de los aminoácidos ácidos o cargados negativamente, mientras que la lisina (K), arginina (R) e histidina (H) son parte de los aminoácidos básicos o cargados positivamente, y se considerarán como tales en el sentido de la presente invención. Están disponibles escalas de hidrofobicidad que utilizan las propiedades hidrofóbicas e hidrofílicas de cada uno de los 20 aminoácidos y asignan una puntuación de hidrofobicidad a cada aminoácido, creando así una clasificación de hidrofobicidad.

Como un ejemplo solamente ilustrativo, se puede utilizar la escala de Kyte y Doolittle (Kyte y col. 1982. J. Mol. Bio. 157: 105-132). Esta escala permite a un experto en la técnica calcular el promedio de hidrofobicidad dentro de un segmento de longitud predeterminada. En consecuencia, las regiones hidrofóbicas en una secuencia de aminoácidos puede identificarlas el experto como dianas potenciales para la mutación de acuerdo con la presente invención. La capacidad de la mutación de dichas regiones para inducir un perfil de producto mejorado de la proteína mutante resultante, es decir, que favorezca la proporción de trimeros monoméricos dentro de la población, se pueden entonces ensayar como se describe posteriormente. La mutación de una región hidrofóbica puede ser una adición, eliminación, o sustitución del aminoácido dentro de la superficie hidrofóbica (por ejemplo, la sustitución de aminoácidos hidrofóbicos por aminoácidos polares).

También se puede determinar si una mutación puede reducir el índice global de hidrofobicidad de la superficie hidrofóbica, por ejemplo, analizando el efecto resultante de dicha mutación sobre el perfil del producto. Por ejemplo, con la expresión recombinante en células huésped, una mutación debería resultar en un perfil enriquecido en trimeros monoméricos mejorado, por ejemplo, al menos un 50 %, al menos un 55 %, al menos un 60 %, al menos un 65 %, al menos un 70 %, al menos un 75 %, al menos un 80 %, al menos un 85 %, al menos un 90 %, o al menos un 95 %, de la gB producida de manera recombinante (o un fragmento inmunogénico de la misma) está en forma de trímero monomérico.

La mutación puede comprender la sustitución de un resto de aminoácido hidrofóbico dentro de la superficie hidrofóbica 1 o 2 con un resto de aminoácido que comprenda una cadena lateral cargada o una cadena lateral polar. El resto de aminoácido hidrofóbico se puede seleccionar de entre el grupo que consiste en: A, V, L, I, P, M, F, G, y W, el resto de aminoácido que comprende una cadena lateral cargada puede seleccionarse de entre el grupo que consiste en D, E, K, R, y H. El resto de aminoácido que comprende una cadena lateral polar se puede seleccionar de entre el grupo que consiste en S, T, C, N, Q, e Y.

La mutación puede comprender la eliminación de un resto de aminoácido hidrofóbico dentro de la superficie hidrofóbica 1 o 2. El resto de aminoácido hidrofóbico se puede seleccionar de entre el grupo que consiste en: A, V, L, I, P, M, F, G, y W.

La mutación puede comprender la inserción de un resto de aminoácido que comprende una cadena lateral cargada o una cadena lateral polar en la superficie hidrofóbica 1 o 2. El resto de aminoácido hidrofóbico se puede seleccionar de entre el grupo que consiste en: A, V, L, I, P, M, F, G, y W. El resto de aminoácido que comprende una cadena lateral cargada puede seleccionarse de entre el grupo que consiste en D, E, K, R, y H. El resto de aminoácido que comprende una cadena lateral polar se puede seleccionar de entre el grupo que consiste en S, T, C, N, Q, e Y.

La mutación puede comprender la sustitución de Y160 por un resto de aminoácido que comprende una cadena lateral cargada o una cadena lateral polar.

El resto de aminoácido que comprende una cadena lateral cargada se puede seleccionar de entre el grupo que consiste en D, E, K, R, y H. La proteína gB o un fragmento inmunogénico de la misma puede comprender una mutación Y160E.

5 El resto de aminoácido que comprende una cadena lateral polar se puede seleccionar de entre el grupo que consiste en S, T, C, N, y Q. La proteína gB o un fragmento inmunogénico de la misma puede comprender una mutación Y160T.

La proteína gB o un fragmento inmunogénico de la misma puede comprender la mutación ¹⁵⁸TTY¹⁶⁰ en ¹⁵⁸NTT¹⁶⁰.

La mutación puede comprender la sustitución de S238 por un resto de aminoácido que comprende una cadena lateral cargada.

10 El resto de aminoácido que comprende una cadena lateral cargada se puede seleccionar de entre el grupo que consiste en D, E, K, R, y H. La proteína gB o un fragmento inmunogénico de la misma puede comprender una mutación S238E.

La mutación puede comprender la sustitución de T239 por un resto de aminoácido que comprende una cadena lateral cargada.

15 El resto de aminoácido que comprende una cadena lateral cargada se puede seleccionar de entre el grupo que consiste en D, E, K, R, y H. La proteína gB o un fragmento inmunogénico de la misma puede comprender una mutación T239E.

La proteína gB o un fragmento inmunogénico de la misma puede comprender una mutación S238E y una mutación T239E.

20 La proteína gB o un fragmento inmunogénico de la misma puede comprender una mutación R236E o una mutación R236D. La proteína gB o un fragmento inmunogénico de la misma puede comprender una mutación R236E.

La proteína gB o un fragmento inmunogénico de la misma puede comprender mutaciones seleccionadas de entre el grupo que consiste en: (i) R236E y S238E; (ii) R236E y T239E; y (iii) R236E, S238E, y T239E.

25 La proteína gB o un fragmento inmunogénico de la misma puede comprender una mutación en I156, H157, o una combinación de las mismas para reducir la hidrofobicidad. El resto se puede sustituir por un resto correspondiente de gB de otras especies de herpesvirus, tal como una gB de HSV-1, HSV-2, o VZV. La proteína gB o un fragmento inmunogénico de la misma puede comprender una mutación I156H. La proteína gB o un fragmento inmunogénico de la misma puede comprender una mutación Y242T.

C. secuencia heteróloga del extremo C

30 En otro aspecto, la divulgación describe adicionalmente una proteína gB de citomegalovirus (MCV), o un fragmento inmunogénico de la misma, en la que (i) dicha proteína gB, o fragmento inmunogénico de la misma, no comprende un dominio transmembrana (TM); (ii) dicha proteína gB, o fragmento inmunogénico de la misma, comprende un ectodominio (restos de aminoácido 69-698); y (iii) dicha proteína gB, o fragmento inmunogénico de la misma, comprende una secuencia heteróloga que tiene al menos 12 restos de longitud en el extremo C. En algunos aspectos, la proteína gB puede ser una proteína de fusión en la que la secuencia heteróloga se fusiona en el extremo C del ectodominio. En algunos aspectos, la secuencia heteróloga puede ser un péptido anfipático.

40 Los inventores descubrieron a partir de la estructura cristalina que la región del extremo C del ectodominio está en vecindad con la superficie hidrofóbica. Adicionalmente, cuando el dominio TM de gB se ha eliminado y el dominio citoplásmico está fusionado directamente con el dominio extracelular, se descubrió que la proteína gB puede formar un trímero monomérico soluble. Se cree que los restos de aminoácido adicionales en la región del extremo C pueden servir como un bloqueo físico para, al menos parcialmente, cubrir o enmascarar la superficie hidrofóbica expuesta, reduciendo de esta manera la agregación y/o adhesión de los trímeros monoméricos.

45 Preferentemente, la secuencia heteróloga tiene al menos aproximadamente 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, o 30 aminoácidos de longitud. Preferentemente, la secuencia heteróloga tendrá al menos 12 aminoácidos de longitud. La secuencia heteróloga puede tener aproximadamente 20 aminoácidos de longitud. La secuencia heteróloga puede tener no más de aproximadamente 50, 45, 40, 35, 30, o 25 de longitud.

Preferentemente, la secuencia heteróloga comprende un péptido anfipático. Un péptido anfipático comprende una parte hidrofóbica que puede interactuar con la superficie hidrofóbica del trímero monomérico, y el péptido también comprende una superficie hidrofílica que se puede exponer a la solución acuosa.

50 Ejemplos de péptidos anfipáticos se pueden encontrar, por ejemplo, en secuencias derivadas de apolipoproteínas. Las apolipoproteínas son proteínas de unión a lípidos que se dividen en seis clases principales (A, B, C, D, E y H) y varias subclases. El diseño y síntesis de péptidos anfipáticos que imitan las propiedades de las apolipoproteínas son conocidos, véase, por ejemplo, Mishra y col. *Biochemistry* 1996, agosto 27; 35(34):11210-20. Ejemplos específicos de péptidos anfipáticos incluyen, por ejemplo, DWLKAFYDKVAEKLKEAFLA (SEQ ID NO: 3); ELLEKWKEALAALAEKLK

(SEQ ID NO. 4); FWLKAFYDKVAEKLKEAF (SEQ ID NO. 5); DWLKAFYDKVAEKLKEAFRLTRKRGLKLA (SEQ ID NO. 6), y DWLKAFYDKVAEKLKEAF (SEQ ID NO. 7).

- 5 Las estrategias de mutaciones descritas en el presente documento se pueden utilizar de manera singular o en cualquier combinación, tal como (i) introduciendo un sitio de glicosilación y reduciendo el índice global de hidrofobicidad de la superficie hidrofóbica; (ii) introduciendo un sitio de glicosilación e introduciendo una secuencia heteróloga en el extremo C; (iii) reduciendo el índice global de hidrofobicidad de la superficie hidrofóbica e introduciendo una secuencia heteróloga en el extremo C; y (iv) introduciendo un sitio de glicosilación, reduciendo el índice global de hidrofobicidad de la superficie hidrofóbica, e introduciendo una secuencia heteróloga en el extremo C.

Las mutaciones específicas desveladas en el presente documento incluyen:

- 10 ²³⁶RGSTW²⁴⁰ (SEQ ID NO: 12) mutada en ²³⁶NGSTW²⁴⁰ (SEQ ID NO: 16);
²³⁶RGSTW²⁴⁰ (SEQ ID NO: 12) mutada en ²³⁶RNSTW²⁴⁰ (SEQ ID NO: 17);
²³⁶RGSTW²⁴⁰ (SEQ ID NO: 12) mutada en ²³⁶EGETW²⁴⁰ (SEQ ID NO: 18);
²³⁶RGSTW²⁴⁰ (SEQ ID NO: 12) mutada en ²³⁶EGEEW²⁴⁰ (SEQ ID NO: 19).
15 ²³⁶RGSTW²⁴⁰ (SEQ ID NO: 12) mutada en ²³⁶RGSTNGTW²⁴³ (SEQ ID NO: 13);
²⁴⁰WLYR²⁴³ (SEQ ID NO: 14) mutada en ²⁴⁰WLYNGTR²⁴⁶ (SEQ ID NO: 15);
¹⁵⁸TTY¹⁶⁰ mutada en ¹⁵⁸NTT¹⁶⁰;
¹⁵⁸TTY¹⁶⁰ mutada en ¹⁵⁸TTE¹⁶⁰; y
¹⁵⁶IH¹⁵⁷ y ²⁴⁰WLY²⁴² mutada en ¹⁵⁶HR¹⁵⁷ y ²⁴⁰NLT²⁴².

- 20 Las mutaciones relativas a los sitios de glicosilación y la hidrofobicidad no se limitan a las mutaciones descritas anteriormente. Se pueden llevar a cabo mutaciones adicionales no descritas en el presente documento, así como combinaciones de las mutaciones descritas en el presente documento. Los mutantes resultantes se pueden analizar, por ejemplo, mediante microscopía electrónica de barrido (SEM), modelado computacional, sedimentación (tal como por ultracentrifugación analítica (AUC), cromatografía, etc., para evaluar la producción del trímero monomérico. Por ejemplo, se puede utilizar una cromatografía de exclusión por tamaño (SEC), tal como la cromatografía de exclusión por tamaño basada en UV (SEC-UV). De manera alternativa, la muestra se puede tratar con un agente de entrecruzamiento, de manera que se formen enlaces covalentes entre dos proteínas. Después del entrecruzamiento, se carga la muestra en un gel en condiciones desnaturalizantes, tal como un SDS-PAGE, y se tiñe el gel para la presencia de proteínas, por ejemplo, con azul de Coomassie o nitrato de plata, los que presentará los agregados, si existen, que se separan según su peso molecular. Por ejemplo, la gB de AD169 de CMV (con el dominio transmembrana eliminado) tiene un peso molecular esperado de 92 kDa. Si se forma un trímero monomérico, el peso molecular medio esperado debería ser aproximadamente 276 kDa.
- 25
- 30

D. Otras modificaciones

También se pueden introducir otras modificaciones para facilitar la producción recombinante de proteína gB, o fragmentos inmunogénicos de la misma.

- 35 En general, el dominio citoplásmico del extremo C original de la proteína gB se puede eliminar en una extensión variable. Por ejemplo, al menos un 10 %, al menos un 20 %, al menos un 30 %, al menos un 40 %, al menos un 50 %, al menos un 60 %, al menos un 70 %, más adecuadamente un 80 %, al menos un 90 %, o el 100 % de los aminoácidos del dominio citoplásmico se pueden eliminar.

- 40 La proteína gB de CMV humana recombinante (o un fragmento inmunogénico de la misma) puede comprender una mutación del sitio de escisión de furina. El ectodominio comprende un sitio de escisión de furina en los restos 457-460 (RTKR (SEQ ID NO: 20) de la cepa Merlin; SEQ ID NO: 1). Dicha mutación puede ser, por ejemplo, R457S, R460S, o las mutaciones dobles R457S/R460S. Dichas mutaciones pueden destruir el sitio de mutación de furina promoviendo de esta manera la producción de una gB o un fragmento inmunogénico de gB intactos (por ejemplo, el ectodominio). Existe otro sitio potencial de escisión de furina en los restos 774-777 (RQRR) (SEQ ID NO: 21), que también se puede mutar si está presente en la proteína gB o el fragmento inmunogénico descrito en el presente documento.
- 45

- La proteína gB de CMV humana recombinante (o un fragmento inmunogénico de la misma) puede comprender una mutación en C246. Dicha mutación puede ser, por ejemplo, C246S, C246A, o C246G. Parece que C246 es una cisteína no emparejada y la mutación de esta cisteína no emparejada puede reducir la formación no deseada de enlaces disulfuro intermoleculares. Existe otra cisteína no emparejada potencial en la región del extremo C (resto 779). Si está presente, esta cisteína también se puede mutar.
- 50

- Opcionalmente, para facilitar la expresión y recuperación, la proteína gB (o fragmento inmunogénico de la misma) puede incluir un péptido de señal en el extremo N. Se puede seleccionar un péptido de señal de entre numerosos péptidos de señal conocidos en la técnica y se escoge normalmente para facilitar la producción y procesamiento en un sistema seleccionado para la expresión recombinante de la proteína gB (o un fragmento inmunogénico de la misma). En general, los péptidos de señal tienen de 5-30 aminoácidos de longitud, y están presentes normalmente en el extremo N de una proteína sintetizada recientemente. El centro del péptido de señal contiene en general un largo tramo de aminoácidos hidrofóbicos que tiene la tendencia de formar una alfa-hélice sencilla. Además, muchos péptidos de señal comienzan con un corto tramo hidrofílico (habitualmente cargado positivamente) de aminoácidos, que pueden
- 55

ayudar a reforzar la topología apropiada del polipéptido durante la translocación. El final del péptido de señal (extremo C), existe normalmente un tramo de aminoácidos hidrofílicos que se reconoce y escinde por la peptidasa de señal. La peptidasa de señal puede escindir durante o después de completar la translocación para generar un péptido de señal libre y una proteína madura.

- 5 El péptido de señal puede ser el que está presente naturalmente en las proteínas gB nativas. Para las cepas Merlin y AD169, el péptido de señal se localiza en los restos 1-22 de la SEQ ID NO: 1 y SEQ ID NO: 2, respectivamente. El péptido de señal de otras cepas se puede identificar mediante alineamiento de secuencias. De manera alternativa, el péptido de señal puede ser una secuencia heteróloga en la que la secuencia se origina de una proteína distinta a la gB. Los péptidos de señal ejemplares adecuados para su uso en el contexto de la proteína gB (o un fragmento inmunogénico de la misma) que se describe en el presente documento incluyen péptidos de señal del activador del plasminógeno tisular (tPA), proteína gD del virus del herpes simple (HSV), endostatina humana, gp120 de VIH, CD33, Her2Neu humano, pg67, o gp350 del virus de Epstein Barr (EBV). El péptido de señal puede no ser nativo y puede comprender mutaciones, tales como sustituciones, inserciones, o eliminaciones de aminoácidos. En particular, se pueden introducir mutaciones en la parte del extremo C del péptido de señal.
- 10
- 15 Opcionalmente, las proteínas gB de CMV (o un fragmento inmunogénico de la misma) de la invención pueden incluir la adición de una secuencia de aminoácidos que constituya un marcador, que pueda facilitar la detección (por ejemplo, un marcador epitópico para la detección por anticuerpos monoclonales) y/o la purificación (por ejemplo, un marcador de polihistidina que permite la purificación en una resina quelante del níquel) de las proteínas. Ejemplos de marcadores de purificación por afinidad incluyen, por ejemplo, un marcador His (hexahistidina (SEQ ID NO: 8), que se une a un ion metálico), proteína de unión a maltosa (MBP) (se une a la amilosa), glutatión-S-transferasa (GST) (se une al glutatión), marcador FLAG (Asp-Tyr-Lys-Asp-Asp-Asp-Lys (SEQ ID NO: 9), se une a un anticuerpo anti-flag), marcador Strep (Ala-Trp-Arg-His-Pro-Gln-Phe-Gly-Gly (SEQ ID NO: 10), o Trp-Ser-His-Pro-Gln-Phe-Glu-Lys (SEQ ID NO: 11), se une a la estreptavidina o un derivado de la misma), marcador HA, marcador MYC, o combinaciones de los mismos.
- 20
- 25 Se pueden utilizar engarces escindibles. Esto permite al marcador separarse del complejo purificado, por ejemplo, mediante la adición de un agente capaz de escindir el engarce. Se conocen varios engarces escindibles diferentes por los expertos en la técnica. Dichos engarces se pueden escindir, por ejemplo, por radiación de un enlace fotolábil o hidrólisis catalizada por ácidos. Existen también engarces polipeptídicos que incorporan un sitio de reconocimiento de proteasas y que se pueden escindir mediante la adición de una enzima proteasa adecuada.
- 30 Puede ser más deseable expresar la gB (o un fragmento inmunogénico de la misma) sin una secuencia marcadora exógena, por ejemplo, por razones de seguridad o eficacia clínicas.

La proteína gB de CMV recombinante (o un fragmento inmunogénico de la misma) que se desvela en el presente documento puede también contener un marcador de trimerización para mejorar la trimerización. Por ejemplo, se puede insertar un plegamiento de fibrina T4 o un dominio de trimerización GCN4 en el extremo C de la proteína gB de CMV (o un fragmento inmunogénico de la misma).

- 35 También se describe en el presente documento un complejo de CMV que comprende la proteína gB recombinante (o un fragmento inmunogénico de la misma) que se describe en el presente documento. El complejo puede ser un trímero monomérico que consiste en tres subunidades de proteína gB.

- 40 Al menos un 50 %, al menos un 55 %, al menos un 60 %, al menos un 65 %, al menos un 70 %, al menos un 75 %, al menos un 80 %, al menos un 85 %, al menos un 90 %, o al menos un 95 %, de la gB producida de manera recombinante (o un fragmento inmunogénico de la misma) puede estar en forma de trímero monomérico. No más de un 50 %, no más de un 45 %, no más de un 40 %, no más de un 35 %, no más de un 30 %, no más de un 25 %, no más de un 20 %, no más de un 15 %, no más de un 10 %, no más de un 5 %, de la gB producida de manera recombinante (o un fragmento inmunogénico de la misma) puede estar en forma de trímero monomérico.

45 **4. EXPRESIÓN RECOMBINANTE DE GB**

También se describen en el presente documento ácidos nucleicos que codifican la gB de CMV y fragmentos inmunogénicos de la misma, como se describe en el presente documento. El ácido nucleico, tal como el ADN, se puede utilizar para la producción recombinante de proteína gB.

- 50 La presente divulgación describe adicionalmente una célula huésped que comprende los ácidos nucleicos descritos en el presente documento. Cuando la célula huésped se cultiva en condiciones adecuadas los ácidos nucleicos pueden expresar una proteína gB (o un fragmento inmunogénico de la misma). Preferentemente, dicha proteína gB (o un fragmento inmunogénico de la misma) forma un trímero monomérico. Preferentemente, el trímero monomérico se puede secretar por la célula huésped.

- 55 Las células huésped adecuadas incluyen, por ejemplo, células de insecto (por ejemplo, *Aedes aegypti*, *Autographa californica*, *Bombyx mori*, *Drosophila melanogaster*, *Spodoptera frugiperda*, y *Trichoplusia ni*), células de mamífero (por ejemplo, humanas, de primate no humano, caballo, vaca, oveja, perro, gato y roedor (por ejemplo, hámster)), células aviares (por ejemplo, de pollo, pato, y gansos), bacterias (por ejemplo, de *E. coli*, *Bacillus subtilis*, y

Streptococcus spp.), células de levadura (por ejemplo, células de *Saccharomyces cerevisiae*, *Candida albicans*, *Candida maltosa*, *Hansenula polymorpha*, *Kluyveromyces fragilis*, *Kluyveromyces lactis*, *Pichia guillermondii*, *Pichia pastoris*, *Schizosaccharomyces pombe* y *Yarrowia lipolytica*), *Tetrahymena* (por ejemplo, *Tetrahymena thermophila*) o combinaciones de las mismas.

- 5 Para las mutantes que comprenden un sitio de glicosilación, la célula huésped debería tener enzimas que intervengan en la glicosilación. Los huéspedes bacterianos no son adecuados en general para dichos mutantes, a menos de que la cepa se modifique para introducir enzimas de glicosilación; en vez de eso, se debería utilizar un huésped eucariota, tal como una célula de insecto, una célula aviar, o una célula de mamífero.

10 Los sistemas de expresión en células de insecto adecuados, tal como los sistemas de baculovirus, son conocidos por los expertos en la técnica y se describen, por ejemplo, en Summers y Smith, Texas Agricultural Experiment Station Bulletin N.º 1555 (1987). Los materiales y procedimientos para los sistemas de expresión en células de insecto/baculovirus están disponibles en el mercado en forma de kit, *inter alia*, en Invitrogen, San Diego, CA. Por ejemplo, para la expresión en células de insecto, se utiliza un vector de expresión de baculovirus adecuado, tal como pFastBac (Invitrogen), para producir partículas de baculovirus recombinantes. Las partículas de baculovirus se amplifican y se utilizan para infectar células de insecto para que expresen la proteína recombinante. Las células de insecto adecuadas incluyen, por ejemplo, las células Sf9, células Sf21, células Tn5, células S2 de Schneider, y células High Five (un aislado clónico derivado de la línea celular BTI-TN-5B1-4 de *Trichoplusia ni* parental (Invitrogen)).

15 Los sistemas de expresión en células aviares también son conocidos por los expertos en la técnica y se describen, por ejemplo, en las Patentes de EE. UU. N.º 5.340.740; 5.656.479; 5.830.510; 6.114.168; y 6.500.668; Patente europea N.º EP 0787180B; Solicitud de Patente europea N.º EP03291813.8; documento WO 03/043415; y documento WO 03/076601. Las células aviares adecuadas incluyen, por ejemplo, células madre embrionarias de pollo (por ejemplo, células EBx®), fibroblastos embrionarios de pollo, células germinales embrionarias de pollo, células de pato (por ejemplo, las líneas celulares AGE1.CR y AGE1.CR.pIX (ProBioGen) que se describen, por ejemplo, en Vaccine 27:4975-4982 (2009) y el documento WO 2005/042728), células EB66, y similares.

25 Preferentemente, las células huésped son células de mamífero (por ejemplo, humanas, de primate no humano, caballo, vaca, oveja, perro, gato, y roedor (por ejemplo, hámster). Las células de mamífero adecuadas incluyen, por ejemplo, células de ovario de hámster chino (CHO), células de riñón embrionario humano (células HEK-293, normalmente transformados por el ADN tipo 5 cortado de adenovirus), células NIH-3T3, célula 293-T, células Vero, células HeLa, células PERC.6 (número de depósito de ECACC 96022940), células Hep G2, MRC-5 (ATCC CCL-171), WI-38 (ATCC CCL-75), células pulmonares de rhesus fetal (ATCC CL-160), células de riñón bovino Madin-Darby ("MDBK"), células de riñón canino de Madin-Darby ("MDCK") (por ejemplo, MDCK (NBL2), ATCC CCL34; o MDCK 33016, DSM ACC 2219), células de riñón de hámster recién nacido (BHK), tal como BHK21-F, células HKCC, y similares.

30 La célula huésped puede ser una célula CHO. El polinucleótido que codifica la proteína gB (o un fragmento inmunogénico de la misma) que se describe en el presente documento puede estar integrado establemente en el genoma de la célula CHO.

35 También hay distintas líneas celulares de CHO en la colección europea de cultivos celulares (ECACC), o la colección americana de cultivos tipo (ATCC), tal como las líneas celulares de CHO hCBE11 (ATCC® PTA-3357™), E77.4 (ATCC® PTA-3765™), hLT-B: R-hG1 CHO #14 (ATCC® CRL-11965™), MOR-CHO- MORAb-003-RCB (ATCC® PTA-7552™), AQ.C2 clon 11B (ATCC® PTA-3274™), AQ.C2 clon 11B (ATCC® PTA-3274™), hsAQC2 en CHO-DG44 (ATCC® PTA-3356™), xrs5 (ATCC® CRL-2348™), CHO-K1 (ATCC® CCL-61™), Lec1 [llamada originalmente Pro-5WgaRI3C] (ATCC® CRL-1735™), Pro-5 (ATCC® CRL-1781™), ACY1-E (ATCC® 65421™), ACY1-E (ATCC® 65420™), pgsE-606 (ATCC® CRL-2246™), CHO-CD36 (ATCC® CRL-2092™), pgsC-605 (ATCC® CRL-2245™), MC2/3 (ATCC® CRL-2143™), CHO-ICAM-1 (ATCC® CRL-2093™), y pgsB-618 (ATCC® CRL-2241™). Se puede utilizar cualquiera de estas líneas celulares de CHO.

Otras líneas celulares de CHO disponibles en el mercado incluyen, por ejemplo, las células CHO-S FreeStyle™ y la línea celular Flp-In™-CHO de Life Technologies.

Los procedimientos para la expresión de proteínas recombinantes en células CHO se han desvelado en general. Véase, por ejemplo, en las Patentes de EE. UU. N.º 4.816.567 y N.º 5.981.214.

50 Los ácidos nucleicos recombinantes pueden tener el codón optimizado para la expresión en la célula huésped procarionota o eucariota seleccionada.

Para facilitar la replicación y expresión, los ácidos nucleicos pueden incorporarse en un vector, tal como un vector de expresión en procarionotas o eucariotas. Los vectores ejemplares incluyen plásmidos que sean capaces de replicarse de manera autónoma o se repliquen en una célula huésped. Los vectores de expresión típicos contienen promotores, amplificadores, y terminadores adecuados, que sean útiles para la regulación de la expresión de las secuencias codificantes en la construcción de expresión. Los vectores también pueden comprender marcadores de selección para proporcionar un rasgo fenotípico de las células huésped transformadas (tales como la transmisión de resistencia a antibióticos tales como la ampicilina o la neomicina).

Los procedimientos ejemplares suficientes para guiar al experto habituado en la técnica a lo largo de la producción de los ácidos nucleicos de gB recombinante de CMV se pueden encontrar en Sambrook y col., *Molecular Cloning: A Laboratory Manual*, 2ª ed., Cold Spring Harbor Laboratory Press, 1989; Sambrook y col., *Molecular Cloning: A Laboratory Manual*, 3ª ed., Cold Spring Harbor Press, 2001; Ausubel y col., *Current Protocols in Molecular Biology*, Greene Publishing Associates, 1992 (y los suplementos hasta 2003); y Ausubel y col., *Short Protocols in Molecular Biology: A Compendium of Methods from Current Protocols in Molecular Biology*, 4ª ed., Wiley & Sons, 1999.

También se describe en el presente documento un cultivo celular que comprende la célula huésped que se describe en el presente documento. El cultivo celular puede ser a gran escala, por ejemplo, de al menos aproximadamente 10 l, al menos aproximadamente 20 l, al menos aproximadamente 30 l, al menos aproximadamente 40 l, al menos aproximadamente 50 l, al menos aproximadamente 60 l, al menos aproximadamente 70 l, al menos aproximadamente 80 l, al menos aproximadamente 90 l, al menos aproximadamente 100 l, al menos aproximadamente 150 l, al menos aproximadamente 200 l, al menos aproximadamente 250 l, al menos aproximadamente 300 l, al menos aproximadamente 400 l, al menos aproximadamente 500 l, al menos aproximadamente 600 l, al menos aproximadamente 700 l, al menos aproximadamente 800 l, al menos aproximadamente 900 l, al menos aproximadamente 1000 l, al menos aproximadamente 2000 l, al menos aproximadamente 3000 l, al menos aproximadamente 4000 l, al menos aproximadamente 5000 l, al menos aproximadamente 6000 l, al menos aproximadamente 10.000 l, al menos aproximadamente 15.000 l, al menos aproximadamente 20.000 l, al menos aproximadamente 25.000 l, al menos aproximadamente 30.000 l, al menos aproximadamente 35.000 l, al menos aproximadamente 40.000 l, al menos aproximadamente 45.000 l, al menos aproximadamente 50.000 l, al menos aproximadamente 55.000 l, al menos aproximadamente 60.000 l, al menos aproximadamente 65.000 l, al menos aproximadamente 70.000 l, al menos aproximadamente 75.000 l, al menos aproximadamente 80.000 l, al menos aproximadamente 85.000 l, al menos aproximadamente 90.000 l, al menos aproximadamente 95.000 l, al menos aproximadamente 100.000 l, etc.

El rendimiento de proteína gB (o un fragmento inmunogénico de la misma) a partir del cultivo celular puede ser de al menos aproximadamente 0,01 g/l, al menos aproximadamente 0,02 g/l, al menos aproximadamente 0,03 g/l, al menos aproximadamente 0,05 g/l, al menos aproximadamente 0,06 g/l, al menos aproximadamente 0,07 g/l, al menos aproximadamente 0,08 g/l, al menos aproximadamente 0,09 g/l, al menos aproximadamente 0,1 g/l, al menos aproximadamente 0,15 g/l, al menos aproximadamente 0,20 g/l, al menos aproximadamente 0,25 g/l, al menos aproximadamente 0,3 g/l, al menos aproximadamente 0,35 g/l, al menos aproximadamente 0,4 g/l, al menos aproximadamente 0,45 g/l, al menos aproximadamente 0,5 g/l, al menos aproximadamente 0,55 g/l, al menos aproximadamente 0,6 g/l, al menos aproximadamente 0,65 g/l, al menos aproximadamente 0,7 g/l, al menos aproximadamente 0,75 g/l, al menos aproximadamente 0,8 g/l, al menos aproximadamente 0,85 g/l, al menos aproximadamente 0,9 g/l, al menos aproximadamente 0,95 g/l, o al menos aproximadamente 1.0 g/l.

También se describe en el presente documento un proceso de producción de proteína gB de citomegalovirus (CMV), o un fragmento inmunogénico de la misma, que comprende: (i) el cultivo de una célula huésped descrita en el presente documento en condiciones adecuadas, la expresión de esta manera de dicha proteína gB, o un fragmento inmunogénico de la misma; y (ii) la recolección de dicha proteína gB o un fragmento inmunogénico de la misma, del cultivo.

La proteína gB (o un fragmento inmunogénico de la misma) descrita en el presente documento puede purificarse. La proteína gB (o un fragmento inmunogénico de la misma) se puede purificar utilizando cualquiera de los procedimientos adecuados, tales como HPLC, distintos tipos de cromatografía (tal como de interacción hidrofóbica, intercambio iónico, afinidad, quelación, y exclusión por tamaño), electroforesis, centrifugación en gradiente de densidad, extracción en disolventes, o similares. Según sea apropiado, la proteína gB (o un fragmento inmunogénico de la misma) se puede purificar adicionalmente, según sea necesario, de manera que se retiren sustancialmente cualquiera de las proteínas que también se secretan en el medio o que resulten de la lisis de las células huésped, de manera que se proporcione un producto que esté al menos sustancialmente, libre de desechos del huésped, por ejemplo, proteínas, lípidos y polisacáridos. Véase, por ejemplo, los que se exponen en Sandana (1997) *Bioseparation of Proteins*, Academic Press, Inc.; y Bollag y col. (1996) *Protein Methods*, 2ª Edición Wiley-Liss, NY; Walker (1996) *The Protein Protocols Handbook Humana Press*, NJ, Harris y Angal (1990) *Protein Purification Applications: A Practical Approach* IRL Press at Oxford, Oxford, R.U.; Scopes (1993) *Protein Purification: Principles and Practice* 3ª Edición Springer Verlag, NY; Janson y Ryden (1998) *Protein Purification: Principles, High Resolution Methods and Applications*, Segunda Edición Wiley-VCH, NY; y Walker (1998) *Protein Protocols on CD-ROM* Humana Press, NJ. Si se desea, la proteína gB (o un fragmento inmunogénico de la misma) puede incluir un "marcador" que facilite la purificación, como se ha descrito anteriormente.

Por ejemplo, los procedimientos de purificación de gB de CMV por cromatografía de inmunoafinidad se han desvelado. Ruiz-Arguello y col., *J. Gen. Virol.*, 85:3677-3687 (2004).

5. COMPOSICIONES FARMACÉUTICAS Y PROCEDIMIENTOS DE TRATAMIENTO

La divulgación describe composiciones farmacéuticas y procedimientos de tratamiento que utilizan la proteína gB de citomegalovirus (CMV) (o fragmentos inmunogénicos de la misma) que se describe en el presente documento, o un ácido nucleico que codifica dicha proteína gB (o fragmentos inmunogénicos de la misma) que se describe en el presente documento. Por ejemplo, las proteínas o fragmentos inmunogénicos se pueden suministrar directamente

como componentes de una composición inmunogénica, o los ácidos nucleicos que codifican las proteínas gB o los fragmentos inmunogénicos se pueden administrar para producir la proteína de CMV o un fragmento inmunogénico *in vivo*. Las formulaciones proteicas, ácidos nucleicos recombinantes (por ejemplo, ARN autorreplicante) y VRP de alfavirus que contienen secuencias que codifican proteínas gB o fragmentos inmunogénicos se describen adicionalmente en el presente documento.

A. Composiciones proteicas

En un aspecto, la divulgación describe una composición inmunogénica que comprende la proteína gB de CMV recombinante (o un fragmento inmunogénico de la misma) que se describe en el presente documento.

La composición inmunogénica puede comprender proteínas de CMV adicionales, tales como gO, gH, gL, pUL128, pUL130, pUL131, un fragmento inmunogénico de las mismas o una combinación de las mismas. Por ejemplo, la gB (o un fragmento inmunogénico de la misma) puede combinarse con un complejo pentamérico de CMV que comprende gH o un fragmento que forma un pentámero de la misma, gL o un fragmento que forma un pentámero de la misma, pUL128 o un fragmento que forma un pentámero de la misma, pUL130 o un fragmento que forma un pentámero de la misma, y pUL131 o un fragmento que forma un pentámero de la misma. La gB (o un fragmento inmunogénico de la misma) también se puede combinar con un complejo trimérico de CMV que comprende: gH o un fragmento que forma un trímero de la misma, gL o un fragmento que forma un trímero de la misma, y gO o un fragmento que forma un trímero de la misma.

La composición inmunogénica puede comprender un adyuvante. Los adyuvantes ejemplares para aumentar la eficacia de la composición incluyen: (1) sales de aluminio (alum), tal como hidróxido de aluminio, fosfato de aluminio, sulfato de aluminio, etc.; (2) formulaciones en emulsión de aceite en agua (con o sin otros adyuvantes específicos tal como muramil péptidos (véase posteriormente) o componentes de la pared celular bacteriana) tales como por ejemplo (a) MF59 (PCT Publ. N.º WO 90/14837), que contiene un 5 % de escualeno, un 0,5 % de Tween 80, y un 0,5 % de Span 85 (opcionalmente que contiene distintas cantidades de MTP-PE (véase posteriormente), aunque no necesariamente) formulados en partículas submicrométricas utilizando un microfluidificador tal como el microfluidificador Modelo 110Y (Microfluidics, Newton, Mass.), (b) SAF, que contiene un 10 % de escualeno, un 0,4 % de Tween 80, un 5 % de polímero bloqueado con pluronic L121, y thr-MDP (véase posteriormente) sea microfluidificado en una emulsión submicrométrica o removido para generar una emulsión de partículas con un tamaño mayor, y (c) sistema de adyuvante Ribi™ (Ras) (Ribi Immunochem, Hamilton, Mont.) que contiene un 2 % de escualeno, un 0,2 % de Tween 80 y uno o más componentes de la pared celular bacteriana de entre el grupo que consiste en un monofosforilípido A (MPL), dimicolato de trehalosa (TDM), y esqueleto de la pared celular (CWS), preferentemente PL + CWS (Detox™); (3) adyuvantes de saponina, tales como Seimulon™ (Cambridge Bioscience, Worcester, Mass.) se pueden utilizar o partículas generadas a partir de estas tales como los ISCOM (complejos inmunoestimulantes); (4) Adyuvante completo de Freund (CFA) y Adyuvante incompleto de Freund (IFA); (5) citocinas, tales como interleucinas (IL-1, IL-2, etc.), factor estimulante de colonias de macrófagos (M-CSF), factor de necrosis tumoral (TNF), etc.; y (6) otras sustancias que actúan como adyuvantes para aumentar la eficacia de la composición.

Los muramil péptidos incluyen, pero no se limitan a, N-acetil-muramil-L-treonil-D-isoglutamina (thr-MDP), N-acetil-normuramil-L-alanil-D-isoglutamina (nor-MDP), N-acetilmuramil-L-alanil-D-isoglutaminil-L-alanina-2-(1'-2'-dipalmitoil-sn-glicer o-3-hidroxisforiloxi)-etilamina (MTP-PE), etc.

El adyuvante puede ser una sal de aluminio. El adyuvante puede ser una emulsión de aceite en agua, tal como MF59.

El adyuvante puede ser un agonista de TLR7, tal como imidazoquinolina o imiquimod.

El adyuvante puede ser una sal de aluminio, tal como, hidróxido de aluminio, fosfato de aluminio, sulfato de aluminio.

Los adyuvantes descritos en el presente documento se pueden utilizar de manera singular o en cualquier combinación, tal como alum/agonista TLR7.

B. VRP de Alfavirus

Las proteínas gB de CMV (o fragmentos inmunogénicos de las mismas) descritas en el presente documento se puede suministrar utilizando partículas de replicón de alfavirus (VRP). Como se utiliza en el presente documento, el término "alfavirus" tiene su significado convencional en la técnica e incluye distintas especies tales como el virus de la encefalitis equina venezolana (VEE; por ejemplo, del burro de Trinidad, TC83CR, etc.), virus del bosque de Semliki (SFV), virus Sindbis, Virus de Ross River, virus de la encefalitis equina oriental, virus Chikungunya, virus S.A. AR86, virus de los Everglades, virus de Mucambo, virus O'nyong-nyong, virus Getah, virus Getah, virus Sagiyama, virus Bebaru, virus Mayaro, virus Una, virus Aura, virus Whataroa, virus Banbanki, virus Kyzylagach, virus Highlands J, virus Fort Morgan, virus Ndumu, and virus Buggy Creek.

Una "partícula de replicón de alfavirus" (VRP) o "partícula de replicón" es un replicón de alfavirus empaquetado con proteínas estructurales de alfavirus.

Un "replicón de alfavirus" (o "replicón") es una molécula de ARN que puede dirigir su propia amplificación *in vivo* en

una célula diana. El replicón codifica las polimerasas que catalizan la amplificación de ARN (nsP1, nsP2, nsP3, nsP4) y contiene secuencias de ARN cis necesarias para la replicación que son reconocidas y utilizadas por las polimerasas codificadas. Un replicón de alfavirus contiene normalmente los siguientes elementos ordenados: Secuencias víricas 5' necesarias en cis para la replicación, secuencias que codifican proteínas no estructurales (sP1, nsP2, nsP3, nsP4) de alfavirus biológicamente activas, secuencias víricas 3' necesarias en cis para la replicación, y un tracto de poliadenilato. Un replicón de alfavirus también puede contener uno más promotores víricos subgenómicos de la "región de unión" direccionados contra la expresión de secuencias de nucleótidos heterólogos que se pueden modificar con el fin de aumentar o reducir la transcripción vírica del fragmento subgenómico y las secuencias heterólogas que se van a expresar. Se pueden utilizar otros elementos de control, tales como secuencias IRES o 2A.

10 **C. Sistemas de suministro de ácido nucleico**

La molécula de ácido nucleico recombinante que codifica las proteínas gB de CMV o fragmentos inmunogénicos descritos en el presente documento se pueden administrar para inducir la producción de las proteínas gB de CMV o fragmentos inmunogénicos y una respuesta inmunitaria contra estas.

15 El ácido nucleico recombinante puede ser ADN (por ejemplo, ADN vírico o plasmídico) o ARN, preferentemente ARN autorreplicante, y puede ser monocistrónico o policistrónico. Se puede utilizar cualquier ADN o ARN adecuado como vector de ácido nucleico que lleva las fases de lectura abierta que codifican las proteínas gB de CMV o fragmentos inmunogénicos de las mismas. Los vectores de ácido nucleico adecuados tienen la capacidad de llevar y dirigir la expresión de una o más proteínas gB de CMV o fragmentos inmunogénicos. Dichos vectores de ácidos nucleicos son conocidos en la técnica e incluyen, por ejemplo, plásmidos, ADN obtenido de virus ADN tales como los vectores de virus vaccinia, (por ejemplo, NYVAC, véase el documento US 5.494.807), y vectores poxvirus (por ejemplo, vector de viruela del canario ALVAC, Sanofi Pasteur), y el ARN obtenidos de virus ARN adecuados tales como alfavirus. Si se desea, la molécula de ácido nucleico recombinante puede modificarse, por ejemplo, contiene nucleobases modificadas y/o enlaces como se describe adicionalmente en el presente documento.

25 Las moléculas de ARN autorreplicantes descritas en el presente documento se basan en ARN genómico o virus ARN, pero carecen de genes que codifican una o más proteínas estructurales. Las moléculas de ARN autorreplicantes son capaces de traducirse para producir proteínas no estructurales del virus ARN y proteínas gB de CMV codificadas por el ARN autorreplicante.

30 El ARN autorreplicante contiene en general al menos uno o más genes seleccionados de entre el grupo que consiste en replicasa vírica, proteasas víricas, helicasas víricas y otras proteínas víricas no estructurales, y también comprenden secuencias de replicación activas cis en el extremo 5' y 3', y una secuencia heteróloga que codifica una o más proteínas gB CMV deseadas. Un promotor subgenómico que dirige la expresión de las secuencias heterólogas se pueden incluir en el ARN autorreplicante. Si se desea, una secuencia heteróloga puede fusionarse en fase con otras regiones codificantes en el ARN autorreplicante y/o puede estar bajo el control de un sitio interno de entrada en el ribosoma (IRES).

35 Las moléculas de ARN autorreplicantes descritas en el presente documento puede diseñarse de manera que la molécula de ARN autorreplicante no pueda inducir la producción de partículas víricas infecciosas. Esto se puede conseguir, por ejemplo, omitiendo uno o más genes víricos que codifican proteínas estructurales que son necesarios para la producción de partículas víricas en el ARN autorreplicante. Por ejemplo, cuando la molécula de ARN autorreplicante se basa en un virus alfa, tal como un virus Sinbis (SIN), virus de la selva de Semliki y el virus de encefalitis equina venezolana (VEE), se pueden omitir uno o más genes que codifican proteínas víricas estructurales, tales como glicoproteínas de la cápside o la envoltura. Si se desea, las moléculas de ARN autorreplicantes se pueden diseñar para inducir la producción de partículas víricas infecciosas que están atenuadas o son virulentas, o para producir partículas víricas que son capaces de una única ronda de infección posterior.

45 Una molécula de ARN autorreplicante puede, cuando se suministra a una célula de vertebrado incluso sin ninguna proteína, da lugar a la producción de múltiples ARN hijos por transcripción de sí mismo (o de una copia antisentido de sí mismo). El ARN autorreplicante puede traducirse directamente después del suministro a una célula, y esta traducción proporciona una ARN polimerasa dependiente de ARN que entonces produce transcripciones del ARN suministrado. Por lo tanto, el ARN suministrado da lugar a la producción de múltiples ARN hijos. Estas transcripciones son antisentido con respecto al ARN suministrado y pueden traducirse en sí mismas para proporcionar la expresión *in situ* de una proteína de CMV codificada, o se puede transcribir para proporcionar transcripciones adicionales con el mismo sentido que el ARN suministrado que se traduce para proporcionar la expresión *in situ* de las proteínas de CMV codificadas.

55 Una molécula de ARN autorreplicante codifica por tanto (i) una ARN polimerasa dependiente de ARN que puede transcribir el ARN de la molécula del ARN autorreplicante y (ii) una o más proteínas gB de CMV o fragmentos inmunogénicos de la misma. La polimerasa puede ser una replicasa de alfavirus, por ejemplo, que comprenden proteínas no estructurales de alfavirus nsP1-nsP4.

Las moléculas de ARN autorreplicante descritas en el presente documento pueden contener uno o más nucleótidos modificados y por lo tanto tiene una estabilidad mejorada y es resistente a la degradación y aclaramiento *in vivo*, y

5 otras ventajas. Existen más de 96 modificaciones de nucleósidos de origen natural que se encuentran en el ARN de mamífero. Véase, por ejemplo, Limbach y col., *Nucleic Acids Research*, 22(12):2183-2196 (1994). La preparación de nucleótidos y nucleótidos modificados y nucleósidos se conoce bien en la técnica, por ejemplo, en las Patentes de EE. UU. Números 4373071, 4458066, 4500707, 4668777, 4973679, 5047524, 5132418, 5153319, 5262530, 5700642, y muchos nucleósidos modificados y nucleótidos modificados están disponibles en el mercado. Si se desea, la molécula de ARN autorreplicante puede contener enlaces fosforamidato, fosforotioato, y/o metilfosfonato.

10 El ARN autorreplicante descrito en el presente documento es adecuado para el suministro en una variedad de modalidades, tal como suministro de ARN desnudo o en combinación con lípidos, polímeros u otros compuestos que facilitan la entrada en las células. Las moléculas de ARN autorreplicante se pueden introducir en células diana o sujetos utilizando cualquier técnica adecuada, por ejemplo, mediante inyección directa, microinyección, electroporación, lipofección, biolística, y similares. La molécula de ARN autorreplicante también puede introducirse en las células por medio de endocitosis mediada por un receptor. Véase, por ejemplo, la Pat. de EE. UU. N.º 6,090,619; Wu y Wu, *J. Biol. Chem.*, 263:14621 (1988); y Curiel y col., *Proc. Natl. Acad. Sci. USA*, 88:8850 (1991). Por ejemplo, la Pat. de EE. UU. N.º 6.083.741 desvela la introducción de un ácido nucleico exógeno en células de mamífero asociando el ácido nucleico a un resto policatiónico (por ejemplo, poli-L-lisina que tienen de 3-100 aminoácidos lisina), que se acopla él mismo a un resto de unión al receptor de integrina (por ejemplo, un péptido cíclico que tiene la secuencia Arg-Gly-Asp).

20 Las moléculas de ARN autorreplicante se pueden suministrar en las células mediante anffilos. Véase, por ejemplo, la Pat. de EE. UU. N.º 6.071.890. Normalmente, una molécula de ácido nucleico puede formar un complejo con el anffilo catiónico. Las células de mamífero que contactan con el complejo pueden captarse fácilmente.

25 El ARN autorreplicante puede suministrarse como ARN desnudo (por ejemplo, simplemente como una solución acuosa de ARN) por, para aumentar la entrada en las células y también los efectos intercelulares posteriores, el ARN autorreplicante se administra preferentemente en combinación con un sistema de suministro, tal como un sistema de suministro en partículas o emulsión. Un gran número de sistemas de suministro es bien conocido por los expertos en la técnica. Tres sistemas de suministro particularmente útiles son (i) liposomas, (ii) micropartículas poliméricas biodegradables y no tóxicas, y (iii) emulsiones de aceite en agua submicrométricas catiónicas.

30 La divulgación también proporciona una composición inmunogénica que comprende el ácido nucleico (por ejemplo, ARN autorreplicante) que se describe en el presente documento. La composición inmunogénica puede comprender un adyuvante, como se ha descrito anteriormente. Los adyuvantes preferidos incluyen, por ejemplo, una sal de aluminio o una emulsión de aceite en agua (tal como MF59).

D. Formulaciones farmacéuticas

35 Cada una de las composiciones inmunogénicas expuestas en el presente documento se puede utilizar sola o en combinación con uno o más antígenos, el último sea del mismo agente patógeno vírico o de otra fuente o fuentes patógenas. Estas formulaciones farmacéuticas puede ser profilácticas (para prevenir la infección) o terapéuticas (para tratar la enfermedad después de la infección).

40 Dichas formulaciones farmacéuticas comprenden una composición inmunogénica, habitualmente en combinación con "vehículos farmacéuticamente aceptables", que incluyen cualquier vehículo que por sí mismo no induce la producción de anticuerpos perjudiciales para el individuo que recibe la composición. Los vehículos adecuados son normalmente macromoléculas grandes, de metabolismo lento tales como proteínas, polisacáridos, ácidos poliláctico, ácidos poliglicólicos, aminoácidos poliméricos, copolímeros de aminoácidos, agregados lipídicos (tales como gotas lipídicas o liposomas), y partículas víricas inactivas. Dichos vehículos son bien conocidos por los expertos habituados en la técnica. De manera adicional, estos vehículos pueden funcionar como adyuvantes. Adicionalmente, el antígeno puede conjugarse con un toxoide bacteriano, tal como un toxoide de difteria, tétanos, cólera, patógenos como *H. pylori*, etc.

Las formulaciones farmacéuticas pueden comprender un adyuvante, como se ha descrito anteriormente.

45 Las formulaciones farmacéuticas (por ejemplo, la composición inmunogénica, vehículo farmacéuticamente aceptable, y adyuvante) contendrán normalmente diluyentes, tales como agua, solución salina, glicerol, etanol, etc. Adicionalmente, pueden estar presentes en dichos vehículos sustancias auxiliares, tal como agentes humectantes o emulsionantes, sustancias tampón del pH, y similares.

50 Normalmente, las formulaciones farmacéuticas se preparan como inyectables, sea como soluciones líquidas o suspensiones; formas sólidas adecuadas para introducir en la solución o la suspensión, también se pueden preparar vehículos líquidos antes de la inyección. La preparación también puede emulsionarse o encapsularse en liposomas para aumentar el efecto adyuvante, como se ha expuesto anteriormente en vehículos farmacéuticamente aceptables.

55 Las formulaciones farmacéuticas comprenden una cantidad inmunológicamente eficaz de los polipéptidos antigénicos, así como cualquier otro componente mencionado anteriormente, según se necesite. Por "cantidad inmunológicamente eficaz", se entiende que la administración de esa cantidad a un individuo sea en una única dosis o como parte de una serie, es eficaz para el tratamiento o prevención de la enfermedad, infección o enfermedad. Esta cantidad varía dependiendo de la salud y estado físico del individuo que se va a tratar, el grupo taxonómico del individuo que se va

a tratar (por ejemplo, un primate no humano, primate, etc.), la capacidad del sistema inmunitario del individuo para sintetizar anticuerpos, el grado de protección deseado, la formulación de la vacuna, la evaluación del médico del tratamiento de la situación médica, y otros factores relevantes. Se espera que la cantidad se encontrará en un intervalo amplio que puede determinarse a través de ensayos de rutina.

- 5 Las formulaciones farmacéuticas se administran convencionalmente por vía parenteral, por ejemplo, mediante inyección, sea subcutánea o intramuscular. Las formulaciones adicionales adecuadas por otros modos de administración incluyen las formulaciones oral y pulmonar, supositorios, y aplicaciones transdérmicas. Las formulaciones orales pueden preferirse para ciertas proteínas víricas. La dosificación del tratamiento puede ser una programación de una única dosis o una programación de múltiples dosis. La composición inmunogénica se puede
10 administrar en conjunción con otros agentes inmunorreguladores.

5. PROCEDIMIENTOS DE TRATAMIENTO

- En otro aspecto, la divulgación describe un procedimiento de inducción de una respuesta inmunitaria contra el citomegalovirus (CMV), que comprende la administración a un sujeto que lo necesite de una cantidad inmunológicamente eficaz de la composición inmunogénica descrita en el presente documento, que comprende las
15 proteínas, moléculas de ADN, moléculas de ARN (por ejemplo, moléculas de ARN autorreplicantes), o VRP como se ha descrito anteriormente.

La respuesta inmunitaria puede comprender la producción de anticuerpos neutralizantes contra el CMV. Los anticuerpos neutralizantes pueden ser independientes del complemento.

- La respuesta inmunitaria puede comprender una respuesta inmunitaria humoral, una respuesta inmunitaria mediada por células, o ambas. Se puede inducir una respuesta inmunitaria contra cada proteína de CMV suministrada. Una respuesta inmunitaria mediada por células puede comprender una respuesta de linfocito T auxiliar (Th), una respuesta de linfocito T citotóxico CD8+ (CTL), o ambas. La respuesta inmunitaria puede comprender una respuesta inmunitaria humoral, y los anticuerpos son anticuerpos neutralizantes. Los anticuerpos neutralizantes bloquean la infección vírica de las células. El CMV infecta células epiteliales y también células fibroblásticas. La respuesta inmunitaria puede
20 reducir o prevenir la infección de ambos tipos celulares. Las respuestas de anticuerpos neutralizantes pueden ser dependientes de complemento o independientes del complemento. La respuesta de anticuerpos neutralizantes puede ser independiente del complemento. La respuesta de anticuerpo neutralizante puede ser de neutralización cruzada; es decir, un anticuerpo generado contra una composición administrada neutraliza un virus CMV o de una cepa distinta de la cepa utilizada en la composición.

- Una medida útil de la potencia del anticuerpo en la técnica es el "título de neutralización del 50 %". Para determinar el título de neutralización del 50 %, se diluye el suero de animales inmunizados para evaluar cuánto del suero disuelto puede aún mantener la capacidad de entrada del 50 % de los virus en las células. Por ejemplo, un título de 700 significa que el suero mantenía la capacidad de neutralizar el 50 % de los virus después de diluirse 700 veces. Por lo tanto, a
30 títulos más altos más potentes respuestas de anticuerpo neutralizantes. Este título puede estar en un intervalo que tenga un límite menor de aproximadamente 200, aproximadamente 400, aproximadamente 600, aproximadamente 800, aproximadamente 1000, aproximadamente 1500, aproximadamente 2000, aproximadamente 2500, aproximadamente 3000, aproximadamente 3500, aproximadamente 4000, aproximadamente 4500, aproximadamente 5000, aproximadamente 5500, aproximadamente 6000, aproximadamente 6500, o aproximadamente 7000. El intervalo de título de neutralización del 50 % puede tener un límite superior de aproximadamente 400, aproximadamente 600, aproximadamente 800, aproximadamente 1000, aproximadamente 1500, aproximadamente 2000, aproximadamente 2500, aproximadamente 3000, aproximadamente 3500, aproximadamente 4000, aproximadamente 4500, aproximadamente 5000, aproximadamente 5500, aproximadamente 6000, aproximadamente 6500, aproximadamente 7000, aproximadamente 8000, aproximadamente 9000, aproximadamente 10000, aproximadamente 11000, aproximadamente 12000, aproximadamente 13000, aproximadamente 14000, aproximadamente 15000, aproximadamente 16000, aproximadamente 17000, aproximadamente 18000, aproximadamente 19000, aproximadamente 20000, aproximadamente 21000, aproximadamente 22000, aproximadamente 23000, aproximadamente 24000, aproximadamente 25000, aproximadamente 26000, aproximadamente 27000, aproximadamente 28000, aproximadamente 29000, o aproximadamente 30000. Por ejemplo, el título del 50 % de neutralización puede ser aproximadamente 3000 a aproximadamente 6500.
45 "Aproximadamente" significa más o menos un 10 % del valor mencionado. El título de neutralización puede medirse como se ha descrito en los ejemplos específicos, posteriormente.

- Se puede estimular una respuesta inmunitaria administrando proteínas, moléculas de ADN, moléculas de ARN (por ejemplo, moléculas de ARN autorreplicantes), o VRP a un individuo, normalmente un mamífero, incluyendo un ser humano. La respuesta inmunitaria inducida puede ser una respuesta inmunitaria protectora, es decir, la respuesta reduce el riesgo o gravedad de la infección por CMV. La estimulación de una respuesta inmunitaria protectora es particularmente deseable en algunas poblaciones particularmente en riesgo de infección por CMV y enfermedad. Por ejemplo, las poblaciones en riesgo incluyen los pacientes de trasplante de órganos sólidos (SOT), pacientes de trasplante de médula ósea, y pacientes de trasplante de células madre hematopoyéticas (HSCT). Las VRP se pueden administrar a un donante de trasplante antes del trasplante, o a un receptor del trasplante antes y/o después del trasplante. Debido a que la transmisión vertical de madre a hijo es una fuente común de infección a bebés, la
60

administración de VRP a mujeres que estén gestando o que puedan estar gestantes es particularmente útil.

Se puede utilizar cualquier vía de administración adecuada. Por ejemplo, se puede administrar una composición por vía intramuscular, intraperitoneal, subcutánea, o transdérmica o mediante una vía intramucosa tal como intraoral, intranasal, intravaginal, e intrarrectal. Las composiciones se pueden administrar de acuerdo con cualquiera calendario adecuado.

También se describe en el presente documento un procedimiento de inhibición de la entrada de un citomegalovirus (CMV) en una célula, que comprende el contacto de la célula con la composición inmunogénica que se describe en el presente documento.

La presente invención se ilustra adicionalmente mediante los siguientes ejemplos que no se deberían considerar como limitantes.

Ejemplos

Ejemplo 1 - Generación de una construcción de gB soluble y resolución de la estructura cristalina de la glicoproteína B de HCMV unida a un fragmento Fab de anticuerpo neutralizante humano

El ectodominio de la gB de CMV, restos 1 a 698 con un marcador 6-His (SEQ ID NO: 8) en el extremo C se expresaba en células 293GnTI (Figura 1). La secuencia de ectodominio WT fallaba en expresarse como una proteína secretada. Para aumentar la secreción de proteína los inventores mutaron tres restos hidrofóbicos en los bucles de fusión con los aminoácidos correspondientes de la gB de HSV-1 (I157H, H158R y W240R), que son más hidrofílicos. Los inventores también mutaron el sitio de escisión de furina para disminuir la heterogeneidad producida por el procesamiento incompleto durante la expresión (R457S / R460S), así como Cys246 en Ser para evitar la formación de enlaces disulfuro espurios (Figura 1). A pesar de estos cambios, la cromatografía de exclusión por tamaño (SEC; Figura 2A) y la cepa negativa EM (datos no mostrados) revelaba que la proteína (gB-698) formaba trímeros diméricos de los característicos trímeros post-fusión tri-lobulados. El análisis de las imágenes EM sugerían que la dimerización estaba mediada por la base del trímero de gB, presumiblemente debido a la hidrofobicidad intrínseca de su superficie. Por lo tanto, los inventores introdujeron un sitio de glicosilación en el bucle de fusión 2 (W240N, Y242T), del que se predecía que estaba expuesto al disolvente en el trímero, para interferir la dimerización de los trímeros. EM y SEC confirmaron que esta construcción, gB-698glyc, no se dimeriza (para formar trímeros diméricos) incluso a alta concentración proteica (Figura 2A y datos no mostrados).

Los intentos iniciales de cristalización de gB-698glyc por sí misma o en complejo con un fragmento Fab de anticuerpo neutralizante fueron insatisfactorios. Los inventores por lo tanto desglicosilaron la proteína con endoglicosidasa H (Endo H) y llevaron a cabo *in situ* una proteólisis limitada con subtilisina E para retirar regiones flexibles que podrían interferir con la cristalización. Este tratamiento daba como resultado cristales que, sin embargo, solo se difractaba hasta una resolución de 4,3 Å. Para mejorar la difracción, los inventores eliminaron 63 restos del extremo N del ectodominio de la gB (Δ NgB, que carece de los restos 25-86) que demostraban ser flexibles en la gB de HSV-1 (Heldwein y col., Science, 313(5784):217-220, 2006). El complejo Δ NgB desglicosilado-1G2 Fab cristalizaba rápidamente sin la necesidad de tratamiento con proteasa. Después de la exploración de varios cristales, se obtuvo un conjunto de datos con una resolución de 3,6 Å y se determinó la estructura por sustitución molecular (Figura 3).

Ejemplo 2 - Construcciones mutantes de gB de CMV

Basándose en la estructura cristalina de gB, los inventores diseñaron mutaciones dentro de la superficie hidrofóbica (englobando los dos bucles de fusión y restos de la vecindad de los dos bucles de fusión) para explorar mutantes que permitieran la expresión de un ectodominio de gB soluble (trímero monomérico). Los mutantes se hicieron y ensayaron en un experimento de expresión. La Figura 4 muestra la transferencia de Western del sobrenadante del cultivo celular con un anticuerpo anti-His. Todas las construcciones excepto el ectodominio de gB de tipo silvestre y R236E/S238E tenían una expresión detectable secretada en condiciones de cocción y reducción (panel de la izquierda).

Se hicieron y ensayaron las siguientes construcciones:

1. WT
2. R236N
3. G237N
4. T158N / Y160T
5. Y160E
6. R236E
7. R236E / S238E
8. NGT insertado antes de W240
9. I156H / H157R / W240N / Y242T (B-698glyc)

Todas las construcciones contenían las siguientes mutaciones adicionales: R457S / R460S (mutaciones del sitio de escisión de fuina) y C246S.

ES 2 765 484 T3

5 T158N/Y160T (calle 4) muestra una banda en el tamaño monomérico en condiciones no de cocción/no reducción similares a gB-698glyc (calle 9). El resto (excepto las calles 1 y 6) parece que se van a formar estructuras oligoméricas de orden superior solo visibles una vez que se rompen con calor y DTT. Estos resultados indican que la inserción de un sitio de glicosilación junto a continuación del bucle de fusión 1 es probablemente suficiente para permitir la expresión soluble de un ectodominio de gB incluso si están presentes restos del bucle de fusión tipo silvestre. Teniendo el sitio de glicosilación justo fuera del bucle de fusión 1 es suficiente para permitir la secreción sin necesidad de mutar los restos hidrofóbicos dentro del FL1. Un sitio de glicosilación fuera de FL1 conseguía resultados comparables a los restos mutados dentro de los bucles de fusión para añadir un sitio de glicosilación.

La secuencia de aminoácidos de la cepa Merlin de la proteína gB de citomegalovirus humano (SEQ ID NO: 1)

```

      10      20      30      40      50      60
MESRIWCLVV CVNLCIVCLG AAVSSSTRG TSATHSHHS HTTSAHSRS GSVSQRVTSS

      70      80      90     100     110     120
QTVSHGVNET IYNTTLKYGD VVGVNTTKYP YRVCSMAQGT DLIRFERNIV CTSMKPINED

     130     140     150     160     170     180
LDEGIMVVYK RNIVAHTFKV RVYQKVLTFR RSYAYIHTTY LLGSNTEYVA PPMWEIHHIN

     190     200     210     220     230     240
SHSQCYSSYS RVIAGTVFVA YHRDSYENKT MQLMPDDYSN THSTRYVTVK DQWHSRGSTW

     250     260     270     280     290     300
LYRETCNLNC MVTITTARSK YPYHFFATST GDVVDISPFY NGTNRNASYF GENADKFFIF

     310     320     330     340     350     360
PNYTIVSDFG RPNSALETHR LVAFLERADS VISWDIQDEK NVTCQLTFWE ASERTIRSEA

     370     380     390     400     410     420
EDSYHFSSAK MTATFLSKKQ EVNMSDSALD CVRDEAINKL QQIFNTSYNQ TYEKYGNVSV

     430     440     450     460     470     480
FETTGGLVVF WQGIKQKSLV ELERLANRSS LNLTHNRTKR STDGNNATHL SNMESVHNLV

     490     500     510     520     530     540
YAQLQFTYDT LRGYINRALA QIAEAWCVDQ RRTLEVFKEL SKINPSAILS AIYNKPIAAR

     550     560     570     580     590     600
FMGDVGLGLAS CVTINQTSVK VLRDMNVKES PGRCYSRPVV IFNFANSSYV QYQQLGEDNE

     610     620     630     640     650     660
ILLGNHRTEE CQLPSLKIFI AGNSAYEYVD YLFKRMIDLS SISTVDSMIA LDIDPLENTD

     670     680     690     700     710     720
FRVLELYSQK ELRSSNVFDL EEIMREFNSY KQRVKYVEDK VVDPLPPYLK GLDDLMSGLG

     730     740     750     760     770     780
AAGKAVGVAI GAVGGAVASV VEGVATFLKN PFGAFTIILV AIAVVIITYL IYTRQRLCT

     790     800     810     820     830     840
QPLQNLFPYL VSADGTTVTS GSTKDTSLQA PPSYEESVYN SGRKGGPPPS SDASTAAPPY

     850     860     870     880     890     900
TNEQAYQMLL ALARLDAEQR AQQNGTDSL D GRTGTQDKGQ KPNLLDRLRH RKNGYRHLKD

     907
SDEEENV

```

10

La secuencia de aminoácidos de la cepa AD169 de la proteína gB de citomegalovirus humano (SEQ ID NO: 2)

ES 2 765 484 T3

10 20 30 40 50 60
 MESRIWCLVĀ CVNLCIVCLĠ AAVSSSSTSH ATSSSTHNGSH TSRTTSAQTR SVYSQHVTSĪ
 70 80 90 100 110 120
 EAVSHRANET IYNTTLKYGD VVGVN¹TKYP YRVCSMAQGT DLIRFERNII CTSMKPINED
 130 140 150 160 170 180
 LDEGIMVVYK RNIVAHTFKV RVYQKVL²TR RSYAYIYTTY LLGSNTEYVA PPMWEIHHIN
 190 200 210 220 230 240
 KFAQCYSSYS RVIGGTVFVA YHRDSYENKT MQLIPDDYSN THSTRYVTVK DQWHSRGSTW
 250 260 270 280 290 300
 LYRETCNLNC MLTITTARSK YPYHFFATST GDVVYISPFY NGTNRNASYF GENADKFFIF
 310 320 330 340 350 360
 PNYTIVSDFĠ RPNAAPETHR LVAFLERADS VISWDIQDEK NVTCQLTFWE ASERTIRSEA
 370 380 390 400 410 420
 EDSYHFSSAK MTATFLSKKQ EVNMSDSALD CVRDEAINKL QQIFNTSYNQ TYEKYGNVSV
 430 440 450 460 470 480
 FETSGGLVVF WQGIKQKSLV ELERLANRSS LNITHRTRRS TSDNNTTHLS SMESVHNLVY
 490 500 510 520 530 540
 AQLQFTYDTL RGYINRALAQ IAEAWCVDQR RTLEVFKELS KINPSAILSĀ IYNKPIAARF
 550 560 570 580 590 600
 MGDVLGLASC VTINQTSVKV LRDMNVKESP GRCYSRPVVI FNFANSSYVQ YGQLGEDNEI
 610 620 630 640 650 660
 LLGNHRTEEC QLPSLKIFIA GNSAYEYVDY LFKRMIDLSS ISTVDSMIAL DIDPLENTDF
 670 680 690 700 710 720
 RVLELYSQKE LRSSNVFDLE EIMREFNSYK QRVKYVEDKV VDPLPPYLKG LDDLMSGLGA
 730 740 750 760 770 780
 AGKAVGVAIG AVGGAVASVĀ EGVATFLKNP FGAF³TIIILVA IAVVIITYLI YTRQRR⁴LCTQ
 790 800 810 820 830 840
 PLQNLFPYLĀ SADGTTVTSG STKDTSLQAP PSYEESVYNS GRKGGPPSS DASTAAPPYT
 850 860 870 880 890 900
 NEQAYQMLLĀ LARLDAEQRĀ QQNGTDSL⁵DG QTGTQDKGQK PNLLDRLRHR KNGYRHLKDS
 906
 DEEENV

LISTADO DE SECUENCIAS

- <110> NOVARTIS AG
- <120> ANTÍGENOS DE CITOMEGALOVIRUS
- 5 <130> PAT056317-EP-EPT
- <140>
- <141>
- <160> 21
- <170> PatentIn versión 3.5

ES 2 765 484 T3

<210> 1
 <211> 907
 <212> PRT
 <213> Citomegalovirus humano

5 <400> 1

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

ES 2 765 484 T3

His His Ile Asn Ser His Ser Gln Cys Tyr Ser Ser Tyr Ser Arg Val
 180 185 190

Ile Ala Gly Thr Val Phe Val Ala Tyr His Arg Asp Ser Tyr Glu Asn
 195 200 205

Lys Thr Met Gln Leu Met Pro Asp Asp Tyr Ser Asn Thr His Ser Thr
 210 215 220

Arg Tyr Val Thr Val Lys Asp Gln Trp His Ser Arg Gly Ser Thr Trp
 225 230 235 240

Leu Tyr Arg Glu Thr Cys Asn Leu Asn Cys Met Val Thr Ile Thr Thr
 245 250 255

Ala Arg Ser Lys Tyr Pro Tyr His Phe Phe Ala Thr Ser Thr Gly Asp
 260 265 270

Val Val Asp Ile Ser Pro Phe Tyr Asn Gly Thr Asn Arg Asn Ala Ser
 275 280 285

Tyr Phe Gly Glu Asn Ala Asp Lys Phe Phe Ile Phe Pro Asn Tyr Thr
 290 295 300

Ile Val Ser Asp Phe Gly Arg Pro Asn Ser Ala Leu Glu Thr His Arg
 305 310 315 320

Leu Val Ala Phe Leu Glu Arg Ala Asp Ser Val Ile Ser Trp Asp Ile
 325 330 335

Gln Asp Glu Lys Asn Val Thr Cys Gln Leu Thr Phe Trp Glu Ala Ser
 340 345 350

Glu Arg Thr Ile Arg Ser Glu Ala Glu Asp Ser Tyr His Phe Ser Ser
 355 360 365

Ala Lys Met Thr Ala Thr Phe Leu Ser Lys Lys Gln Glu Val Asn Met
 370 375 380

Ser Asp Ser Ala Leu Asp Cys Val Arg Asp Glu Ala Ile Asn Lys Leu
 385 390 395 400

Gln Gln Ile Phe Asn Thr Ser Tyr Asn Gln Thr Tyr Glu Lys Tyr Gly
 405 410 415

Asn Val Ser Val Phe Glu Thr Thr Gly Gly Leu Val Val Phe Trp Gln

ES 2 765 484 T3

			420					425					430			
Gly	Ile	Lys	Gln	Lys	Ser	Leu	Val	Glu	Leu	Glu	Arg	Leu	Ala	Asn	Arg	
		435					440					445				
Ser	Ser	Leu	Asn	Leu	Thr	His	Asn	Arg	Thr	Lys	Arg	Ser	Thr	Asp	Gly	
	450					455					460					
Asn	Asn	Ala	Thr	His	Leu	Ser	Asn	Met	Glu	Ser	Val	His	Asn	Leu	Val	
465					470					475					480	
Tyr	Ala	Gln	Leu	Gln	Phe	Thr	Tyr	Asp	Thr	Leu	Arg	Gly	Tyr	Ile	Asn	
				485					490					495		
Arg	Ala	Leu	Ala	Gln	Ile	Ala	Glu	Ala	Trp	Cys	Val	Asp	Gln	Arg	Arg	
			500					505					510			
Thr	Leu	Glu	Val	Phe	Lys	Glu	Leu	Ser	Lys	Ile	Asn	Pro	Ser	Ala	Ile	
		515					520					525				
Leu	Ser	Ala	Ile	Tyr	Asn	Lys	Pro	Ile	Ala	Ala	Arg	Phe	Met	Gly	Asp	
	530					535					540					
Val	Leu	Gly	Leu	Ala	Ser	Cys	Val	Thr	Ile	Asn	Gln	Thr	Ser	Val	Lys	
545					550					555					560	
Val	Leu	Arg	Asp	Met	Asn	Val	Lys	Glu	Ser	Pro	Gly	Arg	Cys	Tyr	Ser	
				565					570					575		
Arg	Pro	Val	Val	Ile	Phe	Asn	Phe	Ala	Asn	Ser	Ser	Tyr	Val	Gln	Tyr	
			580					585					590			
Gly	Gln	Leu	Gly	Glu	Asp	Asn	Glu	Ile	Leu	Leu	Gly	Asn	His	Arg	Thr	
		595					600					605				
Glu	Glu	Cys	Gln	Leu	Pro	Ser	Leu	Lys	Ile	Phe	Ile	Ala	Gly	Asn	Ser	
	610					615					620					
Ala	Tyr	Glu	Tyr	Val	Asp	Tyr	Leu	Phe	Lys	Arg	Met	Ile	Asp	Leu	Ser	
625					630					635					640	
Ser	Ile	Ser	Thr	Val	Asp	Ser	Met	Ile	Ala	Leu	Asp	Ile	Asp	Pro	Leu	
				645					650					655		
Glu	Asn	Thr	Asp	Phe	Arg	Val	Leu	Glu	Leu	Tyr	Ser	Gln	Lys	Glu	Leu	
			660					665					670			

ES 2 765 484 T3

Arg Ser Ser Asn Val Phe Asp Leu Glu Glu Ile Met Arg Glu Phe Asn
675 680 685

Ser Tyr Lys Gln Arg Val Lys Tyr Val Glu Asp Lys Val Val Asp Pro
690 695 700

Leu Pro Pro Tyr Leu Lys Gly Leu Asp Asp Leu Met Ser Gly Leu Gly
705 710 715 720

Ala Ala Gly Lys Ala Val Gly Val Ala Ile Gly Ala Val Gly Gly Ala
725 730 735

Val Ala Ser Val Val Glu Gly Val Ala Thr Phe Leu Lys Asn Pro Phe
740 745 750

Gly Ala Phe Thr Ile Ile Leu Val Ala Ile Ala Val Val Ile Ile Thr
755 760 765

Tyr Leu Ile Tyr Thr Arg Gln Arg Arg Leu Cys Thr Gln Pro Leu Gln
770 775 780

Asn Leu Phe Pro Tyr Leu Val Ser Ala Asp Gly Thr Thr Val Thr Ser
785 790 795 800

Gly Ser Thr Lys Asp Thr Ser Leu Gln Ala Pro Pro Ser Tyr Glu Glu
805 810 815

Ser Val Tyr Asn Ser Gly Arg Lys Gly Pro Gly Pro Pro Ser Ser Asp
820 825 830

Ala Ser Thr Ala Ala Pro Pro Tyr Thr Asn Glu Gln Ala Tyr Gln Met
835 840 845

Leu Leu Ala Leu Ala Arg Leu Asp Ala Glu Gln Arg Ala Gln Gln Asn
850 855 860

Gly Thr Asp Ser Leu Asp Gly Arg Thr Gly Thr Gln Asp Lys Gly Gln
865 870 875 880

Lys Pro Asn Leu Leu Asp Arg Leu Arg His Arg Lys Asn Gly Tyr Arg
885 890 895

His Leu Lys Asp Ser Asp Glu Glu Glu Asn Val
900 905

<210> 2
<211> 906
<212> PRT

<213> Citomegalovirus humano

<400> 2

ES 2 765 484 T3

Leu Tyr Arg Glu Thr Cys Asn Leu Asn Cys Met Leu Thr Ile Thr Thr
 245 250 255
 Ala Arg Ser Lys Tyr Pro Tyr His Phe Phe Ala Thr Ser Thr Gly Asp
 260 265 270
 Val Val Tyr Ile Ser Pro Phe Tyr Asn Gly Thr Asn Arg Asn Ala Ser
 275 280 285
 Tyr Phe Gly Glu Asn Ala Asp Lys Phe Phe Ile Phe Pro Asn Tyr Thr
 290 295 300
 Ile Val Ser Asp Phe Gly Arg Pro Asn Ala Ala Pro Glu Thr His Arg
 305 310 315 320
 Leu Val Ala Phe Leu Glu Arg Ala Asp Ser Val Ile Ser Trp Asp Ile
 325 330 335
 Gln Asp Glu Lys Asn Val Thr Cys Gln Leu Thr Phe Trp Glu Ala Ser
 340 345 350
 Glu Arg Thr Ile Arg Ser Glu Ala Glu Asp Ser Tyr His Phe Ser Ser
 355 360 365
 Ala Lys Met Thr Ala Thr Phe Leu Ser Lys Lys Gln Glu Val Asn Met
 370 375 380
 Ser Asp Ser Ala Leu Asp Cys Val Arg Asp Glu Ala Ile Asn Lys Leu
 385 390 395 400
 Gln Gln Ile Phe Asn Thr Ser Tyr Asn Gln Thr Tyr Glu Lys Tyr Gly
 405 410 415
 Asn Val Ser Val Phe Glu Thr Ser Gly Gly Leu Val Val Phe Trp Gln
 420 425 430
 Gly Ile Lys Gln Lys Ser Leu Val Glu Leu Glu Arg Leu Ala Asn Arg
 435 440 445
 Ser Ser Leu Asn Ile Thr His Arg Thr Arg Arg Ser Thr Ser Asp Asn
 450 455 460
 Asn Thr Thr His Leu Ser Ser Met Glu Ser Val His Asn Leu Val Tyr
 465 470 475 480
 Ala Gln Leu Gln Phe Thr Tyr Asp Thr Leu Arg Gly Tyr Ile Asn Arg
 485 490 495

ES 2 765 484 T3

Ala Leu Ala Gln Ile Ala Glu Ala Trp Cys Val Asp Gln Arg Arg Thr
500 505 510

Leu Glu Val Phe Lys Glu Leu Ser Lys Ile Asn Pro Ser Ala Ile Leu
515 520 525

Ser Ala Ile Tyr Asn Lys Pro Ile Ala Ala Arg Phe Met Gly Asp Val
530 535 540 545

Leu Gly Leu Ala Ser Cys Val Thr Ile Asn Gln Thr Ser Val Lys Val
545 550 555 560

Leu Arg Asp Met Asn Val Lys Glu Ser Pro Gly Arg Cys Tyr Ser Arg
565 570 575

Pro Val Val Ile Phe Asn Phe Ala Asn Ser Ser Tyr Val Gln Tyr Gly
580 585 590

Gln Leu Gly Glu Asp Asn Glu Ile Leu Leu Gly Asn His Arg Thr Glu
595 600 605

Glu Cys Gln Leu Pro Ser Leu Lys Ile Phe Ile Ala Gly Asn Ser Ala
610 615 620

Tyr Glu Tyr Val Asp Tyr Leu Phe Lys Arg Met Ile Asp Leu Ser Ser
625 630 635 640

Ile Ser Thr Val Asp Ser Met Ile Ala Leu Asp Ile Asp Pro Leu Glu
645 650 655

Asn Thr Asp Phe Arg Val Leu Glu Leu Tyr Ser Gln Lys Glu Leu Arg
660 665 670

Ser Ser Asn Val Phe Asp Leu Glu Glu Ile Met Arg Glu Phe Asn Ser
675 680 685

Tyr Lys Gln Arg Val Lys Tyr Val Glu Asp Lys Val Val Asp Pro Leu
690 695 700

Pro Pro Tyr Leu Lys Gly Leu Asp Asp Leu Met Ser Gly Leu Gly Ala
705 710 715 720

Ala Gly Lys Ala Val Gly Val Ala Ile Gly Ala Val Gly Gly Ala Val
725 730 735

Ala Ser Val Val Glu Gly Val Ala Thr Phe Leu Lys Asn Pro Phe Gly
740 745 750

ES 2 765 484 T3

Ala Phe Thr Ile Ile Leu Val Ala Ile Ala Val Val Ile Ile Thr Tyr
755 760 765

Leu Ile Tyr Thr Arg Gln Arg Arg Leu Cys Thr Gln Pro Leu Gln Asn
770 775 780

Leu Phe Pro Tyr Leu Val Ser Ala Asp Gly Thr Thr Val Thr Ser Gly
785 790 795 800

Ser Thr Lys Asp Thr Ser Leu Gln Ala Pro Pro Ser Tyr Glu Glu Ser
805 810 815

Val Tyr Asn Ser Gly Arg Lys Gly Pro Gly Pro Pro Ser Ser Asp Ala
820 825 830

Ser Thr Ala Ala Pro Pro Tyr Thr Asn Glu Gln Ala Tyr Gln Met Leu
835 840 845

Leu Ala Leu Ala Arg Leu Asp Ala Glu Gln Arg Ala Gln Gln Asn Gly
850 855 860

Thr Asp Ser Leu Asp Gly Gln Thr Gly Thr Gln Asp Lys Gly Gln Lys
865 870 875 880

Pro Asn Leu Leu Asp Arg Leu Arg His Arg Lys Asn Gly Tyr Arg His
885 890 895

Leu Lys Asp Ser Asp Glu Glu Glu Asn Val
900 905

<210> 3
<211> 20
<212> PRT
5 <213> Secuencia artificial

<220>
<221> fuente
<223> /nota="Descripción de la secuencia artificial: Péptido sintético"
<400> 3

Asp Trp Leu Lys Ala Phe Tyr Asp Lys Val Ala Glu Lys Leu Lys Glu
1 5 10 15

Ala Phe Leu Ala
20

10

<210> 4
<211> 18
<212> PRT
<213> Secuencia artificial

15

<220>

<221> fuente
 <223> /nota="Descripción de la secuencia artificial: Péptido sintético"

Glu Leu Leu Glu Lys Trp Lys Glu Ala Leu Ala Ala Leu Ala Glu Lys
 1 5 10 15

Leu Lys

Leu Lys

5 <210> 5
 <211> 18
 <212> PRT
 <213> Secuencia artificial

10 <220>
 <221> fuente
 <223> /nota="Descripción de la secuencia artificial: Péptido sintético"

<400> 5
 Phe Trp Leu Lys Ala Phe Tyr Asp Lys Val Ala Glu Lys Leu Lys Glu
 1 5 10 15

Ala Phe

15 <210> 6
 <211> 29
 <212> PRT
 <213> Secuencia artificial

20 <220>
 <221> fuente
 <223> /nota="Descripción de la secuencia artificial: Péptido sintético"

<400> 6
 Asp Trp Leu Lys Ala Phe Tyr Asp Lys Val Ala Glu Lys Leu Lys Glu
 1 5 10 15

Ala Phe Arg Leu Thr Arg Lys Arg Gly Leu Lys Leu Ala
 20 25

25 <210> 7
 <211> 18
 <212> PRT
 <213> Secuencia artificial

30 <220>
 <221> fuente
 <223> /nota="Descripción de la secuencia artificial: Péptido sintético"

<400> 7
 Asp Trp Leu Lys Ala Phe Tyr Asp Lys Val Ala Glu Lys Leu Lys Glu
 1 5 10 15

Ala Phe

35 <210> 8
 <211> 6
 <212> PRT
 <213> Secuencia artificial

ES 2 765 484 T3

<220>
<221> fuente
<223> /nota="Descripción de la secuencia artificial: Marcador sintético 6xHis"

<400> 8

His His His His His His
1 5

5

<210> 9
<211> 8
<212> PRT
<213> Secuencia artificial

10

<220>
<221> fuente
<223> /nota="Descripción de la secuencia artificial: Péptido sintético"

<400> 9

Asp Tyr Lys Asp Asp Asp Asp Lys
1 5

15

<210> 10
<211> 9
<212> PRT
<213> Secuencia artificial

20

<220>
<221> fuente
<223> /nota="Descripción de la secuencia artificial: Péptido sintético"

<400> 10

Ala Trp Arg His Pro Gln Phe Gly Gly
1 5

25

<210> 11
<211> 8
<212> PRT
<213> Secuencia artificial

30

<220>
<221> fuente
<223> /nota="Descripción de la secuencia artificial: Péptido sintético"

<400> 11

Trp Ser His Pro Gln Phe Glu Lys
1 5

35

<210> 12
<211> 5
<212> PRT
<213> Citomegalovirus humano

<400> 12

Arg Gly Ser Thr Trp
1 5

40

<210> 13
<211> 8
<212> PRT
<213> Citomegalovirus humano

ES 2 765 484 T3

<400> 13

Arg Gly Ser Thr Asn Gly Thr Trp
1 5

<210> 14

<211> 4

5 <212> PRT

<213> Citomegalovirus humano

<400> 14

Trp Leu Tyr Arg
1

<210> 15

10 <211> 7

<212> PRT

<213> Citomegalovirus humano

<400> 15

Trp Leu Tyr Asn Gly Thr Arg
1 5

<210> 16

15 <211> 5

<212> PRT

<213> Citomegalovirus humano

<400> 16

Asn Gly Ser Thr Trp
1 5

20

<210> 17

<211> 5

<212> PRT

<213> Citomegalovirus humano

25

<400> 17

Arg Asn Ser Thr Trp
1 5

<210> 18

30 <211> 5

<212> PRT

<213> Citomegalovirus humano

<400> 18

Glu Gly Glu Thr Trp
1 5

<210> 19

35 <211> 5

<212> PRT

<213> Citomegalovirus humano

<400> 19

Glu Gly Glu Glu Trp
1 5

ES 2 765 484 T3

<210> 20
<211> 4
<212> PRT
<213> Citomegalovirus humano

5 <400> 20

Arg Thr Lys Arg
1

<210> 21
<211> 4
<212> PRT
<213> Citomegalovirus humano

10 <400> 21

Arg Gln Arg Arg
1

REIVINDICACIONES

1. Una proteína gB recombinante de citomegalovirus (CMV), en la que
 - (i) dicha proteína gB no comprende un dominio transmembrana (TM); y
 - (ii) dicha proteína gB comprende una región de superficie hidrofóbica formada por los aminoácidos 154-160 y 236-243 de la secuencia que se expone en la SEQ ID NO: 1, o por restos de aminoácido correspondientes en proteínas gB de otros CMV que se identifican mediante alineamiento de la secuencia en cuestión contra la SEQ ID NO: 1, que comprende una mutación que da como resultado un sitio de glicosilación dentro de dicha región.
2. La proteína gB recombinante de la reivindicación 1, en la que dicho sitio de glicosilación es un sitio de N-glicosilación que comprende un motivo N-X-S/T, en el que X es cualquier resto de aminoácido excepto prolina.
3. La proteína gB recombinante de la reivindicación 1 o 2, en la que dicha mutación se selecciona de entre el grupo que consiste en (i) R236N, (ii) G237N, (iii) T158N, (iv) W240N y Y242S, (v) W240N e Y242T, y una combinación de las mismas.
4. La proteína gB recombinante de una cualquiera de las reivindicaciones 1-3 en la que dicha mutación comprende una inserción de la secuencia N-X-S/T, en la que X es cualquier resto de aminoácido excepto prolina, en el bucle de fusión 1 (FL1) que consiste en los restos de aminoácido 155-157 de la secuencia que se expone en la SEQ ID NO: 1, o en restos de aminoácido correspondientes en proteínas gB de otros CMV que se identifican mediante alineamiento de la secuencia en cuestión contra la SEQ ID NO: 1, en el bucle de fusión 2 (FL2) que consiste en los restos de aminoácido 240-242 de la secuencia que se expone en la SEQ ID NO: 1, o en restos de aminoácido correspondientes en proteínas gB de otros CMV que se identifican mediante alineamiento de la secuencia en cuestión contra la SEQ ID NO: 1, o en ambos.
5. La proteína gB recombinante de una cualquiera de las reivindicaciones 1-4, que comprende adicionalmente una mutación que da como resultado una reducción del índice global de hidrofobicidad de dicha superficie hidrofóbica.
6. La proteína gB recombinante de la reivindicación 5, en la que la mutación está en 1156, H157, R236, S238, T239, W240, Y242, o una combinación de las mismas.
7. Una proteína gB recombinante de citomegalovirus (CMV), en la que
 - (i) dicha proteína gB no comprende un dominio transmembrana (TM); y
 - (ii) dicha proteína gB comprende una mutación que da como resultado un sitio de glicosilación, en la que dicho sitio de glicosilación está (1) dentro de una superficie hidrofóbica formada por los aminoácidos 145-167 y 230-252 de la secuencia que se expone en la SEQ ID NO: 1, o por restos de aminoácido correspondientes en proteínas gB de otros CMV que se identifican mediante alineamiento de la secuencia en cuestión contra la SEQ ID NO: 1; o (2) en un resto que esté dentro de los restos de aminoácido 696-698 de la secuencia que se expone en la SEQ ID NO: 1.
8. La proteína gB recombinante de la reivindicación 7, en la que dicho sitio de glicosilación es un sitio de N-glicosilación que comprende un motivo N-X-S/T, en el que X es cualquier resto de aminoácido excepto prolina.
9. Una composición inmunogénica que comprende la proteína gB recombinante de una cualquiera de las reivindicaciones 1-8 y opcionalmente un adyuvante.
10. La composición inmunogénica de la reivindicación 9 para su uso en la inducción de una respuesta inmunitaria contra el citomegalovirus (CMV).

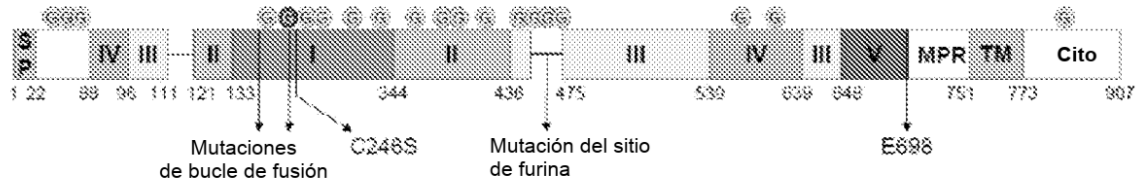


Figura 1

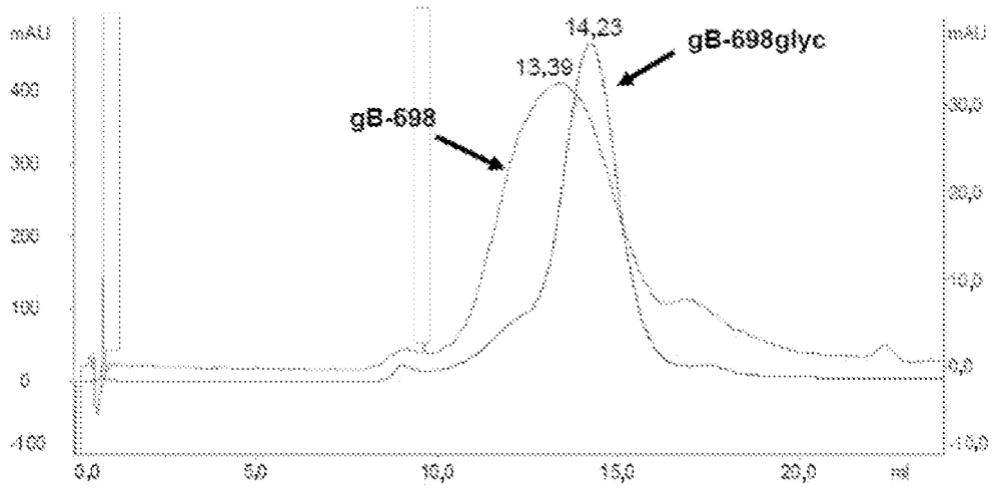


Figura 2

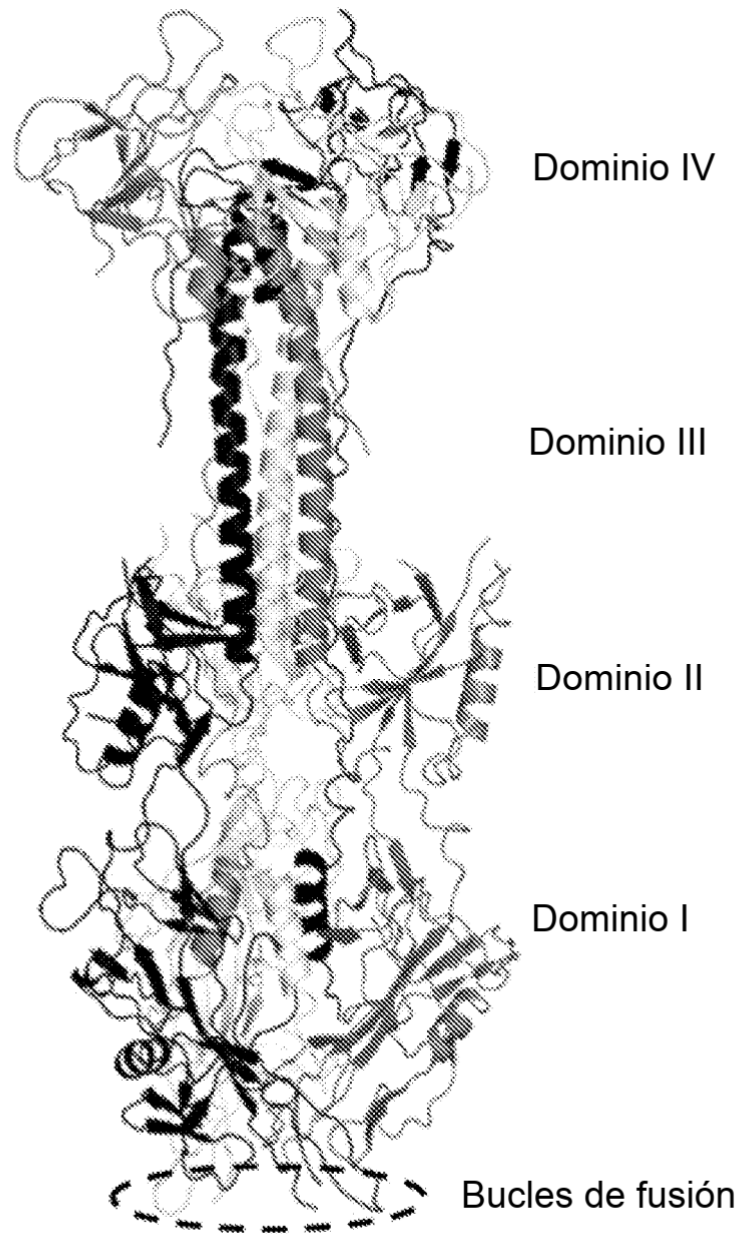


Figura 3

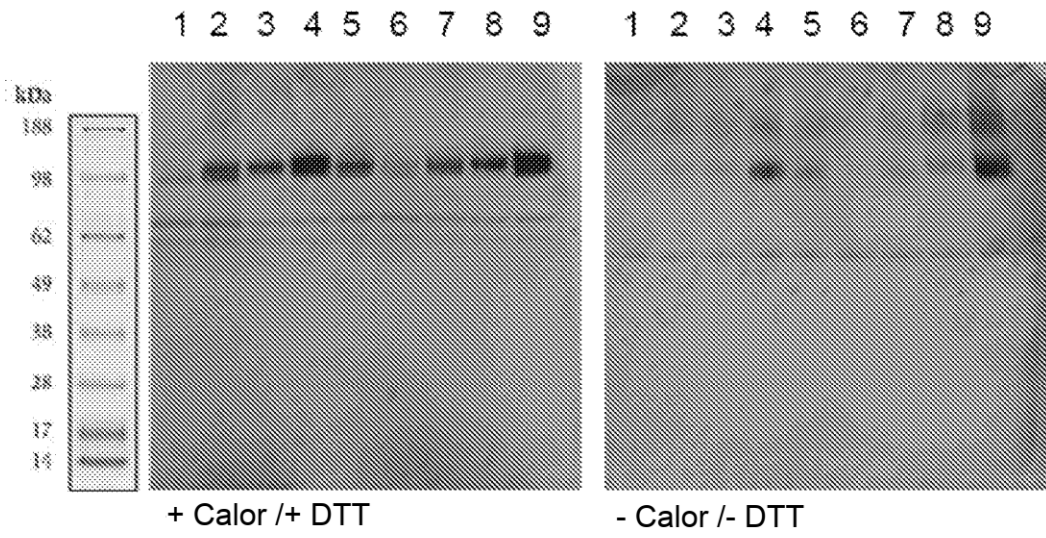


Figura 4