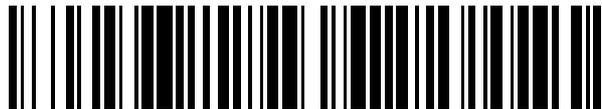


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 799 735**

51 Int. Cl.:

**C07K 14/005** (2006.01)

**A61K 39/245** (2006.01)

**C12N 5/071** (2010.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **29.10.2015 PCT/IB2015/058349**

87 Fecha y número de publicación internacional: **06.05.2016 WO16067239**

96 Fecha de presentación y número de la solicitud europea: **29.10.2015 E 15791768 (3)**

97 Fecha y número de publicación de la concesión europea: **25.03.2020 EP 3212660**

54 Título: **Células de mamífero que expresan antígenos de citomegalovirus**

30 Prioridad:

**31.10.2014 EP 14191385**

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

**21.12.2020**

73 Titular/es:

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

72 Inventor/es:

**CARFI, ANDREA;  
CIFERRI, CLAUDIO;  
HOFMANN, IRMGARD;  
LAUX, HOLGER;  
LILJA, ANDERS y  
WEN, YINGXIA**

74 Agente/Representante:

**GONZÁLEZ PECES, Gustavo Adolfo**

ES 2 799 735 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**

Células de mamífero que expresan antígenos de citomegalovirus

**Campo de la invención**

5 La presente invención se refiere a células hospedadoras que expresan proteínas de citomegalovirus (CMV) adecuadas para usos de vacuna.

**Antecedentes de la invención**

10 El citomegalovirus es un género de virus que pertenece a la familia vírica conocida como Herpesviridae o herpesvirus. La especie que infecta a los humanos se conoce comúnmente como citomegalovirus humano (HCMV) o herpesvirus-5 humano (HHV-5). Dentro de Herpesviridae, el HCMV pertenece a la subfamilia Betaherpesvirinae, que también incluye citomegalovirus de otros mamíferos.

15 Aunque pueden encontrarse en todo el cuerpo, las infecciones por HCMV se asocian frecuentemente a las glándulas salivales. El HCMV infecta entre el 50 % y el 80 % de los adultos en los Estados Unidos (40 % en todo el mundo), como lo indica la presencia de anticuerpos en gran parte de la población general. La infección por HCMV suele pasar desapercibida en personas sanas, pero puede ser mortal para los inmunocomprometidos, tales como las personas infectadas por el VIH, receptores de trasplantes de órganos o recién nacidos. El HCMV es el virus que se transmite con mayor frecuencia a un feto en desarrollo. Después de la infección, el HCMV tiene la capacidad de permanecer latente dentro del cuerpo durante toda la vida del hospedador, con reactivaciones ocasionales por latencia. Dada la gravedad e importancia de esta enfermedad, obtener una vacuna eficaz se considera una prioridad de salud pública (Sung, H., y col., (2010) Expert review of vaccines 9, 1303-1314; Schleiss, Expert Opin Ther Pat. Abr 2010; 20(4): 597-602).

20 Se han secuenciado los genomas de más de 20 cepas diferentes de HCMV, incluyendo aquellos de cepas de laboratorio y aislados clínicos. Por ejemplo, se han secuenciado las siguientes cepas de HCMV: Towne (GL239909366), AD169 (GI:219879600), Toledo (GL290564358) y Merlin (GI: 155573956). Las cepas de HCMV AD169, Towne y Merlin pueden obtener de la American Type Culture Collection (ATCC VR538, ATCC VR977 y ATCC VR1590, respectivamente).

30 El CMV contiene un número desconocido de complejos de proteínas de membrana. De las aproximadamente 30 glucoproteínas conocidas en la envoltura vírica, gH y gL han surgido como particularmente interesantes debido a su presencia en varios complejos diferentes: gH/gL dimérico, gH/gL/gO trimérico (también conocido como el complejo gCIII) y el gH/gL/pUL128/pUL130/pUL131 pentamérico (pUL131 también denominado "pUL131A", "pUL131a" o "UL131A"; las subunidades pUL128, pUL130 y pUL131 a veces también se denominan UL128, UL130, UL131). Se cree que el CMV usa los complejos pentaméricos para ingresar a las células epiteliales y endoteliales por endocitosis y fusión dependiente de pH bajo, pero se cree que ingresa a los fibroblastos por fusión directa en la membrana plasmática en un proceso que implica gH/gL o posiblemente gH/gL/gO. El complejo o complejos gH/gL y/o gH/gL/gO es/son suficientes para la infección por fibroblastos, mientras que el complejo pentamérico es necesario para infectar células endoteliales y epiteliales.

40 El complejo pentamérico se considera una diana principal para la vacunación contra el CMV. Los genes víricos UL128, UL130 y UL131 son necesarios para la entrada endotelial (Hahn, Journal of Virology 2004; 78:10023-33). Las cepas tropicales no endoteliales adaptadas a fibroblastos contienen mutaciones en al menos uno de estos tres genes. La cepa Towne, por ejemplo, contiene una inserción de 2 pares de bases que causa un desplazamiento de marco en el gen UL130, mientras que AD169 contiene una inserción de 1 par de bases en el gen UL131. Tanto Towne como AD169 podrían adaptarse para el crecimiento en células endoteliales y, en ambos casos, se repararon las mutaciones de desplazamiento de marco en los genes UL130 o UL131.

45 El documento US7704510 desvela que se requiere pUL131A para el tropismo de células epiteliales. El documento US7704510 también desvela que pUL128 y pUL130 forman un complejo con gH/gL, que se incorpora a los viriones. Este complejo se requiere para infectar células endoteliales y epiteliales, pero no fibroblastos. Se descubrió que los anticuerpos anti-CD46 inhiben la infección por HCMV de las células epiteliales.

50 Las vacunas contra el CMV probadas en ensayos clínicos incluyen la vacuna de Towne, quimeras Towne-Toledo, un replicón de virus alfa con gB como el antígeno, vacuna gB/MF59, una vacuna gB producida por GlaxoSmithKline y una vacuna de ADN que usa gB y pp65. pp65 es una proteína vírica que es un potente inductor de respuestas CD8+ dirigidas contra el CMV. Todas estas vacunas son inductores pobres de anticuerpos que bloquean la entrada vírica en las células endoteliales/epiteliales (Adler, SP (2013), British Medical Bulletin, 107, 57-68.doi:10.1093/bmb/ldt023).

55 Generalmente se cree que los anticuerpos neutralizantes contra el complejo pentamérico (gH/gLpUL128/pUL130/pUL131) serán significativamente más potentes que los anticuerpos neutralizantes generados contra la subunidad gB del CMV o el complejo dimérico gH/gL. Por lo tanto, para desarrollar una vacuna eficaz contra el CMV, existe una necesidad urgente de producir grandes cantidades (por ejemplo, escalas comerciales) de complejo pentamérico CMV.

Sin embargo, la producción recombinante de complejo pentamérico del CMV sigue siendo un desafío. Las cinco subunidades deben expresarse (preferentemente en una cantidad sustancialmente igual y preferentemente en un período de tiempo sostenido), plegarse correctamente y ensamblarse adecuadamente en un pentámero. Además, también es necesario evitar el ensamblaje no deseado de complejos contaminantes (tales como el dímero gH/gL, el tetrámero gH/gL (dos copias de gH y dos copias de gL), el tetrámero gH/gLpUL128/pUL130, etc.).

La expresión recombinante tal como un complejo de proteínas eucariotas requeriría la identificación de construcciones adecuadas y hospedadores adecuados, para la expresión de cantidades suficientes de proteínas plegadas adecuadamente en un período de tiempo sostenido, que después puede ensamblarse adecuadamente en complejos de proteínas. Adicionalmente, la selección de un hospedador de expresión adecuado tiene un impacto significativo en el rendimiento y la calidad de las proteínas, así como también sobre los costes reales del procedimiento de producción.

### **Sumario de la invención**

En el presente documento se desvelan y ejemplifican células hospedadoras de mamífero, en particular células CHO, en las que la secuencia o secuencias que codifican proteínas del CMV gH, gL, pUL128, pUL130, pUL131 (o un fragmento que forma complejo del mismo) se integran de manera estable en el genoma. Dichas células hospedadoras proporcionan una fuente confiable en la que el complejo pentamérico gH/gL/pUL128/pUL130/pUL131 puede producirse de forma recombinante.

En un aspecto, la invención proporciona una célula CHO recombinante, que comprende: una o más secuencias de polinucleótidos que codifican el complejo pentamérico de citomegalovirus (CMV), en la que dicho complejo pentamérico comprende: gH o un fragmento formador del complejo pentamérico gH/gL/pUL128/pUL130/pUL131 del mismo, gL o un fragmento formador del complejo pentamérico gH/gL/pUL128/pUL130/pUL131 del mismo, pUL128 o un fragmento formador del complejo pentamérico gH/gL/pUL128/pUL130/pUL131 del mismo, pUL130 o un fragmento formador del complejo pentamérico gH/gL/pUL128/pUL130/pUL131 del mismo y pUL131 o un fragmento formador del complejo pentamérico gH/gL/pUL128/pUL130/pUL131 del mismo; en la que dicha una o más secuencias de polinucleótidos están integradas en el ADN genómico de dicha célula CHO; y en la que el nivel de expresión o actividad de la proteína C12orf35 se reduce en dicha célula CHO, en comparación con una célula de tipo silvestre. Debido a que las secuencias de polinucleótidos están integradas de manera estable en el ADN genómico de la célula, pueden transmitirse a su progenie como parte de las secuencias genómicas. Dichas células a menudo se denominan en la técnica líneas celulares estables. Particularmente, la célula no comprende las secuencias víricas de CMV que podrían dar como resultado la producción de virus de CMV infecciosos. Cuando se cultiva en una condición adecuada, dicho complejo pentamérico del CMV se expresa por dicha célula hospedadora.

Las células CHO adecuadas incluyen, por ejemplo, cualquier línea celular CHO disponible en la American Type Culture Collection (ATCC) o en la European Collection of Cell Cultures (ECACC). Las líneas celulares CHO ejemplares incluyen, por ejemplo, célula CHO-K1, CHO-DUXB11, CHO-DG44 o células CHO-S. Para facilitar la selección de clones integrados de forma estable, las células hospedadoras pueden ser deficientes en dihidrofolato reductasa (DHFR). La expresión de proteínas de recombinación en líneas celulares deficientes en DHFR o competentes en DHFR también puede seleccionarse mediante selección con metotrexato (MTX).

La célula hospedadora puede tener una o más modificaciones adicionales para mejorar aún más la producción de complejo pentamérico de CMV. En determinadas realizaciones, el nivel de expresión o actividad de la proteína FAM60A se reduce en la célula hospedadora, en comparación con un control. En determinadas realizaciones, el nivel de expresión o actividad de la matriptasa se reduce en la célula hospedadora, en comparación con un control. Las modificaciones descritas en el presente documento pueden usarse de forma individual o en cualquier combinación.

El complejo pentamérico CMV producido de forma recombinante puede ser soluble (por ejemplo, carece del dominio transmembrana de gH). Para facilitar la producción, el complejo pentamérico CMV producido de forma recombinante puede secretarse de la célula hospedadora al medio de cultivo.

Las células hospedadoras de mamífero descritas en el presente documento son particularmente adecuadas para la producción a gran escala, tales como cultivos que tienen al menos 20 litros (por ejemplo, 50 litros, 100 litros, etc.) de tamaño. En determinadas realizaciones, el rendimiento del complejo pentamérico de CMV es de al menos 0,05 g/l o de al menos 0,1 g/l.

### **Breve descripción de las figuras**

La FIGURA 1 ilustra diversas estrategias de diseño para la coexpresión de cinco componentes del pentámero.

La FIGURA 2 muestra los mapas plasmídicos de los vectores de expresión usados para la transfección de células CHO.

La FIGURA 3 muestra el análisis SDS-PAGE de pentámero purificado producido con los 12 clones principales.

La FIGURA 4 ilustra los diseños experimentales utilizados para estudiar la estabilidad de los 12 clones principales.

La FIGURA 5 muestra el rendimiento de pentámero producido por el clon principal VF7.

### **Descripción detallada de la invención**

#### **1. VISIÓN GENERAL**

5 Una de las dianas principales para la vacunación contra el CMV es el complejo pentamérico (gH/gL/pUL128/pUL130/pUL131). Para producir pentámero de CMV recombinante a escala comercial, existe la necesidad de identificar construcciones adecuadas y hospedadores adecuados, de tal manera que las cinco subunidades del pentámero puedan expresarse en cantidades suficientes durante un período de tiempo sostenido y puedan ensamblarse adecuadamente en un complejo pentamérico. El pentámero de CMV es una diana primaria de anticuerpos neutralizantes contra el CMV humano.

10 El pentámero de CMV se ha producido de forma recombinante utilizando HEK293 como células hospedadoras. Véase, el documento WO2014/005959. Sin embargo, La línea celular HEK293 es generalmente reconocida en la técnica como una de las mejores líneas celulares hospedadoras para la expresión transitoria de genes. Véase, Meyer y col., PLoS One, 17 jul 2013, 8(7):e68674. doi: 10.1371/journal.pone.0068674. Sin embargo, sería difícil mantener la expresión del pentámero durante un período de tiempo sostenido en células HEK293 transfectadas transitoriamente. Además, 15 en los ejemplos desvelados en el documento WO 2014/005959, las cinco subunidades del pentámero se introdujeron en las células HEK293 a través de cinco plásmidos diferentes. Bajo tal configuración, sería difícil mantener la expresión de cada una de las subunidades del pentámero en un nivel sustancialmente igual. Una cantidad desigual de las subunidades puede reducir la eficiencia del ensamblaje del pentámero, especialmente dado que estas subunidades ya tienen tendencia a formar complejos contaminantes, tales como el dímero gH/gL, el tetrámero gH/gL (2 copias de gH y 2 copias de gL), el tetrámero gH/gLpUL128/pUL130. Además, el plásmido que codifica gH se seleccionó con neomicina, los cuatro plásmidos que codifican gL, pUL128, pUL130 y pUL131, respectivamente, se seleccionaron todos con kanamicina. Por lo tanto, debido a que los cuatro plásmidos se seleccionaron con kanamicina, la pérdida de un solo plásmido en una célula hospedadora no se podría detectar fácilmente, pero la pérdida de un solo plásmido impediría el ensamblaje del pentámero.

25 Como se desvela y ejemplifica en el presente documento, los presentes inventores superaron las dificultades de la producción recombinante del pentámero de CMV mediante el uso de células de ovario de hámster chino (CHO). En estas células CHO, las secuencias que codifican gH, gL, pUL128, pUL130, pUL131 (o un fragmento formador de complejos de las mismas) se integran de manera estable en el genoma de las células CHO. Al integrar secuencias codificantes del pentámero en el genoma de la célula CHO, los presentes inventores lograron una expresión genómica estable de pentámero recombinante, con una alta eficiencia y estabilidad genómica. Como se ejemplifica en la sección de Ejemplos, se descubrió que, en comparación con las células HEK293 transfectadas transitoriamente, las líneas estables de CHO lograron consistentemente un rendimiento 100 veces más alto, haciendo a estas líneas celulares CHO particularmente adecuadas para la producción comercial de pentámero de CMV. La integración estable también permite una fácil manipulación del cultivo celular a gran escala, en comparación con las células HEK transfectadas transitoriamente con cinco plásmidos exógenos. De hecho, los clones principales produjeron pentámero con un rendimiento tan alto como 0,1 g/l a 0,5 g/l.

En el presente documento también se proporcionan mejoras adicionales de las células hospedadoras para la producción de pentámeros de CMV. En particular, se ejemplifican tres células hospedadoras modificadas: (i) células hospedadoras en las que se reduce el nivel de expresión o la actividad de la proteína C12orf35, en comparación con un control; (ii) células hospedadoras en las que se reduce el nivel de expresión o la actividad de la proteína FAM60A, en comparación con un control; (iii) células hospedadoras en las que se reduce el nivel de expresión o la actividad de la matriptasa, en comparación con un control. Se descubrió que la reducción del nivel de expresión o actividad de la proteína C12orf35 o la proteína FAM60A da como resultado un aumento significativo en el nivel de expresión de una proteína recombinante. También se descubrió que la reducción del nivel de expresión o actividad de la matriptasa disminuye significativamente la degradación proteolítica ("recorte") de una proteína recombinante. Estas modificaciones pueden usarse de forma individual o en cualquier combinación.

En consecuencia, se describen en el presente documento células hospedadoras de mamíferos, en particular células CHO, en las que la secuencia o secuencias polinucleotídicas que codifican el complejo pentamérico de CMV que comprende gH, gL, pUL128, pUL130, pUL131 (o un fragmento que forma complejo del mismo) se integran de manera estable en el ADN genómico. Cuando se cultiva en una condición adecuada, dicho complejo pentamérico del CMV se expresa por dichas células hospedadoras.

También se describe en el presente documento un complejo pentamérico de citomegalovirus (CMV) producido por las células de mamífero descritas en el presente documento.

55 También se proporciona en el presente documento un procedimiento para producir el complejo pentamérico de citomegalovirus (CMV), en el que dicho complejo pentamérico comprende: gH o un fragmento formador de complejos del mismo, gL o un fragmento formador de complejos del mismo, pUL128 o un fragmento formador de complejos del mismo, pUL130 o un fragmento formador de complejos del mismo y pUL131 o un fragmento formador de complejos del mismo, que comprende: (i) cultivar la célula de mamífero como se describe en el presente documento en una

condición adecuada, expresando de ese modo dicho complejo pentamérico; y cosechar dicho complejo pentamérico del cultivo. El complejo pentamérico puede comprender además estar purificado. También se describe en el presente documento un complejo pentamérico de citomegalovirus (CMV) producido mediante este procedimiento.

5 También se describe en el presente documento una composición que comprende el complejo pentamérico descrito en el presente documento. La composición puede comprender un complejo pentamérico de CMV purificado que sea adecuado para la administración *in vivo*. Por ejemplo, el complejo pentamérico en una composición tal puede tener una pureza de al menos el 85 %, al menos el 86 %, al menos el 87 %, al menos el 88 %, al menos el 89 %, al menos el 90 %, al menos el 91 %, al menos el 92 %, al menos el 93 %, al menos el 94 %, al menos el 95 %, al menos el 96 %, al menos el 97 %, al menos el 98 % o al menos el 99 %, en masa. La composición puede comprender además un adyuvante, tales como una sal de aluminio o MF59.

También se describe en el presente documento una composición para su uso en la inducción de una respuesta inmunitaria contra el CMV. También se describe el uso de la composición descrita en el presente documento para inducir una respuesta inmunitaria contra el CMV y el uso de la composición descrita en el presente documento en la fabricación de un medicamento para inducir una respuesta inmune contra el CMV.

## 15 2. COMPLEJOS PENTAMÉRICOS DE CMV Y SECUENCIAS CODIFICANTES

### A. Complejos pentaméricos del CMV

En un aspecto de la divulgación, se describe una célula hospedadora de mamífero que expresa el complejo pentamérico del CMV, en la que dicho complejo pentamérico comprende (i) gH o un fragmento formador de complejos del mismo, (ii) gL o un fragmento formador de complejos del mismo, (iii) pUL128 o un fragmento formador de complejos del mismo, (iv) pUL130 o un fragmento formador de complejos del mismo y (v) pUL131 o un fragmento formador de complejos del mismo. Las secuencias de polinucleótidos que codifican el complejo pentamérico CMV se integran en el ADN genómico de la célula hospedadora y, cuando se cultiva en una condición adecuada, dicho complejo pentamérico del CMV se expresa por dicha célula hospedadora.

En determinadas realizaciones, dicho complejo pentamérico es soluble. Puede obtenerse el complejo pentamérico soluble, por ejemplo, mediante el uso de un fragmento de gH en el que se elimina el dominio transmembrana de la subunidad gH, como se describe en detalle a continuación.

En determinadas realizaciones, dicho complejo pentamérico se secreta por la célula hospedadora. Se ha informado que la presencia de las cinco subunidades, gH, gL, pUL128, pUL131 y pUL131, es suficiente para el ensamblaje del complejo pentamérico en ER antes de exportarlo al aparato de Golgi. Véase, Ryckman y col., *J Virol.* Ene 2008; 82(1): 60-70. Alternativamente o además, puede usarse un péptido señal apropiado en una o más de las cinco subunidades (por ejemplo, haciendo una proteína de fusión con una señal secretora). Las secuencias señal (y el casete de expresión) para producir proteínas secretoras se conocen en la técnica. En general, los péptidos líderes tienen una longitud de 5-30 aminoácidos y típicamente están presentes en el extremo N de una proteína recién sintetizada. El núcleo del péptido señal generalmente contiene un largo tramo de aminoácidos hidrófobos que tiende a formar una única hélice alfa. Además, muchos péptidos señal comienzan con un corto tramo de aminoácidos con carga positiva, lo que puede ayudar a hacer cumplir la topología apropiada del polipéptido durante la translocación. Al final del péptido señal típicamente hay un tramo de aminoácidos que es reconocido y escindido por la peptidasa señal. La peptidasa señal puede escindir durante o después de la finalización de la translocación para generar un péptido señal libre y una proteína madura.

La glucoproteína H (gH) del CMV humano, codificada por el gen UL75, es una glucoproteína del virión que es esencial para la infectividad y que se conserva entre los miembros de los virus del herpes alfa, beta y gamma. Puede formar un complejo estable con gL y la formación de este complejo facilita la expresión de gH en la superficie celular. Basado en las estructuras cristalinas de los complejos gH/gL del HSV-2 y el EBV, la subunidad gL y los restos N-terminales de gH forman un dominio globular en un extremo de la estructura (la "cabeza"), que está implicado en las interacciones con gB y la activación de la fusión de membrana. El dominio C-terminal de gH, proximal a la membrana vírica (la "cola"), también está implicado en la fusión de membrana. gH muestra determinantes que son reconocidos por el factor del hospedador TLR2 e interactúa directamente con un heterodímero formado entre los factores del hospedador TLR2 y TLR1. TLR2 media la activación de NF- $\kappa$ B y las respuestas inflamatorias de citocinas de las células.

La gH de la cepa Merlin de CMV se muestra como SEQ ID NO: 1 (GI:52139248, 742 restos de aminoácidos). El gH de la cepa Towne de CMV se muestra como SEQ ID NO: 2 (GI:138314, también 742 restos de aminoácidos). La gH de Towne se ha caracterizado por tener: (i) seis sitios de N-glucosilación (en los restos 55, 62, 67, 192, 641 y 700); (ii) una secuencia señal hidrófoba en su extremo N (restos de aminoácidos 1-23); (iii) un ectodominio (restos 24-717) que se proyecta fuera de la célula hacia el espacio extracelular; (iv) un dominio transmembrana hidrófobo (TM) (restos 718-736); y (v) un dominio citoplasmático C-terminal (restos 737-742). La SEQ ID NO: 2 comparte el 99 % y el 96 % de identidad de secuencia de aminoácidos con SEQ ID NO: 1 y la gH de la cepa CMV AD169 (GI:138313, SEQ ID NO: 3), respectivamente.

Habitualmente, la secuencia señal N-terminal de la proteína gH de longitud completa se escinde por una peptidasa señal de la célula hospedadora para producir una proteína gH madura. Como tal, la proteína gH expresada por la

célula hospedadora descrita en el presente documento puede carecer de la secuencia señal N-terminal (por ejemplo, gH está codificada por una secuencia de nucleótidos que carece de la secuencia de codificación para la secuencia señal N-terminal).

5 También se describen en el presente documento fragmentos formadores de complejos de gH, tales como un fragmento de gH que carece del dominio transmembrana (TM) (por ejemplo, restos 718-736 de SEQ ID NO: 2), el dominio C-terminal (por ejemplo, restos 737-742 de SEQ ID NO: 2), la secuencia señal N-terminal (por ejemplo, restos 1-23 de SEQ ID NO: 2), o una combinación de los mismos. Un fragmento formador de complejo de gH puede ser cualquier parte o trozo de la proteína gH que retiene la capacidad de formar un complejo con otra proteína CMV. Por ejemplo, un fragmento formador de complejo de gH forma parte del complejo pentamérico gH/gL/pUL128/pUL130/pUL131. Por ejemplo, la expresión de la secuencia de gH de longitud completa puede dificultar la purificación del complejo pentamérico soluble porque el dominio TM de gH es hidrófobo. En cambio, el complejo pentamérico puede comprender un fragmento de gH con al menos una porción del dominio TM de gH eliminado.

15 Por ejemplo, puede usarse un fragmento de gH que comprende la secuencia señal N-terminal y el ectodominio de gH, pero no el dominio TM. Un fragmento de gH adecuado también puede comprender una porción del ectodominio de gH (por ejemplo, al menos aproximadamente el 70 %, al menos aproximadamente el 80 %, al menos aproximadamente el 85 %, al menos aproximadamente el 90 %, al menos aproximadamente el 95 %, al menos aproximadamente el 96 %, al menos aproximadamente el 97 %, al menos aproximadamente el 98 % o al menos aproximadamente el 99 % de la secuencia de ectodominio de gH), pero ninguno, o solo una pequeña porción del dominio TM. Alternativamente, el fragmento de gH descrito en el presente documento puede carecer entre 1 y 20 restos de aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 o 20 restos de aminoácidos, o carecen de 1-20 restos, 1-15 restos, 1-10 restos, 2-20 restos, 2-15 restos, 2-10 restos, 5-20 restos, 5-15 restos o 5-10 restos) en el N-terminal y/o C-terminal del ectodominio de longitud completa. Se cree que los restos en los dominios C-terminales no son necesarios para la inmunogenicidad. Un ejemplo de fragmento de gH adecuado descrito en el presente documento se muestra como SEQ ID NO: 4, que corresponde a los restos de aminoácidos 1-715 de la SEQ ID NO: 1. Otro ejemplo de fragmento de gH descrito en el presente documento se muestra como SEQ ID NO: 5, que carece de la secuencia señal N-terminal, el dominio TM y el dominio C-terminal de gH y corresponde a los restos de aminoácidos 24-715 de la SEQ ID NO: 1. Otro ejemplo de fragmento de gH descrito comprende la secuencia señal N-terminal completa y el ectodominio, pero carece del dominio C-terminal.

20 El ectodominio de gH corresponde al dominio extracelular de gH. La ubicación y la longitud del ectodominio de una gH (o un homólogo o una variante de la misma) pueden predecirse en función de la alineación por pares de su secuencia con SEQ ID NO: 1, 2, 3, 4 o 5, por ejemplo, alineando la secuencia de aminoácidos de una gH con la SEQ ID NO: 1 e identificando la secuencia que se alinea con los restos 24-717 de la SEQ ID NO: 1. De forma similar, las ubicaciones de la secuencia señal, el dominio TM y el dominio C-terminal pueden predecirse alineando la secuencia de aminoácidos de una gH con las SEQ ID NO: 1, 2, 3, 4 o 5, e identificando las secuencias que se alinean con las regiones correspondientes (por ejemplo, restos 1-23 (secuencia señal), 718-736

30 (TM) y 737-742 (dominio C-terminal) de SEQ ID NO: 1, respectivamente). Alternativamente, la ubicación y la longitud del ectodominio, la secuencia señal, el dominio TM y el dominio C-terminal pueden predecirse basándose en el análisis computacional de la hidrofobicidad a lo largo de una secuencia dada de gH. La secuencia señal y el dominio TM tienen los niveles más altos de hidrofobicidad y estas dos regiones flanquean el ectodominio, que es menos hidrófobo.

40 También puede obtenerse o determinarse un fragmento de gH formador de complejo adecuado mediante ensayos convencionales conocidos en la técnica, tales como el ensayo de coinmunoprecipitación, reticulación o co-localización por tinción fluorescente, etc. También puede usarse SDS-PAGE o transferencia Western (por ejemplo, mostrando que las cinco subunidades están presentes en una electroforesis en gel). Por ejemplo, el fragmento formador de complejo de gH (i) puede formar parte del complejo pentamérico gH/gL/pUL128/pUL130/pUL131; (ii) puede comprender al menos un epítipo de SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4 o SEQ ID NO: 5; y/o (iii) puede provocar anticuerpos *in vivo* que reaccionan inmunológicamente de forma cruzada con un virión de CMV.

45 Otras proteínas gH adecuadas pueden ser variantes de gH que tienen diversos grados de identidad con las SEQ ID NO: 1, 2, 3, 4 o 5, tal como al menos el 60 %, el 70 %, el 80 %, el 85 %, el 90 %, el 91 %, el 92 %, el 93 %, el 94 %, el 95 %, el 96 %, el 97 %, el 98 % o el 99 % idénticas a la secuencia mencionada en SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4 o SEQ ID NO: 5. Por ejemplo, las proteínas variantes de gH: (i) pueden formar parte del complejo pentamérico gH/gL/pUL128/pUL130/pUL131; (ii) pueden comprender al menos un epítipo de SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4 o SEQ ID NO: 5; y/o (iii) pueden provocar anticuerpos *in vivo* que reaccionan inmunológicamente de forma cruzada con un virión de CMV.

55 La glucoproteína L (gL) del CMV humano está codificada por el gen UL115. Se cree que gL es esencial para la replicación vírica y todas las propiedades funcionales conocidas de gL están directamente asociadas con su dimerización con gH. El complejo gL/gH es necesario para la fusión de membranas víricas y plasmáticas que conducen a la entrada del virus en la célula hospedadora. Se ha informado que gL de la cepa Merlin de CMV (GI:39842115, SEQ ID NO: 6) y la cepa Towne de CMV (GI:239909463, SEQ ID NO: 7) tienen 278 aminoácidos de longitud. Se ha informado que gL de la cepa AD169 de HCMV (GI:2506510, SEQ ID NO: 8) tiene 278 aminoácidos de longitud, con una secuencia señal en su N-terminal (restos de aminoácidos 1-35), dos sitios de N-glucosilación (en los restos 74 y

114) y que carece de un dominio TM. Se predice que la secuencia señal N-terminal en SEQ ID NO: 6 comprende los restos de aminoácidos 1-30. La SEQ ID NO: 7 comparte una identidad de secuencia de aminoácidos del 98 % con la SEQ ID NO: 6. La secuenciación del gen gL de longitud completa de 22 a 39 aislados clínicos, así como las cepas de laboratorio AD169, Towne y Toledo revelaron menos del 2 % de variación en las secuencias de aminoácidos entre los aislados.

Habitualmente, la secuencia señal N-terminal de la proteína gL de longitud completa se escinde por una peptidasa señal de la célula hospedadora para producir una proteína gL madura. Como tal, la proteína gL expresada por la célula hospedadora descrita en el presente documento puede carecer de la secuencia señal N-terminal (por ejemplo, gL está codificada por una secuencia de nucleótidos que carece de la secuencia de codificación para la secuencia señal N-terminal). Un ejemplo de una secuencia señal que carece de gL es la SEQ ID NO: 9, que comprende los restos de aminoácidos 31-278 de SEQ ID NO: 6 y carece de una secuencia señal N-terminal de SEQ ID NO: 6. La secuencia de señal de otras proteínas gL puede determinarse mediante herramientas de alineación de secuencia o análisis de secuencia como se describió anteriormente.

También se describen en el presente documento fragmentos formadores de complejos de gL. Un fragmento formador de complejo de gL puede ser cualquier parte o trozo de la proteína gL que retiene la capacidad de formar un complejo con otra proteína CMV. Por ejemplo, un fragmento formador de complejo de gL forma parte del complejo pentamérico gH/gL/pUL128/pUL130/pUL131.

Puede obtenerse o determinarse un fragmento de gL formador de complejo adecuado mediante ensayos convencionales conocidos en la técnica, tales como el ensayo de coimmunoprecipitación, reticulación o co-localización por tinción fluorescente, etc. También puede usarse SDS-PAGE o transferencia Western (por ejemplo, mostrando que las cinco subunidades están presentes en una electroforesis en gel). Por ejemplo, el fragmento formador de complejo de gL (i) puede formar parte del complejo pentamérico gH/gL/pUL128/pUL130/pUL131; (ii) pueden comprender al menos un epítipo de SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8 o SEQ ID NO: 9; y/o (iii) pueden provocar anticuerpos *in vivo* que reaccionan inmunológicamente de forma cruzada con un virión de CMV.

Otras proteínas gL adecuadas pueden ser variantes de gL (y fragmentos de variantes) que tienen diversos grados de identidad con las SEQ ID NO: 6, 7, 8 o 9, tal como al menos el 60 %, el 70 %, el 80 %, el 85 %, el 90 %, el 91 %, el 92 %, el 93 %, el 94 %, el 95 %, el 96 %, el 97 %, el 98 % o el 99 % idénticas a la secuencia mencionada en SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8 o SEQ ID NO: 9. Por ejemplo, las proteínas variantes de gL: (i) pueden formar parte del complejo pentamérico gH/gL/pUL128/pUL130/pUL131; (ii) pueden comprender al menos un epítipo de SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8 o SEQ ID NO: 9; y/o (iii) pueden provocar anticuerpos *in vivo* que reaccionan inmunológicamente de forma cruzada con un virión de CMV.

Se ha informado que la pUL128 de la cepa Merlin de CMV humano (GI:39842124, SEQ ID NO: 10) tiene 130 restos de aminoácidos y con una sustitución de 1 nucleótido que provoca la terminación prematura. Se ha informado que la pUL128 de las cepas Towne (GI:39841882, SEQ ID NO: 11) y AD169 (GI:59803078, SEQ ID NO: 12) de CMV humano tienen 171 restos de aminoácidos. Las SEQ ID NO: 10 y 12 comparten más del 99 % de identidad de secuencia en toda la longitud de la SEQ ID NO: 10; sin embargo, debido a la terminación prematura de la traducción, la SEQ ID NO: 10 no tiene los 41 restos de aminoácidos C-terminales de la SEQ ID NO: 12 (aproximadamente el 75 % de identidad de secuencia en toda la longitud de la SEQ ID NO: 12).

Se predice que pUL128 tiene una secuencia señal N-terminal, que se encuentra en los restos 1-27 de la SEQ ID NO: 10, pero se predice que carece de un dominio TM. La secuencia señal N-terminal de la proteína pUL128 de longitud completa puede escindirse mediante una peptidasa señal de la célula hospedadora para producir una proteína pUL128 madura. Como tal, la proteína pUL128 expresada por la célula hospedadora descrita en el presente documento puede carecer de la secuencia señal N-terminal (por ejemplo, pUL128 está codificada por una secuencia de nucleótidos que carece de la secuencia de codificación para la secuencia señal N-terminal). Un ejemplo de una proteína pUL128 madura es la SEQ ID NO: 13, que carece de una secuencia señal N-terminal y corresponde a los restos de aminoácidos 28-171 de la SEQ ID NO: 11. La SEQ ID NO: 13 también coincide con los restos de aminoácidos 28-171 de la SEQ ID NO: 12.

También se describen en el presente documento fragmentos formadores complejos de pUL128. Un fragmento formador de complejo de pUL128 puede ser cualquier parte o trozo de la proteína pUL128 que retiene la capacidad de formar un complejo con otra proteína CMV. Por ejemplo, un fragmento formador de complejo de pUL128 forma parte del complejo pentamérico gH/gL/pUL128/pUL130/pUL131.

Puede obtenerse o determinarse un fragmento de pUL128 formador de complejo adecuado mediante ensayos convencionales conocidos en la técnica, tales como el ensayo de coimmunoprecipitación, reticulación o co-localización por tinción fluorescente, etc. También puede usarse SDS-PAGE o transferencia Western (por ejemplo, mostrando que las cinco subunidades están presentes en una electroforesis en gel). Por ejemplo, el fragmento formador de complejo de pUL128 (i) puede formar parte del complejo pentamérico gH/gL/pUL128/pUL130/pUL131; (ii) comprende al menos un epítipo de SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12 o SEQ ID NO: 13; y/o (iii) puede provocar anticuerpos *in vivo* que reaccionan inmunológicamente de forma cruzada con un virión de CMV.

Otras proteínas pUL128 adecuadas pueden ser variantes de pUL128 (y fragmentos de variantes) que tienen diversos grados de identidad con las SEQ ID NO: 10, 11, 12 o 13, tal como al menos el 60 %, el 70 %, el 80 %, el 85 %, el 90 %, el 91 %, el 92 %, el 93 %, el 94 %, el 95 %, el 96 %, el 97 %, el 98 % o el 99 % idénticas a la secuencia mencionada en SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12 o SEQ ID NO: 13. Por ejemplo, las proteínas variantes de pUL128: (i) pueden formar parte del complejo pentamérico gH/gL/pUL128/pUL130/pUL131; (ii) pueden comprender al menos un epítipo de SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12 o SEQ ID NO: 13; y/o (iii) pueden provocar anticuerpos *in vivo* que reaccionan inmunológicamente de forma cruzada con un virión de CMV.

UL130 es el gen central y el más grande (214 codones) del locus UL131A-128. La traducción conceptual del gen predice una secuencia señal N-terminal larga (25 aminoácidos) que precede a una proteína hidrófila, con dos sitios potenciales de glucosilación enlazados en N (Asn85 y Asn118) dentro de un supuesto dominio de quimiocinas (aminoácidos 46 a 120) y un sitio adicional de N-glucosilación (Asn201) cerca del final de una región C-terminal única. Se predice que pUL130 carece de un dominio TM. Se ha informado que es una glucoproteína luminal que se secreta de manera ineficiente desde las células infectadas pero se incorpora a la envoltura del virión como una forma madurada por Golgi. Las secuencias de pUL130 de las cepas Merlin y Towne de CMV humano están disponibles públicamente (GI:39842125, SEQ ID NO: 14, 214 restos de aminoácidos; y GI:239909473, SEQ ID NO: 15, 229 restos de aminoácidos, respectivamente). Se ha informado que la SEQ ID NO: 15 contiene una mutación de desplazamiento de marco en la región C-terminal de pUL130 y comparte una identidad del 94 % con la SEQ ID NO: 14 del HCMV en toda la longitud de la SEQ ID NO: 14.

La secuencia señal N-terminal de la proteína pUL130 de longitud completa puede escindirse mediante una peptidasa señal de la célula hospedadora para producir una proteína pUL130 madura. Como tal, la proteína pUL130 expresada por la célula hospedadora descrita en el presente documento puede carecer de la secuencia señal N-terminal (por ejemplo, pUL130 está codificada por una secuencia de nucleótidos que carece de la secuencia de codificación para la secuencia señal N-terminal). Un ejemplo de una proteína pUL130 madura es la SEQ ID NO: 16, que carece de una secuencia señal N-terminal y corresponde a los restos de aminoácidos 26-214 de la SEQ ID NO: 14.

También se describen en el presente documento fragmentos formadores complejos de pUL130. Un fragmento formador de complejo de pUL130 puede ser cualquier parte o trozo de la proteína pUL130 que retiene la capacidad de formar un complejo con otra proteína CMV. Por ejemplo, un fragmento formador de complejo de pUL130 forma parte del complejo pentamérico gH/gL/pUL128/pUL130/pUL131.

Puede obtenerse o determinarse un fragmento de pUL130 formador de complejo adecuado mediante ensayos convencionales conocidos en la técnica, tales como el ensayo de coimmunoprecipitación, reticulación o co-localización por tinción fluorescente, etc. También puede usarse SDS-PAGE o transferencia Western (por ejemplo, mostrando que las cinco subunidades están presentes en una electroforesis en gel). Por ejemplo, el fragmento formador de complejo de pUL130 (i) puede formar parte del complejo pentamérico gH/gL/pUL128/pUL130/pUL131; (ii) comprende al menos un epítipo de SEQ ID NO: 14, SEQ ID NO: 15 o SEQ ID NO: 16; y/o (iii) puede provocar anticuerpos *in vivo* que reaccionan inmunológicamente de forma cruzada con un virión de CMV.

Otras proteínas pUL130 adecuadas pueden ser variantes de pUL130 (y fragmentos de variantes) que tienen diversos grados de identidad con las SEQ ID NO: 14, 15 o 16, tal como al menos el 60 %, el 70 %, el 80 %, el 85 %, el 90 %, el 91 %, el 92 %, el 93 %, el 94 %, el 95 %, el 96 %, el 97 %, el 98 % o el 99 % idénticas a la secuencia mencionada en SEQ ID NO: 14, SEQ ID NO: 15 o SEQ ID NO: 16. Por ejemplo, las proteínas variantes de pUL130: (i) pueden formar parte del complejo pentamérico gH/gL/pUL128/pUL130/pUL131; (ii) pueden comprender al menos un epítipo de SEQ ID NO: 14, SEQ ID NO: 15 o SEQ ID NO: 16; y/o (iii) pueden provocar anticuerpos *in vivo* que reaccionan inmunológicamente de forma cruzada con un virión de CMV.

La función pUL131A es necesaria para la replicación del CMV humano no solo en las células endoteliales sino también en las células epiteliales. Se han informado pUL131A de las cepas Merlin (GI:39842126, SEQ ID NO: 17, 129 aminoácidos) y Towne (GI:239909474, SEQ ID NO: 18, 129 aminoácidos) y AD169 (GI:219879712, SEQ ID NO: 19, 76 aminoácidos) de CMV humano. Se predice que pUL131A contiene una secuencia de señal N-terminal, que se encuentra en los restos 1-18 de SEQ ID NO: 18 y que carece de un dominio TM. Se ha informado que pUL131A de la cepa AD169 contiene una inserción de 1 par de bases, que provoca un desplazamiento de marco. La SEQ ID NO: 17 es un 96 % idéntica a SEQ ID NO: 19 en los 28 aminoácidos N-terminales, pero es solo un 36 % idéntica a SEQ ID NO: 19 sobre la longitud total de SEQ ID NO: 17, debido al desplazamiento de marco en el gen AD169 UL131A.

La secuencia señal N-terminal de la proteína pUL131 de longitud completa puede escindirse mediante una peptidasa señal de la célula hospedadora para producir una proteína pUL131 madura. Como tal, la proteína pUL131 expresada por la célula hospedadora descrita en el presente documento puede carecer de la secuencia señal N-terminal (por ejemplo, pUL131 está codificada por una secuencia de nucleótidos que carece de la secuencia de codificación para la secuencia señal N-terminal). Un ejemplo de una proteína pUL130 madura es la SEQ ID NO: 20, que carece de una secuencia señal N-terminal y corresponde a los restos de aminoácidos 19-129 de la SEQ ID NO: 17. La SEQ ID NO: 35 también corresponde a los restos de aminoácidos 19-129 de la SEQ ID NO: 18.

También se describen en el presente documento fragmentos formadores complejos de pUL131. Un fragmento formador de complejo de pUL131 puede ser cualquier parte o trozo de la proteína pUL131 que retiene la capacidad

de formar un complejo con otra proteína CMV. Por ejemplo, un fragmento formador de complejo de pUL131 forma parte del complejo pentamérico gH/gL/pUL128/pUL130/pUL131.

5 Puede obtenerse o determinarse un fragmento de pUL131 formador de complejo adecuado mediante ensayos convencionales conocidos en la técnica, tales como el ensayo de coinmunoprecipitación, reticulación o co-localización por tinción fluorescente, etc. También puede usarse SDS-PAGE o transferencia Western (por ejemplo, mostrando que las cinco subunidades están presentes en una electroforesis en gel). Por ejemplo, el fragmento formador de complejo de pUL131 (i) puede formar parte del complejo pentamérico gH/gL/pUL128/pUL130/pUL131; (ii) comprende al menos un epítipo de SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19 o SEQ ID NO: 20; y/o (iii) puede provocar anticuerpos *in vivo* que reaccionan inmunológicamente de forma cruzada con un virión de CMV.

10 Otras proteínas pUL131 adecuadas pueden ser variantes de pUL131 (y fragmentos de variantes) que tienen diversos grados de identidad con las SEQ ID NO: 17, 18, 19 o 20, tal como al menos el 60 %, el 70 %, el 80 %, el 85 %, el 90 %, el 91 %, el 92 %, el 93 %, el 94 %, el 95 %, el 96 %, el 97 %, el 98 % o el 99 % idénticas a la secuencia mencionada en SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19 o SEQ ID NO: 20. Por ejemplo, las proteínas variantes de pUL131: (i) pueden formar parte del complejo pentamérico gH/gL/pUL128/pUL130/pUL131; (ii) pueden comprender al menos un epítipo de SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19 o SEQ ID NO: 20; y/o (iii) pueden provocar anticuerpos *in vivo* que reaccionan inmunológicamente de forma cruzada con un virión de CMV.

20 Las gH, gL, pUL128, pUL130, pUL131 (o un fragmento de las mismas) descritas en el presente documento pueden contener restos de aminoácidos adicionales, tales como extensiones N-terminal o C-terminal. Dichas extensiones pueden incluir una o más etiquetas, que pueden facilitar la detección (por ejemplo, una etiqueta de epítipo para la detección por anticuerpos monoclonales) y/o la purificación (por ejemplo, una etiqueta de polihistidina para permitir la purificación en una resina quelante de níquel) de las proteínas. Algunos ejemplos de etiquetas de purificación por afinidad incluyen, por ejemplo, etiqueta His (hexahistidina (SEQ ID NO: 36), se une al ion metálico), proteína de unión a maltosa (MBP) (se une a la amilosa), glutatión-S-transferasa (GST) (se une al glutatión), etiqueta FLAG (Asp-Tyr-Lys-Asp-Asp-Asp-Asp-Lys (SEQ ID NO: 37), se une a un anticuerpo anti-flag), etiqueta Strep (Ala-Trp-Arg-His-Pro-Gln-Phe-Gly-Gly (SEQ ID NO: 38), o Trp-Ser-His-Pro-Gln-Phe-Glu-Lys (SEQ ID NO: 39), se unen a estreptavidina o un derivado de la misma).

30 Pueden usarse enlazadores escindibles. Esto permite que la etiqueta se separe del complejo purificado, por ejemplo mediante la adición de un agente capaz de escindir el enlazador. Los expertos en la materia conocen una cantidad de diferentes enlaces escindibles. Dichos enlazadores pueden escindirse, por ejemplo, por irradiación de un enlace fotolábil o hidrólisis catalizada por ácido. También hay enlazadores polipeptídicos que incorporan un sitio de reconocimiento de proteasa y que pueden escindirse mediante la adición de una enzima proteasa adecuada.

Puede ser más deseable tener proteínas gH, gL, pUL128, pUL130, pUL131 (o un fragmento de las mismas) que no comprendan una secuencia de etiqueta exógena, por ejemplo, por razones de seguridad clínica o eficacia.

35 Aunque las proteínas gH, gL, pUL130 a veces se denominan glucoproteínas, esta nomenclatura no debe entenderse como que estas proteínas deben estar glucosiladas cuando se usan con la invención. Si bien se han mencionado anteriormente cepas específicas, debe entenderse que pueden usarse proteínas de CMV gH, gL, pUL128, pUL130, pUL131 (o fragmentos de las mismas) de diferentes cepas de CMV. A modo de ejemplo no limitante, otra cepa de CMV puede incluir las cepas Towne, Toledo, AD169, Merlin, TB20 y VR1814.

## **B. Ácido nucleico que codifican proteínas y complejos de CMV**

40 También se describen en el presente documento ácidos nucleicos que codifican las proteínas y complejos de CMV para la integración genómica y la posterior expresión de pentámero de CMV.

45 Pueden usarse una o más construcciones de ácido nucleico que codifican las proteínas y complejos de CMV descritos en el presente documento para la integración genómica. Por ejemplo, una sola construcción de ácido nucleico que codifica las cinco subunidades, gH, gL, pUL128, pUL130, pUL131 (o fragmentos de las mismas), puede introducirse en una célula hospedadora. Alternativamente, las secuencias codificantes para las cinco subunidades (o fragmentos de las mismas) pueden transportarse por dos o más construcciones de ácido nucleico, que después se introducen en la célula hospedadora de forma simultánea o secuencial.

50 Por ejemplo, se describe en el presente documento una construcción de ácido nucleico único que codifica: el ectodominio de gH, gL, pUL128, pUL130 y pUL131. Alternativamente, se describen en el presente documento dos construcciones de ácido nucleico que codifican: el ectodominio de gH, gL, pUL128, pUL130 y pUL131. Véase, la Figura 1. En ambos ejemplos, se logró una integración genómica exitosa.

55 La construcción de ácido nucleico puede comprender ADN genómico que comprende uno o más intrones o ADNc. Algunos genes se expresan de manera más eficiente cuando los intrones están presentes. La secuencia genómica nativa que codifica pUL128 comprende dos intrones, la secuencia genómica nativa que codifica pUL131 comprende un exón, mientras que la secuencia genómica nativa que codifica pUL130 no comprende ningún intrón. La secuencia genómica nativa que codifica pUL128 comprende tres exones, la secuencia genómica nativa que codifica pUL131 comprende dos exones y la secuencia genómica nativa que codifica pUL130 comprende un exón. Particularmente la

secuencia de ácido nucleico es adecuada para la expresión de polipéptidos exógenos en dicha célula de mamífero.

También se describen en el presente documento vectores que comprenden secuencias codificantes para gH, gL, pUL128, pUL130 y/o pUL131 (o un fragmento de las mismas). Los vectores ejemplares incluyen plásmidos que pueden replicarse de forma autónoma o replicarse en una célula de mamífero. Los vectores de expresión típicos contienen promotores, potenciadores y terminadores adecuados que son útiles para la regulación de la expresión de la secuencia o secuencias codificantes en la construcción de expresión. Los vectores también pueden comprender marcadores de selección para proporcionar un rasgo fenotípico para la selección de células hospedadoras transformadas (tales como conferir resistencia a antibióticos tales como ampicilina o neomicina).

Algunos promotores adecuados incluyen, por ejemplo, promotor de CMV, adenovirus, EF1 $\alpha$ , promotor de metalotionina GAPDH, promotor temprano de SV-40, promotor tardío de SV-40, promotor del virus del tumor mamario murino, promotor del virus del sarcoma de Rous, promotor de polihedrina, etc. Los promotores pueden ser constitutivos o inducibles. Pueden usarse uno o más vectores (por ejemplo, un vector que codifica las cinco subunidades o fragmentos de las mismas o dos o más vectores que codifican juntos las cinco subunidades o fragmentos de las mismas); véase, por ejemplo, la Figura 1.

Cuando la célula hospedadora es una célula CHO, el promotor, el potenciador o terminador está activo en las células CHO. Un promotor de uso frecuente es el promotor del gen de citomegalovirus humano (hCMV) temprano inmediato (IE). El promotor de este gen dirige altos niveles de expresión transgénica en una amplia diversidad de tipos celulares. La actividad de este promotor depende de una serie de repeticiones imperfectas de 17, 18, 19 y 21 pb, algunas de las cuales se unen a factores de transcripción de la proteína de unión sensible a AMPc NF- $\kappa$ B (CREB) y las familias de factor nuclear-1.

Un promotor fuerte es aquel que hace que los ARNm se inicien a una frecuencia alta igual o mayor que la del fragmento promotor/potenciador central de hCMV (descrito en la Pat. de EE.UU. N.º 5.168.062) en una célula CHO. Dicho promotor puede ser un promotor fuerte dependiente del tipo celular, como se describe en la Pat. de EE.UU. N.º 5.589.392, o un promotor fuerte ubicuamente activo. Los promotores víricos constitutivamente activos ejemplares incluyen, por ejemplo, promotores tempranos y tardíos del virus SV40, el promotor temprano inmediato del citomegalovirus humano (hCMV) o del citomegalovirus murino (mCMV), el promotor de timidina quinasa (TK) del virus del herpes simple o el promotor de repetición terminal larga del virus del sarcoma de Rous (RS-LTR). Otros ejemplos incluyen, por ejemplo, el promotor hCMV-MIE como se define por el fragmento Pst I de 2,1 kb descrito en la Pat. de EE.UU. N.º 5.385.839 y/o EP-323 997-A1 o una parte funcional del mismo que tiene actividad promotora.

También pueden usarse el sitio interno de entrada al ribosoma (I RES) y las secuencias peptídicas 2A. I RES y el péptido 2A proporcionan enfoques alternativos para la co-expresión de secuencias múltiples. IRES es una secuencia de nucleótidos que permite el inicio de la traducción en el medio de una secuencia de ARN mensajero (ARNm) como parte del mayor proceso de síntesis de proteínas. Habitualmente, en eucariotas, la traducción solo puede iniciarse en el extremo 5' de la molécula de ARNm. Los elementos IRES permiten la expresión de múltiples genes en un transcrito. Los vectores policistrónicos basados en IRES, que expresan múltiples proteínas de un transcrito, pueden reducir el escape de clones que no se expresan a partir de la selección.

El péptido 2A permite la traducción de múltiples proteínas en un solo marco de lectura abierto en una poliproteína que posteriormente se divide en proteínas individuales a través de un mecanismo de omisión de ribosomas. El péptido 2A puede proporcionar una expresión más equilibrada de múltiples productos proteicos.

Algunos ejemplos de secuencias IRES incluyen, por ejemplo, EV71 IRES, EMCV IRES, HCV IRES.

Para la integración genómica, la integración puede ser específica del sitio o aleatoria. La recombinación específica del sitio puede lograrse mediante la introducción de secuencias homólogas en las construcciones de ácido nucleico descritas en el presente documento. Dicha secuencia homóloga coincide sustancialmente con la secuencia endógena en un sitio diana específico en el genoma del hospedador. Alternativamente, puede usarse integración aleatoria. A veces, el nivel de expresión de una proteína puede variar dependiendo del sitio de integración. Por lo tanto, puede ser deseable seleccionar un número de clones de acuerdo con el nivel de expresión de proteína recombinante para identificar un clon que logre el nivel de expresión deseado.

### 3. CÉLULAS HOSPEDADORAS

En otro aspecto de la divulgación, se describen células hospedadoras en las que las secuencias codificantes de pentámeros están integradas de manera estable en el genoma de las células hospedadoras y, cuando se cultivan bajo una condición adecuada, expresan el pentámero CMV como se desvela en el presente documento. La célula hospedadora puede ser una célula de mamífero. Por ejemplo, la célula hospedadora puede ser una célula de roedor.

Los ejemplos de líneas celulares de roedores incluyen, por ejemplo, riñón de hámster bebé (BHK) (por ejemplo, BHK21, BH TK), Sertoli de ratón (TM4), hígado de rata búfalo (BRL 3A), tumor mamario de ratón (MMT), hepatoma de rata (HTC), mieloma de ratón (NS0), hibridoma murino (Sp2/0), timoma de ratón (EL4), ovario de hámster chino (CHO) y derivados de células CHO, embrión murino (NIH/3T3, 3T3 Li), miocardio de rata (H9c2), mioblastos de ratón (C2C12) y riñón de ratón (miMCD-3).

Las células CHO adecuadas incluyen, por ejemplo, líneas DUXB11 y DG44. Estas dos líneas celulares son deficientes en la actividad de dihidrofolato reductasa (DHFR) y, por lo tanto, dependen de una fuente exógena de precursores de nucleótidos para el crecimiento. La deficiencia de DHFR es un fenotipo fácilmente manipulable adecuado para seleccionar la integración del genoma y la expresión estable de ADN exógeno. La integración genómica se logra mediante la transfección de las células con casetes de expresión para el gen de interés y un gen DHFR. Después de la transfección, las células se colocan en medios de selección que carecen de precursores de nucleótidos.

La expresión de proteínas de recombinación en líneas celulares deficientes en DHFR puede mejorarse aún más añadiendo metotrexato (MTX) a los cultivos, de tal manera que pueda seleccionarse un número alto de copias del vector de expresión introducido. MTX es un inhibidor competitivo de la enzima DHFR. La aplicación de esta presión de selección adicional sobre la ausencia de precursores de nucleótidos permite la selección y el aislamiento de la población menor de células que han sufrido una amplificación espontánea del vector de expresión integrado que contiene el marcador seleccionable DHFR y, en la mayoría de los casos, el gen de interés. La presencia de múltiples copias de genes ayuda a lograr un alto nivel de expresión de proteínas exógenas. Alternativamente, la selección de MTX puede llevarse a cabo independientemente de la deficiencia de DHFR (es decir, usar MTX para seleccionar una célula hospedadora que sea originalmente competente para DHFR), como se ejemplifica en los Ejemplos desvelados en el presente documento.

Otra línea celular CHO adecuada es la línea celular CHO-K1 de tipo silvestre y su derivado CHO-K1SV.

Un procedimiento de selección comúnmente usado para las líneas celulares CHO-K1 es la selección de glutamina sintetasa (GS). En ausencia de una fuente exógena de glutamina, la supervivencia celular depende de la enzima GS para producir glutamina. Con líneas celulares huésped tales como células NS/0 derivadas de mieloma murino y células CHO, que tienen una actividad enzimática GS relativamente baja endógena, el procedimiento permite un esquema de selección sencillo cuando se usa un marcador seleccionable GS en el vector de expresión y los medios de selección libres de glutamina. Similar al sistema DHFR/MTX, el inhibidor competitivo GS metionina sulfoximina (MSX) puede añadirse a los medios para aplicar presión adicional y seleccionar las células CHO que están impulsando altos niveles de expresión del vector integrado.

Las células CHO-K1, o cualquier otra célula CHO comúnmente usada, también puede seleccionarse en función de la deficiencia de DHFR como se ha descrito anteriormente. Por ejemplo, una célula CHO-K1, o cualquier otro tipo de célula CHO, puede tener deficiencia de DHFR, tal como una delección en la que al menos una copia de la secuencia genómica del gen de dihidrofolato reductasa (DHFR) o al menos el 30 % (por ejemplo, al menos el 40 %, al menos el 50 %, al menos el 60 %, al menos el 70 %, al menos el 80 %, al menos el 90 % o el 100 %) de la secuencia codificante de dicho gen DHFR, está eliminada. Otras formas de introducir la deficiencia de DHFR incluyen la creación de mutaciones en el gen endógeno DHFR. Las líneas celulares pueden mejorarse adicionalmente añadiendo metotrexato (MTX) a los cultivos como se describió anteriormente.

Las células CHO-K1, o cualquier otra célula CHO comúnmente usada, también puede seleccionarse en función de MTX, con o sin deficiencia de DHFR. En los ejemplos proporcionados en el presente documento, las células CHO se seleccionaron basándose en MTX, sin deficiencia de DHFR (es decir, la célula CHO original usada para la integración genómica es competente para DHFR). En un sistema tal, normalmente, el número de copias de secuencias exógenas (por ejemplo, las secuencias que codifican las proteínas de CMV) es generalmente bajo. Se estima que las líneas celulares en los Ejemplos descritos en el presente documento tienen aproximadamente 1-10 copias de secuencias exógenas que codifican las proteínas de CMV, en un número muy limitado de sitios de integración (por ejemplo, 1-2 sitios de integración). En general, cuando se usa una línea celular deficiente en DHFR, el número de copias de secuencias exógenas es típicamente mucho más alto, a veces tan alto como unos cientos de copias. Se espera que ambos procedimientos sean adecuados para producir líneas celulares CHO desveladas en el presente documento, aunque cuando el número de copias es alto, la célula hospedadora puede perder una o más copias de secuencias exógenas durante el pasaje y/o la expansión de la línea celular.

Otras cepas de células CHO adecuadas para la invención descrita en el presente documento incluyen, por ejemplo, células CHO-ICAM-1 y células CHO-hIFN $\gamma$ . Estas células genéticamente modificadas permiten la inserción estable de ADN recombinante en un gen específico o región de expresión de las células, amplificación del ADN insertado y selección de células que exhiben un alto nivel de expresión de la proteína recombinante.

Las líneas celulares CHO ejemplares disponibles en la European Collection of Cell Cultures (ECACC) se enumeran en la Tabla 1. Puede usarse cualquier célula CHO enumerada en la Tabla 1.

Tabla 1

Nombre de línea celular	Palabras clave
CHO	Ovario de hámster chino

(continuación)

Nombre de línea celular	Palabras clave
CHO (SIN PROTEÍNA)	Ovario de hámster chino
CHO-CHRM1	Receptor colinérgico humano muscarínico M1, CHRM1, Receptor acoplado a proteína G, GPCR, Transfectado, InSCREENeX SCREENflex™, Hospedador CHO-K1.
CHO-CHRM2	Receptor colinérgico humano muscarínico M2, CHRM2, Receptor acoplado a proteína G, GPCR, Transfectado, InSCREENeX SCREENflex™, Hospedador CHO-K1.
CHO-CHRM5	Receptor colinérgico humano muscarínico M5, CHRM5, Receptor acoplado a proteína G, GPCR, Transfectado, InSCREENeX SCREENflex™, Hospedador CHO-K1.
CHO-CNR1	Receptor cannabinoide humano I, CNR1 ID de gen 1268, Receptor acoplado a proteína G, GPCR, Transfectado, InSCREENeX SCREENflex™, Hospedador CHO-K1.
CHO-FFAR2	Receptor de ácido graso libre humano 2, FFAR2, Receptor acoplado a proteína G, GPCR, Transfectado, InSCREENeX SCREENflex™, Hospedador CHO-K1.
CHO-GPR120	Receptor GPR120 humano (huérfano), GPR120, Receptor acoplado a proteína G, GPCR, Transfectado, InSCREENeX SCREENflex™, Hospedador CHO-K1.
CHO-K1	Ovario de hámster chino
Sin CHO-K1-AC	Ovario de hámster chino, sin suero
CHO-K1/SF	Ovario de hámster chino (MEM adaptado)
CHO-NPY1R	Receptor de neuropéptido Y humano, NPY1R, ID de gen 4886, Receptor acoplado a proteína G, GPCR, Transfectado, InSCREENeX SCREENflex™, Hospedador CHO-K1.
CHO-OPRL1	Receptor tipo de opiáceos humano 1, OPRL1, Receptor acoplado a proteína G, GPCR, Transfectado, InSCREENeX SCREENflex™, Hospedador CHO-K1.
CHO-SSTR1	Receptor de somatostatina humana 1, Receptor acoplado a proteína G SSTR1, GPCR, Transfectado, InSCREENeX SCREENflex™, Hospedador CHO-K1.
CHO/dhFr-	Ovario de hámster chino
sin CHO/dhFr-AC	Ovario de hámster chino, sin suero
RR-CHOKI	Ovario de hámster chino
T02J-10/10 (CHO-GCGR (GCGR))	Receptor de glucagón humano, GCGR, Receptor acoplado a proteína G, GPCR, Transfectado, InSCREENeX SCREENflex™, Hospedador CHO-K1.

Diversas líneas celulares CHO también están disponibles de la American Type Culture Collection (ATCC), tales como las líneas celulares CHO hCBE11 (ATCC® PTA-3357™), E77.4 (ATCC® PTA-3765™), hLT-B: R-hG1 CHO n.º 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 [originalmente denominado 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™). Puede usarse una cualquiera de estas líneas celulares CHO.

Otras líneas celulares CHO disponibles en el mercado incluyen, por ejemplo, Células CHO-S FreeStyle™ y Línea Celular Flp-In™-CHO de Life Technologies.

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

Además de las células CHO, otras células de mamíferos también pueden usarse como hospedadores. Las células de roedores ejemplares incluyen células BHK21, células NS0, células Sp2/0, células EL4, células NIH/3T3, células 3T3-L1, células ES-D3, células H9c2, células C2C12, YB2/0, células mired 3, etc. Las células humanas ejemplares incluyen: células SH-SY5Y, células IM 32, células LAN, células MCFIOA, células 293T, células SK-BR3, células huvec, células huasmc, células HKB-1, células hmsc, células U293, células HE 293, células PERC6®, células Jurkai, células

HT-29, células Incap.FGC, célula A549, células MDA MB453, células hepg2, células THP-1, células bxp-3, células Capan-1, células DU145 y células PC-3.

5 Por ejemplo, la línea celular PERC6®, célula de mieloma de ratón NS0, célula de riñón de hámster bebé (BHK) y la línea celular de riñón embrionario humano (HEK293) recibieron la aprobación reglamentaria para la producción de proteínas recombinantes.

10 Algunos ejemplos de líneas celulares de primates no humanos útiles en los procedimientos proporcionados en el presente documento incluyen las líneas celulares de riñón de mono (CVI-76), riñón de mono verde africano (VERO-76), fibroblastos de mono verde (COS-1) y células de riñón de mono (CVI) transformadas por SV40 (COS-7). Los expertos en la materia conocen líneas celulares de mamíferos adicionales y están catalogadas en el catálogo de la American Type Culture Collection (Manassas, VA).

15 Las células hospedadoras pueden ser adecuadas para el crecimiento en cultivos en suspensión. Las células hospedadoras competentes en suspensión están generalmente monodispersas o crecen en agregados sueltos sin agregación sustancial. Las células hospedadoras competentes en suspensión incluyen células que son adecuadas para el cultivo en suspensión sin adaptación o manipulación (por ejemplo, células hematopoyéticas, células linfoides) y células que se han hecho competentes en suspensión mediante modificación o adaptación de células dependientes de la unión (por ejemplo, células epiteliales, fibroblastos).

20 Alternativamente, la célula hospedadora puede ser una célula dependiente de unión que se cultiva y se mantiene en cultivo adherente. Los ejemplos de líneas celulares adherentes humanas útiles en los procedimientos proporcionados en el presente documento incluyen las líneas celulares de neuroblastoma humano (SH-SY5 Y, IMR32 y LANS), carcinoma cervical humano (HeLa), epitelial de mama humana (MCF10A), riñón embrionario humano (293T) y carcinoma de mama humano (SK-BR3).

#### *Genes C12orf35 y proteínas C12orf35*

25 La célula hospedadora de la invención es una célula en la que se reduce el nivel de expresión o actividad de la proteína C12orf35, en comparación con un control. La Solicitud de Patente Provisional de EE.UU. N.º 61/919.313, presentada el 20 de diciembre de 2013 proporciona una descripción detallada de las células de mamíferos en las que el nivel de expresión o actividad de la proteína C12orf35 se reduce en comparación con un control.

Puede usarse una diversidad de controles. El nivel de expresión o actividad de la proteína C12orf35 de una célula de tipo silvestre correspondiente puede usarse como control. Alternativamente, un control puede ser un nivel predeterminado o un nivel umbral que puede identificarse en las referencias bibliográficas o en la base de datos.

30 El gen C12orf35 humano se refiere a la secuencia de nucleótidos que codifica el marco de lectura abierta 35 del cromosoma 12. La proteína C12orf35 codificada no está caracterizada. El homólogo *Cricetulus griseus* (hámster chino) del gen humano C12orf35, llamado Kiaa1551, se cree que se encuentra en el cromosoma 8. Se cree que el homólogo de *Mus musculus* del gen humano C12orf35 se encuentra en el cromosoma 6. La ID del gen para el gen CHO C12orf35 se publica como ID del gen de GenBank N.º 100762086; y para el gen C12orf35 humano se publica como ID del gen de GenBank N.º 55196. La información sobre el gen C12orf35 en *Cricetulus griseus*, la secuencia de codificación y la proteína C12orf35 predicha también están disponibles en GenBank con el número de referencia de NCBI: XM\_003512865.

35 El gen C12orf35 se expresa endógenamente en células eucariotas como, por ejemplo, especies de mamíferos tales como humano, ratón y hámster. La proteína predicha codificada por el gen C12orf35 es una proteína grande que excede los 1500 restos. La lista de secuencias muestra ejemplos de secuencias de aminoácidos o secuencias de aminoácidos putativas de la proteína codificada por el gen endógeno C12orf35 de diferentes especies de mamíferos tales como hámster (SEQ ID NO: 21 y 22), humano (SEQ ID NO: 23 y 24) y ratón (SEQ ID NO: 35). La CDS (secuencia de ADN codificante, por sus siglas en inglés) de C12orf35 del hámster chino se muestra como SEQ ID NO: 25. Adicionalmente, se secuenció una sección del 5'UTR (véase la SEQ ID NO: 26) y del 3'UTR (véase la SEQ ID NO: 27) del ARNm C12orf35 del hámster chino.

40 En humanos, el gen C12orf35 también se conoce como KIAA1551. El gen C12orf35 también se conoce como tipo C12orf35 u homólogo C12orf35 en hámster o 2810474O19Rik en ratón. Pueden asignarse diferentes nombres en diferentes especies para la proteína o el gen y los nombres alternativos no limitantes (alias) también se enumeran anteriormente en la Tabla 2. Por simplicidad, en la presente divulgación, los homólogos y ortólogos de diferentes especies se denominan "gen C12orf35" o "proteína C12orf35".

45 Contra este contexto científico, es sorprendente e inesperado que cuando el nivel de expresión o actividad de la proteína C12orf35 se reduce en comparación con un control (por ejemplo, al eliminar el gen C12orf35 o al introducir mutaciones), el rendimiento de la proteína recombinante se mejora significativamente. Como tal, las células hospedadoras de mamífero (por ejemplo, células CHO) con nivel de expresión reducido o actividad de la proteína C12orf35 son particularmente adecuadas para la producción recombinante de complejo pentamérico.

La reducción del nivel de expresión o actividad de una proteína C12orf35 puede lograrse por diversos medios. Por

ejemplo, el nivel de expresión o actividad de una proteína C12orf35 puede reducirse por inactivación génica, mutación génica, delección génica, silenciamiento génico o una combinación de cualquiera de los anteriores. La inactivación génica es una técnica genética mediante la cual un gen se vuelve inoperativo al interrumpir su función. Por ejemplo, puede insertarse un ácido nucleico en la secuencia codificante, interrumpiendo de esta manera la función del gen. Adicionalmente, el gen C12orf35 de longitud completa (o un fragmento del mismo) puede eliminarse, por lo que la expresión de la proteína C12orf35 funcional se elimina sustancialmente. Por ejemplo, la delección puede ser de al menos el 30 %, al menos el 40 %, al menos el 50 %, al menos el 55 %, al menos el 60 %, al menos el 65 %, al menos el 70 %, al menos el 75 %, al menos el 80 %, al menos el 85 %, al menos el 90 %, al menos el 95 % o el 100 % de la secuencia codificante del gen C12orf35. Otra opción es introducir una o más mutaciones en la secuencia codificante, que produce una proteína C12orf35 no funcional o menos funcional. Por ejemplo, pueden introducirse una o más mutaciones de desplazamiento de marco, dando como resultado una proteína C12orf35 no funcional o menos funcional. Alternativa o adicionalmente, pueden introducirse uno o más codones de parada en la secuencia codificante de tal manera que se obtiene una proteína truncada, no funcional o menos funcional. Otras opciones incluyen, pero no se limitan a, una o más mutaciones en el promotor, en el 5'- y/o 3' UTR u otros elementos reguladores, por ejemplo, introduciendo una delección del promotor o introduciendo una construcción entre el promotor y el inicio de la transcripción. Los procedimientos para la interrupción génica para suprimir o eliminar la expresión del gen diana también son bien conocidos por la persona experta.

Como cada célula tiene dos copias del gen C12orf35 en su genoma, en determinadas realizaciones, al menos una copia de la secuencia genómica del gen C12orf35, o al menos el 50 % de la secuencia codificante de dicho gen C12orf35, está eliminada. Por ejemplo, pueden eliminarse ambas copias de las secuencias genómicas del gen C12orf35 (o al menos el 50 % de la secuencia codificante de dicho gen C12orf35 de cada copia).

En determinadas realizaciones, la secuencia eliminada comprende una porción de la región telomérica del cromosoma 8 de una célula CHO. Una región telomérica es una región de secuencias repetitivas de nucleótidos en cada extremo de una cromátida, que protege el extremo del cromosoma del deterioro o de la fusión con los cromosomas vecinos. Por ejemplo, al menos el 30 %, al menos el 40 %, al menos el 50 %, al menos el 55 %, al menos el 60 %, al menos el 65 %, al menos el 70 %, al menos el 75 %, al menos el 80 %, al menos el 85 %, al menos el 90 %, al menos el 95 % o el 100 % de la secuencia de nucleótidos de la región telomérica del cromosoma 8 de una célula CHO puede eliminarse.

Como se describe en el presente documento, la secuencia eliminada puede comprender además la delección de un gen seleccionado del grupo que consiste en: Bicd1, Amn1, proteína similar a la metiltransferasa 20, Dend5b, FAM60A, Caprin2, Ipo8, RPS4Y2 y una combinación de los mismos.

La delección del gen C12orf35 (o un fragmento del mismo) puede deberse a una ruptura cromosómica. La ruptura cromosómica puede inducirse, por ejemplo, tratando las células eucariotas con un agente tóxico que promueve la ruptura cromosómica, tales como por ejemplo MTX, afidicolina o higromicina. Otras opciones para inducir rupturas cromosómicas incluyen, pero no se limitan a, radiación, irradiación, mutágenos, sustancias cancerígenas y bleomicina. Las rupturas cromosómicas también pueden producirse espontáneamente durante la transfección, por ejemplo, electroporación. Los procedimientos para inducir la ruptura cromosómica también son conocidos por la persona experta y, por lo tanto, no necesitan ninguna descripción detallada en este punto. Después de inducir la ruptura cromosómica, las células eucariotas que tienen el punto de ruptura deseado (que da como resultado una delección del gen C12orf35, o un fragmento del mismo) pueden identificarse, por ejemplo, analizando el ADN o usando el procedimiento de acuerdo con el quinto aspecto de la presente divulgación. Por ejemplo, el perfil de expresión de las células tratadas puede analizarse para determinar si el gen C12orf35 o los genes ubicados centroméricos del gen C12orf35 se expresan, si la expresión se reduce o si los genes no se expresan. Por ejemplo, en el caso de las células de ratón o hámster, puede analizarse si el gen C12orf35 se expresa y, alternativamente o además de ello, puede analizarse si uno o más genes seleccionados del grupo que consiste en la proteína 20 similar a la metiltransferasa, Dend5b, FAM60A, Caprin2, Ipo8, Tmtc1 o genes que se encuentran teloméricos de los genes mencionados anteriormente (en el que telomérico significa a este respecto en la dirección del extremo telomérico) se expresan por la célula y/o si la expresión se reduce o se elimina sustancialmente.

La reducción del nivel de expresión de la proteína C12orf35 puede lograrse mediante silenciamiento génico post-transcripcional, por ejemplo, por moléculas de ácido nucleico antisentido, o moléculas que median la interferencia de ARN. Los ejemplos no limitantes incluyen ARNip, ARNh, miARN, oligonucleótidos antisentido, etc., todos los cuales son bien conocidos en la técnica.

El nivel de expresión de la proteína C12orf35 puede evaluarse mediante procedimientos conocidos en la técnica, por ejemplo, midiendo el nivel de ARNm que codifica la proteína C12orf35, o la propia proteína C12orf35. Dichos procedimientos incluyen, por ejemplo transferencia Northern, FACS, ImageStream, transferencia Western, qPCR, RT-PCR, qRT-PCR, ELISA, Luminex, Multiplex, etc.

El nivel de expresión o actividad de la proteína C12orf35 puede reducirse al menos 3 veces, al menos 5 veces, al menos 10 veces, al menos 20 veces, al menos 30 veces, al menos 40 veces, al menos 50 veces, al menos 60 veces, al menos 70 veces, al menos 75 veces, al menos 80 veces, al menos 90 veces, al menos 100 veces, en comparación con un control.

*Genes FAM60A y proteínas FAM60A*

En determinadas realizaciones, la célula hospedadora es una célula en la que se reduce el nivel de expresión o actividad de la proteína FAM60A, en comparación con un control. La Solicitud de Patente Provisional de EE.UU. N.º 61/919.340, presentada el 20 de diciembre de 2013 proporciona una descripción detallada de las células de mamíferos en las que se reduce el nivel de expresión o actividad de la proteína FAM60A.

Puede usarse una diversidad de controles como se analizó anteriormente. El nivel de expresión o actividad de la proteína FAM60A de una célula de tipo silvestre correspondiente puede usarse como control. Alternativamente, un control puede ser un nivel predeterminado o un nivel umbral que puede identificarse en las referencias bibliográficas o en la base de datos.

La proteína FAM60A es una subunidad del complejo SIN3-histona desacetilasa (HDAC) (complejo SIN3/HDAC) que funciona en la represión transcripcional (Munoz y col., 2012, THE Journal of Biological Chemistry VOL. 287, N.º 39, pp. 32346-32353; Smith y col., 2012, Mol Cell Proteomics 11 (12): 1815-1828). Las desacetilasas de histonas (HDAC) catalizan la retirada de grupos acetilo de las histonas. La acetilación de histonas en lisinas es un mecanismo principal para modular la conformación de la cromatina. La acetilación de histonas promueve un estado de cromatina relajado, transcripcionalmente activo, mientras que la desacetilación catalizada por histona desacetilasas (HDAC) favorece un estado inactivo, silente. El análisis de la base de datos reveló la presencia de al menos un ortólogo FAM60A en la mayoría de los metazoos, pero no en nematodos. El gen FAM60A se conserva en metazoos y puede encontrarse en todos los genomas de vertebrados y la mayoría de invertebrados que se han secuenciado por completo. La investigación de similitud de secuencia de los homólogos de FAM60A indica que predominantemente, solo hay un miembro representativo de esta familia en el genoma. Solo hay algunas excepciones. De acuerdo con Smith y col., 2012, la proteína FAM60A tiene una secuencia única que carece de cualquier dominio de proteína conocido. Además, se describió por Smith y col. 2012, que no exhibe ninguna homología de secuencia con otras proteínas conocidas en el proteoma humano. La comparación de secuencias entre las proteínas FAM60A de diferentes especies mostró que la proteína FAM60A generalmente comprende tres regiones: (1) un N-terminal que comprende segmentos altamente conservados en todos los metazoos (2) una región media que está altamente conservada a través de vertebrados, mientras que en los invertebrados consiste en un espaciador no conservado de longitud variable (3) un C-terminal que comprende segmentos altamente conservados en todos los metazoos. Por lo tanto, la conservación más alta se observó en las regiones N- y C-terminales de FAM60A.

Los estudios indican que la proteína FAM60A se asocia con los complejos SIN3/HDAC en diversos tipos de células eucariotas, como en particular las células de mamíferos. Sin embargo, hasta la fecha, la información funcional sobre la proteína FAM60A es limitada. Estudios funcionales recientes (véase Smith y col., 2012) indican que la proteína FAM60A puede reprimir la expresión génica y regula un subconjunto específico de genes. Smith y col. 2012 informan un papel de la proteína FAM60A en la regulación de la ruta de señalización de TGF-beta, que juega un papel fundamental en procesos como la progresión del cáncer, la metástasis, la migración celular y la vigilancia inmunitaria. Hay hallazgos que indican que la proteína FAM60A actúa como un represor transcripcional de los componentes de la ruta de señalización de TGF-beta, mientras que esta función de la proteína FAM60A parece estar permitida a través de su papel en el complejo SIN3-HDAC. El agotamiento de la proteína FAM60A en diferentes líneas celulares de cáncer usando ARNip contra la secuencia codificante de FAM60A dio como resultado un cambio de la morfología celular normal del cáncer. Adicionalmente, se descubrió que los niveles de proteína FAM60A cambian periódicamente en el transcurso del ciclo celular en las células U2OS (Muñoz y col., 2012). Los experimentos de inactivación de FAM60A usando ARNip de FAM60A en células de osteosarcoma de hueso humano U2OS revelaron que la proteína FAM60A restringe la expresión del gen de ciclina D1.

Contra este contexto científico, fue sorprendente descubrir que reducir la expresión o actividad de la proteína FAM60A en una célula de mamífero aumenta la estabilidad de la expresión de proteínas recombinantes, sin afectar negativamente a otras características de la célula que son importantes para la expresión recombinante. Esta correlación entre los efectos de la proteína FAM60A y la estabilidad de la expresión durante el cultivo prolongado de las células fue inesperada. Como tal, las células hospedadoras de mamífero (por ejemplo, células CHO) con nivel de expresión reducido o actividad de la proteína FAM60A son particularmente adecuadas para la producción recombinante de complejo pentamérico.

El gen FAM60A se expresa endógenamente en metazoos y, por lo tanto, en especies de mamíferos tales como humano, ratón, rata y hámster, y la secuencia de aminoácidos de FAM60A está altamente conservada en especies de mamíferos, así como en vertebrados. Pueden asignarse diferentes nombres en diferentes especies para la proteína o el gen y los nombres alternativos no limitantes (alias) también se enumeran anteriormente en la Tabla 2 (a continuación). Por simplicidad, en la presente divulgación, los homólogos y ortólogos de diferentes especies se denominan "gen FAM60A" o "proteína FAM60A".

El listado de secuencias muestra secuencias de aminoácidos ejemplares de proteínas FAM60A conocidas y/o predichas de diferentes especies de vertebrados, a saber, *Homo sapiens* (SEQ ID NO: 28), *Mus musculus* (SEQ ID NO: 29), *Cricetulus griseus* (SEQ ID NO: 30). El ADNc de FAM60A predicho de *Cricetulus griseus* se muestra en SEQ ID NO: 31 (secuencia codificante de 14-679; véase también la Secuencia de Referencia NCBI: XM\_003505482.1). La proteína FAM60A no se ha descrito en detalle en las referencias bibliográficas. Por lo tanto, fue sorprendente que la

estabilidad de la expresión de una célula hospedadora recombinante pueda mejorarse, si el genoma de la célula hospedadora se altera de tal manera que se reduce el nivel de expresión o la actividad de la proteína endógena FAM60A, en comparación con un control. Fue inesperado que la proteína FAM60A influya en la estabilidad de la expresión de un producto recombinante de interés.

5 Se han informado secuencias del gen FAM60A que codifica la proteína FAM60A en *Homo sapiens* (NCBI Gene-ID: 58516); *Rattus norvegicus* (ID de gen NCBI: 686611); *Mus musculus* (ID de gen NCBI: 56306); *Bos taurus* (ID de gen NCBI: 538649) y otros. Las variantes de transcripción pueden existir de una manera dependiente de la especie y en diferentes números (por ejemplo, el gen FAM60A humano expresa tres isoformas de transcripción putativas que difieren en las UTR pero codifican la misma proteína).

10 La reducción del nivel de expresión o actividad de una proteína FAM60A puede lograrse por diversos medios. Por ejemplo, el nivel de expresión o actividad de una proteína FAM60A puede reducirse por inactivación génica, mutación génica, delección génica, silenciamiento génico o una combinación de cualquiera de los anteriores. La inactivación génica es una técnica genética mediante la cual un gen se vuelve inoperativo al interrumpir su función. Por ejemplo, puede insertarse un ácido nucleico en la secuencia codificante, interrumpiendo de esta manera la función del gen.

15 Adicionalmente, el gen FAM60A de longitud completa (o un fragmento del mismo) puede eliminarse, por lo que la expresión de la proteína FAM60A funcional se elimina sustancialmente. Por ejemplo, la delección puede ser de al menos el 30 %, al menos el 40 %, al menos el 50 %, al menos el 55 %, al menos el 60 %, al menos el 65 %, al menos el 70 %, al menos el 75 %, al menos el 80 %, al menos el 85 %, al menos el 90 %, al menos el 95 % o el 100 % de la secuencia codificante del gen FAM60A. Otra opción es introducir una o más mutaciones en la secuencia codificante, que produce una proteína FAM60A no funcional o menos funcional. Por ejemplo, pueden introducirse una o más mutaciones de desplazamiento de marco, dando como resultado una proteína FAM60A no funcional o menos funcional. Alternativa o adicionalmente, pueden introducirse uno o más codones de parada en la secuencia codificante de tal manera que se obtiene una proteína truncada, no funcional o menos funcional. Otras opciones incluyen, pero no se limitan a, una o más mutaciones en el promotor, en el 5'- y/o 3' UTR u otros elementos reguladores, por ejemplo, introduciendo una delección del promotor o introduciendo una construcción entre el promotor y el inicio de la transcripción. Los procedimientos para la interrupción génica para suprimir o eliminar la expresión del gen diana también son bien conocidos por la persona experta.

30 Como cada célula tiene dos copias del gen FAM60A en su genoma, en determinadas realizaciones, al menos una copia de la secuencia genómica del gen FAM60A, o al menos el 50 % de la secuencia codificante de dicho gen FAM60A, está eliminada. Por ejemplo, ambas copias de las secuencias genómicas del gen FAM60A (o al menos el 50 % de la secuencia codificante de dicho gen FAM60A de cada copia) se eliminan.

35 En determinadas realizaciones, la secuencia eliminada comprende una porción de la región telomérica del cromosoma 8 de una célula CHO. Una región telomérica es una región de secuencias repetitivas de nucleótidos en cada extremo de una cromátida, que protege el extremo del cromosoma del deterioro o de la fusión con los cromosomas vecinos. Por ejemplo, al menos el 30 %, al menos el 40 %, al menos el 50 %, al menos el 55 %, al menos el 60 %, al menos el 65 %, al menos el 70 %, al menos el 75 %, al menos el 80 %, al menos el 85 %, al menos el 90 %, al menos el 95 % o el 100 % de la secuencia de nucleótidos de la región telomérica del cromosoma 8 de una célula CHO se elimina.

La secuencia eliminada puede comprender además una delección del gen seleccionado del grupo que consiste en: Caprin2 e lpo8, y una combinación de los mismos.

40 La eliminación del gen FAM60A se debe a una ruptura cromosómica. La ruptura cromosómica puede inducirse por los procedimientos descritos anteriormente. Después de inducir la ruptura cromosómica, las células que tienen el punto de corte deseado (que da como resultado una delección del gen FAM60A) pueden identificarse, por ejemplo, analizando el ADN o usando el procedimiento de acuerdo con el quinto aspecto de la presente divulgación. Por ejemplo, el perfil de expresión de las células tratadas puede analizarse para determinar si el gen FAM60A o los genes ubicados centroméricos del gen FAM60A se expresan, si la expresión se reduce o si los genes no se expresan. Por ejemplo, en el caso de células de ratón o hámster, puede analizarse si el gen FAM60A se expresa y, alternativamente o además de ello, puede analizarse si uno o más genes seleccionados del grupo que consiste en Bcd1, C12orf35, proteína similar a la metiltransferasa 20, Dennd5b, Caprin2, lpo8, Tmtc1 o genes que se encuentran teloméricos de los genes mencionados anteriormente (en el que telomérico significa a este respecto en la dirección del extremo telomérico) se expresan por la célula y/o si la expresión se reduce o se elimina sustancialmente.

50 La reducción del nivel de expresión de la proteína FAM60A puede lograrse mediante el silenciamiento génico post-transcripcional, por ejemplo, por moléculas de ácido nucleico antisentido, o moléculas que median la interferencia de ARN. Los ejemplos no limitantes incluyen ARNip, ARNhc, miARN, oligonucleótidos antisentido, etc., todos los cuales son bien conocidos en la técnica.

55 El nivel de expresión de la proteína FAM60A puede evaluarse mediante procedimientos conocidos en la técnica, por ejemplo, midiendo el nivel de ARNm que codifica la proteína FAM60A, o la propia proteína FAM60A. Dichos procedimientos incluyen, por ejemplo transferencia Northern, FACS, ImageStream, transferencia Western, qPCR, RT-PCR, qRT-PCR, ELISA, Luminex, Multiplex, etc.

El nivel de expresión o actividad de la proteína FAM60A puede reducirse al menos 3 veces, al menos 5 veces, al menos 10 veces, al menos 20 veces, al menos 30 veces, al menos 40 veces, al menos 50 veces, al menos 60 veces, al menos 70 veces, al menos 75 veces, al menos 80 veces, al menos 90 veces, al menos 100 veces, en comparación con un control.

## 5 Genes de Matriptasa y Matriptasas

En determinadas realizaciones, la célula hospedadora es una célula en la que se reduce el nivel de expresión o actividad de la matriptasa, en comparación con un control. La solicitud de patente provisional de EE.UU. n.º 61/985.589, presentada el 29 de abril de 2014 y la Solicitud de Patente Provisional de EE.UU. N.º 61/994.310, presentada el 16 de mayo de 2014 proporciona una descripción detallada de las células de mamíferos en las que se reduce el nivel de expresión o la actividad de la matriptasa.

Puede usarse una diversidad de controles como se analizó anteriormente. El nivel de expresión o actividad de la matriptasa de una célula de tipo silvestre correspondiente puede usarse como control. Alternativamente, un control puede ser un nivel predeterminado o un nivel umbral que puede identificarse en las referencias bibliográficas o en la base de datos.

La matriptasa se describió por primera vez en 1993 como una nueva actividad gelatinolítica en células de cáncer de mama cultivadas. La matriptasa pertenece a la familia de las serina proteasas transmembrana de tipo II (TTSP). Los ortólogos de matriptasa están presentes en diferentes especies de vertebrados, incluyendo especies de mamíferos, y se identificaron, por ejemplo, en humano, chimpancé, perro, ratón, rata, pollo, pez cebra, pez globo verde manchado y pez globo tigre lo que sugiere una función evolutiva conservada. La matriptasa figura en la nomenclatura de enzimas IUBMB como EC 3.4.21.109. La matriptasa también se conoce como serina proteasa 1 de tipo membrana (MT-SP1) y supresor de tumorigenicidad-14 (ST14) (véase Chen y col., *The Transmembrane Serine Protease Matriptase: Implications for Cellular Biology and Human Diseases* J Med Sci 2012; 32 (3): 097-108). Es una proteína de membrana integral con un dominio transmembrana de un solo tramo cerca del N-terminal citoplasmático. La parte extracelular consiste en una región madre (que incluye un solo SEA, 2 dominios CUB y 4 dominios LDLRA) y el dominio de serina proteasa C-terminal que es estructuralmente muy similar a otros TTSP e incluye una tríada catalítica histidina/ácido aspártico/serina (HDS) conservada esencial para la actividad catalítica (véase, por ejemplo, List y col., *Matriptase: Potent Proteolysis on the cell Surface*; Mol Med 12 (1-3) 1-7, Enero-Marzo de 2006 y Chen y col., *The Transmembrane Serine Protease Matriptase: Implications for Cellular Biology and Human Diseases* J Med Sci 2012; 32 (3): 097-108). La matriptasa se describe como estando expresada en los epitelios en muchos sistemas de órganos tales como piel, mama, pulmón, epidermis, córnea, glándula salival, cavidades orales y nasales, tiroides, timo, esófago, tráquea, bronquiolos, alveolos, estómago, páncreas, vesícula biliar, duodeno, intestino delgado, colon, recto, riñón, adrenales, vejiga urinaria, uréter, vesículas seminales, epidídimo, próstata, ovarios, útero y vagina (véase List y col., 2006 y Chen y col., 2012). La matriptasa se sintetiza como un zimógeno inactivo y se convierte a su forma activa a través de un procedimiento complicado. Los detalles relacionados con el procedimiento de activación que implica la escisión endoproteolítica se describen para la matriptasa humana en List y col. 2006 y Chen y col. 2012. La matriptasa se une a la membrana como proteína transmembrana de tipo II con el dominio catalítico orientado hacia el espacio extracelular. Adicionalmente, se describe en las referencias bibliográficas que un desprendimiento significativo de matriptasa, respectivamente su parte extracelular, se produce *in vivo* (véase List y col., 2006 y Chen y col. 2012). Se describe en las referencias bibliográficas que la matriptasa se desprende en forma de un complejo, por ejemplo, complejo con el inhibidor de la serina proteasa de tipo Kunitz HAI-1. Diferentes estudios sugieren que en las células humanas el inhibidor específico HAI-1 facilita el transporte de la matriptasa a la membrana celular, ya que se demostró que la eliminación o incluso mutaciones puntuales en HAI-1 conducen a una acumulación de la matriptasa en el compartimento de Golgi. En las referencias bibliográficas, se han descrito varios inhibidores endógenos diferentes de la matriptasa además de HAI-1 tales como HAI-2, antitrombina, alfa-1 antitripsina y alfa -2-antiplasmina. Adicionalmente, también se han descrito otros inhibidores de la matriptasa (véase, por ejemplo, Chen y col., 2012). En las referencias bibliográficas se describe que la matriptasa puede desempeñar numerosos papeles en la fisiología normal tales como la función de barrera cutánea, integridad epitelial, desarrollo del folículo piloso y homeostasis del timo, y en patologías humanas, tales como osteoartritis, aterosclerosis y progresión tumoral, invasión y metástasis.

Contra este contexto científico que no está relacionado con la producción recombinante de una proteína, el hallazgo actual de que la matriptasa es una proteasa implicada en el recorte de proteínas producidas de forma recombinante que son secretadas por las células hospedadoras en el medio de cultivo celular fue muy sorprendente. Considerando la gran cantidad y variedad de proteasas expresadas en células de vertebrados, tales como en particular las células de mamíferos, fue aún más sorprendente que la reducción del nivel de expresión o actividad de esta proteína puede reducir significativamente el recorte del polipéptido secretado de interés en el medio de cultivo celular. Estos efectos ventajosos no se ven con otras proteasas incluso estrechamente relacionadas. Como tal, las células hospedadoras de mamífero (por ejemplo, células CHO) con nivel de expresión reducido o actividad de matriptasa son particularmente adecuadas para la producción recombinante de complejo pentamérico.

La lista de secuencias muestra secuencias de aminoácidos ejemplares de matriptasa de diferentes especies de vertebrados tales como el hámster (SEQ ID NO: 32 - secuencia de referencia NCBI: XP\_003495890), humano (SEQ ID NO: 33 - Secuencia de referencia NCBI: NP\_068813), ratón (SEQ ID NO: 34 - Secuencia de referencia NCBI: NP\_035306).

También se han informado las secuencias de nucleótidos que codifican matriptasa de diferentes especies de mamíferos. Véase, por ejemplo, hámster chino (ID de gen NCBI: 100755225); *Homo sapiens* (ID de gen NCBI: 6768); *Mus musculus* (ID de gen NCBI: 19143); *Rattus norvegicus* (ID de gen NCBI: 114093); Pan Troglodytes (ID de gen NCBI: 100188950) y otros. Los sinónimos de algunos de los genes de la matriptasa se enumeran en la Tabla 3, es comúnmente usado "ST14" o "St14".

Como se muestra en la Tabla 3, la matriptasa también se conoce como "supresor de la proteína 14 de tumorigenicidad" (por ejemplo, para humanos) y "supresor de la proteína homogénea 14 de tumorigenicidad" (por ejemplo, en ratón y hámster chino). Por simplicidad, en la presente divulgación, Los homólogos y ortólogos de diferentes especies se denominan "gen de matriptasa" o "matriptasa" (proteína).

También se describen en el presente documento variantes funcionales de una matriptasa (variantes que tienen sustancialmente las mismas actividades catalíticas que una matriptasa de tipo silvestre). Por ejemplo, una variante de matriptasa puede comprender una secuencia que es al menos el 50 %, al menos el 60 %, al menos el 70 %, al menos el 75 %, al menos el 80 %, al menos el 85 %, al menos el 90 %, al menos el 95 %, al menos el 96 %, al menos el 97 %, al menos el 98 % o al menos el 99 % idéntica a cualquiera de las secuencias de SEQ ID NO: 32-34 y tiene las mismas, o sustancialmente las mismas actividades catalíticas que una proteína de matriptasa de tipo silvestre. La actividad catalítica de una variante de matriptasa puede evaluarse, por ejemplo, por la reacción química para escindir varios sustratos sintéticos con Arg o Lys en la posición P1 y prefieren aminoácidos de cadena lateral pequeña, tales como Ala y Gly, en la posición P2 (véase EC 3.4.21.109).

También se describen en el presente documento fragmentos funcionales de una matriptasa (fragmentos que tienen sustancialmente las mismas actividades catalíticas de una matriptasa de longitud completa). Un fragmento funcional de una matriptasa puede ser un subconjunto de aminoácidos contiguos de la matriptasa de longitud completa desvelada en el presente documento, y tiene la misma actividad catalítica, o sustancialmente la misma, que la secuencia proteica de longitud completa. La actividad catalítica de un fragmento de matriptasa puede evaluarse, por ejemplo, por la reacción química para escindir varios sustratos sintéticos con Arg o Lys en la posición P1 y prefieren aminoácidos de cadena lateral pequeña, tales como Ala y Gly, en la posición P2 (véase EC 3.4.21.109).

La reducción del nivel de expresión o actividad de una matriptasa puede lograrse por diversos medios. Por ejemplo, el nivel de expresión o actividad de una matriptasa puede reducirse por la inactivación de genes, mutación génica, delección génica, silenciamiento génico o una combinación de cualquiera de los anteriores. La inactivación génica es una técnica genética mediante la cual un gen se vuelve inoperativo al interrumpir su función. Por ejemplo, puede insertarse un ácido nucleico en la secuencia codificante, interrumpiendo de esta manera la función del gen. Adicionalmente, el gen de la matriptasa de longitud completa (o un fragmento del mismo) puede eliminarse, por lo que la expresión de matriptasa funcional se elimina sustancialmente. Por ejemplo, la delección puede ser de al menos el 30 %, al menos el 40 %, al menos el 50 %, al menos el 55 %, al menos el 60 %, al menos el 65 %, al menos el 70 %, al menos el 75 %, al menos el 80 %, al menos el 85 %, al menos el 90 %, al menos el 95 % o el 100 % de la secuencia de codificación del gen de la matriptasa. Otra opción es introducir una o más mutaciones en la secuencia codificante, que produce una matriptasa no funcional o menos funcional. Por ejemplo, pueden introducirse una o más mutaciones de desplazamiento de marco, dando como resultado una matriptasa no funcional o menos funcional. Alternativa o adicionalmente, pueden introducirse uno o más codones de parada en la secuencia codificante de tal manera que se obtiene una proteína truncada, no funcional o menos funcional. Otras opciones incluyen, pero no se limitan a, una o más mutaciones en el promotor, en el 5'- y/o 3' UTR u otros elementos reguladores, por ejemplo, introduciendo una delección del promotor o introduciendo una construcción entre el promotor y el inicio de la transcripción. Los procedimientos para la interrupción génica para suprimir o eliminar la expresión del gen diana también son bien conocidos por la persona experta.

Como cada célula tiene dos copias del gen de la matriptasa en su genoma, en determinadas realizaciones, al menos una copia de la secuencia genómica del gen de la matriptasa, o al menos el 50 % de la secuencia de codificación de dicho gen de la matriptasa, o un fragmento funcional de dicho gen de la matriptasa, está eliminada. Por ejemplo, ambas copias de las secuencias genómicas del gen de la matriptasa (o al menos el 50 % de la secuencia codificante de dicho gen de la matriptasa, o un fragmento funcional de dicho gen de la matriptasa, de cada copia) se eliminan.

En determinadas realizaciones, la célula hospedadora comprende una mutación en el exón 2 del gen de la matriptasa. El exón 2 es particularmente adecuado como una diana para alterar la actividad de la matriptasa porque existen varias variantes de corte y empalme funcional diferentes. Por lo tanto, los exones cercanos al N-terminal de la matriptasa tales como, por ejemplo, el exón 1, el exón 2 y el exón 3, son ventajosos para introducir una o más mutaciones, en particular una o más mutaciones de desplazamiento de marco. Una mutación de desplazamiento de fotograma en uno de estos exones muy probablemente conduce a un codón de detención temprano en la secuencia. Las mutaciones también pueden introducirse en uno de los exones posteriores, por ejemplo, seleccionado de los exones 4 a 19.

La matriptasa puede comprender como una mutación en el dominio catalítico. El dominio catalítico es la región de una enzima que interactúa con su sustrato para provocar la reacción enzimática. Pueden introducirse una o más mutaciones en este dominio para que la actividad catalítica de la proteína se reduzca en comparación con un control. El dominio catalítico comprende secuencias codificadas por los exones 16, 17, 18 y 19. Las mutaciones pueden ser una delección, una inserción, una sustitución o una combinación de los mismos. Las mutaciones pueden causar una

mutación de desplazamiento de marco, una mutación puntual específica, una mutación de codón de parada o una recombinación de las mismas, en la secuencia que codifica el dominio catalítico. También se han descrito mutantes catalíticos inactivos de matriptasa tales como, por ejemplo, G827R-matriptasa o S805A-matriptasa (véase Desilets y col., The Journal of Biological Chemistry Vol. 283, N.º 16, pp. 10535-10542, 2008). Adicionalmente, se conoce la estructura cristalina del dominio catalítico de una matriptasa recombinante. A partir de esta estructura y datos de secuencia, la persona experta puede derivar dianas específicas adicionales para mutaciones que perjudiquen la función catalítica de la matriptasa.

La reducción del nivel de expresión de la matriptasa puede lograrse mediante el silenciamiento génico post-transcripcional, por ejemplo, por moléculas de ácido nucleico antisentido, o moléculas que median la interferencia de ARN. Los ejemplos no limitantes incluyen ARNip, ARNhc, miARN, oligonucleótidos antisentido, etc., todos los cuales son bien conocidos en la técnica.

El nivel de expresión de matriptasa puede evaluarse mediante procedimientos conocidos en la técnica, por ejemplo, midiendo el nivel de ARNm que codifica la matriptasa o la propia matriptasa. Dichos procedimientos incluyen, por ejemplo transferencia Northern, FACS, ImageStream, transferencia Western, qPCR, RT-PCR, qRT-PCR, ELISA, Luminex, Multiplex, etc. La actividad de la matriptasa puede evaluarse, por ejemplo, de acuerdo con su actividad enzimática.

El nivel de expresión o actividad de la matriptasa puede reducirse al menos 3 veces, al menos 5 veces, al menos 10 veces, al menos 20 veces, al menos 30 veces, al menos 40 veces, al menos 50 veces, al menos 60 veces, al menos 70 veces, al menos 75 veces, al menos 80 veces, al menos 90 veces, al menos 100 veces, en comparación con un control.

**Tabla 2:** Abreviaturas y nombres alternativos (alias) de productos codificados por genes ubicados en el cromosoma 8 del hámster chino o el cromosoma 6 del ratón.

Abreviaturas	En una anotación pública del hámster chino, el producto génico se anota como	En la anotación pública del ratón el producto genético se anota como	Alias (véase <a href="http://www.genecards.org">www.genecards.org</a> )
<b>Ccdc91</b>	Proteína que contiene un dominio en espiral 91	Dominio en espiral que contiene 91	Dominio en espiral que contiene 91 P56 Compañero de unión de GGA Proteína accesoria P56 DKFZp779L1558 FLJ11088 Proteína que contiene un dominio en espiral 91 Compañero de unión de GGA GGABP
<b>Far2</b>	Acil graso CoA reductasa 1 isoforma 1	Acil graso CoA reductasa 2	Acil Graso CoA Reductasa 2 MLSTD1 SDR10E2 Proteína 1 que contiene el dominio de esterilidad masculina EC 1.2.1.N2 FLJ10462 Dominio que contiene esterilidad masculina 1 Acil-CoA Graso Reductasa 2 Cadena corta Familia de deshidrogenasa/reductasa 10E, Miembro 2

(continuación)

Abreviaturas	En una anotación pública del hámster chino, el producto génico se anota como	En la anotación pública del ratón el producto genético se anota como	Alias (véase <a href="http://www.genecards.org">www.genecards.org</a> )
<b>Ergic2</b>	proteína 2 de compartimento intermedio de Retículo endoplásmico-Golgi	ERGIC y golgi 2	ERGIC Y Golgi 2 PTX1 Erv41 Cd002 Proteína CD14 Retículo endoplásmico-Golgi Proteína 2 de compartimento intermedio ERV41 CDA14
<b>RPS4Y2</b>	Proteína ribosómica 40S S4, Tipo Isoforma X	Proteína ribosómica S4, enlazada a Y 2	Proteína Ribosómica S4, Enlazada a Y 2 RPS4Y2P Proteína Ribosómica S4, Enlazada a Y 2 Pseudogén Proteína ribosómica 40S S4, Y Proteína ribosómica 40S S4, Y isoforma 2
<b>Tmtc1</b>	Proteína 1 que contiene repetición transmembrana y TPR	1 que contiene repetición transmembrana y tetratricopéptido	1 Que Contiene Repetición Transmembrana y Tetratricopéptido OLF ARG99 FLJ31400 FLJ41625 TMTC1A 1A que contiene Repetición transmembrana y tetratricopéptido Proteína 1 Que Contiene Repetición Transmembrana y TPR
<b>Zfp 1</b>	Proteína 1 que contiene un dominio HIT de dedo de cinc		Proteína de dedo de zinc ZFP1 ZNF475 Proteína de dedo de zinc 475 FLJ34243 Proteína de dedo de zinc 1 Proteína de dedo de zinc 1 homólogo (ratón) Zfp-1 Proteína de dedo de zinc 1 homólogo
<b>IPO8</b>	Tipo Importina 8	Importina 8	Importina 8 RANBP8 Proteína de unión RAN 8 Proteína de unión Ran 8 IMP8 Imp8 Importina 8 RanBP8

(continuación)

Abreviaturas	En una anotación pública del hámster chino, el producto génico se anota como	En la anotación pública del ratón el producto genético se anota como	Alias (véase <a href="http://www.genecards.org">www.genecards.org</a> )
<b>Caprin2</b>	Proteína similar a caprina 2	Miembro de la familia de caprina 2	<p>Miembro de la familia de caprina 2 C1QDC1 EEG1 RNG140 Caprina 2 Proteína 2 de Activación citoplasmática/asociada a la proliferación</p> <p>Proteína asociada a Cáncer gástrico de resistencia a múltiples fármacos</p> <p>Proteína que contiene dominio C1q 1 Proteína 140 de gránulo de ARN FLJ11391 FLJ22569 1 que contiene dominio C1q EEG-1 KIAA1873 Proteína EEG-1</p>
<b>FAM60A</b>	Proteína tipo FAM60A	Familia con similitud de secuencia 60, miembro A	<p>Familia con similitud de secuencia 60, Miembro A C12orf14 TERA Homólogo de proteína Tera Marco de lectura abierto 14 del cromosoma 12 Proteína FAM60A</p>
<b>Dennd5b</b>	Similar a proteína 5B que contiene dominio Denn	5B que contiene dominio DENN/MADD	<p>Dominio DENN/MADD que contiene la proteína tipo 5B Rab6IP1 MGC24039 Proteína 5B que contiene el dominio DENN</p>
<b>METTL20</b>	Proteína similar a la metiltransferasa 20	4833442J19Rik	<p>METTL20 C12orf72 DKFZp451L235 MGC50559 Marco de lectura abierto 72 del cromosoma 12 Proteína similar a la metiltransferasa 20 EC 2.1.1.</p>
<b>AMN1</b>	Proteína similar a AMN1 proteína putativa	Antagonista de la red de salida mitótica 1	<p>Antagonista De La Red De Salida Mitótica 1 Homólogo AMN1 Proteína Homóloga (S. Cerevisiae)</p>
	Proteína similar al receptor del factor de crecimiento opioide	Proteína similar al receptor del factor de crecimiento opioide	

(continuación)

Abreviaturas	En una anotación pública del hámster chino, el producto génico se anota como	En la anotación pública del ratón el producto genético se anota como	Alias (véase <a href="http://www.genecards.org">www.genecards.org</a> )
<b>C12orf35</b>	Homólogo C12orf35 de proteína no caracterizada	ortólogo probable del marco de lectura abierto 35 del cromosoma 12 de H. sapiens (C12orf35); 2810474O19Rik	KIAA1551, C12orf35, FLJ10652, FLJ20696 Marco de lectura abierto 35 del cromosoma 12 Proteína no caracterizada C12orf35 Proteína no caracterizada KIAA1551
<b>Bicd 1</b>	Proteína putativa bicaudal D	Homólogo 1 de Bicaudal D	Homólogo 1 de Bicaudal D (Drosophila) Bic-D 1 Bicaudal D (Drosophila) Homólogo 1 BICD Similar al citoesqueleto Proteína Bicaudal D Homólogo 1 Proteína Bicaudal D Homólogo 1

**Tabla 3:** Ejemplos de nombres alternativos del gen de matriptasa y/o el producto de proteína codificado matriptasa usados en las referencias bibliográficas (orden alfabético)

Proteasa de 80 kDa de cáncer de mama
CAP3
Proteasa activadora de canales 3
EC 3.4.21.109
Epitina
HAI
Matriptasa
Matriptasa-1
Serina proteasa 1 de tipo membrana
MT-SP1; MTSP1
Prostamina
PRSS14 g.p. (Homo sapiens), PRSS14
Serina endopeptidasa SNC19
Serina proteasa 14
Serina proteasa TADG-15; TADG-15; TADG15
SNC19
ST14 (nombre oficial del gen en humanos de acuerdo con HGNC)
St14
Proteína 14 de supresión de tumorigenicidad; supresión de tumorigenicidad 14 (carcinoma de colon, matriptasa, epitina)
Proteína 14 supresora de tumorigenicidad
Homólogo (ratón) de proteína 14 supresora de tumorigenicidad
TMPRSS14
Tumor asociado a la proteína del gen 15 expresada diferencialmente

Una célula hospedadora adecuada puede tener cualquier combinación de modificaciones descritas en el presente documento, por ejemplo, una célula en la que tanto el nivel de expresión o actividad de la proteína C12orf35 se reduce en dicha célula, en comparación con un control, y el nivel de expresión o actividad de la proteína FAM60A se reduce en dicha célula, en comparación con un control. Otras combinaciones incluyen, por ejemplo, reducción del nivel o actividad de expresión de proteína C12orf35, y reducción del nivel o actividad de expresión de matriptasa; reducción del nivel o actividad de expresión de proteína FAM60A, y reducción del nivel o actividad de expresión de matriptasa; reducción del nivel o actividad de expresión de proteína C12orf35, reducción del nivel o actividad de expresión de proteína FAM60A, y reducción del nivel o actividad de expresión de matriptasa; reducción del nivel o actividad de expresión de la proteína C12orf35 e inclusión de la secuencia de dihidrofolato reductasa (DHFR) como un marcador de selección, etc.

Las células hospedadoras descritas en el presente documento son adecuadas para cultivo a gran escala. Por ejemplo, los cultivos celulares pueden ser de 10 l, 30 l, 50 l, 100 l, 150 l, 200 l, 300 l, 500 l, 1000 l, 2000 l, 3000 l, 4000 l, 5000 l, 10.000 l o más grandes. Por ejemplo, el tamaño del cultivo celular puede variar de 10 l a 5000 l, de 10 l a 10.000 l, de 10 l a 20.000 l, de 10 l a 50.000 l, de 40 l a 50.000 l, de 100 l a 50.000 l, de 500 l a 50.000 l, de 1000 l a 50.000 l, de 2000 l a 50.000 l, de 3000 l a 50.000 l, de 4000 l a 50.000 l, de 4500 l a 50.000 l, de 1000 l a 10.000 l, de 1000 l a 20.000 l, de 1000 l a 25.000 l, de 1000 l a 30.000 l, de 15 l a 2000 l, de 40 l a 1000 l, de 100 l a 500 l, de 200 l a 400 l, o cualquier número entero entre ellos.

Los componentes de medios para el cultivo celular son conocidos en la técnica y pueden incluir, por ejemplo, tampones, contenido de aminoácidos, contenido de vitaminas, contenido de sal, contenido mineral, contenido de suero, contenido de fuente de carbono, contenido de lípidos, contenido de ácido nucleico, contenido hormonal, contenido de elementos traza, contenido de amoníaco, contenido de cofactor, contenido del indicador, contenido de molécula pequeña, contenido de hidrolizado y contenido de modulador enzimático.

#### 4. PURIFICACIÓN DEL COMPLEJO PENTAMÉRICO DEL CULTIVO CELULAR

El complejo pentamérico producido de acuerdo con los procedimientos descritos en el presente documento puede cosecharse de células hospedadoras y purificarse usando cualquier procedimiento adecuado. Los procedimientos adecuados incluyen precipitación y diversos tipos de cromatografía, tales como de interacción hidrófoba, de intercambio iónico, de afinidad, de quelación y de exclusión por tamaño, todas las cuales se conocen en la técnica. Pueden crearse esquemas de purificación adecuados usando dos o más de estos u otros procedimientos adecuados. Si se desea, una o más de las subunidades del complejo pentamérico pueden comprender una "etiqueta" que facilita la purificación, tales como una etiqueta epítipo o una etiqueta HIS, etiqueta Strep. Tales polipéptidos etiquetados pueden purificarse convenientemente, por ejemplo de medios condicionados, por cromatografía quelante o cromatografía de afinidad. Opcionalmente, la secuencia de etiqueta puede escindirse después de la purificación.

Por ejemplo, El documento WO2014/005959 desvela procedimientos ejemplares de purificación de complejo pentamérico por cromatografía de afinidad.

Como se describe en el presente documento, una o más subunidades del complejo pentamérico pueden comprender una etiqueta para la purificación por afinidad. Algunos ejemplos de etiquetas de purificación por afinidad incluyen, por ejemplo, etiqueta His (se une al ion metálico), un anticuerpo (se une a la proteína A o proteína G), proteína de unión a maltosa (MBP) (se une a la amilosa), glutatión-S-transferasa (GST) (se une al glutatión), etiqueta FLAG (Asp-Tyr-Lys-Asp-Asp-Asp-Asp-Lys (SEQ ID NO: 37)) (se une a un anticuerpo anti-flag), etiqueta Strep (se une a estreptavidina o un derivado de la misma).

Se desconoce la estructura del complejo pentamérico gH/gL/pUL128/pUL130/pUL131. Sin embargo, si la etiqueta de purificación por afinidad está adherida a un sitio que interfiere con la formación del complejo pentamérico o un sitio que está enterrado dentro del complejo, la purificación por afinidad no sería exitosa. Se cree que los siguientes sitios son adecuados para unir una etiqueta de purificación por afinidad, ya que la etiqueta no parece interferir con la formación del complejo pentamérico y parece estar expuesta en la superficie de un pentámero ensamblado: (i) región C-terminal de pUL130, (ii) región N-terminal de pUL130, (iii) región C-terminal de pUL131, (iv) región N-terminal de pUL131, (v) región C-terminal de pUL128, (vi) región N-terminal de pUL128 o una combinación de los mismos.

El complejo pentamérico puede no comprender una etiqueta de purificación.

Otro procedimiento adecuado es la cromatografía de intercambio iónico. Los ejemplos de materiales útiles en la cromatografía de intercambio iónico incluyen DEAE-celulosa, QAE-celulosa, DEAE-cefalosa, QAE-cefalosa, DEAE-Toyopearl, QAE-Toyopearl, Mono Q, Mono S, Q sepharose, SP sepharose, etc. En particular, el procedimiento puede usar la columna Mono S. Alternativamente, el procedimiento puede usar la columna Mono Q.

El rendimiento del complejo pentamérico de CMV 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.

5 **5. DEFINICIONES**

La expresión "fragmento formador complejo" de una proteína CMV (tales como gH, gL, pUL128, pUL130 o pUL131) se refiere a cualquier parte o porción de la proteína que retiene la capacidad de formar un complejo con otra proteína CMV. Dichos complejos incluyen, por ejemplo, complejo pentamérico gH/gL/pUL128/pUL130/pUL131. Los fragmentos que retienen la capacidad de formar el complejo pentamérico también se denominan "fragmentos formadores de pentámero".

Un "cultivo a gran escala" se refiere a un cultivo que tiene al menos aproximadamente 10 litros de tamaño, (por ejemplo, un volumen 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, a 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.).

Un complejo pentamérico "soluble" se refiere al complejo gH/gL/pUL128/pUL130/pUL131 en el que la subunidad gH no comprende el dominio transmembrana.

A lo largo de la memoria descriptiva, incluyendo las reivindicaciones, cuando el contexto lo permite, la expresión "que comprende" y variantes como "comprende" o "que comprenden", deben interpretarse como que incluyen el número entero o los números enteros indicados sin excluir necesariamente ningún otro número entero.

La identidad de secuencia se calcula de acuerdo con el porcentaje de coincidencias de restos entre dos secuencias de polinucleótidos, o coincidencias de nucleótidos entre dos secuencias de polinucleótidos, alineados usando un algoritmo normalizado. Un algoritmo tal puede insertar, de manera normalizada y reproducible, huecos en las secuencias que se comparan para optimizar la alineación entre dos secuencias y, por lo tanto, lograr una comparación más significativa de las dos secuencias. El porcentaje de identidad puede medirse sobre la longitud de una secuencia definida completa, o puede medirse sobre una longitud más corta, por ejemplo, a lo largo de un fragmento tomado de una secuencia definida más grande, por ejemplo, un fragmento de al menos 45, al menos 60, al menos 90, al menos 120, al menos 150, al menos 210 o al menos 450 restos o nucleótidos contiguos. Si no se especifica la longitud, la identidad de secuencia se calcula a través de la longitud completa de la más corta de las dos secuencias.

La presente invención se ilustra adicionalmente mediante los siguientes ejemplos que no deben interpretarse como limitantes.

**Ejemplos**

Este ejemplo se refiere a la producción del complejo de la proteína pentamérica del CMV mediante líneas celulares CHO en las que las secuencias codificantes de las subunidades pentaméricas se integraron de manera estable en el cromosoma. Las líneas celulares también se denominan líneas celulares estables CHO.

Como se muestra en este ejemplo, estas líneas celulares CHO estables produjeron pentámero de CMV funcional, con las cinco subunidades ensambladas en la conformación natural. El rendimiento es alto, permitiendo de esta manera la producción de complejos pentaméricos a gran escala comercial. Estas líneas celulares son particularmente adecuadas para la fabricación a gran escala de vacunas contra el CMV usando pentámero.

Como se muestra en este ejemplo, las secuencias de nucleótidos que codifican el ectodominio gH, gL, pUL128, pUL130 y pUL131 se clonaron en vectores de expresión únicos o dobles, con promotores subgenómicos e IRES para impulsar la expresión del componente individual (Figura 1). Específicamente, las secuencias codificantes para el ectodominio gH, gL de longitud completa, pUL128, pUL130 y pUL131 fueron codones optimizados para la expresión de cultivos de mamíferos, sintetizados y clonados en un solo vector de expresión (con selección Neo y DHFR) dirigido por el promotor CMV (para gH y UL128) y EV71 IRES (para gL, UL130, UL131) para expresión; o vectores de doble expresión, uno con selección Neo y DHFR para gH impulsado por el promotor CMV, y gL impulsado por el promotor

CMV o EV71 IRES, y uno con selección Hyg y Fra para UL128 dirigido por el promotor CMV, UL130 y UL131 dirigidos por EV71 IRES (Figura 1). Las etiquetas de purificación por afinidad, tales como la etiqueta His y/o la etiqueta Strep, se introdujeron en el extremo C del ectodominio gH y/o UL130 para facilitar la purificación. El codón de parada y el dominio transmembrana de gH también se introdujeron en el extremo C del ectodominio de gH para facilitar la subclonación de FACS (Figura 1).

El plásmido o plásmidos de expresión se transfectaron en un panel de células hospedadoras CHO, incluyendo (i) CHOC8TD, (ii) CHOHP3 y (iii) líneas celulares CHOC8TD matriptasa KO.

CHOC8TD deriva de una línea celular CHO K1 y se modifica adicionalmente eliminando la región telomérica del cromosoma, entonces esta línea celular tiene las siguientes características: (i) telómero del cromosoma 8 eliminado; (ii) alta productividad y buen crecimiento celular; y (iii) problemas potenciales de alta degradación proteolítica. CHOHP3 es una línea celular CHO K1 que muestra actividad proteolítica reducida, entonces esta línea celular tiene las siguientes características: (i) menor degradación proteolítica que CHOC8TD y (ii) menor crecimiento, productividad y eficiencia de clonación que CHOC8TD. CHOC8TD matriptasa KO derivó de CHOC8TD, con matriptasa inactivada, entonces esta línea celular tiene las siguientes características: (i) Matriptasa (serina proteasa) inactivada; y (ii) menos degradación proteolítica.

El plásmido o plásmidos de expresión que comprenden secuencias codificantes para gH, gL, UL128, UL130 y UL131 se transfectaron en las tres células hospedadoras con nucleofección AMAXA. Para la estrategia de un solo vector, las células transfectadas se seleccionaron por G418 y MTX secuencialmente. Para la estrategia de vector doble, las células transfectadas se seleccionaron por Hyg y después MTX secuencialmente.

Los grupos de células supervivientes se cultivaron en lotes para producir proteína pentámero para evaluar el rendimiento de la proteína con ELISA indirecto, la contaminación de gH/gL y degradación de gL con transferencia Western y SDS-PAGE. Basado en el rendimiento del pentámero y el crecimiento celular, el grupo de transfecto CHO con Vector Strategy 2 (Figura 1), usando células hospedadoras CHOC8TD, se seleccionaron para la clonación de células individuales (plásmidos de expresión mostrados en la Figura 2). Los clones seleccionados muestran alto rendimiento y menos contaminación por gH/gL.

Los clones individuales se clasificaron con FACS usando mAb específico de pentámero (-200). Los 30 clones principales se seleccionaron primero a base del título de pentámero evaluado con ELISA indirecto y después se seleccionaron a los 12 clones principales. El cultivo discontinuo en lotes seguido de purificación por afinidad con Strep mostró que los 12 principales pueden producir pentámero con rendimiento > 300 mg/l y alta pureza (Figura 3). De hecho, los clones superiores produjeron pentámero con un rendimiento de aproximadamente 0,3 g/l a aproximadamente 0,5 g/l. Se evaluó la unión del complejo proteico purificado con un panel de mAb contra pentámero, y se mostró que presentaba todos los epítopos clave.

Los 12 clones principales se evaluaron adicionalmente a través de estudios de estabilidad de 14 semanas (Figura 4). Se evaluaron los lotes controlados y no controlados para determinar el crecimiento celular y el título de pentámero y se seleccionó el Clon VF7.

También se evaluó la producción de biorreactor de pentámero con el clon superior VF7. Este clon produjo > 100 mg/l de pentámero purificado (Figura 5).

Como se ejemplifica en el presente documento, pueden usarse células CHO estables para expresar pentámero. Esto está en contraste con la expresión transitoria por las células HEK293. Las líneas celulares estables de CHO de la invención pueden producir proteína pentámero consistentemente con un rendimiento 100 veces mayor.

La capacidad de producir líneas celulares CHO estables en las que las secuencias de codificación de las cinco subunidades del pentámero se integran en el cromosoma, como se ejemplifica en el presente documento, es bastante notable. Siempre existen incertidumbres sobre si se puede generar una línea celular estable, incluso para una sola proteína. Por ejemplo, cuando se introdujo una secuencia codificante para IGF-1 humano en células CHO para generar una línea estable, se descubrió que las líneas celulares IGF-1/CHO resultantes mostraban inhibición del crecimiento celular y títulos bajos. Esta medida de título máximo fue de aproximadamente 8 ug/ml, que corresponde a 100 mg/l de un título de anticuerpos (basado en la masa molar). En comparación, las mediciones de título promedio de un anticuerpo recombinante en el procedimiento de biorreactor son alrededor de 3 g/l. Una causa del bajo título de IGF-1 fue la reducción del crecimiento celular y la baja viabilidad celular de las células que expresan IGF-1. Durante un procedimiento de expresión de anticuerpos, Las células celulares derivadas de CHO-K1 crecen hasta  $2 \times 10^7$  células/ml y la viabilidad celular es superior al 97 % durante las primeras 230-260 h de cultivo. En contraste, las células derivadas de CHO-K1 que expresan IGF-1 crecieron solo hasta  $0,5 \times 10^7$  células/ml y la viabilidad celular ya había caído a menos del 97 % después de dos días. Para detalles adicionales, véase, la Solicitud provisional de EE.UU. N.º 61/738466, presentada el 18 de diciembre de 2012 y la publicación de solicitud PCT N.º WO/2014/097113, presentada el 16 de diciembre de 2013.

Por lo tanto, los presentes inventores enfrentaron dificultades adicionales significativas porque las secuencias de codificación para las cinco subunidades pentámeras de CMV deben integrarse de manera estable en el genoma de la célula CHO. Los inventores superaron estas dificultades, como lo demuestran los clones seleccionados que muestran

altos rendimientos, con pentámero producido recombinantemente en su conformación natural y con todos los epítomos clave.

LISTADO DE SECUENCIAS

<110> GLAXOSMITHKLINE BIOLOGICALS sa

5 <120> CÉLULAS DE MAMÍFEROS QUE EXPRESAN ANTÍGENOS DE CITOMEGALOVIRUS

<130> VN56323

<140> 14191385.5

<141> 31/10/2014

<160> 39

10 <170> PatentIn versión 3.5

<210> 1

<211> 742

<212> PRT

<213> Herpesvirus humano 5

15 <220>

<221> MISC\_FEATURE

<222> (1)..(742)

<223> /nota="Cepa Merlin"

<400> 1

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

20

ES 2 799 735 T3

Gln Asp Ser Leu Gly Glu Gln Pro Thr Thr Val Pro Pro Pro Ile Asp  
145 150 155 160

Leu Ser Ile Pro His Val Trp Met Pro Pro Gln Thr Thr Pro His Gly  
165 170 175

Trp Thr Glu Ser His Thr Thr Ser Gly Leu His Arg Pro His Phe Asn  
180 185 190

Gln Thr Cys Ile Leu Phe Asp Gly His Asp Leu Leu Phe Ser Thr Val  
195 200 205

Thr Pro Cys Leu His Gln Gly Phe Tyr Leu Ile Asp Glu Leu Arg Tyr  
210 215 220

Val Lys Ile Thr Leu Thr Glu Asp Phe Phe Val Val Thr Val Ser Ile  
225 230 235 240

Asp Asp Asp Thr Pro Met Leu Leu Ile Phe Gly His Leu Pro Arg Val  
245 250 255

Leu Phe Lys Ala Pro Tyr Gln Arg Asp Asn Phe Ile Leu Arg Gln Thr  
260 265 270

Glu Lys His Glu Leu Leu Val Leu Val Lys Lys Asp Gln Leu Asn Arg  
275 280 285

His Ser Tyr Leu Lys Asp Pro Asp Phe Leu Asp Ala Ala Leu Asp Phe  
290 295 300

Asn Tyr Leu Asp Leu Ser Ala Leu Leu Arg Asn Ser Phe His Arg Tyr  
305 310 315 320

Ala Val Asp Val Leu Lys Ser Gly Arg Cys Gln Met Leu Asp Arg Arg  
325 330 335

Thr Val Glu Met Ala Phe Ala Tyr Ala Leu Ala Leu Phe Ala Ala Ala  
340 345 350

Arg Gln Glu Glu Ala Gly Ala Gln Val Ser Val Pro Arg Ala Leu Asp  
355 360 365

Arg Gln Ala Ala Leu Leu Gln Ile Gln Glu Phe Met Ile Thr Cys Leu  
370 375 380

Ser Gln Thr Pro Pro Arg Thr Thr Leu Leu Leu Tyr Pro Thr Ala Val  
385 390 395 400

ES 2 799 735 T3

Asp Leu Ala Lys Arg Ala Leu Trp Thr Pro Asn Gln Ile Thr Asp Ile  
 405 410 415  
 Thr Ser Leu Val Arg Leu Val Tyr Ile Leu Ser Lys Gln Asn Gln Gln  
 420 425 430  
 His Leu Ile Pro Gln Trp Ala Leu Arg Gln Ile Ala Asp Phe Ala Leu  
 435 440 445  
 Lys Leu His Lys Thr His Leu Ala Ser Phe Leu Ser Ala Phe Ala Arg  
 450 455 460  
 Gln Glu Leu Tyr Leu Met Gly Ser Leu Val His Ser Met Leu Val His  
 465 470 475 480  
 Thr Thr Glu Arg Arg Glu Ile Phe Ile Val Glu Thr Gly Leu Cys Ser  
 485 490 495  
 Leu Ala Glu Leu Ser His Phe Thr Gln Leu Leu Ala His Pro His His  
 500 505 510  
 Glu Tyr Leu Ser Asp Leu Tyr Thr Pro Cys Ser Ser Ser Gly Arg Arg  
 515 520 525  
 Asp His Ser Leu Glu Arg Leu Thr Arg Leu Phe Pro Asp Ala Thr Val  
 530 535 540  
 Pro Ala Thr Val Pro Ala Ala Leu Ser Ile Leu Ser Thr Met Gln Pro  
 545 550 555 560  
 Ser Thr Leu Glu Thr Phe Pro Asp Leu Phe Cys Leu Pro Leu Gly Glu  
 565 570 575  
 Ser Phe Ser Ala Leu Thr Val Ser Glu His Val Ser Tyr Ile Val Thr  
 580 585 590  
 Asn Gln Tyr Leu Ile Lys Gly Ile Ser Tyr Pro Val Ser Thr Thr Val  
 595 600 605  
 Val Gly Gln Ser Leu Ile Ile Thr Gln Thr Asp Ser Gln Thr Lys Cys  
 610 615 620  
 Glu Leu Thr Arg Asn Met His Thr Thr His Ser Ile Thr Val Ala Leu  
 625 630 635 640  
 Asn Ile Ser Leu Glu Asn Cys Ala Phe Cys Gln Ser Ala Leu Leu Glu  
 645 650 655

ES 2 799 735 T3

Tyr Asp Asp Thr Gln Gly Val Ile Asn Ile Met Tyr Met His Asp Ser  
660 665 670

Asp Asp Val Leu Phe Ala Leu Asp Pro Tyr Asn Glu Val Val Val Ser  
675 680 685

Ser Pro Arg Thr His Tyr Leu Met Leu Leu Lys Asn Gly Thr Val Leu  
690 695 700

Glu Val Thr Asp Val Val Val Asp Ala Thr Asp Ser Arg Leu Leu Met  
705 710 715 720

Met Ser Val Tyr Ala Leu Ser Ala Ile Ile Gly Ile Tyr Leu Leu Tyr  
725 730 735

Arg Met Leu Lys Thr Cys  
740

<210> 2  
<211> 742  
<212> PRT  
<213> Herpesvirus humano 5

5

<220>  
<221> MISC\_FEATURE  
<222> (1)..(742)  
<223>/nota = "CepaTowne"

10

<400> 2

Met Arg Pro Gly Leu Pro Ser Tyr Leu Ile Val Leu Ala Val Cys Leu  
1 5 10 15

Leu Ser His Leu Leu Ser Ser Arg Tyr Gly Ala Glu Ala Ile Ser Glu  
20 25 30

Pro Leu Asp Lys Ala Phe His Leu Leu Leu Asn Thr Tyr Gly Arg Pro  
35 40 45

Ile Arg Phe Leu Arg Glu Asn Thr Thr Gln Cys Thr Tyr Asn Ser Ser  
50 55 60

Leu Arg Asn Ser Thr Val Val Arg Glu Asn Ala Ile Ser Phe Asn Phe  
65 70 75 80

Phe Gln Ser Tyr Asn Gln Tyr Tyr Val Phe His Met Pro Arg Cys Leu  
85 90 95

Phe Ala Gly Pro Leu Ala Glu Gln Phe Leu Asn Gln Val Asp Leu Thr



ES 2 799 735 T3

Arg Gln Glu Glu Ala Gly Ala Gln Val Ser Val Pro Arg Ala Leu Asp  
 355 360 365

Arg Gln Ala Ala Leu Leu Gln Ile Gln Glu Phe Met Ile Thr Cys Leu  
 370 375 380

Ser Gln Thr Pro Pro Arg Thr Thr Leu Leu Leu Tyr Pro Thr Ala Val  
 385 390 395 400

Asp Leu Ala Lys Arg Ala Leu Trp Thr Pro Asn Gln Ile Thr Asp Ile  
 405 410 415

Thr Ser Leu Val Arg Leu Val Tyr Ile Leu Ser Lys Gln Asn Gln Gln  
 420 425 430

His Leu Ile Pro Gln Trp Ala Leu Arg Gln Ile Ala Asp Phe Ala Leu  
 435 440 445

Lys Leu His Lys Thr His Leu Ala Ser Phe Leu Ser Ala Phe Ala Arg  
 450 455 460

Gln Glu Leu Tyr Leu Met Gly Ser Leu Val His Ser Met Leu Val His  
 465 470 475 480

Thr Thr Glu Arg Arg Glu Ile Phe Ile Val Glu Thr Gly Leu Cys Ser  
 485 490 495

Leu Ala Glu Leu Ser His Phe Thr Gln Leu Leu Ala His Pro His His  
 500 505 510

Glu Tyr Leu Ser Asp Leu Tyr Thr Pro Cys Ser Ser Ser Gly Arg Arg  
 515 520 525

Asp His Ser Leu Glu Arg Leu Thr Arg Leu Phe Pro Asp Ala Thr Val  
 530 535 540

Pro Thr Thr Val Pro Ala Ala Leu Ser Ile Leu Ser Thr Met Gln Pro  
 545 550 555 560

Ser Thr Leu Glu Thr Phe Pro Asp Leu Phe Cys Leu Pro Leu Gly Glu  
 565 570 575

Ser Phe Ser Ala Leu Thr Val Ser Glu His Val Ser Tyr Val Val Thr  
 580 585 590

Asn Gln Tyr Leu Ile Lys Gly Ile Ser Tyr Pro Val Ser Thr Thr Val  
 595 600 605

ES 2 799 735 T3

Val Gly Gln Ser Leu Ile Ile Thr Gln Thr Asp Ser Gln Thr Lys Cys  
610 615 620

Glu Leu Thr Arg Asn Met His Thr Thr His Ser Ile Thr Ala Ala Leu  
625 630 635 640

Asn Ile Ser Leu Glu Asn Cys Ala Phe Cys Gln Ser Ala Leu Leu Glu  
645 650 655

Tyr Asp Asp Thr Gln Gly Val Ile Asn Ile Met Tyr Met His Asp Ser  
660 665 670

Asp Asp Val Leu Phe Ala Leu Asp Pro Tyr Asn Glu Val Val Val Ser  
675 680 685

Ser Pro Arg Thr His Tyr Leu Met Leu Leu Lys Asn Gly Thr Val Leu  
690 695 700

Glu Val Thr Asp Val Val Val Asp Ala Thr Asp Ser Arg Leu Leu Met  
705 710 715 720

Met Ser Val Tyr Ala Leu Ser Ala Ile Ile Gly Ile Tyr Leu Leu Tyr  
725 730 735

Arg Met Leu Lys Thr Cys  
740

<210> 3  
<211> 743  
<212> PRT  
5 <213> Herpesvirus humano 5

<220>  
<221> MISC\_FEATURE  
<222> (1)..(743)  
<223>/nota = "Cepa AD169"

10 <400> 3

Met Arg Pro Gly Leu Pro Pro Tyr Leu Thr Val Phe Thr Val Tyr Leu  
1 5 10 15

Leu Ser His Leu Pro Ser Gln Arg Tyr Gly Ala Asp Ala Ala Ser Glu  
20 25 30

Ala Leu Asp Pro His Ala Phe His Leu Leu Leu Asn Thr Tyr Gly Arg  
35 40 45

Pro Ile Arg Phe Leu Arg Glu Asn Thr Thr Gln Cys Thr Tyr Asn Ser  
50 55 60

ES 2 799 735 T3

Ser Leu Arg Asn Ser Thr Val Val Arg Glu Asn Ala Ile Ser Phe Asn  
65 70 75 80

Phe Phe Gln Ser Tyr Asn Gln Tyr Tyr Val Phe His Met Pro Arg Cys  
85 90 95

Leu Phe Ala Gly Pro Leu Ala Glu Gln Phe Leu Asn Gln Val Asp Leu  
100 105 110

Thr Glu Thr Leu Glu Arg Tyr Gln Gln Arg Leu Asn Thr Tyr Ala Leu  
115 120 125

Val Ser Lys Asp Leu Ala Ser Tyr Arg Ser Phe Ser Gln Gln Leu Lys  
130 135 140

Ala Gln Asp Ser Leu Gly Gln Gln Pro Thr Thr Val Pro Pro Pro Ile  
145 150 155 160

Asp Leu Ser Ile Pro His Val Trp Met Pro Pro Gln Thr Thr Pro His  
165 170 175

Asp Trp Lys Gly Ser His Thr Thr Ser Gly Leu His Arg Pro His Phe  
180 185 190

Asn Gln Thr Cys Ile Leu Phe Asp Gly His Asp Leu Leu Phe Ser Thr  
195 200 205

Val Thr Pro Cys Leu His Gln Gly Phe Tyr Leu Met Asp Glu Leu Arg  
210 215 220

Tyr Val Lys Ile Thr Leu Thr Glu Asp Phe Phe Val Val Thr Val Ser  
225 230 235 240

Ile Asp Asp Asp Thr Pro Met Leu Leu Ile Phe Gly His Leu Pro Arg  
245 250 255

Val Leu Phe Lys Ala Pro Tyr Gln Arg Asp Asn Phe Ile Leu Arg Gln  
260 265 270

Thr Glu Lys His Glu Leu Leu Val Leu Val Lys Lys Ala Gln Leu Asn  
275 280 285

Arg His Ser Tyr Leu Lys Asp Ser Asp Phe Leu Asp Ala Ala Leu Asp  
290 295 300

Phe Asn Tyr Leu Asp Leu Ser Ala Leu Leu Arg Asn Ser Phe His Arg



ES 2 799 735 T3

Pro Ser Thr Leu Glu Thr Phe Pro Asp Leu Phe Cys Leu Pro Leu Gly  
565 570 575

Glu Ser Phe Ser Ala Leu Thr Val Ser Glu His Val Ser Tyr Val Val  
580 585 590

Thr Asn Gln Tyr Leu Ile Lys Gly Ile Ser Tyr Pro Val Ser Thr Thr  
595 600 605

Val Val Gly Gln Ser Leu Ile Ile Thr Gln Thr Asp Ser Gln Thr Lys  
610 615 620

Cys Glu Leu Thr Arg Asn Met His Thr Thr His Ser Ile Thr Ala Ala  
625 630 635 640

Leu Asn Ile Ser Leu Glu Asn Cys Ala Phe Cys Gln Ser Ala Leu Leu  
645 650 655

Glu Tyr Asp Asp Thr Gln Gly Val Ile Asn Ile Met Tyr Met His Asp  
660 665 670

Ser Asp Asp Val Leu Phe Ala Leu Asp Pro Tyr Asn Glu Val Val Val  
675 680 685

Ser Ser Pro Arg Thr His Tyr Leu Met Leu Leu Lys Asn Gly Thr Val  
690 695 700

Leu Glu Val Thr Asp Val Val Val Asp Ala Thr Asp Ser Arg Leu Leu  
705 710 715 720

Met Met Ser Val Tyr Ala Leu Ser Ala Ile Ile Gly Ile Tyr Leu Leu  
725 730 735

Tyr Arg Met Leu Lys Thr Cys  
740

<210> 4  
<211> 715  
<212> PRT  
5 <213> Herpesvirus humano 5

<220>  
<221> MISC\_FEATURE  
<222> (1)..(715)  
<223> /nota="Cepa Merlin"

10 <400> 4

Met Arg Pro Gly Leu Pro Ser Tyr Leu Ile Ile Leu Ala Val Cys Leu  
1 5 10 15

ES 2 799 735 T3

Phe Ser His Leu Leu Ser Ser Arg Tyr Gly Ala Glu Ala Val Ser Glu  
 20 25 30

Pro Leu Asp Lys Ala Phe His Leu Leu Leu Asn Thr Tyr Gly Arg Pro  
 35 40 45

Ile Arg Phe Leu Arg Glu Asn Thr Thr Gln Cys Thr Tyr Asn Ser Ser  
 50 55 60

Leu Arg Asn Ser Thr Val Val Arg Glu Asn Ala Ile Ser Phe Asn Phe  
 65 70 75 80

Phe Gln Ser Tyr Asn Gln Tyr Tyr Val Phe His Met Pro Arg Cys Leu  
 85 90 95

Phe Ala Gly Pro Leu Ala Glu Gln Phe Leu Asn Gln Val Asp Leu Thr  
 100 105 110

Glu Thr Leu Glu Arg Tyr Gln Gln Arg Leu Asn Thr Tyr Ala Leu Val  
 115 120 125

Ser Lys Asp Leu Ala Ser Tyr Arg Ser Phe Ser Gln Gln Leu Lys Ala  
 130 135 140

Gln Asp Ser Leu Gly Glu Gln Pro Thr Thr Val Pro Pro Pro Ile Asp  
 145 150 155 160

Leu Ser Ile Pro His Val Trp Met Pro Pro Gln Thr Thr Pro His Gly  
 165 170 175

Trp Thr Glu Ser His Thr Thr Ser Gly Leu His Arg Pro His Phe Asn  
 180 185 190

Gln Thr Cys Ile Leu Phe Asp Gly His Asp Leu Leu Phe Ser Thr Val  
 195 200 205

Thr Pro Cys Leu His Gln Gly Phe Tyr Leu Ile Asp Glu Leu Arg Tyr  
 210 215 220

Val Lys Ile Thr Leu Thr Glu Asp Phe Phe Val Val Thr Val Ser Ile  
 225 230 235 240

Asp Asp Asp Thr Pro Met Leu Leu Ile Phe Gly His Leu Pro Arg Val  
 245 250 255

Leu Phe Lys Ala Pro Tyr Gln Arg Asp Asn Phe Ile Leu Arg Gln Thr  
 260 265 270

ES 2 799 735 T3

Glu Lys His Glu Leu Leu Val Leu Val Lys Lys Asp Gln Leu Asn Arg  
 275 280 285  
 His Ser Tyr Leu Lys Asp Pro Asp Phe Leu Asp Ala Ala Leu Asp Phe  
 290 295 300  
 Asn Tyr Leu Asp Leu Ser Ala Leu Leu Arg Asn Ser Phe His Arg Tyr  
 305 310 315 320  
 Ala Val Asp Val Leu Lys Ser Gly Arg Cys Gln Met Leu Asp Arg Arg  
 325 330 335  
 Thr Val Glu Met Ala Phe Ala Tyr Ala Leu Ala Leu Phe Ala Ala Ala  
 340 345 350  
 Arg Gln Glu Glu Ala Gly Ala Gln Val Ser Val Pro Arg Ala Leu Asp  
 355 360 365  
 Arg Gln Ala Ala Leu Leu Gln Ile Gln Glu Phe Met Ile Thr Cys Leu  
 370 375 380  
 Ser Gln Thr Pro Pro Arg Thr Thr Leu Leu Leu Tyr Pro Thr Ala Val  
 385 390 395 400  
 Asp Leu Ala Lys Arg Ala Leu Trp Thr Pro Asn Gln Ile Thr Asp Ile  
 405 410 415  
 Thr Ser Leu Val Arg Leu Val Tyr Ile Leu Ser Lys Gln Asn Gln Gln  
 420 425 430  
 His Leu Ile Pro Gln Trp Ala Leu Arg Gln Ile Ala Asp Phe Ala Leu  
 435 440 445  
 Lys Leu His Lys Thr His Leu Ala Ser Phe Leu Ser Ala Phe Ala Arg  
 450 455 460  
 Gln Glu Leu Tyr Leu Met Gly Ser Leu Val His Ser Met Leu Val His  
 465 470 475 480  
 Thr Thr Glu Arg Arg Glu Ile Phe Ile Val Glu Thr Gly Leu Cys Ser  
 485 490 495  
 Leu Ala Glu Leu Ser His Phe Thr Gln Leu Leu Ala His Pro His His  
 500 505 510  
 Glu Tyr Leu Ser Asp Leu Tyr Thr Pro Cys Ser Ser Ser Gly Arg Arg

ES 2 799 735 T3

515	520	525																		
Asp	His	Ser	Leu	Glu	Arg	Leu	Thr	Arg	Leu	Phe	Pro	Asp	Ala	Thr	Val					
530						535					540									
Pro	Ala	Thr	Val	Pro	Ala	Ala	Leu	Ser	Ile	Leu	Ser	Thr	Met	Gln	Pro					
545					550					555					560					
Ser	Thr	Leu	Glu	Thr	Phe	Pro	Asp	Leu	Phe	Cys	Leu	Pro	Leu	Gly	Glu					
				565					570					575						
Ser	Phe	Ser	Ala	Leu	Thr	Val	Ser	Glu	His	Val	Ser	Tyr	Ile	Val	Thr					
			580					585					590							
Asn	Gln	Tyr	Leu	Ile	Lys	Gly	Ile	Ser	Tyr	Pro	Val	Ser	Thr	Thr	Val					
		595					600						605							
Val	Gly	Gln	Ser	Leu	Ile	Ile	Thr	Gln	Thr	Asp	Ser	Gln	Thr	Lys	Cys					
	610					615					620									
Glu	Leu	Thr	Arg	Asn	Met	His	Thr	Thr	His	Ser	Ile	Thr	Val	Ala	Leu					
625					630					635					640					
Asn	Ile	Ser	Leu	Glu	Asn	Cys	Ala	Phe	Cys	Gln	Ser	Ala	Leu	Leu	Glu					
				645					650					655						
Tyr	Asp	Asp	Thr	Gln	Gly	Val	Ile	Asn	Ile	Met	Tyr	Met	His	Asp	Ser					
			660					665					670							
Asp	Asp	Val	Leu	Phe	Ala	Leu	Asp	Pro	Tyr	Asn	Glu	Val	Val	Val	Ser					
		675					680					685								
Ser	Pro	Arg	Thr	His	Tyr	Leu	Met	Leu	Leu	Lys	Asn	Gly	Thr	Val	Leu					
	690					695						700								
Glu	Val	Thr	Asp	Val	Val	Val	Asp	Ala	Thr	Asp										
705					710					715										

<210> 5  
 <211> 692  
 <212> PRT  
 <213> Herpesvirus humano 5

5

<220>  
 <221> MISC\_FEATURE  
 <222> (1)..(692)  
 <223> /nota="Cepa Merlin"

10

<400> 5

ES 2 799 735 T3

Arg Tyr Gly Ala Glu Ala Val Ser Glu Pro Leu Asp Lys Ala Phe His  
1 5 10 15

Leu Leu Leu Asn Thr Tyr Gly Arg Pro Ile Arg Phe Leu Arg Glu Asn  
20 25 30

Thr Thr Gln Cys Thr Tyr Asn Ser Ser Leu Arg Asn Ser Thr Val Val  
35 40 45

Arg Glu Asn Ala Ile Ser Phe Asn Phe Phe Gln Ser Tyr Asn Gln Tyr  
50 55 60

Tyr Val Phe His Met Pro Arg Cys Leu Phe Ala Gly Pro Leu Ala Glu  
65 70 75 80

Gln Phe Leu Asn Gln Val Asp Leu Thr Glu Thr Leu Glu Arg Tyr Gln  
85 90 95

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

Arg Ser Phe Ser Gln Gln Leu Lys Ala Gln Asp Ser Leu Gly Glu Gln  
115 120 125

Pro Thr Thr Val Pro Pro Pro Ile Asp Leu Ser Ile Pro His Val Trp  
130 135 140

Met Pro Pro Gln Thr Thr Pro His Gly Trp Thr Glu Ser His Thr Thr  
145 150 155 160

Ser Gly Leu His Arg Pro His Phe Asn Gln Thr Cys Ile Leu Phe Asp  
165 170 175

Gly His Asp Leu Leu Phe Ser Thr Val Thr Pro Cys Leu His Gln Gly  
180 185 190

Phe Tyr Leu Ile Asp Glu Leu Arg Tyr Val Lys Ile Thr Leu Thr Glu  
195 200 205

Asp Phe Phe Val Val Thr Val Ser Ile Asp Asp Asp Thr Pro Met Leu  
210 215 220

Leu Ile Phe Gly His Leu Pro Arg Val Leu Phe Lys Ala Pro Tyr Gln  
225 230 235 240

Arg Asp Asn Phe Ile Leu Arg Gln Thr Glu Lys His Glu Leu Leu Val  
245 250 255

ES 2 799 735 T3

Leu Val Lys Lys Asp Gln Leu Asn Arg His Ser Tyr Leu Lys Asp Pro  
 260 265 270  
 Asp Phe Leu Asp Ala Ala Leu Asp Phe Asn Tyr Leu Asp Leu Ser Ala  
 275 280 285  
 Leu Leu Arg Asn Ser Phe His Arg Tyr Ala Val Asp Val Leu Lys Ser  
 290 295 300  
 Gly Arg Cys Gln Met Leu Asp Arg Arg Thr Val Glu Met Ala Phe Ala  
 305 310 315 320  
 Tyr Ala Leu Ala Leu Phe Ala Ala Ala Arg Gln Glu Glu Ala Gly Ala  
 325 330 335  
 Gln Val Ser Val Pro Arg Ala Leu Asp Arg Gln Ala Ala Leu Leu Gln  
 340 345 350  
 Ile Gln Glu Phe Met Ile Thr Cys Leu Ser Gln Thr Pro Pro Arg Thr  
 355 360 365  
 Thr Leu Leu Leu Tyr Pro Thr Ala Val Asp Leu Ala Lys Arg Ala Leu  
 370 375 380  
 Trp Thr Pro Asn Gln Ile Thr Asp Ile Thr Ser Leu Val Arg Leu Val  
 385 390 395 400  
 Tyr Ile Leu Ser Lys Gln Asn Gln Gln His Leu Ile Pro Gln Trp Ala  
 405 410 415  
 Leu Arg Gln Ile Ala Asp Phe Ala Leu Lys Leu His Lys Thr His Leu  
 420 425 430  
 Ala Ser Phe Leu Ser Ala Phe Ala Arg Gln Glu Leu Tyr Leu Met Gly  
 435 440 445  
 Ser Leu Val His Ser Met Leu Val His Thr Thr Glu Arg Arg Glu Ile  
 450 455 460  
 Phe Ile Val Glu Thr Gly Leu Cys Ser Leu Ala Glu Leu Ser His Phe  
 465 470 475 480  
 Thr Gln Leu Leu Ala His Pro His His Glu Tyr Leu Ser Asp Leu Tyr  
 485 490 495  
 Thr Pro Cys Ser Ser Ser Gly Arg Arg Asp His Ser Leu Glu Arg Leu  
 500 505 510

ES 2 799 735 T3

Thr Arg Leu Phe Pro Asp Ala Thr Val Pro Ala Thr Val Pro Ala Ala  
515 520 525

Leu Ser Ile Leu Ser Thr Met Gln Pro Ser Thr Leu Glu Thr Phe Pro  
530 535 540

Asp Leu Phe Cys Leu Pro Leu Gly Glu Ser Phe Ser Ala Leu Thr Val  
545 550 555 560

Ser Glu His Val Ser Tyr Ile Val Thr Asn Gln Tyr Leu Ile Lys Gly  
565 570 575

Ile Ser Tyr Pro Val Ser Thr Thr Val Val Gly Gln Ser Leu Ile Ile  
580 585 590

Thr Gln Thr Asp Ser Gln Thr Lys Cys Glu Leu Thr Arg Asn Met His  
595 600 605

Thr Thr His Ser Ile Thr Val Ala Leu Asn Ile Ser Leu Glu Asn Cys  
610 615 620

Ala Phe Cys Gln Ser Ala Leu Leu Glu Tyr Asp Asp Thr Gln Gly Val  
625 630 635 640

Ile Asn Ile Met Tyr Met His Asp Ser Asp Asp Val Leu Phe Ala Leu  
645 650 655

Asp Pro Tyr Asn Glu Val Val Val Ser Ser Pro Arg Thr His Tyr Leu  
660 665 670

Met Leu Leu Lys Asn Gly Thr Val Leu Glu Val Thr Asp Val Val Val  
675 680 685

Asp Ala Thr Asp  
690

<210> 6  
<211> 278  
<212> PRT  
5 <213> Herpesvirus humano 5

<220>  
<221> MISC\_FEATURE  
<222> (1)..(278)  
<223> /nota="Cepa Merlin"

10 <400> 6

Met Cys Arg Arg Pro Asp Cys Gly Phe Ser Phe Ser Pro Gly Pro Val

ES 2 799 735 T3

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

ES 2 799 735 T3

Leu Lys Gln Thr Arg Val Asn Leu Pro Ala His Ser Arg Tyr Gly Pro  
 260 265 270

Gln Ala Val Asp Ala Arg  
 275

<210> 7  
 <211> 278  
 <212> PRT  
 <213> Herpesvirus humano 5

5

<220>  
 <221> MISC\_FEATURE  
 <222> (1)..(278)  
 <223>/nota = "CepaTowne"

10

<400> 7

Met Cys Arg Arg Pro Asp Cys Gly Phe Ser Phe Ser Pro Gly Pro Val  
 1 5 10 15

Ala Leu Leu Trp Cys Cys Leu Leu Leu Pro Ile Val Ser Ser Ala Thr  
 20 25 30

Val Ser Val Ala Pro Thr Val Ala Glu Lys Val Pro Ala Glu Cys Pro  
 35 40 45

Glu Leu Thr Arg Arg Cys Leu Leu Gly Glu Val Phe Gln Gly Asp Lys  
 50 55 60

Tyr Glu Ser Trp Leu Arg Pro Leu Val Asn Val Thr Arg Arg Asp Gly  
 65 70 75 80

Pro Leu Ser Gln Leu Ile Arg Tyr Arg Pro Val Thr Pro Glu Ala Ala  
 85 90 95

Asn Ser Val Leu Leu Asp Asp Ala Phe Leu Asp Thr Leu Ala Leu Leu  
 100 105 110

Tyr Asn Asn Pro Asp Gln Leu Arg Ala Leu Leu Thr Leu Leu Ser Ser  
 115 120 125

Asp Thr Ala Pro Arg Trp Met Thr Val Met Arg Gly Tyr Ser Glu Cys  
 130 135 140

Gly Asp Gly Ser Pro Ala Val Tyr Thr Cys Val Asp Asp Leu Cys Arg  
 145 150 155 160

Gly Tyr Asp Leu Thr Arg Leu Ser Tyr Gly Arg Ser Ile Phe Thr Glu  
 165 170 175

ES 2 799 735 T3

His Val Leu Gly Phe Glu Leu Val Pro Pro Ser Leu Phe Asn Val Val  
 180 185 190

Val Ala Ile Arg Asn Glu Ala Thr Arg Thr Asn Arg Ala Val Arg Leu  
 195 200 205

Pro Val Ser Thr Ala Ala Ala Pro Glu Gly Ile Thr Leu Phe Tyr Gly  
 210 215 220

Leu Tyr Asn Ala Val Lys Glu Phe Cys Leu Arg His Gln Leu Asp Pro  
 225 230 235 240

Pro Leu Leu Arg His Leu Asp Lys Tyr Tyr Ala Gly Leu Pro Pro Glu  
 245 250 255

Leu Lys Gln Thr Arg Val Asn Leu Pro Ala His Ser Arg Tyr Gly Pro  
 260 265 270

Gln Ala Val Asp Ala Arg  
 275

<210> 8  
 <211> 278  
 <212> PRT  
 <213> Herpesvirus humano 5  
 <220>  
 <221> MISC\_FEATURE  
 <222> (1)..(278)  
 <223>/nota = "Cepa AD169"

5

10

<400> 8

Met Cys Arg Arg Pro Asp Cys Gly Phe Ser Phe Ser Pro Gly Pro Val  
 1 5 10 15

Val Leu Leu Trp Cys Cys Leu Leu Leu Pro Ile Val Ser Ser Val Ala  
 20 25 30

Val Ser Val Ala Pro Thr Ala Ala Glu Lys Val Pro Ala Glu Cys Pro  
 35 40 45

Glu Leu Thr Arg Arg Cys Leu Leu Gly Glu Val Phe Gln Gly Asp Lys  
 50 55 60

Tyr Glu Ser Trp Leu Arg Pro Leu Val Asn Val Thr Arg Arg Asp Gly  
 65 70 75 80

Pro Leu Ser Gln Leu Ile Arg Tyr Arg Pro Val Thr Pro Glu Ala Ala  
 85 90 95

ES 2 799 735 T3

Asn Ser Val Leu Leu Asp Asp Ala Phe Leu Asp Thr Leu Ala Leu Leu  
 100 105 110

Tyr Asn Asn Pro Asp Gln Leu Arg Ala Leu Leu Thr Leu Leu Ser Ser  
 115 120 125

Asp Thr Ala Pro Arg Trp Met Thr Val Met Arg Gly Tyr Ser Glu Cys  
 130 135 140

Gly Asp Gly Ser Pro Ala Val Tyr Thr Cys Val Asp Asp Leu Cys Arg  
 145 150 155 160

Gly Tyr Asp Leu Thr Arg Leu Ser Tyr Gly Arg Ser Ile Phe Thr Glu  
 165 170 175

His Val Leu Gly Phe Glu Leu Val Pro Pro Ser Leu Phe Asn Val Val  
 180 185 190

Val Ala Ile Arg Asn Glu Ala Thr Arg Thr Asn Arg Ala Val Arg Leu  
 195 200 205

Pro Val Ser Thr Ala Ala Ala Pro Glu Gly Ile Thr Leu Phe Tyr Gly  
 210 215 220

Leu Tyr Asn Ala Val Lys Glu Phe Cys Leu Arg His Gln Leu Asp Pro  
 225 230 235 240

Pro Leu Leu Arg His Leu Asp Lys Tyr Tyr Ala Gly Leu Pro Pro Glu  
 245 250 255

Leu Lys Gln Thr Arg Val Asn Leu Pro Ala His Ser Arg Tyr Gly Pro  
 260 265 270

Gln Ala Val Asp Ala Arg  
 275

<210> 9  
 <211> 248  
 <212> PRT  
 <213> Herpesvirus humano 5

5

<220>  
 <221> MISC\_FEATURE  
 <222> (1)..(248)  
 <223> /nota="Cepa Merlin"

10

<400> 9

Ala Ala Val Ser Val Ala Pro Thr Ala Ala Glu Lys Val Pro Ala Glu

ES 2 799 735 T3

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

ES 2 799 735 T3

<210> 10  
 <211> 130  
 <212> PRT  
 <213> Herpesvirus humano 5

5

<220>  
 <221> MISC\_FEATURE  
 <222> (1)..(130)  
 <223> /nota="Cepa Merlin"

<400> 10

Met Ser Pro Lys Asp Leu Thr Pro Phe Leu Thr Ala Leu Trp Leu Leu  
 1 5 10 15

Leu Gly His Ser Arg Val Pro Arg Val Arg Ala Glu Glu Cys Cys Glu  
 20 25 30

Phe Ile Asn Val Asn His Pro Pro Glu Arg Cys Tyr Asp Phe Lys Met  
 35 40 45

Cys Asn Arg Phe Thr Val Ala Leu Arg Cys Pro Asp Gly Glu Val Cys  
 50 55 60

Tyr Ser Pro Glu Lys Thr Ala Glu Ile Arg Gly Ile Val Thr Thr Met  
 65 70 75 80

Thr His Ser Leu Thr Arg Gln Val Val His Asn Lys Leu Thr Ser Cys  
 85 90 95

Asn Tyr Asn Pro Leu Tyr Leu Glu Ala Asp Gly Arg Ile Arg Cys Gly  
 100 105 110

Lys Val Asn Asp Lys Ala Gln Tyr Leu Leu Gly Ala Ala Gly Ser Val  
 115 120 125

Pro Tyr  
 130

10

<210> 11  
 <211> 171  
 <212> PRT  
 <213> Herpesvirus humano 5

15

<220>  
 <221> MISC\_FEATURE  
 <222> (1)..(171)  
 <223> /nota = "CepaTowne"

<400> 11

ES 2 799 735 T3

```

Met Ser Pro Lys Asn Leu Thr Pro Phe Leu Thr Ala Leu Trp Leu Leu
1           5           10           15

Leu Gly His Ser Arg Val Pro Arg Val Arg Ala Glu Glu Cys Cys Glu
           20           25           30

Phe Ile Asn Val Asn His Pro Pro Glu Arg Cys Tyr Asp Phe Lys Met
           35           40           45

Cys Asn Arg Phe Thr Val Ala Leu Arg Cys Pro Asp Gly Glu Val Cys
           50           55           60

Tyr Ser Pro Glu Lys Thr Ala Glu Ile Arg Gly Ile Val Thr Thr Met
65           70           75           80

Thr His Ser Leu Thr Arg Gln Val Val His Asn Lys Leu Thr Ser Cys
           85           90           95

Asn Tyr Asn Pro Leu Tyr Leu Glu Ala Asp Gly Arg Ile Arg Cys Gly
           100          105          110

Lys Val Asn Asp Lys Ala Gln Tyr Leu Leu Gly Ala Ala Gly Ser Val
           115          120          125

Pro Tyr Arg Trp Ile Asn Leu Glu Tyr Asp Lys Ile Thr Arg Ile Val
130          135          140

Gly Leu Asp Gln Tyr Leu Glu Ser Val Lys Lys His Lys Arg Leu Asp
145          150          155          160

Val Cys Arg Ala Lys Met Gly Tyr Met Leu Gln
           165          170

```

5 <210> 12  
 <211> 171  
 <212> PRT  
 <213> Herpesvirus humano 5

<220>  
 <221> MISC\_FEATURE  
 <222> (1)..(171)  
 <223>/nota = "Cepa AD169"

10 <400> 12

ES 2 799 735 T3

Met Ser Pro Lys Asp Leu Thr Pro Phe Leu Thr Thr Leu Trp Leu Leu  
 1 5 10 15

Leu Gly His Ser Arg Val Pro Arg Val Arg Ala Glu Glu Cys Cys Glu  
 20 25 30

Phe Ile Asn Val Asn His Pro Pro Glu Arg Cys Tyr Asp Phe Lys Met  
 35 40 45

Cys Asn Arg Phe Thr Val Ala Leu Arg Cys Pro Asp Gly Glu Val Cys  
 50 55 60

Tyr Ser Pro Glu Lys Thr Ala Glu Ile Arg Gly Ile Val Thr Thr Met  
 65 70 75 80

Thr His Ser Leu Thr Arg Gln Val Val His Asn Lys Leu Thr Ser Cys  
 85 90 95

Asn Tyr Asn Pro Leu Tyr Leu Glu Ala Asp Gly Arg Ile Arg Cys Gly  
 100 105 110

Lys Val Asn Asp Lys Ala Gln Tyr Leu Leu Gly Ala Ala Gly Ser Val  
 115 120 125

Pro Tyr Arg Trp Ile Asn Leu Glu Tyr Asp Lys Ile Thr Arg Ile Val  
 130 135 140

Gly Leu Asp Gln Tyr Leu Glu Ser Val Lys Lys His Lys Arg Leu Asp  
 145 150 155 160

Val Cys Arg Ala Lys Met Gly Tyr Met Leu Gln  
 165 170

<210> 13

<211> 144

<212> PRT

5 <213> Herpesvirus humano 5

<400> 13

ES 2 799 735 T3

Glu Glu Cys Cys Glu Phe Ile Asn Val Asn His Pro Pro Glu Arg Cys  
 1 5 10 15

Tyr Asp Phe Lys Met Cys Asn Arg Phe Thr Val Ala Leu Arg Cys Pro  
 20 25 30

Asp Gly Glu Val Cys Tyr Ser Pro Glu Lys Thr Ala Glu Ile Arg Gly  
 35 40 45

Ile Val Thr Thr Met Thr His Ser Leu Thr Arg Gln Val Val His Asn  
 50 55 60

Lys Leu Thr Ser Cys Asn Tyr Asn Pro Leu Tyr Leu Glu Ala Asp Gly  
 65 70 75 80

Arg Ile Arg Cys Gly Lys Val Asn Asp Lys Ala Gln Tyr Leu Leu Gly  
 85 90 95

Ala Ala Gly Ser Val Pro Tyr Arg Trp Ile Asn Leu Glu Tyr Asp Lys  
 100 105 110

Ile Thr Arg Ile Val Gly Leu Asp Gln Tyr Leu Glu Ser Val Lys Lys  
 115 120 125

His Lys Arg Leu Asp Val Cys Arg Ala Lys Met Gly Tyr Met Leu Gln  
 130 135 140

<210> 14  
 <211> 214  
 <212> PRT  
 <213> Herpesvirus humano 5

5

<220>  
 <221> MISC\_FEATURE  
 <222> (1)..(214)  
 <223> /nota="Cepa Merlin"

10

<400> 14

ES 2 799 735 T3

Met Leu Arg Leu Leu Leu Arg His His Phe His Cys Leu Leu Leu Cys  
 1 5 10 15

Ala Val Trp Ala Thr Pro Cys Leu Ala Ser Pro Trp Ser Thr Leu Thr  
 20 25 30

Ala Asn Gln Asn Pro Ser Pro Pro Trp Ser Lys Leu Thr Tyr Ser Lys  
 35 40 45

Pro His Asp Ala Ala Thr Phe Tyr Cys Pro Phe Leu Tyr Pro Ser Pro  
 50 55 60

Pro Arg Ser Pro Leu Gln Phe Ser Gly Phe Gln Arg Val Ser Thr Gly  
 65 70 75 80

Pro Glu Cys Arg Asn Glu Thr Leu Tyr Leu Leu Tyr Asn Arg Glu Gly  
 85 90 95

Gln Thr Leu Val Glu Arg Ser Ser Thr Trp Val Lys Lys Val Ile Trp  
 100 105 110

Tyr Leu Ser Gly Arg Asn Gln Thr Ile Leu Gln Arg Met Pro Arg Thr  
 115 120 125

Ala Ser Lys Pro Ser Asp Gly Asn Val Gln Ile Ser Val Glu Asp Ala  
 130 135 140

Lys Ile Phe Gly Ala His Met Val Pro Lys Gln Thr Lys Leu Leu Arg  
 145 150 155 160

Phe Val Val Asn Asp Gly Thr Arg Tyr Gln Met Cys Val Met Lys Leu  
 165 170 175

Glu Ser Trp Ala His Val Phe Arg Asp Tyr Ser Val Ser Phe Gln Val  
 180 185 190

Arg Leu Thr Phe Thr Glu Ala Asn Asn Gln Thr Tyr Thr Phe Cys Thr  
 195 200 205

His Pro Asn Leu Ile Val  
 210

<210> 15  
 <211> 229  
 <212> PRT  
 <213> Herpesvirus humano 5

5

<220>  
 <221> MISC\_FEATURE  
 <222> (1)..(229)

ES 2 799 735 T3

<223>/nota = "CepaTowne"

<400> 15

Met Leu Arg Leu Leu Leu Arg His His Phe His Cys Leu Leu Leu Cys  
 1 5 10 15

Ala Val Trp Ala Thr Pro Cys Leu Ala Ser Pro Trp Ser Thr Leu Thr  
 20 25 30

Ala Asn Gln Asn Pro Ser Pro Pro Trp Ser Lys Leu Thr Tyr Ser Lys  
 35 40 45

Pro His Asp Ala Ala Thr Phe Tyr Cys Pro Phe Leu Tyr Pro Ser Pro  
 50 55 60

Pro Arg Ser Pro Leu Gln Phe Ser Gly Phe Gln Arg Val Leu Thr Gly  
 65 70 75 80

Pro Glu Cys Arg Asn Glu Thr Leu Tyr Leu Leu Tyr Asn Arg Glu Gly  
 85 90 95

Gln Thr Leu Val Glu Arg Ser Ser Thr Trp Val Lys Lys Val Ile Trp  
 100 105 110

Tyr Leu Ser Gly Arg Asn Gln Thr Ile Leu Gln Arg Met Pro Arg Thr  
 115 120 125

Ala Ser Lys Pro Ser Asp Gly Asn Val Gln Ile Ser Val Glu Asp Ala

ES 2 799 735 T3

130

135

140

Lys Ile Phe Gly Ala His Met Val Pro Lys Gln Thr Lys Leu Leu Arg  
 145 150 155 160

Phe Val Val Asn Asp Gly Thr Arg Tyr Gln Met Cys Val Met Lys Leu  
 165 170 175

Glu Ser Trp Ala His Val Phe Arg Asp Tyr Ser Val Ser Phe Gln Val  
 180 185 190

Arg Leu Thr Phe Thr Glu Ala Asn Asn Gln Thr Phe Thr Pro Ser Ala  
 195 200 205

Pro Ile Pro Ile Ser Ser Phe Glu Pro Val Ala Arg Ala Gly Asn Phe  
 210 215 220

Glu Asn Arg Ala Ser  
 225

<210> 16  
 <211> 189  
 <212> PRT  
 <213> Herpesvirus humano 5

5

<220>  
 <221> MISC\_FEATURE  
 <222> (1)..(189)  
 <223> /nota="Cepa Merlin"

10

<400> 16

Ser Pro Trp Ser Thr Leu Thr Ala Asn Gln Asn Pro Ser Pro Pro Trp  
 1 5 10 15

Ser Lys Leu Thr Tyr Ser Lys Pro His Asp Ala Ala Thr Phe Tyr Cys  
 20 25 30

Pro Phe Leu Tyr Pro Ser Pro Pro Arg Ser Pro Leu Gln Phe Ser Gly  
 35 40 45

Phe Gln Arg Val Ser Thr Gly Pro Glu Cys Arg Asn Glu Thr Leu Tyr  
 50 55 60

Leu Leu Tyr Asn Arg Glu Gly Gln Thr Leu Val Glu Arg Ser Ser Thr  
 65 70 75 80

Trp Val Lys Lys Val Ile Trp Tyr Leu Ser Gly Arg Asn Gln Thr Ile  
 85 90 95

ES 2 799 735 T3

Leu Gln Arg Met Pro Arg Thr Ala Ser Lys Pro Ser Asp Gly Asn Val  
 100 105 110

Gln Ile Ser Val Glu Asp Ala Lys Ile Phe Gly Ala His Met Val Pro  
 115 120 125

Lys Gln Thr Lys Leu Leu Arg Phe Val Val Asn Asp Gly Thr Arg Tyr  
 130 135 140

Gln Met Cys Val Met Lys Leu Glu Ser Trp Ala His Val Phe Arg Asp  
 145 150 155 160

Tyr Ser Val Ser Phe Gln Val Arg Leu Thr Phe Thr Glu Ala Asn Asn  
 165 170 175

Gln Thr Tyr Thr Phe Cys Thr His Pro Asn Leu Ile Val  
 180 185

<210> 17

<211> 129

<212> PRT

5 <213> Herpesvirus humano 5

<220>

<221> MISC\_FEATURE

<222> (1)..(129)

<223> /nota="Cepa Merlin"

10 <400> 17

ES 2 799 735 T3

Met Arg Leu Cys Arg Val Trp Leu Ser Val Cys Leu Cys Ala Val Val  
 1 5 10 15

Leu Gly Gln Cys Gln Arg Glu Thr Ala Glu Lys Asn Asp Tyr Tyr Arg  
 20 25 30

Val Pro His Tyr Trp Asp Ala Cys Ser Arg Ala Leu Pro Asp Gln Thr  
 35 40 45

Arg Tyr Lys Tyr Val Glu Gln Leu Val Asp Leu Thr Leu Asn Tyr His  
 50 55 60

Tyr Asp Ala Ser His Gly Leu Asp Asn Phe Asp Val Leu Lys Arg Ile  
 65 70 75 80

Asn Val Thr Glu Val Ser Leu Leu Ile Ser Asp Phe Arg Arg Gln Asn  
 85 90 95

Arg Arg Gly Gly Thr Asn Lys Arg Thr Thr Phe Asn Ala Ala Gly Ser  
 100 105 110

Leu Ala Pro His Ala Arg Ser Leu Glu Phe Ser Val Arg Leu Phe Ala  
 115 120 125

Asn

<210> 18  
 <211> 129  
 <212> PRT  
 5 <213> Herpesvirus humano 5

<220>  
 <221> MISC\_FEATURE  
 <222> (1)..(129)  
 <223>/nota = "CepaTowne"

10 <400> 18

ES 2 799 735 T3

Met Arg Leu Cys Arg Val Trp Leu Ser Val Cys Leu Cys Ala Val Val  
 1 5 10 15

Leu Gly Gln Cys Gln Arg Glu Thr Ala Glu Lys Asn Asp Tyr Tyr Arg  
 20 25 30

Val Pro His Tyr Trp Asp Ala Cys Ser Arg Ala Leu Pro Asp Gln Thr  
 35 40 45

Arg Tyr Lys Tyr Val Glu Gln Leu Val Asp Leu Thr Leu Asn Tyr His  
 50 55 60

Tyr Asp Ala Ser His Gly Leu Asp Asn Phe Asp Val Leu Lys Arg Ile  
 65 70 75 80

Asn Val Thr Glu Val Ser Leu Leu Ile Ser Asp Phe Arg Arg Gln Asn  
 85 90 95

Arg Arg Gly Gly Thr Asn Lys Arg Thr Thr Phe Asn Ala Ala Gly Ser  
 100 105 110

Leu Ala Pro His Ala Arg Ser Leu Glu Phe Ser Val Arg Leu Phe Ala  
 115 120 125

Asn

<210> 19  
 <211> 74  
 <212> PRT  
 5 <213> Herpesvirus humano 5

<220>  
 <221> MISC\_FEATURE  
 <222> (1)..(74)  
 <223>/nota = "Cepa AD169"

10 <400> 19

ES 2 799 735 T3

Met Arg Leu Cys Arg Val Trp Leu Ser Val Cys Leu Cys Ala Val Val  
 1 5 10 15

Leu Gly Gln Cys Gln Arg Glu Thr Ala Glu Lys Lys Arg Leu Leu Pro  
 20 25 30

Ser Thr Ala Leu Leu Gly Arg Val Leu Ser Arg Ala Ala Arg Pro Asn  
 35 40 45

Pro Leu Gln Val Cys Gly Thr Ala Arg Gly Pro His Val Glu Leu Pro  
 50 55 60

Leu Arg Cys Glu Pro Arg Leu Gly Gln Leu  
 65 70

<210> 20  
 <211> 111  
 <212> PRT  
 <213> Herpesvirus humano 5

5

<400> 20

Gln Cys Gln Arg Glu Thr Ala Glu Lys Asn Asp Tyr Tyr Arg Val Pro  
 1 5 10 15

His Tyr Trp Asp Ala Cys Ser Arg Ala Leu Pro Asp Gln Thr Arg Tyr  
 20 25 30

Lys Tyr Val Glu Gln Leu Val Asp Leu Thr Leu Asn Tyr His Tyr Asp  
 35 40 45

Ala Ser His Gly Leu Asp Asn Phe Asp Val Leu Lys Arg Ile Asn Val  
 50 55 60

Thr Glu Val Ser Leu Leu Ile Ser Asp Phe Arg Arg Gln Asn Arg Arg  
 65 70 75 80

Gly Gly Thr Asn Lys Arg Thr Thr Phe Asn Ala Ala Gly Ser Leu Ala  
 85 90 95

Pro His Ala Arg Ser Leu Glu Phe Ser Val Arg Leu Phe Ala Asn  
 100 105 110

<210> 21  
 <211> 1547  
 <212> PRT  
 <213> Cricetulus griseus

10

<400> 21

ES 2 799 735 T3

Met Asn Trp Asn Ala Lys Pro Glu Asn Ala Ala Pro Asn Pro Pro Tyr  
1 5 10 15

Ser Lys Ser Gln Ser Ser Leu Leu Gln Gln Phe Leu Met Pro Ser Thr  
20 25 30

Thr Ser Gln Ser Ser Phe Ser Cys Leu Pro His Asn Gln Glu Ala Cys  
35 40 45

Ile Tyr Pro Thr Asn Ser Asn Ser Val Ser Gln Pro Leu Leu Asn Val  
50 55 60

Arg Ser Phe Ile Asn Pro Pro Ile Ser Val Ser Asn Val His Asn Arg  
65 70 75 80

Thr Val Val Ala Ser Gln Thr Ser Val Glu Arg Val Thr Tyr Thr Asn  
85 90 95

Val Lys Gly Ala Gln Gln Pro Asn His Asn Leu Gln Thr Val Ser Ser  
100 105 110

Gly Val Val Gln Asn Ala Trp Met Asn Ser Thr Met Arg Asn Phe Met  
115 120 125

Pro Ser Leu Thr Glu Ala Thr Ile Ser His Lys Pro Asp Gly Gly Pro  
130 135 140

Ser Met Pro Tyr Met His Ala Pro Gln Ser His Leu Val Thr Ser Asp  
145 150 155 160

Thr Tyr Ser Val Gln Leu Gln Met Thr Pro Ser Asn Ser Val Arg Gly  
165 170 175

Pro Val Thr Tyr Gln Gly Asn Tyr Gln Gly Asn Pro Gly Leu Asn His  
180 185 190

Ser Met Ala Gly Glu Leu Gly Trp Val Gln Cys Ala Ser Ser Glu Leu  
195 200 205

Thr Tyr Pro Asp Tyr Arg Pro Pro Pro Lys Gln Tyr Pro Tyr Leu Pro  
210 215 220

Gln Ser Phe Val Gln Asp Thr Ser Val Gln Lys Gln Asn Phe Val Ser  
225 230 235 240

ES 2 799 735 T3

Ser Thr Ser Leu Gln Val Lys Asn Asn Gln Leu Pro Pro Ser Thr Gln  
 245 250 255

Thr Leu Pro Ser Lys Arg Pro Val Pro Val Ser Ser Tyr Gln Tyr Ala  
 260 265 270

Ala Glu Thr Ser Lys Arg Leu Pro Pro Pro Tyr Ser Cys Arg Tyr  
 275 280 285

Gly Ser Gln His Val Gln Asn Ser Gln Ser Val Ser Arg His Leu Pro  
 290 295 300

Val Glu Val Pro Gln Ser Ser Glu Met His Ser Ser Glu Lys Lys Lys  
 305 310 315 320

Asp Ala Tyr Lys Val Phe Gln Gln Gln Trp Gln Ser Thr Ser Lys Asn  
 325 330 335

Val Ser Thr Ile Gly Lys Phe Cys Glu Leu Lys Ile Asn Thr Lys Gln  
 340 345 350

Ser Tyr Asn Asp Ser Ala Gly Ser Ser Gly Asp Gly Val His Thr Leu  
 355 360 365

Val Gln Asn Asn Gln Glu Glu Arg Lys Tyr Ser Tyr Asn Pro Ser Thr  
 370 375 380

Asn Gln Ile Leu Asp Thr Asn Val Thr Lys Lys Glu Lys Leu Val Arg Asp  
 385 390 395 400

Ile Lys Ser Leu Val Glu Ile Lys Lys Lys Phe Ser Glu Leu Ala Arg  
 405 410 415

Lys Ile Lys Ile Asn Lys Lys Leu Leu Met Ala Ala Gly Cys Ser Lys  
 420 425 430

Thr Ala Asn Thr Ser Tyr Thr Glu Pro Thr Arg His Ser Glu Phe Ser  
 435 440 445

Ala Lys Glu Met Ser Ala Lys Arg Asp Asn Gln Cys Ser Met Glu Leu  
 450 455 460

Leu Ala Thr Cys Leu Ser Leu Trp Lys Asn Gln Pro Pro Lys Thr Thr  
 465 470 475 480

Glu Glu Asn Val Ser Lys Pro Leu Glu Glu Lys Gln Tyr Asn Ala Ser  
 485 490 495

ES 2 799 735 T3

Arg Thr Ser Thr Thr Ala Val Gly Pro Ser Asn Pro Met Asn Glu Val  
500 505 510

His Val Lys Asn Phe Cys Ser Gly Val Arg Asn Ser Gln Lys Ile Thr  
515 520 525

Thr Ser Ser Gln Thr Val Leu Ser Val Leu Thr Pro Val Tyr Asp Ser  
530 535 540

Ser Asp Val Ala Val Gly Lys Gly Thr Glu Leu Gln Ile Ala Val Val  
545 550 555 560

Ser Pro Leu Ile Leu Ser Asp Val Ser Thr Val Pro Gly Lys Glu Leu  
565 570 575

Ala Pro Glu Val Val Ser Glu Thr Val Tyr Pro Val Val Lys Glu Gly  
580 585 590

Ser Val Cys Ser Leu Gln Asn Gln Gln Ala Glu Asn Ala Thr Val Thr  
595 600 605

Ala Gly Leu Pro Phe Asp Val Ile Arg Ala Val Ala Ser Ala Thr Val  
610 615 620

Ser Ala Glu Leu Ser Leu Pro Gly His Lys Glu Lys Gln His Lys Pro  
625 630 635 640

Thr Gln Ser Asp Leu Asp Ile Ala Asp Gly Ser Leu Gly Lys His Ser  
645 650 655

Pro Gln Gly Ala Glu Ala Leu Pro Asn Pro Arg Asp Ser Thr Ile Val  
660 665 670

Ser Gly Pro Ile Leu Gln Ile Glu Ser Ile Cys Ser Leu Ala Glu Gly  
675 680 685

Asp Val Ser Tyr Asn Ser Gln Ile Ala Glu Ile Phe Asn Ser Val Gln  
690 695 700

Asn Glu Pro Gln Lys Pro Ser Pro Asp Gln Gln Val Ile Asn Ser Gln  
705 710 715 720

Gln Glu Glu Gln Val Asp Lys Val Ala Glu Asn Lys Asp Leu Ser Phe  
725 730 735

Leu Lys Asp Lys Cys Met Gln Cys Thr Asp Val Pro His Glu Val Thr  
740 745 750

ES 2 799 735 T3

Glu Gln Pro Glu Pro Leu Gln Pro Leu Glu Thr Thr Ser Asp Glu Tyr  
 755 760 765

Val Glu Ala Asn Gly Glu Ile Leu Glu Glu Ser Ser Lys Glu Asn Pro  
 770 775 780

Gly Glu Lys Glu Met Thr Lys Asp Ile Leu Cys Ser Pro Ala Ala Val  
 785 790 795 800

Gln Gln Asp Pro Gln Pro Gln Glu Ile Asp Thr Ala Ser Ser Lys Ser  
 805 810 815

Gly His Ser Phe Ser Thr Val Asn Glu Ile Asn Asp Glu Asn Glu Pro  
 820 825 830

Val Ser Tyr Leu His Asp Gln Leu Leu Glu Leu Leu Lys Glu Phe Pro  
 835 840 845

Tyr Gly Ile Glu Thr Ile Ala Arg Pro Glu Val Tyr Val Gly Gln Gln  
 850 855 860

Lys Thr His Glu Ile Leu Glu Asn Gln Thr Gly Ser Lys Thr Gly Asn  
 865 870 875 880

Val Ser Gly Asp Asn Thr Asp Gln Ile Lys Ile Thr Val Leu Asn Ser  
 885 890 895

Glu Gln Ile Lys Glu Leu Phe Pro Glu Glu Asp Gln Pro Cys Asp Val  
 900 905 910

Asp Lys Leu Ala Glu Pro Glu Asn Thr Lys Ile Ile Ala Glu Val Lys  
 915 920 925

Ser Leu Cys Asp Ser Gln Val Pro Arg Glu Glu Ser His Asn Pro Gly  
 930 935 940

Met Leu Asp Leu Glu Lys Asp Lys Ile His Cys Cys Ala Leu Gly Trp  
 945 950 955 960

Leu Ser Met Val Tyr Glu Gly Val Pro Gln Cys Gln Cys Ser Ser Met  
 965 970 975

Glu Glu Lys Glu Lys Asp Gln Cys Ser Leu Glu Ile Ser Asn Cys Lys  
 980 985 990

Gln Gly Glu Gln Ala Cys Asn Ser Gly Ile Thr Ile Phe Glu Ile Asn

ES 2 799 735 T3

995				1000				1005						
Pro	Ile	Ser	Asn	Asn	Ser	Lys	Ser	Pro	Leu	Ile	Gln	Glu	Ser	Glu
	1010					1015					1020			
Lys	Gly	His	Phe	Ser	Asp	Ile	His	Gly	Glu	Lys	Ile	Lys	Thr	Ser
	1025					1030					1035			
Glu	Thr	Lys	Asn	Ser	Ser	Ser	Pro	Arg	Val	Glu	Gln	Glu	Leu	Thr
	1040					1045					1050			
Gly	His	Phe	Ser	Met	Lys	Cys	Tyr	Gln	Lys	Asp	Lys	Ser	Thr	Thr
	1055					1060					1065			
Lys	Gln	Asp	Ser	Ser	Leu	Lys	Thr	Glu	Gln	Lys	Ile	Lys	Asn	Leu
	1070					1075					1080			
Ser	Ser	Lys	Cys	Asp	Lys	Pro	Asn	Pro	Leu	Lys	Ser	Ser	Lys	Ile
	1085					1090					1095			
Pro	Thr	Pro	Glu	Thr	Phe	Asn	Val	Val	Thr	Ser	Asn	Ser	Asp	Lys
	1100					1105					1110			
Asn	Met	Pro	Ala	Phe	Ser	Lys	Gln	Asp	Ser	Gln	Gly	Ser	Leu	Gln
	1115					1120					1125			
Lys	Lys	His	Leu	Phe	Gln	Asp	Ser	Asp	Pro	Val	Lys	Gly	His	Val
	1130					1135					1140			
Trp	Leu	Leu	Pro	Asn	Lys	Asp	Pro	Arg	Arg	Arg	Asn	Thr	Phe	Leu
	1145					1150					1155			
Val	Gln	Ser	Val	Ser	Pro	Glu	Lys	Lys	Lys	Leu	Lys	Phe	Lys	Ser
	1160					1165					1170			
Gly	Ser	Ser	Lys	Leu	Lys	Tyr	Phe	Glu	Lys	Arg	Lys	Met	Asp	His
	1175					1180					1185			
Leu	Leu	Ile	Ser	Asp	Val	Glu	Ile	Lys	Lys	Lys	Lys	Tyr	Glu	Lys
	1190					1195					1200			
Gln	Glu	Gln	Asn	Lys	Asn	Ala	Gly	Gly	Thr	Leu	Lys	Leu	Cys	Ser
	1205					1210					1215			
Thr	Leu	Thr	Glu	Pro	Asn	Glu	Arg	Ala	Cys	Ala	Lys	Glu	Lys	Ile
	1220					1225					1230			

ES 2 799 735 T3

Val Thr Asn Ser Glu Pro Ser Asp Ser Lys Gly Ser Ser Ser Lys  
1235 1240 1245

Ser Thr Arg Val Ile Thr Val Gln Glu Tyr Leu Gln Arg Lys Lys  
1250 1255 1260

Asp Lys His Val Ile Gly Asn Asn Ala Ser Lys Asn Ile Cys Val  
1265 1270 1275

Glu Asn Val Pro Cys Asp Ser Glu Pro Met Lys Ser Ser Lys His  
1280 1285 1290

Ser Ala Ser Pro Ser Leu Gly Lys Leu Ile Glu Gly Gln Gly Val  
1295 1300 1305

Ser Ala Glu Thr Leu Lys Glu Val Glu His Asn Ser Thr Ser His  
1310 1315 1320

Gly Lys Asn Leu Lys Thr His Arg Ser Glu Glu Thr Arg Pro Tyr  
1325 1330 1335

Ser Val Ser Asn Ser Lys Glu Lys Phe Tyr Arg Thr His Pro Asp  
1340 1345 1350

Lys Ser Tyr Ile Asp Lys Ala Lys Leu Glu Arg Leu Thr Ser Met  
1355 1360 1365

Ser Ser Lys Ser Ser Gln Leu Gln Val Lys Glu Lys Arg Lys Gln  
1370 1375 1380

Tyr Leu Asn Arg Val Ala Phe Lys Cys Thr Glu Gln Glu Ser Ile  
1385 1390 1395

Cys Leu Thr Lys Leu Asp Ser Ala Ser Lys Lys Leu Ser Lys Glu  
1400 1405 1410

Lys Glu Lys Ser Thr Ala Cys Ala Pro Met Thr Lys Asp Tyr Thr  
1415 1420 1425

His Lys Pro Met Leu Glu Phe Lys Leu Cys Pro Asp Val Leu Leu  
1430 1435 1440

Lys Asn Thr Ser Ser Ile Asp Lys Gly Asp Asp Pro Arg Pro Gly  
1445 1450 1455

Pro Glu Lys Glu Arg Ala Pro Val Gln Val Ser Gly Ile Lys Thr  
1460 1465 1470

ES 2 799 735 T3

Thr Lys Glu Asp Trp Leu Lys Cys Ile Pro Thr Arg Thr Lys Met  
 1475 1480 1485

Pro Glu Ser Ser Glu Gln Thr Asp Arg Ala Asp Ser Arg Leu Ser  
 1490 1495 1500

Lys Arg Ser Phe Ser Ala Asp Glu Phe Glu Thr Leu Gln Asn Pro  
 1505 1510 1515

Val Lys Asp Ser Asn Val Met Phe Arg Thr Phe Lys Lys Met Tyr  
 1520 1525 1530

Leu Glu Lys Arg Ser Arg Ser Leu Gly Ser Ser Pro Val Lys  
 1535 1540 1545

<210> 22  
 <211> 1515  
 <212> PRT  
 <213> Cricetinae sp.

5

<220>  
 <221> MOD\_RES  
 <222> (879)..(879)  
 <223> Qualquer aminoácido

10

<400> 22

Met Asn Trp Asn Ala Lys Pro Glu Asn Ala Ala Pro Asn Pro Pro Tyr  
 1 5 10 15

Ser Lys Ser Gln Ser Ser Leu Leu Gln Gln Phe Leu Met Pro Ser Thr  
 20 25 30

Thr Ser Gln Ser Ser Phe Ser Cys Leu Pro His Asn Gln Glu Ala Cys  
 35 40 45

Ile Tyr Pro Thr Asn Ser Asn Ser Val Ser Gln Pro Leu Leu Asn Val  
 50 55 60

Arg Ser Phe Ile Asn Pro Pro Ile Ser Val Ser Asn Val His Asn Arg  
 65 70 75 80

Thr Val Val Ala Ser Gln Thr Ser Val Glu Arg Val Thr Tyr Thr Asn  
 85 90 95

Val Lys Gly Ala Gln Gln Pro Asn His Asn Leu Gln Thr Val Ser Ser  
 100 105 110

Gly Val Val Gln Asn Ala Trp Met Asn Ser Thr Met Arg Asn Phe Met  
 115 120 125

ES 2 799 735 T3

Pro Ser Leu Thr Glu Ala Thr Ile Ser His Lys Pro Asp Gly Gly Pro  
 130 135 140

Ser Met Pro Tyr Met His Ala Pro Gln Ser His Leu Val Thr Ser Asp  
 145 150 155 160

Thr Tyr Ser Val Gln Leu Gln Met Thr Pro Ser Asn Ser Val Arg Gly  
 165 170 175

Pro Val Thr Tyr Gln Gly Asn Tyr Gln Gly Asn Pro Gly Leu Asn His  
 180 185 190

Ser Met Ala Gly Glu Leu Gly Trp Val Gln Cys Ala Ser Ser Glu Leu  
 195 200 205

Thr Tyr Pro Asp Tyr Arg Pro Pro Pro Lys Gln Tyr Pro Tyr Leu Pro  
 210 215 220

Gln Ser Phe Val Gln Asp Thr Ser Val Gln Lys Gln Asn Phe Val Ser  
 225 230 235 240

Ser Thr Ser Leu Gln Val Lys Asn Asn Gln Leu Pro Pro Ser Thr Gln  
 245 250 255

Thr Leu Pro Ser Lys Arg Pro Val Pro Val Ser Ser Tyr Gln Tyr Ala  
 260 265 270

Ala Glu Thr Ser Lys Arg Leu Pro Pro Pro Pro Tyr Ser Cys Arg Tyr  
 275 280 285

Gly Ser Gln His Val Gln Asn Ser Gln Ser Val Ser Arg His Leu Pro  
 290 295 300

Val Glu Val Pro Gln Ser Ser Glu Met His Ser Ser Glu Lys Lys Lys  
 305 310 315 320

Asp Ala Tyr Lys Val Phe Gln Gln Gln Trp Gln Ser Thr Ser Lys Asn  
 325 330 335

Val Ser Thr Ile Gly Lys Phe Cys Glu Leu Lys Ile Asn Thr Lys Gln  
 340 345 350

Ser Tyr Asn Asp Ser Ala Gly Ser Ser Gly Asp Gly Val His Thr Leu  
 355 360 365

Val Gln Asn Asn Gln Glu Glu Arg Lys Tyr Ser Tyr Asn Pro Ser Thr

ES 2 799 735 T3

370		375		380											
Asn	Gln	Ile	Leu	Asp	Thr	Asn	Val	Thr	Lys	Glu	Lys	Leu	Val	Arg	Asp
385						390				395					400
Ile	Lys	Ser	Leu	Val	Glu	Ile	Ser	Trp	Ala	Met	Val	Ala	His	Ser	Glu
				405					410					415	
Phe	Ser	Ala	Lys	Glu	Met	Ser	Ala	Lys	Arg	Asp	Asn	Gln	Cys	Ser	Met
			420					425					430		
Glu	Leu	Leu	Ala	Thr	Cys	Leu	Ser	Leu	Trp	Lys	Asn	Gln	Pro	Pro	Lys
		435					440					445			
Thr	Thr	Glu	Glu	Asn	Val	Ser	Lys	Pro	Leu	Glu	Glu	Lys	Gln	Tyr	Asn
	450					455						460			
Ala	Ser	Arg	Thr	Ser	Thr	Thr	Ala	Val	Gly	Pro	Ser	Asn	Pro	Met	Asn
465					470					475					480
Glu	Val	His	Val	Lys	Asn	Phe	Cys	Ser	Gly	Val	Arg	Asn	Ser	Gln	Lys
				485					490					495	
Ile	Thr	Thr	Ser	Ser	Gln	Thr	Val	Leu	Ser	Val	Leu	Thr	Pro	Val	Tyr
			500					505					510		
Asp	Ser	Ser	Asp	Val	Ala	Val	Gly	Lys	Gly	Thr	Glu	Leu	Gln	Ile	Ala
		515					520					525			
Val	Val	Ser	Pro	Leu	Ile	Leu	Ser	Asp	Val	Ser	Thr	Val	Pro	Gly	Lys
	530					535					540				
Glu	Leu	Ala	Pro	Glu	Val	Val	Ser	Glu	Thr	Val	Tyr	Pro	Val	Val	Lys
545					550					555					560
Glu	Gly	Ser	Val	Cys	Ser	Leu	Gln	Asn	Gln	Gln	Ala	Glu	Asn	Ala	Thr
				565					570					575	
Val	Thr	Ala	Gly	Leu	Pro	Phe	Asp	Val	Ile	Arg	Ala	Val	Ala	Ser	Ala
			580					585					590		
Thr	Val	Ser	Ala	Glu	Leu	Ser	Leu	Pro	Gly	His	Lys	Glu	Lys	Gln	His
		595					600					605			
Lys	Pro	Thr	Gln	Thr	Asp	Leu	Asp	Thr	Ala	Asp	Gly	Ser	Leu	Gly	Lys
	610					615					620				

ES 2 799 735 T3

His Ser Pro Gln Gly Ala Glu Ala Leu Pro Asn Pro Arg Asp Ser Thr  
 625 630 635 640

Ile Val Ser Gly Pro Ile Leu Gln Ile Glu Ser Ile Cys Ser Leu Ala  
 645 650 655

Glu Gly Asp Val Ser Tyr Asn Ser Gln Ile Ala Glu Ile Phe Asn Ser  
 660 665 670

Val Gln Asn Glu Pro Gln Lys Pro Ser Pro Asp Gln Gln Val Ile Asn  
 675 680 685

Ser Gln Gln Glu Glu Gln Val Asp Lys Val Ala Glu Asn Lys Asp Leu  
 690 695 700

Ser Phe Leu Lys Asp Lys Cys Met Gln Cys Thr Asp Val Pro His Glu  
 705 710 715 720

Val Thr Glu Gln Pro Glu Pro Leu Gln Pro Leu Glu Thr Thr Ser Asp  
 725 730 735

Glu Tyr Val Glu Ala Asn Gly Glu Ile Leu Glu Glu Ser Ser Lys Glu  
 740 745 750

Asn Pro Gly Glu Lys Glu Met Thr Lys Asp Ile Leu Cys Ser Pro Ala  
 755 760 765

Ala Val Gln Gln Asp Pro Gln Pro Gln Glu Ile Asp Thr Ala Ser Ser  
 770 775 780

Lys Ser Gly His Ser Phe Ser Thr Val Asn Glu Ile Asn Asp Glu Asn  
 785 790 795 800

Glu Pro Val Ser Tyr Leu His Asp Gln Leu Leu Glu Leu Leu Lys Glu  
 805 810 815

Phe Pro Tyr Gly Ile Glu Thr Ile Ala Arg Pro Glu Val Tyr Val Gly  
 820 825 830

Gln Gln Lys Thr His Glu Ile Leu Glu Asn Gln Thr Gly Ser Lys Thr  
 835 840 845

Gly Asn Val Ser Gly Asp Asn Thr Asp Gln Ile Lys Ile Thr Val Leu  
 850 855 860

Asn Ser Glu Gln Ile Lys Glu Leu Phe Pro Glu Glu Asp Gln Xaa Val  
 865 870 875 880

ES 2 799 735 T3

Asp Lys Leu Ala Glu Pro Glu Asn Thr Lys Ile Ile Ala Glu Val Lys  
 885 890 895

Ser Leu Cys Asp Ser Gln Val Pro Arg Glu Glu Ser His Asn Pro Gly  
 900 905 910

Met Leu Asp Leu Glu Lys Asp Lys Ile His Cys Cys Ala Leu Gly Trp  
 915 920 925

Leu Ser Met Val Tyr Glu Gly Val Pro Gln Cys Gln Cys Ser Ser Met  
 930 935 940

Glu Glu Lys Glu Lys Asp Gln Cys Ser Leu Glu Ile Ser Asn Cys Lys  
 945 950 955 960

Gln Gly Glu Gln Ala Cys Asn Ser Gly Ile Thr Ile Phe Glu Ile Asn  
 965 970 975

Pro Ile Ser Asn Asn Ser Lys Ser Pro Leu Ile Gln Glu Ser Glu Lys  
 980 985 990

Gly His Phe Ser Asp Ile His Gly Glu Lys Ile Lys Thr Ser Glu Thr  
 995 1000 1005

Lys Asn Ser Ser Ser Pro Arg Val Glu Gln Glu Leu Thr Gly His  
 1010 1015 1020

Phe Ser Met Lys Cys Tyr Gln Lys Asp Lys Ser Thr Thr Lys Gln  
 1025 1030 1035

Asp Ser Ser Leu Lys Thr Glu Gln Lys Ile Lys Asn Leu Ser Ser  
 1040 1045 1050

Lys Cys Asp Lys Pro Asn Pro Leu Lys Ser Ser Lys Ile Pro Thr  
 1055 1060 1065

Pro Glu Thr Phe Asn Val Val Thr Ser Asn Ser Asp Lys Asn Met  
 1070 1075 1080

Pro Ala Phe Ser Lys Gln Asp Ser Gln Gly Ser Leu Gln Lys Lys  
 1085 1090 1095

His Leu Phe Gln Asp Ser Asp Pro Val Lys Gly His Val Trp Leu  
 1100 1105 1110

Leu Pro Asn Lys Asp Pro Arg Arg Arg Asn Thr Phe Leu Val Gln  
 1115 1120 1125

ES 2 799 735 T3

Ser Val Ser Pro Glu Lys Lys Lys Leu Lys Phe Lys Ser Gly Ser  
1130 1135 1140

Ser Lys Leu Lys Tyr Phe Glu Lys Arg Lys Met Asp His Leu Leu  
1145 1150 1155

Ile Ser Asp Val Glu Ile Lys Lys Lys Tyr Glu Lys Gln Glu  
1160 1165 1170

Gln Asn Lys Asn Ala Gly Gly Thr Leu Lys Leu Cys Ser Thr Leu  
1175 1180 1185

Thr Glu Pro Asn Glu Arg Ala Cys Ala Lys Glu Lys Ile Val Thr  
1190 1195 1200

Asn Ser Glu Pro Ser Asp Ser Lys Gly Ser Ser Ser Lys Ser Thr  
1205 1210 1215

Arg Val Ile Thr Val Gln Glu Tyr Leu Gln Arg Lys Lys Asp Lys  
1220 1225 1230

His Val Ile Gly Asn Asn Ala Ser Lys Asn Ile Cys Val Glu Asn  
1235 1240 1245

Val Pro Cys Asp Ser Glu Pro Met Lys Ser Ser Lys His Ser Ala  
1250 1255 1260

Ser Pro Ser Leu Gly Lys Leu Ile Glu Gly Gln Gly Val Ser Ala  
1265 1270 1275

Glu Thr Leu Lys Glu Val Glu His Asn Ser Ser Ser His Gly Lys  
1280 1285 1290

Asn Leu Lys Thr His Arg Ser Glu Glu Thr Arg Pro Tyr Ser Val  
1295 1300 1305

Ser Asn Ser Lys Glu Lys Phe Tyr Arg Thr His Pro Asp Lys Ser  
1310 1315 1320

Tyr Ile Asp Lys Ala Lys Leu Glu Arg Leu Thr Ser Met Ser Ser  
1325 1330 1335

Lys Ser Ser Gln Leu Gln Val Lys Glu Lys Arg Lys Gln Tyr Leu  
1340 1345 1350

Asn Arg Val Ala Phe Lys Cys Thr Glu Gln Glu Ser Ile Cys Leu

ES 2 799 735 T3

1355		1360		1365
Thr Lys Leu Asp Ser Ala Ser Lys Lys Leu Ser Lys Glu Lys Glu				
1370		1375		1380
Lys Ser Thr Ala Cys Ala Pro Met Thr Lys Asp Tyr Thr His Lys				
1385		1390		1395
Pro Met Leu Glu Phe Lys Leu Cys Pro Asp Val Leu Leu Lys Asn				
1400		1405		1410
Thr Ser Ser Ile Asp Lys Gly Asp Asp Pro Arg Pro Gly Pro Glu				
1415		1420		1425
Lys Glu Arg Ala Pro Val Gln Val Ser Gly Ile Lys Thr Thr Lys				
1430		1435		1440
Glu Asp Trp Leu Lys Cys Ile Pro Thr Arg Thr Lys Met Pro Glu				
1445		1450		1455
Ser Ser Glu Gln Thr Asp Arg Ala Asp Ser Arg Leu Ser Lys Arg				
1460		1465		1470
Ser Phe Ser Ala Asp Glu Phe Glu Thr Leu Gln Asn Pro Val Lys				
1475		1480		1485
Asp Ser Asn Val Met Phe Arg Thr Phe Lys Lys Met Tyr Leu Glu				
1490		1495		1500
Lys Arg Ser Arg Ser Leu Gly Ser Ser Pro Val Lys				
1505		1510		1515

<210> 23  
 <211> 1747  
 <212> PRT  
 <213> *Homo sapiens*

5

<400> 23

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

ES 2 799 735 T3

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

ES 2 799 735 T3

305					310						315					320
Gln	Gln	Gln	Trp	Gln	Asn	Pro	Asn	Glu	Asn	Val	Ser	Thr	Ile	Gly	Asn	
				325					330					335		
Phe	Thr	Asn	Leu	Lys	Val	Asn	Thr	Asn	Ser	Lys	Gln	Pro	Phe	Asn	Ser	
			340					345					350			
Pro	Ile	Arg	Ser	Ser	Val	Asp	Gly	Val	Gln	Thr	Leu	Ala	Gln	Thr	Asn	
		355					360					365				
Glu	Glu	Lys	Ile	Met	Asp	Ser	Cys	Asn	Pro	Thr	Ser	Asn	Gln	Val	Leu	
	370					375					380					
Asp	Thr	Ser	Val	Ala	Lys	Glu	Lys	Leu	Val	Arg	Asp	Ile	Lys	Thr	Leu	
385					390					395					400	
Val	Glu	Ile	Lys	Gln	Lys	Phe	Ser	Glu	Leu	Ala	Arg	Lys	Ile	Lys	Ile	
				405					410					415		
Asn	Lys	Asp	Leu	Leu	Met	Ala	Ala	Gly	Cys	Ile	Lys	Met	Thr	Asn	Thr	
			420					425					430			
Ser	Tyr	Ser	Glu	Pro	Ala	Gln	Asn	Ser	Lys	Leu	Ser	Leu	Lys	Gln	Thr	
		435					440					445				
Ala	Lys	Ile	Gln	Ser	Gly	Pro	Gln	Ile	Thr	Pro	Val	Met	Pro	Glu	Asn	
	450					455					460					
Ala	Glu	Arg	Gln	Thr	Pro	Thr	Val	Val	Glu	Ser	Ala	Glu	Thr	Asn	Lys	
465					470					475					480	
Thr	Gln	Cys	Met	Leu	Asn	Ser	Asp	Ile	Gln	Glu	Val	Asn	Cys	Arg	Arg	
				485					490					495		
Phe	Asn	Gln	Val	Asp	Ser	Val	Leu	Pro	Asn	Pro	Val	Tyr	Ser	Glu	Lys	
			500					505					510			
Arg	Pro	Met	Pro	Asp	Ser	Ser	His	Asp	Val	Lys	Val	Leu	Thr	Ser	Lys	
		515					520					525				
Thr	Ser	Ala	Val	Glu	Met	Thr	Gln	Ala	Val	Leu	Asn	Thr	Gln	Leu	Ser	
	530					535					540					
Ser	Glu	Asn	Val	Thr	Lys	Val	Glu	Gln	Asn	Ser	Pro	Ala	Val	Cys	Glu	
545					550					555					560	

ES 2 799 735 T3

Thr Ile Ser Val Pro Lys Ser Met Ser Thr Glu Glu Tyr Lys Ser Lys  
 565 570 575

Ile Gln Asn Glu Asn Met Leu Leu Leu Ala Leu Leu Ser Gln Ala Arg  
 580 585 590

Lys Thr Gln Lys Thr Val Leu Lys Asp Ala Asn Gln Thr Ile Gln Asp  
 595 600 605

Ser Lys Pro Asp Ser Cys Glu Met Asn Pro Asn Thr Gln Met Thr Gly  
 610 615 620

Asn Gln Leu Asn Leu Lys Asn Met Glu Thr Pro Ser Thr Ser Asn Val  
 625 630 635 640

Ser Gly Arg Val Leu Asp Asn Ser Phe Cys Ser Gly Gln Glu Ser Ser  
 645 650 655

Thr Lys Gly Met Pro Ala Lys Ser Asp Ser Ser Cys Ser Met Glu Val  
 660 665 670

Leu Ala Thr Cys Leu Ser Leu Trp Lys Lys Gln Pro Ser Asp Thr Ala  
 675 680 685

Lys Glu Lys Glu Cys Asp Lys Leu Arg Thr Asn Thr Thr Ala Val Gly  
 690 695 700

Ile Ser Lys Pro Ala Asn Ile His Val Lys Ser Pro Cys Ser Val Val  
 705 710 715 720

Gly Asn Ser Asn Ser Gln Asn Lys Ile Ser Asn Pro Ser Gln Gln Thr  
 725 730 735

Ala Leu Ser Met Val Met His Asn Tyr Glu Ser Ser Gly Ile Asn Ile  
 740 745 750

Thr Lys Gly Thr Glu Leu Gln Ile Ala Val Val Ser Pro Leu Val Leu  
 755 760 765

Ser Glu Val Lys Thr Leu Ser Val Lys Gly Ile Thr Pro Ala Val Leu  
 770 775 780

Pro Glu Thr Val Tyr Pro Val Ile Lys Glu Gly Ser Val Cys Ser Leu  
 785 790 795 800

Gln Asn Gln Leu Ala Glu Asn Ala Lys Ala Thr Ala Ala Leu Lys Val  
 805 810 815

ES 2 799 735 T3

Asp Val Ser Gly Pro Val Ala Ser Thr Ala Thr Ser Thr Lys Ile Phe  
 820 825 830  
 Pro Leu Thr Gln Lys Glu Lys Gln Asn Glu Ser Thr Asn Gly Asn Ser  
 835 840 845  
 Glu Val Thr Pro Asn Val Asn Gln Gly Lys His Asn Lys Leu Glu Ser  
 850 855 860  
 Ala Ile His Ser Pro Met Asn Asp Gln Gln Ile Ser Gln Glu Ser Arg  
 865 870 875 880  
 Asn Ser Thr Val Val Ser Ser Asp Thr Leu Gln Ile Asp Asn Ile Cys  
 885 890 895  
 Ser Leu Val Glu Gly Asp Thr Ser Tyr Asn Ser Gln Ile Ala Lys Ile  
 900 905 910  
 Phe Ser Ser Leu Pro Leu Lys Met Val Glu Pro Gln Lys Pro Ser Leu  
 915 920 925  
 Pro Asn Gln Gln Gly Ile Gly Ser Arg Glu Pro Glu Lys Gln Leu Asp  
 930 935 940  
 Asn Thr Thr Glu Asn Lys Asp Phe Gly Phe Gln Lys Asp Lys Pro Val  
 945 950 955 960  
 Gln Cys Thr Asp Val Ser His Lys Ile Cys Asp Gln Ser Lys Ser Glu  
 965 970 975  
 Pro Pro Leu Glu Ser Ser Phe Asn Asn Leu Glu Thr Asn Arg Val Ile  
 980 985 990  
 Leu Glu Lys Ser Ser Leu Glu His Ala Thr Glu Lys Ser Thr Ala Asn  
 995 1000 1005  
 Asp Thr Cys Ser Ser Ala Ala Ile Gln Glu Asp Ile Tyr Pro Gln  
 1010 1015 1020  
 Glu Ile Asp Ala Ser Ser Asn Tyr Thr Pro Gln Asp Pro Ala Arg  
 1025 1030 1035  
 Asn Glu Ile His Ser Asp Lys Ala Pro Val Leu Tyr Leu His Asp  
 1040 1045 1050  
 Gln Leu Ser Glu Leu Leu Lys Glu Phe Pro Tyr Gly Ile Glu Ala  
 1055 1060 1065

ES 2 799 735 T3

Val	Asn	Thr	Arg	Glu	Gly	Ser	Val	Gly	Gln	Gln	Thr	Thr	Tyr	Gln
1070						1075					1080			
Thr	Ser	Glu	Asp	Gln	Thr	Ala	Asp	Lys	Thr	Ser	Ser	Asp	Ser	Lys
1085						1090					1095			
Asp	Pro	Ala	Asp	Gln	Ile	Gln	Ile	Thr	Ile	Leu	Ser	Ser	Glu	Gln
1100						1105					1110			
Met	Lys	Glu	Ile	Phe	Pro	Glu	Gln	Asp	Asp	Gln	Pro	Tyr	Val	Val
1115						1120					1125			
Asp	Lys	Leu	Ala	Glu	Pro	Gln	Lys	Glu	Glu	Pro	Ile	Thr	Glu	Val
1130						1135					1140			
Val	Ser	Gln	Cys	Asp	Leu	Gln	Ala	Pro	Ala	Ala	Gly	Gln	Ser	Arg
1145						1150					1155			
Asp	Ser	Val	Ile	Leu	Asp	Ser	Glu	Lys	Asp	Asp	Ile	His	Cys	Cys
1160						1165					1170			
Ala	Leu	Gly	Trp	Leu	Ser	Met	Val	Tyr	Glu	Gly	Val	Pro	Gln	Cys
1175						1180					1185			
Gln	Cys	Asn	Ser	Ile	Lys	Asn	Ser	Ser	Ser	Glu	Glu	Glu	Lys	Gln
1190						1195					1200			
Lys	Glu	Gln	Cys	Ser	Pro	Leu	Asp	Thr	Asn	Ser	Cys	Lys	Gln	Gly
1205						1210					1215			
Glu	Arg	Thr	Ser	Asp	Arg	Asp	Val	Thr	Val	Val	Gln	Phe	Lys	Ser
1220						1225					1230			
Leu	Val	Asn	Asn	Pro	Lys	Thr	Pro	Pro	Asp	Gly	Lys	Ser	His	Phe
1235						1240					1245			
Pro	Glu	Leu	Gln	Asp	Asp	Ser	Arg	Lys	Asp	Thr	Pro	Lys	Thr	Lys
1250						1255					1260			
His	Lys	Ser	Leu	Pro	Arg	Thr	Glu	Gln	Glu	Leu	Val	Ala	Gly	Gln
1265						1270					1275			
Phe	Ser	Ser	Lys	Cys	Asp	Lys	Leu	Asn	Pro	Leu	Gln	Asn	His	Lys
1280						1285					1290			
Arg	Lys	Lys	Leu	Arg	Phe	His	Glu	Val	Thr	Phe	His	Ser	Ser	Asn

ES 2 799 735 T3

1295						1300						1305			
Lys	Met	Thr	Ala	Ser	Tyr	Glu	Gln	Ala	Ser	Gln	Glu	Thr	Arg	Gln	
	1310					1315					1320				
Lys	Lys	His	Val	Thr	Gln	Asn	Ser	Arg	Pro	Leu	Lys	Thr	Lys	Thr	
	1325					1330					1335				
Ala	Phe	Leu	Pro	Asn	Lys	Asp	Val	Tyr	Lys	Lys	His	Ser	Ser	Leu	
	1340					1345					1350				
Gly	Gln	Ser	Leu	Ser	Pro	Glu	Lys	Ile	Lys	Leu	Lys	Leu	Lys	Ser	
	1355					1360					1365				
Val	Ser	Phe	Lys	Gln	Lys	Arg	Lys	Leu	Asp	Gln	Gly	Asn	Val	Leu	
	1370					1375					1380				
Asp	Met	Glu	Val	Lys	Lys	Lys	Lys	His	Asp	Lys	Gln	Glu	Gln	Lys	
	1385					1390					1395				
Gly	Ser	Val	Gly	Ala	Thr	Phe	Lys	Leu	Gly	Asp	Ser	Leu	Ser	Asn	
	1400					1405					1410				
Pro	Asn	Glu	Arg	Ala	Ile	Val	Lys	Glu	Lys	Met	Val	Ser	Asn	Thr	
	1415					1420					1425				
Lys	Ser	Val	Asp	Thr	Lys	Ala	Ser	Ser	Ser	Lys	Phe	Ser	Arg	Ile	
	1430					1435					1440				
Leu	Thr	Pro	Lys	Glu	Tyr	Leu	Gln	Arg	Gln	Lys	His	Lys	Glu	Ala	
	1445					1450					1455				
Leu	Ser	Asn	Lys	Ala	Ser	Lys	Lys	Ile	Cys	Val	Lys	Asn	Val	Pro	
	1460					1465					1470				
Cys	Asp	Ser	Glu	His	Met	Arg	Pro	Ser	Lys	Leu	Ala	Val	Gln	Val	
	1475					1480					1485				
Glu	Ser	Cys	Gly	Lys	Ser	Asn	Glu	Lys	His	Ser	Ser	Gly	Val	Gln	
	1490					1495					1500				
Thr	Ser	Lys	Glu	Ser	Leu	Asn	Gly	Leu	Thr	Ser	His	Gly	Lys	Asn	
	1505					1510					1515				
Leu	Lys	Ile	His	His	Ser	Gln	Glu	Ser	Lys	Thr	Tyr	Asn	Ile	Leu	
	1520					1525					1530				

ES 2 799 735 T3

Arg Asn Val Lys Glu Lys Val Gly Gly Lys Gln Pro Asp Lys Ile  
 1535 1540 1545

Trp Ile Asp Lys Thr Lys Leu Asp Lys Leu Thr Asn Ile Ser Asn  
 1550 1555 1560

Glu Ala Gln Phe Ser Gln Met Pro Pro Gln Val Lys Asp Gln Lys  
 1565 1570 1575

Lys Leu Tyr Leu Asn Arg Val Gly Phe Lys Cys Thr Glu Arg Glu  
 1580 1585 1590

Ser Ile Ser Leu Thr Lys Leu Glu Ser Ser Pro Arg Lys Leu His  
 1595 1600 1605

Lys Asp Lys Arg Gln Glu Asn Lys His Lys Thr Phe Leu Pro Val  
 1610 1615 1620

Lys Gly Asn Thr Glu Lys Ser Asn Met Leu Glu Phe Lys Leu Cys  
 1625 1630 1635

Pro Asp Ile Leu Leu Lys Asn Thr Asn Ser Val Glu Glu Arg Lys  
 1640 1645 1650

Asp Val Lys Pro His Pro Arg Lys Glu Gln Ala Pro Leu Gln Val  
 1655 1660 1665

Ser Gly Ile Lys Ser Thr Lys Glu Asp Trp Leu Lys Phe Val Ala  
 1670 1675 1680

Thr Lys Lys Arg Thr Gln Lys Asp Ser Gln Glu Arg Asp Asn Val  
 1685 1690 1695

Asn Ser Arg Leu Ser Lys Arg Ser Phe Ser Ala Asp Gly Phe Glu  
 1700 1705 1710

Met Leu Gln Asn Pro Val Lys Asp Ser Lys Glu Met Phe Gln Thr  
 1715 1720 1725

Tyr Lys Gln Met Tyr Leu Glu Lys Arg Ser Arg Ser Leu Gly Ser  
 1730 1735 1740

Ser Pro Val Lys  
 1745

<210> 24  
 <211> 1747  
 <212> PRT

ES 2 799 735 T3

<213> *Homo sapiens*

<400> 24

Met Asn Trp Asn Glu Lys Pro Lys Ser Ala Thr Leu Pro Pro Leu Tyr  
 1 5 10 15

Pro Lys Ser Gln Pro Pro Phe Leu His Gln Ser Leu Ile Asn Gln Ile  
 20 25 30

Thr Thr Thr Ser Gln Ser Ser Phe Ser Tyr Pro Gly Ser Asn Gln Glu  
 35 40 45

Ala Cys Met Tyr Pro Gly Asn Ser Asn Pro Ile Ser Gln Pro Leu Leu  
 50 55 60

Asn Ile Gln Asn Tyr Pro Gln Gln Ile Ser Val Ser Asp Met His Asn  
 65 70 75 80

Gly Thr Val Val Ala Ser His Thr Ser Val Glu Arg Ile Thr Tyr Ala  
 85 90 95

Asn Val Asn Gly Pro Lys Gln Leu Thr His Asn Leu Gln Met Ser Ser  
 100 105 110

Gly Val Thr Gln Asn Val Trp Leu Asn Ser Pro Met Arg Asn Pro Val  
 115 120 125

His Ser His Ile Gly Ala Thr Val Ser His Gln Thr Asp Phe Gly Ala  
 130 135 140

Asn Val Pro Asn Met Pro Ala Leu Gln Ser Gln Leu Ile Thr Ser Asp  
 145 150 155 160

Thr Tyr Ser Met Gln Met Gln Met Ile Pro Ser Asn Ser Thr Arg Leu  
 165 170 175

Pro Val Ala Tyr Gln Gly Asn Gln Gly Leu Asn Gln Ser Phe Ser Glu  
 180 185 190

Gln Gln Val Asp Trp Thr Gln Gln Cys Ile Ser Lys Gly Leu Thr Tyr  
 195 200 205

Pro Asp Tyr Arg Pro Pro Pro Lys Leu Tyr Arg Tyr Ser Pro Gln Ser  
 210 215 220

Phe Leu Pro Asp Ser Thr Ile Gln Lys Gln Asn Phe Ile Pro His Thr  
 225 230 235 240

ES 2 799 735 T3

Ser Leu Gln Val Lys Asn Ser Gln Leu Leu Asn Ser Val Leu Thr Leu  
245 250 255

Pro Ser Arg Gln Thr Ser Ala Val Pro Ser Gln Gln Tyr Ala Thr Gln  
260 265 270

Thr Asp Lys Arg Pro Pro Pro Pro Tyr Asn Cys Arg Tyr Gly Ser  
275 280 285

Gln Pro Leu Gln Ser Thr Gln His Ile Thr Lys His Leu Ser Met Glu  
290 295 300

Val Pro Gln Ser Arg Glu Met Leu Ser Ser Glu Ile Arg Thr Ser Phe  
305 310 315 320

Gln Gln Gln Trp Gln Asn Pro Asn Glu Asn Val Ser Thr Ile Gly Asn  
325 330 335

Phe Thr Asn Leu Lys Val Asn Thr Asn Ser Lys Gln Pro Phe Asn Ser  
340 345 350

Pro Ile Arg Ser Ser Val Asp Gly Val Gln Thr Leu Ala Gln Thr Asn  
355 360 365

Glu Glu Lys Ile Met Asp Ser Cys Asn Pro Thr Ser Asn Gln Val Leu  
370 375 380

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

Val Glu Ile Lys Gln Lys Phe Ser Glu Leu Ala Arg Lys Ile Lys Ile  
405 410 415

Asn Lys Asp Leu Leu Met Ala Ala Gly Cys Ile Lys Met Thr Asn Thr  
420 425 430

Ser Tyr Ser Glu Pro Ala Gln Asn Ser Lys Leu Ser Leu Lys Gln Thr  
435 440 445

Ala Lys Ile Gln Ser Gly Pro Gln Ile Thr Pro Val Met Pro Glu Asn  
450 455 460

Ala Glu Arg Gln Thr Pro Thr Val Val Glu Ser Ala Glu Thr Asn Lys  
465 470 475 480

Thr Gln Cys Met Leu Asn Ser Asp Ile Gln Glu Val Asn Cys Arg Arg  
485 490 495

ES 2 799 735 T3

Phe Asn Gln Val Asp Ser Val Leu Pro Asn Pro Val Tyr Ser Glu Lys  
 500 505 510

Arg Pro Met Pro Asp Pro Ser His Asp Val Lys Val Leu Thr Ser Lys  
 515 520 525

Thr Ser Ala Val Glu Met Thr Gln Ala Val Leu Asn Thr Gln Leu Ser  
 530 535 540

Ser Glu Asn Val Thr Lys Val Glu Gln Asn Ser Pro Ala Val Cys Glu  
 545 550 555 560

Thr Ile Ser Val Pro Lys Ser Met Ser Thr Glu Glu Tyr Lys Ser Lys  
 565 570 575

Ile Gln Asn Glu Asn Met Leu Leu Leu Ala Leu Leu Ser Gln Ala Arg  
 580 585 590

Lys Thr Gln Lys Thr Val Leu Lys Asp Ala Asn Gln Thr Ile Gln Asp  
 595 600 605

Ser Lys Pro Asp Ser Cys Glu Met Asn Pro Asn Thr Gln Met Thr Gly  
 610 615 620

Asn Gln Leu Asn Leu Lys Asn Met Glu Thr Pro Ser Thr Ser Asn Val  
 625 630 635 640

Ser Gly Arg Val Leu Asp Asn Ser Phe Cys Ser Gly Gln Glu Ser Ser  
 645 650 655

Thr Lys Gly Met Pro Ala Lys Ser Asp Ser Ser Cys Ser Met Glu Val  
 660 665 670

Leu Ala Thr Cys Leu Ser Leu Trp Lys Lys Gln Pro Ser Asp Thr Ala  
 675 680 685

Lys Glu Lys Glu Cys Asp Lys Leu Arg Thr Asn Thr Thr Ala Val Gly  
 690 695 700

Ile Ser Lys Pro Ala Asn Ile His Val Lys Ser Pro Cys Ser Val Val  
 705 710 715 720

Gly Asn Ser Asn Ser Gln Asn Lys Ile Ser Asn Pro Ser Gln Gln Thr  
 725 730 735

Ala Leu Ser Met Val Met His Asn Tyr Glu Ser Ser Gly Ile Asn Ile  
 740 745 750

ES 2 799 735 T3

Thr Lys Gly Thr Glu Leu Gln Ile Ala Val Val Ser Pro Leu Val Leu  
 755 760 765

Ser Glu Val Lys Thr Leu Ser Val Lys Gly Ile Thr Pro Ala Val Leu  
 770 775 780

Pro Glu Thr Val Tyr Pro Val Ile Lys Glu Gly Ser Val Cys Ser Leu  
 785 790 795 800

Gln Asn Gln Leu Ala Glu Asn Ala Lys Ala Thr Ala Ala Leu Lys Val  
 805 810 815

Asp Val Ser Gly Pro Val Ala Ser Thr Ala Thr Ser Thr Lys Ile Phe  
 820 825 830

Pro Leu Thr Gln Lys Glu Lys Gln Asn Glu Ser Thr Asn Gly Asn Ser  
 835 840 845

Glu Val Thr Pro Asn Val Asn Gln Gly Lys His Asn Lys Leu Glu Ser  
 850 855 860

Ala Ile His Ser Pro Met Asn Asp Gln Gln Ile Ser Gln Glu Ser Arg  
 865 870 875 880

Asn Ser Thr Val Val Ser Ser Asp Thr Leu Gln Ile Asp Asn Ile Cys  
 885 890 895

Ser Leu Val Glu Gly Asp Thr Ser Tyr Asn Ser Gln Ile Ala Lys Ile  
 900 905 910

Phe Ser Ser Leu Pro Leu Lys Met Val Glu Pro Gln Lys Pro Ser Leu  
 915 920 925

Pro Asn Gln Gln Gly Ile Gly Ser Arg Glu Pro Glu Lys Gln Leu Asp  
 930 935 940

Asn Thr Thr Glu Asn Lys Asp Phe Gly Phe Gln Lys Asp Lys Pro Val  
 945 950 955 960

Gln Cys Thr Asp Val Ser His Lys Ile Cys Asp Gln Ser Lys Ser Glu  
 965 970 975

Pro Pro Leu Glu Ser Ser Phe Asn Asn Leu Glu Thr Asn Arg Val Ile  
 980 985 990

Leu Glu Lys Ser Ser Leu Glu His Ala Thr Glu Lys Ser Thr Ala Asn

ES 2 799 735 T3

995					1000					1005				
Asp	Thr	Cys	Ser	Ser	Ala	Ala	Ile	Gln	Glu	Asp	Ile	Tyr	Pro	Gln
1010						1015					1020			
Glu	Ile	Asp	Ala	Ser	Ser	Asn	Tyr	Thr	Pro	Gln	Asp	Pro	Ala	Arg
1025						1030					1035			
Asn	Glu	Ile	His	Ser	Asp	Lys	Ala	Pro	Val	Leu	Tyr	Leu	His	Asp
1040						1045					1050			
Gln	Leu	Ser	Glu	Leu	Leu	Lys	Glu	Phe	Pro	Tyr	Gly	Ile	Glu	Ala
1055						1060					1065			
Val	Asn	Thr	Arg	Glu	Gly	Ser	Val	Gly	Gln	Gln	Thr	Thr	Tyr	Gln
1070						1075					1080			
Thr	Ser	Glu	Asp	Gln	Thr	Ala	Asp	Lys	Thr	Ser	Ser	Asp	Ser	Lys
1085						1090					1095			
Asp	Pro	Ala	Asp	Gln	Ile	Gln	Ile	Thr	Ile	Leu	Ser	Ser	Glu	Gln
1100						1105					1110			
Met	Lys	Glu	Ile	Phe	Pro	Glu	Gln	Asp	Asp	Gln	Pro	Tyr	Val	Val
1115						1120					1125			
Asp	Lys	Leu	Ala	Glu	Pro	Gln	Lys	Glu	Glu	Pro	Ile	Thr	Glu	Val
1130						1135					1140			
Val	Ser	Gln	Cys	Asp	Leu	Gln	Ala	Pro	Ala	Ala	Gly	Gln	Ser	Arg
1145						1150					1155			
Asp	Ser	Val	Ile	Leu	Asp	Ser	Glu	Lys	Asp	Asp	Ile	His	Cys	Cys
1160						1165					1170			
Ala	Leu	Gly	Trp	Leu	Ser	Met	Val	Tyr	Glu	Gly	Val	Pro	Gln	Cys
1175						1180					1185			
Gln	Cys	Asn	Ser	Ile	Lys	Asn	Ser	Ser	Ser	Glu	Glu	Glu	Lys	Gln
1190						1195					1200			
Lys	Glu	Gln	Cys	Ser	Pro	Leu	Asp	Thr	Asn	Ser	Cys	Lys	Gln	Gly
1205						1210					1215			
Glu	Arg	Thr	Ser	Asp	Arg	Asp	Val	Thr	Val	Val	Gln	Phe	Lys	Ser
1220						1225					1230			

ES 2 799 735 T3

Leu Val Asn Asn Pro Lys Thr Pro Pro Asp Gly Lys Ser His Phe  
 1235 1240 1245  
  
 Pro Glu Leu Gln Asp Asp Ser Arg Lys Asp Thr Pro Lys Thr Lys  
 1250 1255 1260  
  
 His Lys Ser Leu Pro Arg Thr Glu Gln Glu Leu Val Ala Gly Gln  
 1265 1270 1275  
  
 Phe Ser Ser Lys Cys Asp Lys Leu Asn Pro Leu Gln Asn His Lys  
 1280 1285 1290  
  
 Arg Lys Lys Leu Arg Phe His Glu Val Thr Phe His Ser Ser Asn  
 1295 1300 1305  
  
 Lys Met Thr Ala Ser Tyr Glu Gln Ala Ser Gln Glu Thr Arg Gln  
 1310 1315 1320  
  
 Lys Lys His Val Thr Gln Asn Ser Arg Pro Leu Lys Thr Lys Thr  
 1325 1330 1335  
  
 Ala Phe Leu Pro Asn Lys Asp Val Tyr Lys Lys His Ser Ser Leu  
 1340 1345 1350  
  
 Gly Gln Ser Leu Ser Pro Glu Lys Ile Lys Leu Lys Leu Lys Ser  
 1355 1360 1365  
  
 Val Ser Phe Lys Gln Lys Arg Lys Leu Asp Gln Gly Asn Val Leu  
 1370 1375 1380  
  
 Asp Met Glu Val Lys Lys Lys Lys His Asp Lys Gln Glu Gln Lys  
 1385 1390 1395  
  
 Gly Ser Val Gly Ala Thr Phe Lys Leu Gly Asp Ser Leu Ser Asn  
 1400 1405 1410  
  
 Pro Asn Glu Arg Ala Ile Val Lys Glu Lys Met Val Ser Asn Thr  
 1415 1420 1425  
  
 Lys Ser Val Asp Thr Lys Ala Ser Ser Ser Lys Phe Ser Arg Ile  
 1430 1435 1440  
  
 Leu Thr Pro Lys Glu Tyr Leu Gln Arg Gln Lys His Lys Glu Ala  
 1445 1450 1455  
  
 Leu Ser Asn Lys Ala Ser Lys Lys Ile Cys Val Lys Asn Val Pro  
 1460 1465 1470

ES 2 799 735 T3

Cys Asp Ser Glu His Met Arg Pro Ser Lys Leu Ala Val Gln Val  
 1475 1480 1485  
  
 Glu Ser Cys Gly Lys Ser Asn Glu Lys His Ser Ser Gly Val Gln  
 1490 1495 1500  
  
 Thr Ser Lys Glu Ser Leu Asn Gly Leu Thr Ser His Gly Lys Asn  
 1505 1510 1515  
  
 Leu Lys Ile His His Ser Gln Glu Ser Lys Thr Tyr Asn Ile Leu  
 1520 1525 1530  
  
 Arg Asn Val Lys Glu Lys Val Gly Gly Lys Gln Pro Asp Lys Ile  
 1535 1540 1545  
  
 Trp Ile Asp Lys Thr Lys Leu Asp Lys Leu Thr Asn Ile Ser Asn  
 1550 1555 1560  
  
 Glu Ala Gln Phe Ser Gln Met Pro Pro Gln Val Lys Asp Gln Lys  
 1565 1570 1575  
  
 Lys Leu Tyr Leu Asn Arg Val Gly Phe Lys Cys Thr Glu Arg Glu  
 1580 1585 1590  
  
 Ser Ile Ser Leu Thr Lys Leu Glu Ser Ser Pro Arg Lys Leu His  
 1595 1600 1605  
  
 Lys Asp Lys Arg Gln Glu Asn Lys His Lys Thr Phe Leu Pro Val  
 1610 1615 1620  
  
 Lys Gly Asn Thr Glu Lys Ser Asn Met Leu Glu Phe Lys Leu Cys  
 1625 1630 1635  
  
 Pro Asp Ile Leu Leu Lys Asn Thr Asn Ser Val Glu Glu Arg Lys  
 1640 1645 1650  
  
 Asp Val Lys Pro His Pro Arg Lys Glu Gln Ala Pro Leu Gln Val  
 1655 1660 1665  
  
 Ser Gly Ile Lys Ser Thr Lys Glu Asp Trp Leu Lys Phe Val Ala  
 1670 1675 1680  
  
 Thr Lys Lys Arg Thr Gln Lys Asp Ser Gln Glu Arg Asp Asn Val  
 1685 1690 1695  
  
 Asn Ser Arg Leu Ser Lys Arg Ser Phe Ser Ala Asp Gly Phe Glu  
 1700 1705 1710

ES 2 799 735 T3

Met Leu Gln Asn Pro Val Lys Asp Ser Lys Glu Met Phe Gln Thr  
 1715 1720 1725

Tyr Lys Gln Met Tyr Leu Glu Lys Arg Ser Arg Ser Leu Gly Ser  
 1730 1735 1740

Ser Pro Val Lys  
 1745

<210> 25  
 <211> 4644  
 <212> ADN  
 <213> Cricetulus griseus

5

<400> 25

atgaattgga atgcaaaacc agagaatgct gccccaaacc caccatattc taaaagccag 60  
 tcgtctcttt tgcagcagtt tttaatgcoct tccacaactt ctcaaagttc tttcagctgt 120  
 ctcccacata accaagaagc atgcatatat ccactaatt caaattcagt ttcacagcca 180  
 cttctgaacg tcaggagttt cataaatcct ccgatctctg tttctaagt gcataatag 240  
 acagttgtgg cctcacagac ctcaagtaga agagtcacat atacaaatgt taaaggagcc 300  
 caacaaccaa accacaattt gcaaacagtg tcttctggag ttgtgcaaaa tgcctggatg 360  
 aattcaacaa tgaggaattt tatgccttct cttacagagg caacatata tcataaacct 420  
 gatggtgggc ctagtatgcc atatatgcat gcaccacaga gtcatcttgt cacatcagac 480  
 acctactctg tgcaactaca gatgactcct tcaaactctg taagaggccc tgtaacttac 540  
 caaggaaatt atcaaggaaa tccgggactt aaccactcga tggcaggtga gcttggctgg 600  
 gtacaatgtg catccagtga acttacttat ccagattaca gaccacctcc aaagcaatat 660  
 ccttatttac cacaagctt tgtgcaagac acttctgttc agaaacaaaa ctttgtgtca 720  
 tctacatcat tacaagttaa aaataatcag cttccacctt ctacacagac cttaccatca 780  
 aagcgccttg tacctgtgtc gtcatatcag tatgctgcag aaaccagcaa aagactccct 840  
 cccccctt acagctgtag atatggaagc caacatgtgc aaaattctca gtctgtttct 900  
 agacacttgc ctgtggaagt tcctcagagt tcagaaatgc actcgtctga aaaaaagaaa 960  
 gatgcttaca aagtctttca acagcagtg cagagcacta gtaaaaatgt cagtacaata 1020  
 ggaaaattct gtgagttgaa aattaatata aaacagtctt acaatgactc tgctggctct 1080  
 tctggggatg gtgttcatac tcttgttcaa aataatcaag aagaaagaaa gtattcttat 1140  
 aatccaagta caaatcaaat actagacaca aatgtcacia aagaaaagct ggtgagggat 1200  
 attaaatcac tagtagaaat taaaaagaaa ttttcagaac ttgcaagaaa aattaaaatc 1260  
 aacaaaaagc ttttgatggc agctggttgc agtaaaacag ctaatacttc ttatactgaa 1320

ES 2 799 735 T3

ccaactcggc attctgaatt ttcagcaaaa gaaatgtctg ctaaaagggg caatcagtgc 1380  
 tccatggaat tgctagcaac atgcctttct ctttgaaaa accaacctcc aaaaaccaca 1440  
 gaagaaaatg tttcaaaacc tttagaagaa aaacaatata atgcatcaag aactagtaca 1500  
 acagcggttg gcccttcaaa tcccatgaat gaagttcatg tgaagaattt ttgttcaggt 1560  
 gttagaaatt ctcagaaaat aaccacctcg tcacaaacag tcttgtcagt tctcacacca 1620  
 gtttacgatt cttcagatgt agctgttggg aaaggaacag agcttcagat tgctgtgggt 1680  
 tcacctttaa ttctttcaga tgtcagtaact gtacctggga aagagttagc tcctgaagtc 1740  
 gtatctgaaa ctgtatatcc agttgtgaag gaaggcagt tttgtagctt acaaaaccag 1800  
 caggcagaaa atgcaacagt aactgctggt ttgccctttg atgttatcag agcagtagca 1860  
 agtgctactg tatcagctga gctatcactg cctgggcata aagaaaagca gcacaaacca 1920  
 acacagagtg atctagacat cgctgatggc agcctaggga aacactctcc ccagggtgct 1980  
 gaagctttgc ctaaccctag ggacagcacc attgtgagtg ggcctatatt acagattgaa 2040  
 agtatctggt ctcttgcaga aggtgatgta tcttacaatt cccaaatagc agagatattc 2100  
 aactctgtac aaaatgagcc ccagaaacct tcacctgatc agcaagtaat taatagtcaa 2160  
 caagaagaac aagtagataa ggttgctgaa aataaagact taagttttct gaaagacaag 2220  
 tgtatgcagt gtacagatgt tcctcatgaa gtcactgaac agccagagcc actgcagcct 2280  
 ttagagacaa catctgatga gtatgttgaa gcaaacggag aaatcctaga ggaaagcagt 2340  
 aaggagaatc ctggtgaaaa agagatgact aaggacatat tgtgttcacc agctgctggt 2400  
 cagcaagatc ctcaacctca ggaaattgac acagccagca gtaagtcagg acacagtttt 2460  
 tctacagtaa atgagattaa tgatgaaaat gaacctgtct catacctaca tgaccagctg 2520  
 ttagaacttc taaaagagtt tccttatggc attgaaacta ttgccaggcc tgaagtttat 2580  
 gtgggccaac aaaagacaca tgaaatctta gaaatcaaa ctggtagtaa aactggtaat 2640  
 gtgtctgggg ataacacaga ccaaataaaa attacagtat taaactcaga acaaatcaaa 2700  
 gaattatttc ctgaagagga tcagccatgt gatgtagaca aattggcaga acccgagaat 2760  
 acaaaaatca ttgcagaagt aaagagcctg tgtgattcac aggtccccag agaagaaagt 2820  
 cacaaccctg gaatgttggg tctggagaaa gataaaatcc attgctgtgc cttgggctgg 2880  
 ctctcaatgg tttatgaagg tgtgccacag tgtcagtgca gttccatgga agagaaagag 2940  
 aaagaccagt gttctttgga aatctctaata tgcaacaag gagagcaggc ctgcaatagt 3000  
 ggaatcacta tttttgaaat taatcctatt tctaataact caaaaagtcc tctgatccaa 3060  
 gaatctgaga aaggccattt ttctgacata catggtgaaa agataaaaac atctgaaaca 3120  
 5 aaaaacagca gctcaccaag ggtagaacag gaattaactg gtcatttttc aatgaaatgt 3180

ES 2 799 735 T3

taccagaaag ataatctac aacaaaacag gatagctcac tgaaaacaga gcaaaaaata 3240  
 aaaaatcttt cttctaaatg tgacaaacca aatcccttaa aaagcagtaa aataccaacc 3300  
 cctgaaacat ttaatgtggt aacttccaac tctgataaaa atatgccagc attttctaaa 3360  
 caagattctc agggaagcct gcagaagaaa cacctattcc aagactcaga tccagtataaa 3420  
 ggacatgtat ggcttttgcc aaataaagat ccacgcagga ggaatacctt tttagtacag 3480  
 tcagtatcac cagaaaagaa aaagttaaaa ttcaaactcg gtagctccaa actgaaatat 3540  
 tttgaaaaaa gaaaaatgga ccatttgctt atctcagatg tggaaataaa aaagaagaaa 3600  
 tacgaaaaac aagagcagaa caaaaatgct ggaggcacac tcaaattatg tagtactctg 3660  
 actgaaccaa atgaaagagc ctgtgctaaa gaaaagatag tgacaaattc tgagccctca 3720  
 gactcaaagg gaagctcctc taagagtact agagttataa ctgtgcagga atatttacag 3780  
 cggaaaaaag acaaacatgt aataggaaat aatgcctcca aaaacatctg tgtagaaaat 3840  
 gtgccatgtg actctgaacc catgaagtcc agtaaacatt ctgcatcacc tagtttgga 3900  
 aaattaattg agggccaggg tgtcagtgca gagactttaa aagaagtaga acataattcc 3960  
 accagccatg gcaaaaatct caagaccac cgttctgagg agactaggcc atacagtgtg 4020  
 tcaaatagta aagagaaatt ttataggaca catccagaca aatcttacat tgataaagct 4080  
 aaattagaaa gattgaccag tatgagtagt aagtccagcc agctccaggt aaaggaaaaa 4140  
 aggaaacagt acctgaatcg agttgcattc aaatgcacag aacaggaaag catttgtctc 4200  
 accaaattgg acagtgcac caagaagctt agtaaagaga aagaaaagag tacagcatgt 4260  
 gcacccatga caaaagacta cacacacaag cccatggtgg agtttaaaatt atgtccagat 4320  
 gtgctattga agaatacaag ctccattgac aaaggggatg atccaaggcc tgggcctgag 4380  
 aaggagcgag cacctgtgca agtttcagga ataaaaacta caaaagaaga ctggttaaaa 4440  
 tgtatcccaa caaggacaaa gatgcccga tcaagtgaac aacagatcg ggctgactca 4500  
 agactctcta agagaagctt cagtgcagat gaatttgaaa ctctacaaa cccagtataaa 4560  
 gactcaaatg tcatgttccg gactttcaaa aagatgtacc tggagaagag aagcaggagc 4620  
 ctggggagca gtccagtgaa gtag 4644

<210> 26  
 <211> 112  
 <212> ADN  
 <213> *Cricetulus griseus*  
 <400> 26

5

tgctgggatt taaggggaaa gctttaataa aagatcttta tttgtatttc ttgcagattt 60  
 gtgacattca aaaccacaga ctatgcaaca ctactactaa accaggtcaa at 112

<210> 27

ES 2 799 735 T3

<211> 36  
<212> ADN  
<213> *Cricetulus griseus*

5 <400> 27  
cttcgggata gagtggtttt gctttacca ccagga 36

<210> 28  
<211> 221  
<212> PRT  
<213> *Homo sapiens*

10 <400> 28

ES 2 799 735 T3

Met Phe Gly Phe His Lys Pro Lys Met Tyr Arg Ser Ile Glu Gly Cys  
 1 5 10 15

Cys Ile Cys Arg Ala Lys Ser Ser Ser Ser Arg Phe Thr Asp Ser Lys  
 20 25 30

Arg Tyr Glu Lys Asp Phe Gln Ser Cys Phe Gly Leu His Glu Thr Arg  
 35 40 45

Ser Gly Asp Ile Cys Asn Ala Cys Val Leu Leu Val Lys Arg Trp Lys  
 50 55 60

Lys Leu Pro Ala Gly Ser Lys Lys Asn Trp Asn His Val Val Asp Ala  
 65 70 75 80

Arg Ala Gly Pro Ser Leu Lys Thr Thr Leu Lys Pro Lys Lys Val Lys  
 85 90 95

Thr Leu Ser Gly Asn Arg Ile Lys Ser Asn Gln Ile Ser Lys Leu Gln  
 100 105 110

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

Ala Ser Pro Ala Gln Ser Pro Cys Tyr Ser Asn Gln Ser Asp Asp Gly  
 130 135 140

Ser Asp Thr Glu Met Ala Ser Gly Ser Asn Arg Thr Pro Val Phe Ser  
 145 150 155 160

Phe Leu Asp Leu Thr Tyr Trp Lys Arg Gln Lys Ile Cys Cys Gly Ile  
 165 170 175

Ile Tyr Lys Gly Arg Phe Gly Glu Val Leu Ile Asp Thr His Leu Phe  
 180 185 190

Lys Pro Cys Cys Ser Asn Lys Lys Ala Ala Ala Glu Lys Pro Glu Glu

195

200

205

Gln Gly Pro Glu Pro Leu Pro Ile Ser Thr Gln Glu Trp  
 210 215 220

<210> 29  
 <211> 221  
 <212> PRT  
 <213> Mus musculus  
 <400> 29

5

ES 2 799 735 T3

Met Phe Gly Phe His Lys Pro Lys Met Tyr Arg Ser Ile Glu Gly Cys  
 1 5 10 15

Cys Ile Cys Arg Ala Lys Ser Ser Ser Ser Arg Phe Thr Asp Ser Lys  
 20 25 30

Arg Tyr Glu Lys Asp Phe Gln Ser Cys Phe Gly Leu His Glu Thr Arg  
 35 40 45

Ser Gly Asp Ile Cys Asn Ala Cys Val Leu Leu Val Lys Arg Trp Lys  
 50 55 60

Lys Leu Pro Ala Gly Ser Lys Lys Asn Trp Asn His Val Val Asp Ala  
 65 70 75 80

Arg Ala Gly Pro Ser Leu Lys Thr Thr Leu Lys Pro Lys Lys Val Lys  
 85 90 95

Thr Leu Ser Gly Asn Arg Met Lys Ser Asn Gln Ile Ser Lys Leu Gln  
 100 105 110

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

Ala Ser Pro Ala Gln Ser Pro Cys Tyr Ser Asn Gln Ser Asp Glu Gly  
 130 135 140

Ser Asp Thr Glu Met Ala Ser Ser Ser Asn Arg Thr Pro Val Phe Ser  
 145 150 155 160

Phe Leu Asp Leu Thr Tyr Trp Lys Arg Gln Lys Ile Cys Cys Gly Ile  
 165 170 175

Ile Tyr Lys Gly Arg Phe Gly Glu Val Leu Ile Asp Thr His Leu Phe  
 180 185 190

Lys Pro Cys Cys Ser Ser Lys Lys Ala Ala Ala Glu Lys Pro Glu Glu  
 195 200 205

Gln Gly Pro Ala Pro Leu Pro Ile Ser Thr Gln Glu Trp  
 210 215 220

<210> 30  
 <211> 221  
 <212> PRT  
 <213> Cricetulus griseus  
 <400> 30

5

ES 2 799 735 T3

Met Phe Gly Phe His Lys Pro Lys Met Tyr Arg Ser Ile Glu Gly Cys  
 1 5 10 15

Cys Ile Cys Arg Ala Lys Ser Ser Ser Ser Arg Phe Thr Asp Ser Lys  
 20 25 30

Arg Tyr Glu Lys Asp Phe Gln Ser Cys Phe Gly Leu His Glu Thr Arg  
 35 40 45

Ser Gly Asp Ile Cys Asn Ala Cys Val Leu Leu Val Lys Arg Trp Lys  
 50 55 60

Lys Leu Pro Ala Gly Ser Lys Lys Asn Trp Asn His Val Ser His Ser  
 65 70 75 80

Arg Ala Gly Pro Ser Leu Lys Thr Thr Leu Lys Pro Lys Lys Val Lys  
 85 90 95

Thr Leu Ser Gly Asn Arg Met Lys Ser Asn Gln Ile Ser Lys Leu Gln  
 100 105 110

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

Ala Ser Pro Ala Gln Ser Pro Cys Tyr Ser Asn Gln Ser Asp Asp Gly  
 130 135 140

Ser Asp Thr Glu Met Ala Ser Ser Ser Asn Arg Thr Pro Val Phe Ser  
 145 150 155 160

Phe Leu Asp Leu Thr Tyr Trp Lys Arg Gln Lys Ile Cys Cys Gly Ile  
 165 170 175

Ile Tyr Lys Gly Arg Phe Gly Glu Val Leu Ile Asp Thr His Leu Phe  
 180 185 190

Lys Pro Cys Cys Ser Ser Lys Lys Ala Ala Pro Glu Lys Pro Glu Glu  
 195 200 205

Gln Gly Pro Ala Pro Leu Pro Ile Ser Thr Gln Glu Trp  
 210 215 220

<210> 31  
 <211> 1236  
 <212> ADN  
 <213> Cricetulus griseus

5

<400> 31

ES 2 799 735 T3

ctcagaagaa aagatgtttg gttttcacia gccaaagatg taccgaagta tagagggctg 60  
ctgtatctgc agagccaagt cctccagctc tcggttcacg gacagtaaac gttatgaaaa 120  
ggacttccag agctgttttg ggttgcacga gactcgctca ggagatatct gcaatgcctg 180  
tgtgctgctt gtgaaaagat ggaagaagtt gccagcagga tcaaaaaaaaa actggaatca 240  
tgtgtcacac tcaagggcag gacccagtct aaagacaaca ttgaaaccaa agaaagtgaa 300  
aactctatct ggaaacagga tgaaaagcaa ccagatcagt aaactgcaga aggagttaa 360  
acgccacaac tctgatgctc acagtaccac ctcaagtgcc tcgccagccc agtctccctg 420  
ctacagtaac cagtcagatg atggctcaga cacagagatg gcttccagct ctaacagaac 480  
tccagttttt tccttcttag atcttaccta ctggaaaaga cagaaaatat gttgtgggat 540  
catctataag ggccgttttg gggaagtcct catcgacacg catctcttca agccttgctg 600  
cagcagtaag aaggcagctc ctgagaagcc tgaggaacag ggaccagcgc ctctgcccatt 660  
ctctactcag gagtggtgac tgaggttcat gcagaaggga acaaagagca atttaaactt 720  
tgaaaagacc acaaagcaac agactgacct tcctatTTTT aacttggata cctgctattc 780  
tgccaaaaga cattttctag aatagttttt aatgggttac ccatcccccc atccaacaaa 840  
ctcggaagcc agttctagct tactgcaaga agagagtgta cataatattt aatatgctga 900  
gtatttcata ggaaggctga atgctgctgt aaagtgctct ttaagtcttt tttttttttt 960  
aatcccctct aatgaatgag attagggggg tttcagggga cagagatggg atttgttgtg 1020  
tgataaacca tatgtagttt agtctttctg tggagaggca gtggttgggg cattttaaatt 1080  
ggctggctac acttgttttc ccctcatggt aatttgctcat aactcagtag cacgacctgc 1140  
ccctagaagt agttaaagat ttttaaagtc taaggcgttg ccaaggttct gatgattcag 1200  
acctgtacta ctgattatta agcaggacag actgag 1236

<210> 32  
<211> 855  
<212> PRT  
<213> Cricetulus griseus  
<400> 32

5

Met Gly Ser Asn Arg Gly Arg Lys Ala Gly Gly Ser Ser Lys Asp Phe  
1 5 10 15

ES 2 799 735 T3

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

ES 2 799 735 T3

Asn Asp Leu Val Thr Val Tyr Asp Thr Leu Ser Pro Met Glu Pro His  
 275 280 285

Ala Val Val Arg Leu Cys Gly Thr Tyr Pro Pro Ser Tyr Asn Leu Thr  
 290 295 300

Phe Leu Ser Ser Gln Asn Val Phe Leu Val Thr Leu Ile Thr Asn Thr  
 305 310 315 320

Asp Arg Arg His Pro Gly Phe Glu Ala Thr Phe Phe Gln Leu Pro Lys  
 325 330 335

Met Arg Ser Cys Gly Gly Ser Leu Ser Glu Ala Gln Gly Leu Phe Ser  
 340 345 350

Ser Pro Tyr Tyr Pro Gly His Tyr Pro Pro Asn Ile Asp Cys Thr Trp  
 355 360 365

Asn Ile Lys Val Pro Asn Asn Arg Asn Val Lys Val Arg Phe Lys Leu  
 370 375 380

Phe Tyr Leu Val Asp Pro Asn Ile Pro Leu Gly Thr Cys Pro Lys Asp  
 385 390 395 400

Tyr Val Glu Ile Asn Gly Glu Arg Tyr Cys Gly Glu Lys Ser Gln Phe  
 405 410 415

Val Val Ser Ser Asn Ser Ser Lys Ile Thr Val Arg Phe His Ser Asp  
 420 425 430

His Ser Tyr Thr Asp Thr Gly Phe Leu Ala Glu Tyr Leu Ser Tyr Asp  
 435 440 445

Ser Asn Asp Pro Cys Pro Gly Met Phe Met Cys Asn Thr Gly Arg Cys  
 450 455 460

Ile Arg Lys Asp Leu Arg Cys Asp Gly Trp Ala Asp Cys Pro Asp Tyr  
 465 470 475 480

Ser Asp Glu His Phe Cys Arg Cys Asn Thr Thr His Gln Phe Met Cys  
 485 490 495

Lys Asn Lys Leu Cys Lys Pro Leu Phe Trp Val Cys Asp Asn Ile Asn  
 500 505 510

Asp Cys Gly Asp Gly Ser Asp Glu Glu Gly Cys Ser Cys Pro Ala Glu

ES 2 799 735 T3

	515		520		525														
Thr	Phe	Lys	Cys	Ser	Asn	Gly	Lys	Cys	Leu	Pro	Gln	Ser	Gln	Lys	Cys				
	530					535					540								
Asp	Gly	Lys	Asp	Asn	Cys	Gly	Asp	Gly	Ser	Asp	Glu	Ala	Ser	Cys	Asp				
545					550					555					560				
Arg	Val	Lys	Val	Val	Ser	Cys	Thr	Lys	Tyr	Thr	Tyr	Arg	Cys	His	Asn				
				565					570					575					
Gly	Leu	Cys	Leu	Ser	Lys	Gly	Asn	Pro	Glu	Cys	Asp	Gly	Lys	Lys	Asp				
			580					585					590						
Cys	Ser	Asp	Gly	Ser	Asp	Glu	Lys	Asn	Cys	Asp	Cys	Gly	Leu	Arg	Ser				
		595					600					605							
Phe	Thr	Lys	Gln	Ala	Arg	Val	Val	Gly	Gly	Thr	Asn	Ala	Asp	Glu	Gly				
	610					615					620								
Glu	Trp	Pro	Trp	Gln	Val	Ser	Leu	His	Ala	Leu	Gly	Gln	Gly	His	Leu				
625					630					635					640				
Cys	Gly	Ala	Ser	Leu	Ile	Ser	Pro	Asn	Trp	Leu	Val	Ser	Ala	Ala	His				
				645					650					655					
Cys	Phe	Met	Asp	Asp	Arg	Asn	Phe	Lys	Tyr	Ser	Asp	His	Thr	Lys	Trp				
			660					665					670						
Thr	Ala	Phe	Leu	Gly	Leu	Leu	Asp	Gln	Ser	Lys	Arg	Ser	Ser	Thr	Gly				
		675					680						685						
Val	Gln	Glu	His	Lys	Leu	Lys	Arg	Ile	Ile	Thr	His	Pro	Leu	Phe	Asn				
	690					695					700								
Glu	Ile	Thr	Phe	Asp	Tyr	Asp	Ile	Ala	Leu	Leu	Glu	Leu	Glu	Lys	Pro				
705					710					715					720				
Ala	Glu	Tyr	Ser	Thr	Val	Val	Arg	Pro	Ile	Cys	Leu	Pro	Asp	Thr	Thr				
				725					730					735					
His	Val	Phe	Pro	Ala	Gly	Lys	Ala	Ile	Trp	Val	Thr	Gly	Trp	Gly	His				
			740					745					750						
Thr	Gln	Glu	Gly	Gly	Thr	Gly	Ala	Leu	Ile	Leu	Gln	Lys	Gly	Glu	Ile				
		755					760					765							

ES 2 799 735 T3

Arg Val Ile Asn Gln Thr Thr Cys Glu Asp Leu Met Pro Gln Gln Ile  
770 775 780

Thr Pro Arg Met Met Cys Val Gly Phe Leu Ser Gly Gly Val Asp Ser  
785 790 795 800

Cys Gln Gly Asp Ser Gly Gly Pro Leu Ser Ser Val Glu Thr Glu Gly  
805 810 815

Arg Ile Phe Gln Ala Gly Val Val Ser Trp Gly Glu Gly Cys Ala Gln  
820 825 830

Arg Asn Lys Pro Gly Val Tyr Thr Arg Leu Pro Ala Val Arg Asp Trp  
835 840 845

Ile Lys Glu Gln Thr Gly Val  
850 855

<210> 33  
<211> 855  
<212> PRT  
<213> *Homo sapiens*

5

<400> 33

Met Gly Ser Asp Arg Ala Arg Lys Gly Gly Gly Gly Pro Lys Asp Phe  
1 5 10 15

Gly Ala Gly Leu Lys Tyr Asn Ser Arg His Glu Lys Val Asn Gly Leu  
20 25 30

Glu Glu Gly Val Glu Phe Leu Pro Val Asn Asn Val Lys Lys Val Glu  
35 40 45

Lys His Gly Pro Gly Arg Trp Val Val Leu Ala Ala Val Leu Ile Gly  
50 55 60

Leu Leu Leu Val Leu Leu Gly Ile Gly Phe Leu Val Trp His Leu Gln  
65 70 75 80

Tyr Arg Asp Val Arg Val Gln Lys Val Phe Asn Gly Tyr Met Arg Ile  
85 90 95

Thr Asn Glu Asn Phe Val Asp Ala Tyr Glu Asn Ser Asn Ser Thr Glu  
100 105 110

Phe Val Ser Leu Ala Ser Lys Val Lys Asp Ala Leu Lys Leu Leu Tyr  
115 120 125

Ser Gly Val Pro Phe Leu Gly Pro Tyr His Lys Glu Ser Ala Val Thr

ES 2 799 735 T3

130						135						140					
Ala	Phe	Ser	Glu	Gly	Ser	Val	Ile	Ala	Tyr	Tyr	Trp	Ser	Glu	Phe	Ser		
145					150				155						160		
Ile	Pro	Gln	His	Leu	Val	Glu	Glu	Ala	Glu	Arg	Val	Met	Ala	Glu	Glu		
				165					170					175			
Arg	Val	Val	Met	Leu	Pro	Pro	Arg	Ala	Arg	Ser	Leu	Lys	Ser	Phe	Val		
			180					185					190				
Val	Thr	Ser	Val	Val	Ala	Phe	Pro	Thr	Asp	Ser	Lys	Thr	Val	Gln	Arg		
		195					200					205					
Thr	Gln	Asp	Asn	Ser	Cys	Ser	Phe	Gly	Leu	His	Ala	Arg	Gly	Val	Glu		
210						215					220						
Leu	Met	Arg	Phe	Thr	Thr	Pro	Gly	Phe	Pro	Asp	Ser	Pro	Tyr	Pro	Ala		
225					230					235					240		
His	Ala	Arg	Cys	Gln	Trp	Ala	Leu	Arg	Gly	Asp	Ala	Asp	Ser	Val	Leu		
				245					250					255			
Ser	Leu	Thr	Phe	Arg	Ser	Phe	Asp	Leu	Ala	Ser	Cys	Asp	Glu	Arg	Gly		
			260					265					270				
Ser	Asp	Leu	Val	Thr	Val	Tyr	Asn	Thr	Leu	Ser	Pro	Met	Glu	Pro	His		
		275					280					285					
Ala	Leu	Val	Gln	Leu	Cys	Gly	Thr	Tyr	Pro	Pro	Ser	Tyr	Asn	Leu	Thr		
	290					295					300						
Phe	His	Ser	Ser	Gln	Asn	Val	Leu	Leu	Ile	Thr	Leu	Ile	Thr	Asn	Thr		
305					310					315					320		
Glu	Arg	Arg	His	Pro	Gly	Phe	Glu	Ala	Thr	Phe	Phe	Gln	Leu	Pro	Arg		
				325					330					335			
Met	Ser	Ser	Cys	Gly	Gly	Arg	Leu	Arg	Lys	Ala	Gln	Gly	Thr	Phe	Asn		
			340					345					350				
Ser	Pro	Tyr	Tyr	Pro	Gly	His	Tyr	Pro	Pro	Asn	Ile	Asp	Cys	Thr	Trp		
		355					360					365					
Asn	Ile	Glu	Val	Pro	Asn	Asn	Gln	His	Val	Lys	Val	Arg	Phe	Lys	Phe		
370						375					380						

ES 2 799 735 T3

Phe Tyr Leu Leu Glu Pro Gly Val Pro Ala Gly Thr Cys Pro Lys Asp  
 385 390 395 400

Tyr Val Glu Ile Asn Gly Glu Lys Tyr Cys Gly Glu Arg Ser Gln Phe  
 405 410 415

Val Val Thr Ser Asn Ser Asn Lys Ile Thr Val Arg Phe His Ser Asp  
 420 425 430

Gln Ser Tyr Thr Asp Thr Gly Phe Leu Ala Glu Tyr Leu Ser Tyr Asp  
 435 440 445

Ser Ser Asp Pro Cys Pro Gly Gln Phe Thr Cys Arg Thr Gly Arg Cys  
 450 455 460

Ile Arg Lys Glu Leu Arg Cys Asp Gly Trp Ala Asp Cys Thr Asp His  
 465 470 475 480

Ser Asp Glu Leu Asn Cys Ser Cys Asp Ala Gly His Gln Phe Thr Cys  
 485 490 495

Lys Asn Lys Phe Cys Lys Pro Leu Phe Trp Val Cys Asp Ser Val Asn  
 500 505 510

Asp Cys Gly Asp Asn Ser Asp Glu Gln Gly Cys Ser Cys Pro Ala Gln  
 515 520 525

Thr Phe Arg Cys Ser Asn Gly Lys Cys Leu Ser Lys Ser Gln Gln Cys  
 530 535 540

Asn Gly Lys Asp Asp Cys Gly Asp Gly Ser Asp Glu Ala Ser Cys Pro  
 545 550 555 560

Lys Val Asn Val Val Thr Cys Thr Lys His Thr Tyr Arg Cys Leu Asn  
 565 570 575

Gly Leu Cys Leu Ser Lys Gly Asn Pro Glu Cys Asp Gly Lys Glu Asp  
 580 585 590

Cys Ser Asp Gly Ser Asp Glu Lys Asp Cys Asp Cys Gly Leu Arg Ser  
 595 600 605

Phe Thr Arg Gln Ala Arg Val Val Gly Gly Thr Asp Ala Asp Glu Gly  
 610 615 620

Glu Trp Pro Trp Gln Val Ser Leu His Ala Leu Gly Gln Gly His Ile  
 625 630 635 640

ES 2 799 735 T3

Cys Gly Ala Ser Leu Ile Ser Pro Asn Trp Leu Val Ser Ala Ala His  
645 650 655

Cys Tyr Ile Asp Asp Arg Gly Phe Arg Tyr Ser Asp Pro Thr Gln Trp  
660 665 670

Thr Ala Phe Leu Gly Leu His Asp Gln Ser Gln Arg Ser Ala Pro Gly  
675 680 685

Val Gln Glu Arg Arg Leu Lys Arg Ile Ile Ser His Pro Phe Phe Asn  
690 695 700

Asp Phe Thr Phe Asp Tyr Asp Ile Ala Leu Leu Glu Leu Glu Lys Pro  
705 710 715 720

Ala Glu Tyr Ser Ser Met Val Arg Pro Ile Cys Leu Pro Asp Ala Ser  
725 730 735

His Val Phe Pro Ala Gly Lys Ala Ile Trp Val Thr Gly Trp Gly His  
740 745 750

Thr Gln Tyr Gly Gly Thr Gly Ala Leu Ile Leu Gln Lys Gly Glu Ile  
755 760 765

Arg Val Ile Asn Gln Thr Thr Cys Glu Asn Leu Leu Pro Gln Gln Ile  
770 775 780

Thr Pro Arg Met Met Cys Val Gly Phe Leu Ser Gly Gly Val Asp Ser  
785 790 795 800

Cys Gln Gly Asp Ser Gly Gly Pro Leu Ser Ser Val Glu Ala Asp Gly  
805 810 815

Arg Ile Phe Gln Ala Gly Val Val Ser Trp Gly Asp Gly Cys Ala Gln  
820 825 830

Arg Asn Lys Pro Gly Val Tyr Thr Arg Leu Pro Leu Phe Arg Asp Trp  
835 840 845

Ile Lys Glu Asn Thr Gly Val  
850 855

<210> 34  
<211> 855  
<212> PRT  
<213> Mus musculus  
  
<400> 34

ES 2 799 735 T3

Met Gly Ser Asn Arg Gly Arg Lys Ala Gly Gly Gly Ser Gln Asp Phe  
1 5 10 15

Gly Ala Gly Leu Lys Tyr Asn Ser Arg Leu Glu Asn Met Asn Gly Phe  
20 25 30

Glu Glu Gly Val Glu Phe Leu Pro Ala Asn Asn Ala Lys Lys Val Glu  
35 40 45

Lys Arg Gly Pro Arg Arg Trp Val Val Leu Val Ala Val Leu Phe Ser  
50 55 60

Phe Leu Leu Leu Ser Leu Met Ala Gly Leu Leu Val Trp His Phe His  
65 70 75 80

Tyr Arg Asn Val Arg Val Gln Lys Val Phe Asn Gly His Leu Arg Ile  
85 90 95

Thr Asn Glu Ile Phe Leu Asp Ala Tyr Glu Asn Ser Thr Ser Thr Glu  
100 105 110

Phe Ile Ser Leu Ala Ser Gln Val Lys Glu Ala Leu Lys Leu Leu Tyr  
115 120 125

Asn Glu Val Pro Val Leu Gly Pro Tyr His Lys Lys Ser Ala Val Thr  
130 135 140

Ala Phe Ser Glu Gly Ser Val Ile Ala Tyr Tyr Trp Ser Glu Phe Ser  
145 150 155 160

Ile Pro Pro His Leu Ala Glu Glu Val Asp Arg Ala Met Ala Val Glu  
165 170 175

Arg Val Val Thr Leu Pro Pro Arg Ala Arg Ala Leu Lys Ser Phe Val  
180 185 190

Leu Thr Ser Val Val Ala Phe Pro Ile Asp Pro Arg Met Leu Gln Arg  
195 200 205

Thr Gln Asp Asn Ser Cys Ser Phe Ala Leu His Ala His Gly Ala Ala  
210 215 220

Val Thr Arg Phe Thr Thr Pro Gly Phe Pro Asn Ser Pro Tyr Pro Ala  
225 230 235 240

His Ala Arg Cys Gln Trp Val Leu Arg Gly Asp Ala Asp Ser Val Leu  
245 250 255

ES 2 799 735 T3

Ser Leu Thr Phe Arg Ser Phe Asp Val Ala Pro Cys Asp Glu His Gly  
 260 265 270

Ser Asp Leu Val Thr Val Tyr Asp Ser Leu Ser Pro Met Glu Pro His  
 275 280 285

Ala Val Val Arg Leu Cys Gly Thr Phe Ser Pro Ser Tyr Asn Leu Thr  
 290 295 300

Phe Leu Ser Ser Gln Asn Val Phe Leu Val Thr Leu Ile Thr Asn Thr  
 305 310 315

Asp Arg Arg His Pro Gly Phe Glu Ala Thr Phe Phe Gln Leu Pro Lys  
 325 330

Met Ser Ser Cys Gly Gly Phe Leu Ser Asp Thr Gln Gly Thr Phe Ser  
 340 345

Ser Pro Tyr Tyr Pro Gly His Tyr Pro Pro Asn Ile Asn Cys Thr Trp  
 355 360 365

Asn Ile Lys Val Pro Asn Asn Arg Asn Val Lys Val Arg Phe Lys Leu  
 370 375 380

Phe Tyr Leu Val Asp Pro Asn Val Pro Val Gly Ser Cys Thr Lys Asp  
 385 390 395 400

Tyr Val Glu Ile Asn Gly Glu Lys Tyr Cys Gly Glu Arg Ser Gln Phe  
 405 410 415

Val Val Ser Ser Asn Ser Ser Lys Ile Thr Val His Phe His Ser Asp  
 420 425 430

His Ser Tyr Thr Asp Thr Gly Phe Leu Ala Glu Tyr Leu Ser Tyr Asp  
 435 440 445

Ser Asn Asp Pro Cys Pro Gly Met Phe Met Cys Lys Thr Gly Arg Cys  
 450 455 460

Ile Arg Lys Glu Leu Arg Cys Asp Gly Trp Ala Asp Cys Pro Asp Tyr  
 465 470 475 480

Ser Asp Glu Arg Tyr Cys Arg Cys Asn Ala Thr His Gln Phe Thr Cys  
 485 490 495

Lys Asn Gln Phe Cys Lys Pro Leu Phe Trp Val Cys Asp Ser Val Asn  
 500 505 510

ES 2 799 735 T3

Asp Cys Gly Asp Gly Ser Asp Glu Glu Gly Cys Ser Cys Pro Ala Gly  
 515 520 525

Ser Phe Lys Cys Ser Asn Gly Lys Cys Leu Pro Gln Ser Gln Lys Cys  
 530 535 540

Asn Gly Lys Asp Asn Cys Gly Asp Gly Ser Asp Glu Ala Ser Cys Asp  
 545 550 555 560

Ser Val Asn Val Val Ser Cys Thr Lys Tyr Thr Tyr Arg Cys Gln Asn  
 565 570 575

Gly Leu Cys Leu Ser Lys Gly Asn Pro Glu Cys Asp Gly Lys Thr Asp  
 580 585 590

Cys Ser Asp Gly Ser Asp Glu Lys Asn Cys Asp Cys Gly Leu Arg Ser  
 595 600 605

Phe Thr Lys Gln Ala Arg Val Val Gly Gly Thr Asn Ala Asp Glu Gly  
 610 615 620

Glu Trp Pro Trp Gln Val Ser Leu His Ala Leu Gly Gln Gly His Leu  
 625 630 635 640

Cys Gly Ala Ser Leu Ile Ser Pro Asp Trp Leu Val Ser Ala Ala His  
 645 650 655

Cys Phe Gln Asp Asp Lys Asn Phe Lys Tyr Ser Asp Tyr Thr Met Trp  
 660 665 670

Thr Ala Phe Leu Gly Leu Leu Asp Gln Ser Lys Arg Ser Ala Ser Gly  
 675 680 685

Val Gln Glu Leu Lys Leu Lys Arg Ile Ile Thr His Pro Ser Phe Asn  
 690 695 700

Asp Phe Thr Phe Asp Tyr Asp Ile Ala Leu Leu Glu Leu Glu Lys Ser  
 705 710 715 720

Val Glu Tyr Ser Thr Val Val Arg Pro Ile Cys Leu Pro Asp Ala Thr  
 725 730 735

His Val Phe Pro Ala Gly Lys Ala Ile Trp Val Thr Gly Trp Gly His  
 740 745 750

Thr Lys Glu Gly Gly Thr Gly Ala Leu Ile Leu Gln Lys Gly Glu Ile



ES 2 799 735 T3

Met Asn Trp Asn Thr Lys Gln Glu Asn Val Pro Lys Pro Pro Pro Tyr  
 1 5 10 15

Ser Lys Thr Gln Ser Ser Ile Leu Gln His Phe Leu Met Thr Ser Thr  
 20 25 30

Thr Ser Gln Ser Ser Phe Asn Tyr Ser Pro His Asn Gln Glu Ala Ser  
 35 40 45

Gln Thr Ser Phe Asn Tyr Ser Leu His Asn Gln Glu Ala Cys Met Tyr  
 50 55 60

Ser Gly Asn Ser Asn Ser Val Ser Gln Pro Leu Leu Ser Gly Arg Asn  
 65 70 75 80

Tyr Ile Thr Pro Gln Thr Gln Ile Ser Val Ser Asn Met Pro Thr Arg  
 85 90 95

Thr Ile Val Ala Ser Gln Ser Ser Met Glu Arg Val Val Ser Thr Asn  
 100 105 110

Gly Lys Gly Pro Gln Gln Pro Asn His Asn Leu Gln Thr Val Ser Ser  
 115 120 125

ES 2 799 735 T3

Gly Ile Met Gln Asn Val Trp Leu Pro Ser His Thr Glu Ala Thr Ile  
 130 135 140

Ser His Asn Pro Asp Gly Gly Thr Asn Met Pro Tyr Met His Pro Pro  
 145 150 155 160

Gln Asn Gln Leu Val Thr Ser Asp Thr Tyr Ser Met Gln Leu Gln Met  
 165 170 175

Ala Pro Leu His Ser Gly Lys Val Pro Met Thr His Gln Gly Ser Gln  
 180 185 190

Gly Leu Asn His Phe Ile Pro Asp Gln Leu Val Asp Trp Thr Gln Tyr  
 195 200 205

Thr Ser Asn Glu Leu Ser Tyr Pro Glu Tyr Arg Pro Pro Pro Lys Gln  
 210 215 220

Tyr Ser Tyr Ile Leu Pro Ala Thr Thr Ser Leu Gln Val Lys Asn Asn  
 225 230 235 240

Gln Leu Pro Thr Tyr Thr Gln Ser Leu Gln Ser Lys His Ser Val Pro  
 245 250 255

Leu Ser Ser His Gln Tyr Ala Ala Glu Ala Ser Lys Arg Leu Ser Ala  
 260 265 270

Leu Pro Tyr Ser Cys Arg Tyr Glu Asn Gln His Val Gln Asn Ala Gln  
 275 280 285

Pro Val Ser Lys His Leu Pro Met Glu Val Pro Gln Ser Ser Glu Val  
 290 295 300

His Ser Ser Glu Lys Lys Lys Asp Thr Tyr Arg Gly Phe Lys Gln Gln  
 305 310 315 320

Trp Gln Asn Pro Asn Glu Lys Val Ser Ile Gly Gln Phe Ser Glu Val  
 325 330 335

Lys Ile Asn Ile Lys Gln Pro Tyr Ser Glu Ser Val Arg Pro Ser Gly  
 340 345 350

Asp Gly Val Gln Ala Leu Val Gln Asn Asn Gln Glu Lys Arg Lys Tyr  
 355 360 365

Thr Tyr Asn Pro Asn Thr Asn Gln Val Ile Asp Thr Asn Ala Thr Lys

ES 2 799 735 T3

370						375						380					
Glu	Lys	Leu	Val	Arg	Asp	Ile	Lys	Ser	Leu	Val	Glu	Ile	Lys	Lys	Lys		
385					390					395					400		
Phe	Ser	Glu	Leu	Ala	Arg	Lys	Ile	Lys	Ile	Asn	Lys	Ser	Leu	Leu	Met		
				405					410					415			
Ala	Ala	Gly	Cys	Ser	Lys	Thr	Ala	Asn	Thr	Ser	Tyr	Thr	Glu	Pro	Ile		
			420					425					430				
Gln	His	Ser	Glu	Phe	Ser	Ala	Lys	Glu	Met	Ser	Ala	Lys	Asn	Gly	Asn		
		435					440					445					
Asp	Cys	Ser	Met	Glu	Leu	Leu	Ala	Thr	Cys	Leu	Ser	Leu	Trp	Lys	Asn		
	450					455						460					
Gln	Pro	Ser	Lys	Thr	Thr	Glu	Glu	Asn	Val	Pro	Lys	Pro	Leu	Glu	Glu		
465					470					475					480		
Lys	Gln	Cys	Asn	Thr	Ser	Arg	Ile	Ser	Thr	Thr	Val	Val	Gly	Ser	Ala		
				485					490					495			
Asn	Pro	Thr	Asn	Glu	Val	His	Val	Lys	Ser	Leu	Cys	Ser	Gly	Val	Gly		
			500					505					510				
Asn	Ser	Gln	Lys	Met	Met	Ser	Ser	Ser	Gln	Thr	Val	Leu	Pro	Val	Leu		
		515					520					525					
Ile	Pro	Ser	Cys	Glu	Ser	Ser	Gly	Val	Ala	Val	Gly	Lys	Gly	Thr	Glu		
	530					535					540						
Leu	Gln	Ile	Ala	Val	Val	Ser	Pro	Leu	Val	Leu	Ser	Asp	Thr	Asn	Thr		
545					550					555					560		
Leu	Pro	Gly	Lys	Asp	Ser	Val	Pro	Glu	Val	Leu	Pro	Glu	Thr	Leu	Tyr		
				565					570					575			
Pro	Val	Val	Lys	Glu	Gly	Ser	Val	Cys	Ser	Leu	Gln	Thr	Gln	Pro	Thr		
			580					585					590				
Glu	Thr	Val	Ala	Leu	Pro	Phe	Asp	Val	Ile	Gly	Ala	Val	Ala	Ser	Asn		
		595					600						605				
Asn	Ile	Ser	Ala	Glu	Ile	Pro	Leu	Pro	Val	Asp	Lys	Glu	Lys	Gln	His		
	610					615					620						

ES 2 799 735 T3

Lys Pro Ile Gln Gly Asp Pro Asp Ile Ala Asp Ser Ser Leu Gly Lys  
 625 630 635 640

His Ser Pro Leu Gly Thr Glu Val Leu Pro Lys Pro Met Asp Ser Thr  
 645 650 655

Ile Val Ser Gly Pro Met Leu Gln Ile Glu Ser Ile Cys Ser Leu Ala  
 660 665 670

Glu Gly Asp Val Ser Tyr Asn Ser Gln Ile Ala Glu Ile Phe Asn Ser  
 675 680 685

Val Gln Thr Glu Pro Gln Lys Pro Ser Pro Asn Gln Val Ile Asp Ser  
 690 695 700

Gln Gln Glu Gln Val Tyr Asp Thr Thr Glu Asn Lys Asp Phe Ser Leu  
 705 710 715 720

Gln Lys Asp Lys Cys Val Gln Cys Thr Asp Val Pro His Glu Val Pro  
 725 730 735

Glu Gln Pro Glu Pro Leu Gln Pro Glu Glu Pro Ala Ser Ser Glu Tyr  
 740 745 750

Val Glu Ala Asn Arg Glu Ala Thr Glu Glu Ser Cys Arg Glu Tyr Thr  
 755 760 765

Gly Arg Lys Glu Ser Thr Ala Lys Asp Val Cys Leu Pro Ala Ala Ile  
 770 775 780

Gln Gln Asp Pro His Pro Arg Glu Thr Asp Met Phe Ser Lys Ser Asp  
 785 790 795 800

His Ser Leu Pro Ala Ile Asn Glu Ile Asn Asp Glu Ser Glu Pro Ile  
 805 810 815

Ser Tyr Leu His Asp Gln Leu Ser Glu Leu Leu Lys Glu Phe Pro Tyr  
 820 825 830

Gly Ile Glu Thr Phe Asn Arg His Glu Val Ser Leu Asp Gln Gln Lys  
 835 840 845

Thr His Lys Ile Val Glu Asn Gln Thr Gly Gly Lys Thr Ser Asn Val  
 850 855 860

Ser Gly Asp Ser Thr Asp Gln Ile Lys Ile Thr Val Leu Asn Ser Glu  
 865 870 875 880

ES 2 799 735 T3

Gln Ile Lys Glu Leu Phe Pro Glu Asp Asp Gln Pro Cys Asp Lys Leu  
885 890 895

Ala Glu Pro Glu Asn Lys Glu Ile Val Ala Glu Val Lys Ser Pro Cys  
900 905 910

Asp Ser Gln Ile Pro Arg Glu Glu Ser His Asp Leu Gly Met Leu Asp  
915 920 925

Pro Glu Lys Asp Lys Ile His Cys Cys Ala Leu Gly Trp Leu Ser Met  
930 935 940

Val Tyr Glu Gly Val Pro Gln Cys His Cys Ser Ser Thr Glu Lys Lys  
945 950 955 960

Glu Lys Asp Gln Cys Leu Asp Ile Asn Ser Ser Lys Gln Gly Glu Gln  
965 970 975

Pro Cys Asn Ser Gly Ile Thr Ile Phe Glu Ile Asn Pro Val Ser Asn  
980 985 990

Asn Ser Lys Thr Pro Leu Thr Gln Ala Thr Glu Glu Gly His Phe Ser  
995 1000 1005

Ala Val His Gly Glu Lys Thr Lys Ala Ser Lys Thr Lys Asp Asn  
1010 1015 1020

Arg Glu Gly Gln Glu Leu Ala Cys His Phe Ser Ala Lys Cys Tyr  
1025 1030 1035

Lys Lys Asp Lys Lys Gly Asn Phe Lys Ile Arg His Asp Thr Ser  
1040 1045 1050

Leu Lys Met Glu Gln Lys Leu Lys Asn Ile Ser Ser Lys Cys Asp  
1055 1060 1065

Ile Pro Asn Pro Ser Lys Cys Asn Lys Ile Ala Ala Pro Glu Ile  
1070 1075 1080

Leu His Val Thr Thr Ser Asn Ser Ala Lys Asn Met Pro Phe Ser  
1085 1090 1095

Lys Gln Ala Ser Gln Glu Ser Leu Gln Lys Lys His Thr Ser Gln  
1100 1105 1110

Asp Leu Gly Pro Val Lys Ala Pro Ile Glu Leu Ser Ser Asn Thr  
1115 1120 1125

ES 2 799 735 T3

Asp	Pro	Cys	Arg	Ser	Asn	Thr	Ser	Ser	Val	Gln	Ser	Val	Ser	Pro
	1130					1135					1140			
Glu	Lys	Lys	Lys	Leu	Lys	Phe	Lys	Ala	Gly	Gly	Ser	Arg	Leu	Lys
	1145					1150					1155			
Tyr	Phe	Glu	Lys	Arg	Lys	Thr	Asp	His	Val	Ile	Ile	Pro	Asp	Val
	1160					1165					1170			
Glu	Ile	Lys	Lys	Lys	Lys	Tyr	Glu	Lys	Gln	Glu	Gln	Asn	Lys	Asn
	1175					1180					1185			
Ala	Gly	Asp	Thr	Leu	Lys	Leu	Cys	Ser	Ile	Leu	Thr	Glu	Ser	Asn
	1190					1195					1200			
Glu	Arg	Ala	Ser	Val	Gln	Glu	Lys	Thr	Val	Pro	Ser	Pro	Glu	Ser
	1205					1210					1215			
Ser	Asp	Pro	Lys	Gly	Ser	Ser	Ser	Lys	Ser	Thr	Arg	Val	Ile	Thr
	1220					1225					1230			
Val	Gln	Glu	Tyr	Leu	Gln	Arg	Gln	Lys	Asp	Lys	Gln	Ile	Thr	Gly
	1235					1240					1245			
Asn	Asn	Ala	Ser	Arg	Asn	Ile	Cys	Val	Glu	Thr	Val	Leu	Cys	Asp
	1250					1255					1260			
Ser	Gly	His	Thr	Lys	Thr	Ser	Lys	His	Ser	Ala	Ala	Val	Ser	Trp
	1265					1270					1275			
Gly	Lys	Leu	Val	Glu	Gly	Gln	Ser	Ile	Ser	Ala	Glu	Thr	Ala	Lys
	1280					1285					1290			
Glu	Leu	Glu	His	Asn	Ser	Ser	Ser	His	Gly	Lys	Asp	Phe	Lys	Ile
	1295					1300					1305			
His	His	Ser	Glu	Ala	Ser	Arg	Thr	His	Ser	Val	Ser	Asn	Asn	Asn
	1310					1315					1320			
Lys	Gly	Lys	Phe	Asp	Gly	Lys	Gln	Pro	Asp	Lys	Met	Phe	Lys	Asn
	1325					1330					1335			
Lys	Thr	Ser	Met	Asn	Asn	Glu	Ser	Asn	Gln	Met	Pro	Leu	Gln	Val
	1340					1345					1350			
Lys	Glu	Gln	Arg	Lys	Gln	Tyr	Leu	Asn	Arg	Val	Ala	Phe	Lys	Cys

ES 2 799 735 T3

1355		1360		1365										
Thr	Glu	Arg	Glu	Ser	Ile	Cys	Leu	Thr	Lys	Leu	Asp	Ser	Ala	Ser
	1370					1375					1380			
Lys	Lys	Leu	Ser	Ile	Glu	Lys	Lys	Ser	Gly	Glu	Tyr	Thr	Ser	Lys
	1385					1390					1395			
Thr	Lys	Asp	Thr	Asp	Lys	Pro	Ser	Met	Leu	Glu	Phe	Lys	Leu	Cys
	1400					1405					1410			
Pro	Asp	Val	Leu	Leu	Lys	Asn	Thr	Ser	Thr	Val	Asp	Lys	Gln	Asp
	1415					1420					1425			
Cys	Pro	Gly	Pro	Gly	Pro	Glu	Lys	Glu	Gln	Ala	Pro	Val	Gln	Val
	1430					1435					1440			
Ser	Gly	Ile	Lys	Ser	Thr	Lys	Glu	Asp	Trp	Leu	Lys	Cys	Ile	Pro
	1445					1450					1455			
Thr	Arg	Thr	Lys	Met	Pro	Glu	Ser	Ser	Gln	Arg	Asp	Ser	Ala	Asp
	1460					1465					1470			
Ser	Arg	Leu	Ser	Lys	Arg	Ser	Leu	Ser	Ala	Asp	Glu	Phe	Glu	Ile
	1475					1480					1485			
Leu	Gln	Asn	Pro	Val	Lys	Glu	Ser	Asn	Ile	Met	Phe	Arg	Thr	Tyr
	1490					1495					1500			
Lys	Lys	Met	Tyr	Leu	Glu	Lys	Arg	Ser	Arg	Ser	Leu	Gly	Ser	Ser
	1505					1510					1515			
Pro	Val	Lys												
	1520													

<210> 36  
 <211> 6  
 <212> PRT  
 <213> Secuencia artificial  
 <220>  
 <221> fuente  
 <223> /nota="Descripción de secuencia artificial: Etiqueta 6xHis sintética"  
 <400> 36

His His His His His His  
 1 5

<210> 37  
 <211> 8  
 <212> PRT  
 <213> Secuencia artificial

ES 2 799 735 T3

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

<400> 37

5                                    **Asp Tyr Lys Asp Asp Asp Asp Lys**  
    1                                    5

<210> 38  
<211> 9  
<212> PRT  
<213> Secuencia artificial

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

<400> 38

**Ala Trp Arg His Pro Gln Phe Gly Gly**  
    1                                    5

15                                    <210> 39  
    <211> 8  
    <212> PRT  
    <213> Secuencia artificial

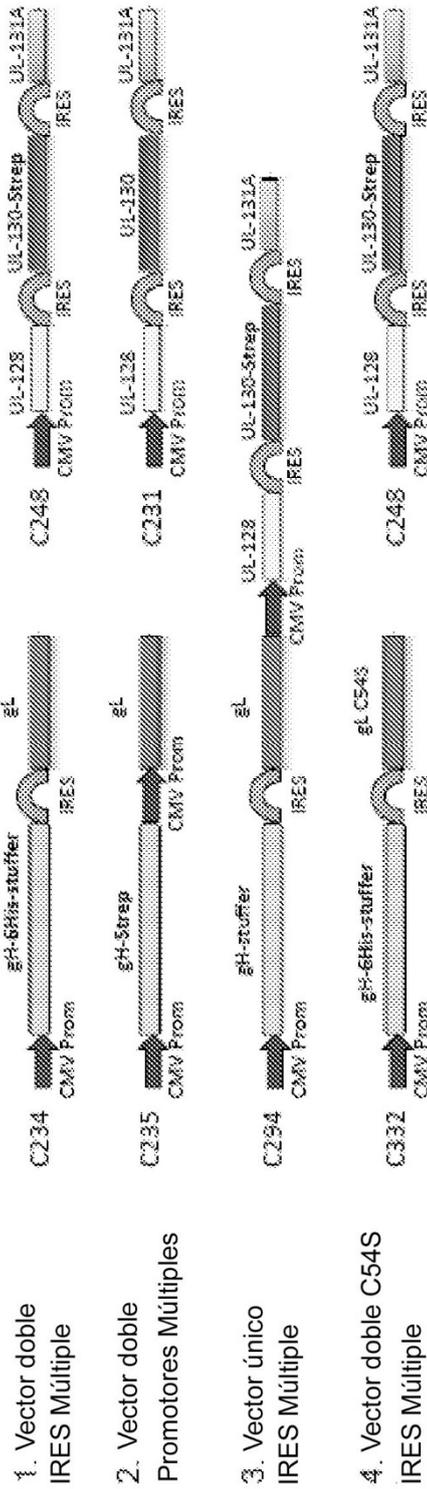
20                                    <220>  
    <221> fuente  
    <223> /nota="Descripción de secuencia artificial: péptido sintético"

<400> 39

**Trp Ser His Pro Gln Phe Glu Lys**  
    1                                    5

## REIVINDICACIONES

1. Una célula CHO recombinante, que comprende:
  - una o más secuencias de polinucleótidos que codifican el complejo pentamérico de citomegalovirus (CMV), en la que dicho complejo pentamérico comprende: gH o un fragmento formador del complejo pentamérico gH/gL/pUL128/pUL130/pUL131 del mismo, gL o un fragmento formador del complejo pentamérico gH/gL/pUL128/pUL130/pUL131 del mismo, pUL128 o un fragmento formador del complejo pentamérico gH/gL/pUL128/pUL130/pUL131 del mismo, pUL130 o un fragmento formador del complejo pentamérico gH/gL/pUL128/pUL130/pUL131 del mismo y pUL131 o un fragmento formador del complejo pentamérico gH/gL/pUL128/pUL130/pUL131 del mismo;
  - en la que dicha una o más secuencias de polinucleótidos están integradas en el ADN genómico de dicha célula CHO; y en la que el nivel de expresión o actividad de la proteína C12orf35 se reduce en dicha célula CHO, en comparación con una célula de tipo silvestre.
2. La célula CHO de la reivindicación 1, en la que dicha célula CHO es una célula CHO-K1, CHO-DUXB11, célula CHO-DG44 o célula CHO-S.
3. La célula CHO de la reivindicación 2, en la que al menos una copia de la secuencia genómica del gen C12orf35, o al menos el 50 % de la secuencia codificante de dicho gen C12orf35, está eliminada.
4. La célula CHO de la reivindicación 3, en la que dicha secuencia eliminada comprende una porción de la región telomérica del cromosoma 8 de una célula CHO.
5. La célula CHO de una cualquiera de las reivindicaciones 1-4, en la que el nivel de expresión o actividad de la proteína FAM60A se reduce en dicha célula, en comparación con una célula de tipo silvestre.
6. La célula CHO de la reivindicación 5, en la que al menos una copia de la secuencia genómica del gen FAM60A, o al menos el 50 % de la secuencia codificante de dicho gen FAM60A, está eliminada.
7. La célula CHO de una cualquiera de las reivindicaciones 1-6, en la que el nivel de expresión o actividad de matriptasa se reduce en dicha célula, en comparación con una célula de tipo silvestre.
8. La célula CHO de la reivindicación 7, en la que dicha célula comprende una mutación en el exón 2 del gen de la matriptasa.
9. La célula CHO de una cualquiera de las reivindicaciones 1-8, en la que dicho fragmento de formación de complejo de gH comprende el ectodominio de una proteína gH de longitud completa.
10. La célula CHO de una cualquiera de las reivindicaciones 1-9, en la que dicho complejo pentamérico es soluble.
11. La célula CHO de una cualquiera de las reivindicaciones 1-10, en la que dicho complejo pentamérico se secreta de la célula hospedadora.
12. Un cultivo a gran escala que comprende la célula CHO de una cualquiera de las reivindicaciones 1-11, en el que dicho cultivo tiene al menos 20 litros de tamaño.
13. El cultivo a gran escala de la reivindicación 12, en el que el rendimiento del complejo pentamérico CMV es de al menos 0,05 g/l.
14. Un procedimiento de producción de complejo pentamérico de citomegalovirus (CMV), en el que dicho complejo pentamérico comprende: gH o un fragmento formador del complejo pentamérico gH/gL/pUL128/pUL130/pUL131 del mismo, gL o un fragmento formador del complejo pentamérico gH/gL/pUL128/pUL130/pUL131 del mismo, pUL128 o un fragmento formador de complejos del mismo, pUL130 o un fragmento formador de complejos del mismo y pUL131 o un fragmento formador de complejos del mismo, que comprende:
  - (i) cultivar la célula CHO de una cualquiera de las reivindicaciones 1-11 en una condición adecuada, expresando de ese modo dicho complejo pentamérico; y
  - (ii) cosechar dicho complejo pentamérico del cultivo.



pBWKn: Kan/neo/DHFR\*      pBWah: Amp/Hyg/Fra

Figura 1

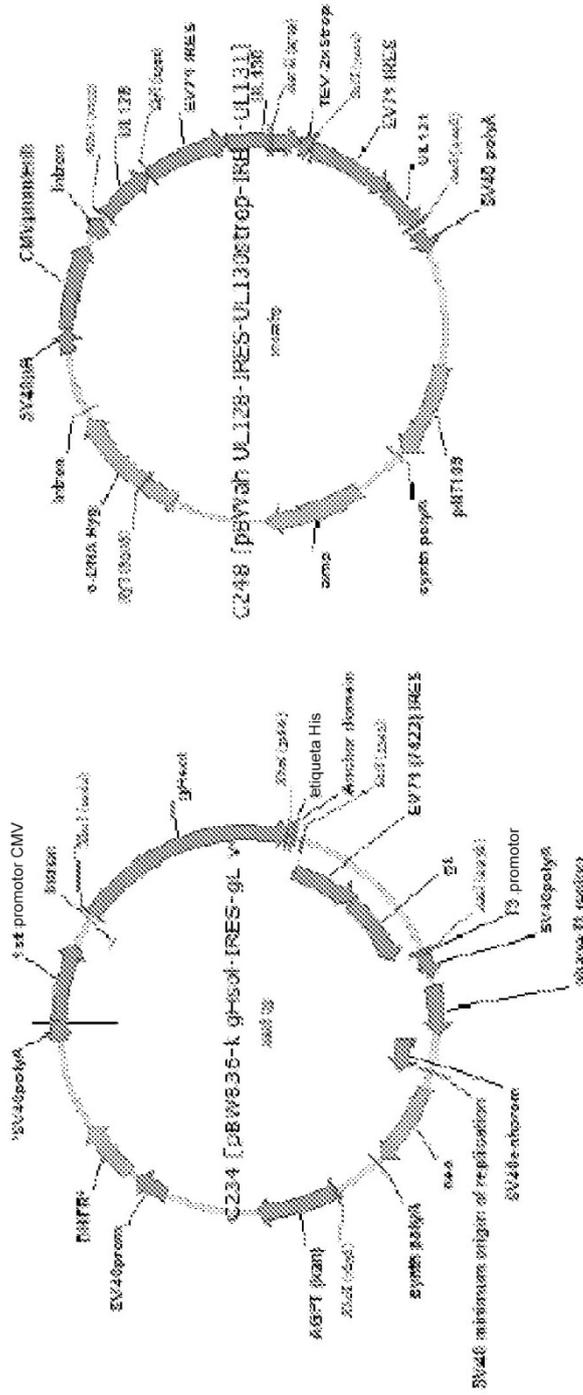


Figure 2

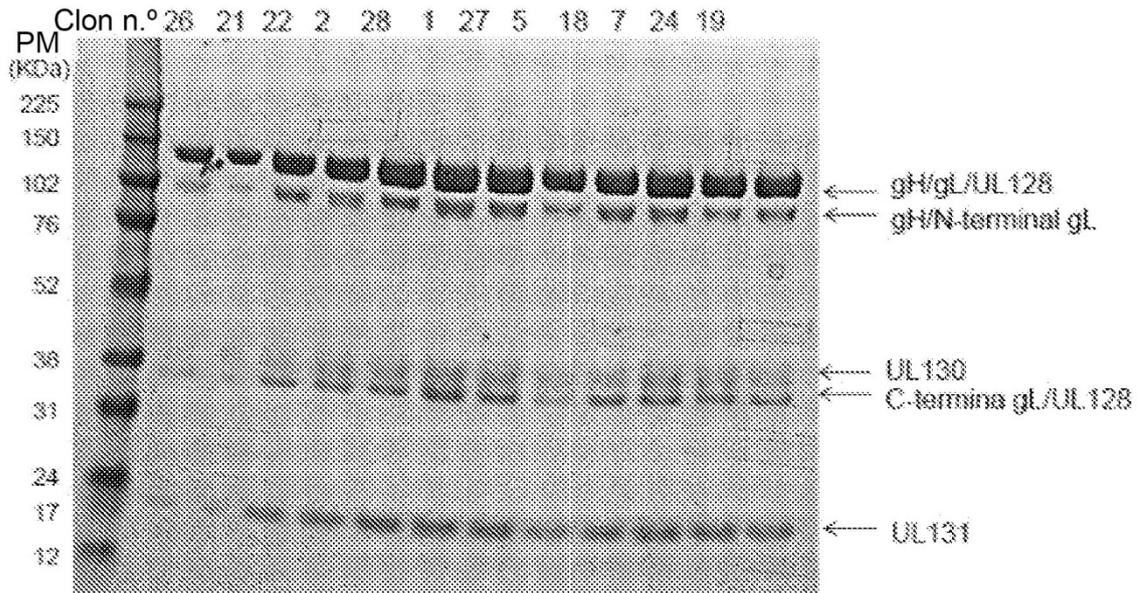


Figura 3

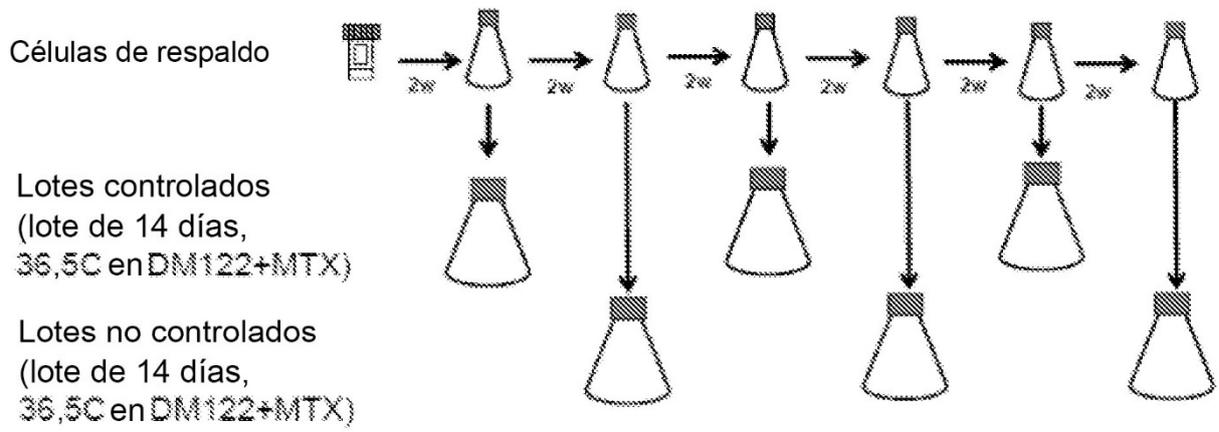


Figura 4

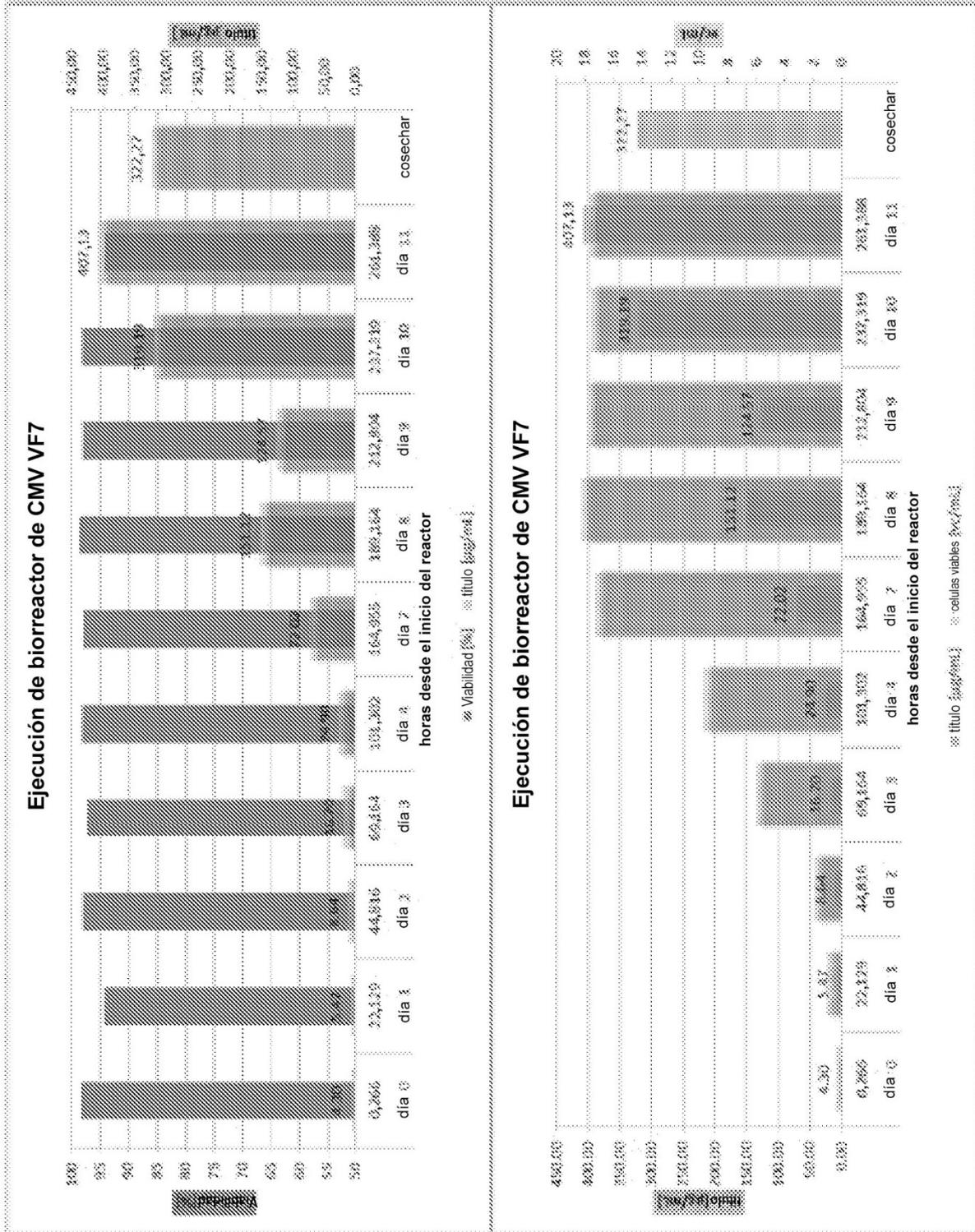


Figura 5