



SPF Economie, PME, Classes
Moyennes & Energie
Office de la Propriété intellectuelle

(11) 1023841 B1

(47) Date de délivrance : 10/08/2017

(12) BREVET D'INVENTION BELGE

(47) Date de publication : 10/08/2017

(21) Numéro de demande : BE2016/5877

(22) Date de dépôt : 24/11/2016

(62) Divisé de la demande de base : BE2015/5701

(62) Date de dépôt demande de base : 29/10/2015

(51) Classification internationale : C07K 14/005, A61K 39/245, C12N 5/071

(30) Données de priorité :

31/10/2014 EP 14191385.5

(73) Titulaire(s) :

GLAXOSMITHKLINE BIOLOGICALS S.A.
1330, RIXENSART
Belgique

(72) Inventeur(s) :

LILJA Anders
1160 VIENNE
Autriche

CARFI Andrea
02139 CAMBRIDGE
États-Unis

CIFERRI Claudio
94602 OAKLAND
États-Unis

HOFMANN Irmgard
1150 WIEN
Autriche

LAUX Holger
4055 BASEL
Suisse

WEN Yingxia
01720 ACTON
États-Unis

(54) CELLULES DE MAMMIFERES EXPRIMANT DES ANTIGENES DU CYTOMEGALOVIRUS

(57) Cette invention concerne des protéines du cytomégalovirus (CMV) appropriées pour des utilisations vaccinales. Sont fournies ici des cellules hôtes de mammifères, en particulier des cellules CHO, dans lesquelles la ou les séquences codant pour les protéines du CMV gH, gL, pUL128, pUL130, pUL131 (ou 5 l'un de leurs fragments formant un complexe) sont intégrées de façon stable dans le génome.

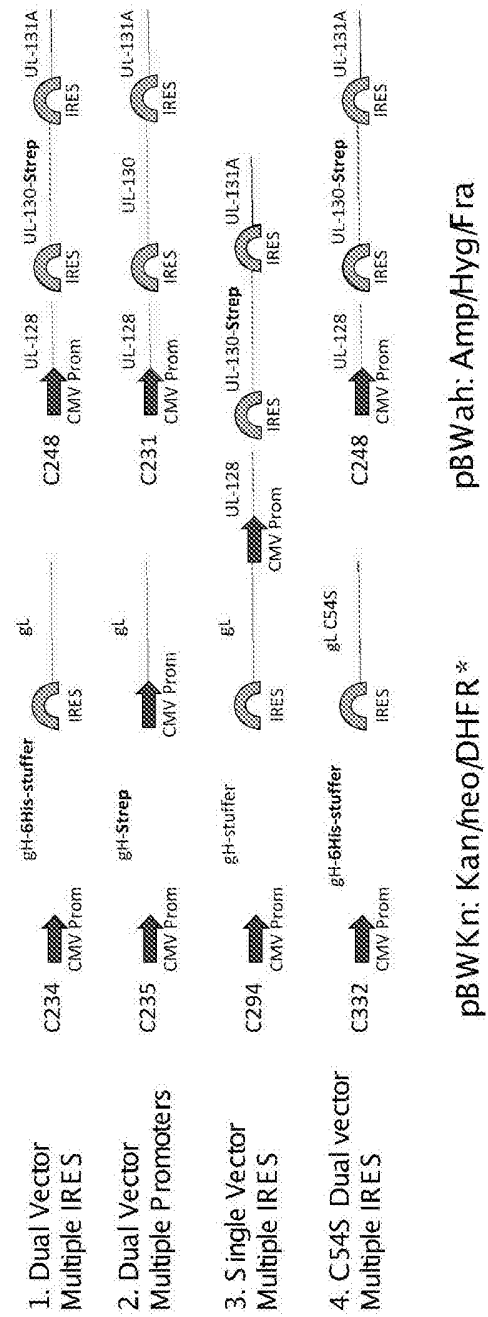


Figure 1

BREVET D'INVENTION BELGE

SPF Economie, PME, Classes
Moyennes & Energie

Numéro de publication : 1023841
Numéro de dépôt : BE2016/5877

Office de la Propriété intellectuelle

Classification Internationale : C07K 14/005 A61K 39/245 C12N
5/071
Date de délivrance : 10/08/2017

Le Ministre de l'Economie,

Vu la Convention de Paris du 20 mars 1883 pour la Protection de la propriété industrielle ;

Vu la loi du 28 mars 1984 sur les brevets d'invention, l'article 22, pour les demandes de brevet introduites avant le 22 septembre 2014 ;

Vu le Titre Ier "Brevets d'invention" du Livre XI du Code de droit économique, l'article XI.24, pour les demandes de brevet introduites à partir du 22 septembre 2014 ;

Vu l'arrêté royal du 2 décembre 1986 relatif à la demande, à la délivrance et au maintien en vigueur des brevets d'invention, l'article 28 ;

Vu la demande de brevet d'invention reçue par l'Office de la Propriété intellectuelle en date du 24/11/2016.

Considérant que pour les demandes de brevet tombant dans le champ d'application du Titre Ier, du Livre XI du Code de Droit économique (ci-après CDE), conformément à l'article XI. 19, §4, alinéa 2, du CDE, si la demande de brevet a fait l'objet d'un rapport de recherche mentionnant un défaut d'unité d'invention au sens du §1er de l'article XI.19 précité et dans le cas où le demandeur n'effectue ni une limitation de sa demande ni un dépôt d'une demande divisionnaire conformément aux résultats du rapport de recherche, le brevet délivré sera limité aux revendications pour lesquelles le rapport de recherche a été établi.

Arrête :

Article premier. - Il est délivré à

GLAXOSMITHKLINE BIOLOGICALS S.A., Rue de l'Institut 89, 1330 RIXENSART Belgique;

représenté par

PRONOVEM - Office Van Malderen, Avenue Josse Goffin 158, 1082, BRUXELLES;

un brevet d'invention belge d'une durée de 20 ans, sous réserve du paiement des taxes annuelles visées à l'article XI.48, §1 du Code de droit économique, pour : CELLULES DE MAMMIFERES EXPRIMANT DES

ANTIGENES DU CYTOMEGALOVIRUS.

INVENTEUR(S) :

LILJA Anders, Baumeistergasse 44/D/6, 1160, VIENNE ;

CARFI Andrea, 350 Massachusetts Avenue, 02139 , CAMBRIDGE;

CIFERRI Claudio, 4224 Norton Ave , 94602, OAKLAND;

HOFMANN Irmgard, Sturzgasse 28/24, 1150, WIEN ;

LAUX Holger, Hegenheimerstr.240, 4055, BASEL;

WEN Yingxia, 6 Centennial Lane , 01720, ACTON;

PRIORITE(S) :

31/10/2014 EP 14191385.5;

DIVISION :

divisé de la demande de base : BE2015/5701

date de dépôt de la demande de base : 29/10/2015

Article 2. – Ce brevet est délivré sans examen préalable de la brevetabilité de l'invention, sans garantie du mérite de l'invention ou de l'exactitude de la description de celle-ci et aux risques et périls du (des) demandeur(s).

Bruxelles, le 10/08/2017,

Par délégation spéciale :

CELLULES DE MAMMIFERES EXPRIMANT DES ANTIGENES DU
CYTOMEGALOVIRUS

Domaine de l'invention

[0001] Cette invention concerne des cellules hôtes exprimant des protéines du cytomégalovirus (CMV) appropriées pour des utilisations vaccinales.

5 Contexte de l'invention

[0002] Le cytomégalovirus est un genre de virus qui appartient à la famille virale connue sous le nom de Herpèsviridés ou herpèsvirus. L'espèce qui infecte les êtres humains est couramment connue sous le nom de cytomégalovirus humain (HCMV) ou d'herpèsvirus-5 (HHV-5) humain. Parmi les Herpèsviridés, le HCMV appartient à la sous-famille des Bêta-herpesvirinés, qui comprend également les cytomégalovirus d'autres mammifères.

10 **[0003]** Bien qu'on puisse les trouver dans tout le corps, les infections à HCMV sont fréquemment associées aux glandes salivaires. Le HCMV infecte entre 50 % et 80 % des adultes aux Etats-Unis (40 % dans le monde), comme cela est indiqué par la présence d'anticorps chez une grande proportion de la population générale. Le HCMV passe généralement inaperçu chez
15 les personnes en bonne santé, mais il peut menacer la vie des immunocompromis, comme les personnes infectées par le VIH, les receveurs de transplantation d'organe ou les nourrissons nouveau-nés. Le HCMV est le virus le plus fréquemment transmis aux fœtus en développement. Après une infection, le HCMV a la
20 capacité de demeurer latent au sein du corps la vie durant de l'hôte, avec des réactivations occasionnelles depuis l'état
25

latent. Etant donné la gravité et l'importance de cette maladie, l'obtention d'un vaccin efficace est considérée comme une priorité absolue pour la santé publique (Sung, H., et al., (2010) Expert review of vaccines 9, 1303-1314 ; Schleiss, Expert Opin Ther Pat. Apr 2010; 20(4): 597-602).

[0004] Les génomes de plus de 20 souches différentes du HCMV ont été séquencés, y compris ceux à la fois de souches de laboratoire et d'isolats. Par exemple, les souches suivantes du HCMV ont été séquencées : Towne (GL239909366), AD169 (GI: 219879600), Toledo (GL290564358) et Merlin (GI: 155573956). Les souches du HCMV AD169, Towne et Merlin peuvent être obtenues auprès de l'American Type Culture Collection (ATCC VR538, ATCC VR977 et ATCC VR1590, respectivement).

[0005] Le CMV contient un nombre inconnu de complexes de protéines membranaires. Parmi les approximativement 30 glycoprotéines connues dans l'enveloppe virale, la gH et la gL sont apparues comme étant particulièrement intéressantes en raison de leur présence dans plusieurs complexes différents : le dimère gH/gL, le trimère gH/gL/gO (également connu comme le complexe gCIII), et le pentamère gH/gL/pUL128/pUL130/pUL131 (la pUL131 est également appelée « pUL131A », « pUL131a », ou « UL131A » ; les sous-unités pUL128, pUL130, et pUL131 sont parfois également appelées UL128, UL130, UL131). On pense que le CMV utilise les complexes pentamères pour pénétrer dans les cellules épithéliales et endothéliales par endocytose et fusion dépendante d'un pH bas mais on pense qu'il pénètre dans les fibroblastes par fusion directe à la membrane plasmique dans un processus impliquant le complexe gH/gL ou éventuellement gH/gL/gO. Le(s) complexe(s) gH/gL et/ou gH/gL/gO est/sont suffisant(s) pour l'infection des fibroblastes, tandis que le complexe pentamère est requis pour infecter les cellules endothéliales et épithéliales.

[0006] Le complexe pentamère est considéré comme une cible majeure pour la vaccination contre le CMV. Les gènes viraux UL128, UL130 et UL131 sont nécessaires à la pénétration dans les cellules endothéliales (Hahn, Journal of Virology 2004; 78: 10023-33). Les souches à tropisme non endothélial adaptées aux fibroblastes contiennent des mutations dans au moins trois de ces gènes. La souche Towne, par exemple, contient une insertion de 2 paires de bases provoquant un décalage du cadre dans le gène UL130, tandis que AD169 contient une insertion de 1 paire de bases dans le gène UL131. A la fois Towne et AD169 ont pu être adaptées pour une croissance dans des cellules endothéliales, et dans les deux cas, les mutations de décalage du cadre dans les gènes UL130 ou UL131 ont été réparées.

[0007] Le brevet US 7 704 510 divulgue que la pUL131A est nécessaire pour un tropisme des cellules épithéliales. Le brevet US 7 704 510 divulgue également que la pUL128 et la pUL130 forment un complexe avec gH/gL, qui est incorporé dans les virions. Ce complexe est nécessaire pour infecter les cellules endothéliales et épithéliales mais pas les fibroblastes. Il a été découvert que des anticorps anti-CD46 inhibent l'infection par le HCMV des cellules épithéliales.

[0008] Des vaccins contre le CMV testés dans des essais cliniques comprennent le vaccin à base de Towne, des chimères Towne-Toledo, un réplicon d'alphavirus avec la gB comme antigène, un vaccin à base de gB/MF59, un vaccin à base de gB produit par GlaxoSmithKline, et un vaccin à base d'ADN utilisant la gB et la pp65. La pp65 est une protéine virale qui est un inducteur puissant des réponses CD8+ dirigées contre le CMV. Ces vaccins sont tous des inducteurs médiocres d'anticorps qui bloquent la pénétration du virus dans les cellules endothéliales/épithéliales (Adler, S. P. (2013), British Medical Bulletin, 107, 57-68. doi:10.1093/bmb/ldt023).

[0009] On pense de façon générale que des anticorps neutralisants dirigés contre le complexe pentamère (gH/gLpUL128/pUL130/pUL131) seront significativement plus puissants que des anticorps neutralisants dirigés contre la sous-unité gB, ou le complexe dimère gH/gL du CMV. Par conséquent, pour développer un vaccin efficace contre le CMV, il existe un besoin urgent de production de grandes quantités (par exemple, à des échelles commerciales) du complexe pentamère du CMV.

[0010] Toutefois, la production recombinante du complexe pentamère du CMV demeure un défi. Les cinq sous-unités doivent être exprimées (de préférence dans une quantité substantiellement égale, et de préférence pendant une période de temps prolongée), repliées correctement, et assemblées de façon correcte en un pentamère. En outre, il existe également un besoin d'éviter l'assemblage non souhaité de complexes contaminants (comme le dimère gH/gL, le tétramère gH/gL (deux copies de gH et deux copies de gL), le tétramère gH/gLpUL128/pUL130, etc.).

[0011] L'expression recombinante d'un tel complexe de protéines eucaryotes nécessitera l'identification de constructions appropriées et d'hôtes appropriés, pour l'expression de quantités suffisantes de protéines repliées correctement pendant une période de temps prolongée, qui peuvent ensuite s'assembler correctement en complexes de protéines. En outre, la sélection d'un hôte d'expression approprié a un impact significatif sur le rendement et la qualité des protéines, ainsi que sur les coûts véritables du procédé de production.

Résumé de l'invention

[0012] Il est divulgué et exemplifié ici des cellules hôtes de mammifères, en particulier des cellules CHO, dans lesquelles la/les séquence(s) codant pour les protéines du

CMV gH, gL, pUL128, pUL130, pUL131 (ou l'un de leurs fragments formant un complexe) est/sont intégrée(s) de façon stable dans le génome. De telles cellules hôtes fournissent une source fiable où le complexe pentamère gH/gL/pUL128/pUL130/pUL131 peut être produit de façon recombinante.

[0013] Dans un aspect, l'invention propose une cellule de mammifère recombinante, comme une cellule de rongeur (par exemple, une cellule CHO) comprenant une ou plusieurs séquences polynucléotidiques codant pour le complexe pentamère du cytomégalovirus (CMV), où ledit complexe pentamère comprend (i) la gH ou l'un de ses fragments formant un complexe, (ii) la gL ou l'un de ses fragments formant un complexe, (iii) la pUL128 ou l'un de ses fragments formant un complexe, (iv) la pUL130 ou l'un de ses fragments formant un complexe, et (v) la pUL131 ou l'un de ses fragments formant un complexe, où lesdites une ou plusieurs séquences polynucléotidiques sont intégrées dans l'ADN génomique de ladite cellule de mammifère. Parce que les séquences nucléotidiques sont intégrées de façon stable dans l'ADN génomique de la cellule, elles peuvent être transmises à la descendance comme faisant partie des séquences génomiques. De telles cellules sont souvent qualifiées dans l'art de lignées cellulaires stables. En particulier, la cellule ne comprend pas les séquences virales du CMV qui pourraient entraîner la production de virus CMV infectieux. Dans certains modes de réalisation, la cellule de mammifère est une cellule CHO. Lorsqu'elle est cultivée dans des conditions appropriées, ledit complexe pentamère du CMV est exprimé par ladite cellule hôte.

[0014] Les cellules CHO appropriées comprennent, par exemple, toutes les lignées cellulaires CHO disponibles à l'American Type Culture Collection (ATCC) ou à l'European Collection of Cell Cultures (ECACC). Les exemples de lignées

cellulaires CHO comprennent, par exemple, les cellules CHO-K1, CHO-DUXB11, CHO-DG44, ou CHO-S. Pour faciliter la sélection des clones intégrés de façon stable, les cellules hôtes peuvent être déficientes pour la dihydrofolate réductase (DHFR). L'expression de protéines de recombinaison dans des lignées cellulaires déficientes pour la DHFR ou compétentes pour la DHFR peut être également criblée par une sélection au méthotrexate (MTX).

[0015] La cellule hôte peut comporter une ou plusieurs modifications supplémentaires pour encore amplifier la production du complexe pentamère du CMV. Dans certains modes de réalisation, le taux d'expression ou l'activité de la protéine Cl2orf35 est réduit dans la cellule hôte, comparativement à un témoin. Dans certains modes de réalisation, le taux d'expression ou l'activité de la protéine FAM60A est réduit dans la cellule hôte, comparativement à un témoin. Dans certains modes de réalisation, le taux d'expression ou l'activité de la matriptase est réduit dans la cellule hôte, comparativement à un témoin. Les modifications décrites ici peuvent être utilisées seules ou dans n'importe quelle combinaison.

[0016] Le complexe pentamère du CMV produit de façon recombinante peut être soluble (par exemple, auquel il manque le domaine transmembranaire de la gH). Pour faciliter la production, le complexe pentamère du CMV produit de façon recombinante peut être sécrété à partir de la cellule hôte dans le milieu de culture.

[0017] Les cellules hôtes de mammifères décrites ici sont particulièrement appropriées pour une production à grande échelle, comme des cultures qui ont une taille d'au moins 20 litres (par exemple, 50 litres, 100 litres, etc.). Dans certains modes de réalisation, le rendement du complexe pentamère du CMV est d'au moins 0,05 g/l, ou d'au moins 0,1 g/l.

Brève description des figures

5 [0018] La figure 1 illustre diverses stratégies de conception pour la coexpression des cinq composants du pentamère.

[0019] La figure 2 représente les cartes plasmidiques des vecteurs d'expression utilisés pour la transfection de cellules CHO.

10 [0020] La figure 3 représente l'analyse SDS-PAGE du pentamère purifié produit avec les 12 principaux clones.

[0021] La figure 4 illustre les conceptions expérimentales utilisées pour étudier la stabilité des 12 principaux clones.

15 [0022] La figure 5 représente le rendement du pentamère produit par le clone principal VF7.

Description détaillée de l'invention

1. Généralités

20 [0023] L'une des cibles principales pour une vaccination contre le CMV est le complexe pentamère (gH/gL/pUL128/pUL130/pUL131). Pour produire un pentamère recombinant du CMV à une échelle commerciale, il existe le besoin d'identifier des constructions appropriées et des
25 hôtes appropriés, de telle façon que les cinq sous-unités du pentamère peuvent être exprimées dans des quantités suffisantes pendant une période de temps prolongée, et peuvent s'assembler correctement en un complexe pentamère. Le pentamère du CMV est une cible primaire des anticorps
30 neutralisants dirigés contre le CMV humain.

[0024] Le pentamère du CMV a été produit de façon recombinante en utilisant des cellules HEK293 en tant que cellules hôtes. Voir, le document WO 2014/005959. Toutefois, la lignée cellulaire HEK293 est généralement reconnue dans

l'art comme l'une des meilleures lignées de cellules hôtes pour une expression transitoire de gènes. Voir, Meyer et al., PLoS One, 2013 Jul 17;8(7):e68674. doi: 10.1371/journal.pone.0068674. Toutefois, il sera difficile de maintenir l'expression du pentamère pendant une période de temps prolongée dans des cellules HEK293 transfectées de façon transitoire. En outre, dans les exemples divulgués dans le document WO 2014/005959, les cinq sous-unités du pentamère ont été introduites dans des cellules HEK293 par l'intermédiaire de cinq plasmides différents. Dans cette configuration, il sera facile de maintenir l'expression de chaque sous-unité du pentamère à un taux substantiellement égal. Une quantité inégale des sous-unités peut réduire l'efficacité de l'assemblage du pentamère, d'autant plus que ces sous-unités ont déjà une tendance à former des complexes contaminants, comme le dimère gH/gL, le tétramère gH/gL (2 copies de gH et 2 copies de gL), le tétramère gH/gLpUL128/pUL130. En outre, le plasmide codant pour la gH a été sélectionné par la néomycine, les quatre plasmides codant pour les gL, pUL128, pUL130, et pUL131, respectivement, ont tous été sélectionnés par la kanamycine. Par conséquent, parce que les quatre plasmides ont été sélectionnés par la kanamycine, la perte d'un seul plasmide dans une cellule hôte ne pourra pas être facilement détectée, mais la perte d'un seul plasmide empêchera l'assemblage du pentamère.

[0025] Comme cela est divulgué et exemplifié ici, les inventeurs ont surmonté les difficultés de la production recombinante du pentamère du CMV en utilisant des cellules d'ovaire de hamster de Chine (CHO). Dans ces cellules CHO, les séquences codant pour les gH, gL, pUL128, pUL130, pUL131 (ou l'un de leurs fragments formant un complexe) sont intégrées de façon stable dans le génome des cellules CHO. En intégrant les séquences codant pour le pentamère dans le génome de la cellule CHO, les inventeurs ont obtenu une

expression génomique stable du pentamère recombinant, avec une efficacité et une stabilité génomique élevées. Comme cela est exemplifié dans la section des exemples, il a été découvert que, comparativement aux cellules HEK293 transfectées de façon transitoire, les lignées CHO stables ont obtenu de façon régulière des rendements 100 fois supérieurs, rendant ces lignées cellulaires CHO particulièrement appropriées pour la production commerciale du pentamère du CMV. L'intégration stable permet également une manipulation aisée d'une culture cellulaire à grande échelle, comparativement aux cellules HEK transfectées de façon transitoire avec cinq plasmides exogènes. En fait, les clones principaux ont produit le pentamère avec un rendement aussi élevé que 0,1 g/l à 0,5 g/l.

[0026] D'autres améliorations des cellules hôtes pour la production du pentamère du CMV sont également fournies ici. En particulier, trois cellules hôtes modifiées sont exemplifiées : (i) des cellules hôtes dans lesquelles le taux d'expression ou l'activité de la protéine C12orf35 est réduit, comparativement à un témoin ; (ii) des cellules hôtes dans lesquelles le taux d'expression ou l'activité de la protéine FAM60A est réduit, comparativement à un témoin ; (iii) des cellules hôtes dans lesquelles le taux d'expression ou l'activité de la matriptase est réduit comparativement à un témoin. Il a été découvert que la réduction du taux d'expression ou de l'activité de la protéine C12orf35 ou de la protéine FAM60A entraîne une augmentation significative du taux d'expression d'une protéine recombinante. Il a été également découvert que la réduction du taux d'expression ou de l'activité de la matriptase diminue significativement la dégradation protéolytique (« coupure ») d'une protéine recombinante. Ces modifications peuvent être utilisées seules ou dans n'importe quelle combinaison.

[0027] Par conséquent, il est proposé ici des cellules hôtes de mammifères, en particulier des cellules CHO, dans lesquelles la/les séquence(s) polynucléotidique(s) codant pour le complexe pentamère du CMV comprenant les gH, gL, pUL128, pUL130, pUL131 (ou l'un de leurs fragments formant un complexe) sont intégrées de façon stable dans l'ADN génomique. Lorsqu'elles sont cultivées dans des conditions appropriées, ledit complexe pentamère du CMV est exprimé par lesdites cellules hôtes.

[0028] Il est également proposé ici un complexe pentamère du cytomégalovirus (CMV) produit par les cellules de mammifères décrites ici.

[0029] Il est également proposé ici un procédé de production du complexe pentamère du cytomégalovirus (CMV), dans lequel ledit complexe pentamère comprend : la gH ou l'un de ses fragments formant un complexe, la gL ou l'un de ses fragments formant un complexe, la pUL128 ou l'un de ses fragments formant un complexe, la pUL130 ou l'un de ses fragments formant un complexe, et la pUL131 ou l'un de ses fragments formant un complexe, comprenant : (i) la culture de la cellule de mammifère telle que décrite ici dans des conditions appropriées, exprimant de cette façon ledit complexe pentamère ; et la récolte dudit complexe pentamère à partir de la culture. Le complexe pentamère peut être en outre purifié. Il est également proposé ici un complexe pentamère du cytomégalovirus (CMV) produit par ce procédé.

[0030] Il est également proposé ici une composition comprenant le complexe pentamère décrit ici. La composition peut comprendre un complexe pentamère du CMV purifié qui est approprié pour une administration *in vivo*. Par exemple, le complexe pentamère dans une telle composition peut avoir une pureté d'au moins 85 %, d'au moins 86 %, d'au moins 87 %, d'au moins 88 %, d'au moins 89 %, d'au moins 90 %, d'au moins 91 %, d'au moins 92 %, d'au moins 93 %, d'au moins 94 %, d'au

moins 95 %, d'au moins 96 %, d'au moins 97 %, d'au moins 98 %, ou d'au moins 99 %, en masse. La composition peut comprendre en outre un adjuvant, comme un sel d'aluminium, ou le MF59.

[0031] Il est également proposé ici une composition pour une utilisation dans l'induction d'une réponse immunitaire contre le CMV. L'utilisation de la composition décrite ici pour l'induction d'une réponse immunitaire contre le CMV et l'utilisation de la composition décrite ici dans la fabrication d'un médicament destiné à l'induction d'une réponse immunitaire contre le CMV sont également fournies.

2. Complexes pentamères du CMV et séquences codantes

A. Complexes pentamères du CMV

[0032] Dans un aspect, l'invention propose une cellule hôte de mammifère qui exprime le complexe pentamère du CMV, où ledit complexe pentamère comprend (i) la gH ou l'un de ses fragments formant un complexe, (ii) la gL ou l'un de ses fragments formant un complexe, (iii) la pUL128 ou l'un de ses fragments formant un complexe, (iv) la pUL130 ou l'un de ses fragments formant un complexe, et (v) la pUL131 ou l'un de ses fragments formant un complexe. Les séquences polynucléotidiques codant pour le complexe pentamère du CMV sont intégrées dans l'ADN génomique de la cellule hôte, et, lorsqu'elles sont cultivées dans des conditions appropriées, ledit complexe pentamère du CMV est exprimé par ladite cellule hôte.

[0033] Dans certains modes de réalisation, ledit complexe pentamère est soluble. Le complexe pentamère soluble peut être obtenu, par exemple, en utilisant un fragment de la gH dans lequel le domaine transmembranaire de la sous-unité gH est délété, comme cela est décrit en détail ci-dessous.

[0034] Dans certains modes de réalisation, ledit complexe pentamère est sécrété à partir de la cellule hôte. Il a été rapporté que la présence des cinq sous-unités, gH, gL, pUL128,

pUL131, et pUL131, est suffisante pour l'assemblage du complexe pentamère dans le RE avant son exportation vers l'appareil de Golgi. Voir, Ryckman et al., J Virol. Jan 2008; 82(1): 60-70. En variante ou en outre, un peptide signal approprié peut être utilisé dans une ou plusieurs des cinq sous-unités (par exemple, en fabriquant une protéine de fusion avec un signal sécrétoire). Les séquences signal (et la cassette d'expression) pour la production de protéine sécrétoire sont connues dans l'art. En général, les peptides de tête ont une longueur de 5 à 30 acides aminés, et ils sont généralement présents à l'extrémité N-terminale d'une protéine nouvellement synthétisée. Le cœur du peptide signal contient généralement une longue séquence d'acides aminés hydrophobes qui a tendance à former une seule hélice alpha. En outre, de nombreux peptides signal débutent par une séquence d'acides aminés chargés positivement, qui peut aider à l'application de la topologie correcte du polypeptide durant la translocation. A la fin du peptide signal, il y a généralement une séquence d'acides aminés qui est reconnue et clivée par une signal peptidase. La signal peptidase peut cliver soit durant soit après la complétion de la translocation pour produire un peptide signal libre et une protéine mature.

[0035] La glycoprotéine H (gH) du CMV humain, codée par le gène UL75, est une glycoprotéine de virion qui est essentielle à l'infectivité et qui est conservée parmi les membres des alpha-, bêta- et gamma-herpèsvirus. Elle peut former un complexe stable avec la gL, et la formation de ce complexe facilite l'expression à la surface cellulaire de la gH. En se basant sur les structures cristallines des complexes gH/gL du HSV-2 et de l'EBV, la sous-unité gL et les résidus N-terminaux de la gH forment un domaine globulaire à une extrémité de la structure (la « tête »), qui est impliqué dans les interactions avec la gB et l'activation de la fusion

membranaire. Le domaine C-terminal de la gH, proximal à la membrane virale (la « queue »), est également impliqué dans la fusion membranaire. La gH affiche des déterminants qui sont reconnus par le facteur de l'hôte TLR2, et elle interagit directement avec un hétérodimère formé entre les facteurs de l'hôte TLR2 et TLR1. Le TLR2 médie l'activation du NF- κ B et les réponses des cytokines inflammatoires à partir des cellules.

[0036] La gH provenant de la souche du CMV Merlin est représentée par SEQ ID NO : 1 (GI:52139248, 742 résidus d'acides aminés). La gH provenant de la souche du CMV Towne est représentée par SEQ ID NO : 2 (GI:138314, également 742 résidus d'acides aminés). La gH de Towne a été caractérisée comme comportant : (i) six sites de N-glycosylation (au niveau des résidus 55, 62, 67, 192, 641 et 700) ; (ii) une séquence signal hydrophobe à son extrémité N-terminale (résidus d'acides aminés 1 à 23) ; (iii) un ectodomaine (résidus 24 à 717) qui se projette hors de la cellule dans l'espace extracellulaire ; (iv) un domaine transmembranaire (TM) hydrophobe (résidus 718 à 736) ; et (v) un domaine cytoplasmique C-terminal (résidus 737 à 742). SEQ ID NO : 2 partage 99 % et 96 % d'identité de séquence d'acides aminés avec SEQ ID NO : 1, et la gH provenant de la souche du CMV AD169 (GI:138313, SEQ ID NO : 3), respectivement.

[0037] Généralement, la séquence signal N-terminale de la protéine gH pleine longueur est clivée par une signal peptidase de la cellule hôte pour produire une protéine gH mature. En tant que telle, à la protéine gH exprimée par la cellule hôte décrite ici, il peut manquer la séquence signal N-terminale (par exemple, la gH est codée par une séquence nucléotidique à laquelle il manque la séquence codant pour la séquence signal N-terminale).

[0038] Sont également englobés dans l'invention, des fragments formant un complexe de la gH, comme un fragment de la gH auquel il manque le domaine transmembranaire (TM) (par exemple, les résidus 718 à 736 de SEQ ID NO : 2), le domaine C-terminal (par exemple, les résidus 737 à 742 de SEQ ID NO : 2), la séquence signal N-terminale (par exemple, les résidus 1 à 23 de SEQ ID NO : 2), ou l'une de leurs combinaisons. Un fragment formant un complexe de la gH peut être toute partie ou portion de la protéine gH qui conserve la capacité de former un complexe avec une autre protéine du CMV. Dans certains modes de réalisation, un fragment formant un complexe de la gH forme une partie du complexe pentamère gH/gL/pUL128/pUL130/pUL131. Par exemple, l'expression de la séquence de la gH pleine longueur peut gêner la purification du complexe pentamère soluble parce que le domaine TM de la gH est hydrophobe. A la place, le complexe pentamère peut comprendre un fragment de la gH avec au moins une portion du domaine TM de la gH délétée.

[0039] Par exemple, un fragment de la gH comprenant la séquence signal N-terminale et l'ectodomaine de la gH, mais pas le domaine TM, peut être utilisé. Un fragment approprié de la gH peut également comprendre une portion de l'ectodomaine de la gH (par exemple, au moins environ 70 %, au moins environ 80 %, au moins environ 85 %, au moins environ 90 %, au moins environ 95 %, au moins environ 96 %, au moins environ 97 %, au moins environ 98 %, ou au moins environ 99 % de la séquence de l'ectodomaine de la gH), mais aucune, ou seulement une petite portion du domaine TM. En variante, au fragment de la gH décrit ici, il peut manquer entre 1 et 20 résidus d'acides aminés (par exemple 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 ou 20 résidus d'acides aminés, ou il manque 1 à 20 résidus, 1 à 15 résidus, 1 à 10 résidus, 2 à 20 résidus, 2 à 15 résidus, 2 à 10 résidus, 5 à 20 résidus, 5 à 15 résidus, ou 5 à 10 résidus)

à l'extrémité N-terminale et/ou à l'extrémité C-terminale de l'ectodomaine pleine longueur. On pense que les résidus au niveau des domaines C-terminaux ne sont pas nécessaires à l'immunogénicité. Un exemple de fragment approprié de la gH décrit ici est représenté par SEQ ID NO : 4, qui correspond aux résidus d'acides aminés 1 à 715 de SEQ ID NO : 1. Un autre exemple de fragment de la gH décrit ici est représenté par SEQ ID NO : 5, auquel il manque la séquence signal N-terminale, le domaine TM et le domaine C-terminal de la gH, et correspond aux résidus d'acides aminés 24 à 715 de SEQ ID NO : 1. Un autre exemple de fragment de la gH décrit comprend la séquence signal N-terminale entière et l'ectodomaine, mais auquel il manque le domaine C-terminal.

[0040] L'ectodomaine de la gH correspond au domaine extracellulaire de la gH. La localisation et la longueur de l'ectodomaine d'une gH (ou de l'un de ses homologues ou de ses variants) peuvent être prédites en se basant sur un alignement par paires de sa séquence avec les SEQ ID NO : 1, 2, 3, 4, ou 5, par exemple en alignant la séquence d'acides aminés d'une gH avec SEQ ID NO : 1, et en identifiant la séquence qui s'aligne avec les résidus 24 à 717 de SEQ ID NO : 1. De façon similaire, les localisations de la séquence signal, du domaine TM, et du domaine C-terminal peuvent être prédites en alignant la séquence d'acides aminés d'une gH avec les SEQ ID NO : 1, 2, 3, 4, ou 5, et en identifiant les séquences qui s'alignent avec les régions correspondantes (par exemple, les résidus 1 à 23 (séquence signal), 718 à 736 (TM) et 737 à 742 (domaine C-terminal) de SEQ ID NO : 1, respectivement). En variante, la localisation et la longueur de l'ectodomaine, de la séquence signal, du domaine TM, et du domaine C-terminal peuvent être prédites en se basant sur une analyse informatique de l'hydrophobicité tout le long d'une séquence de gH donnée. La séquence signal et le domaine TM présentent les taux les plus élevés

d'hydrophobicité et ces deux régions flanquent l'ectodomaine, qui est moins hydrophobe.

[0041] Un fragment formant un complexe approprié de la gH peut être également obtenu ou déterminé par des tests classiques connus dans l'art, comme un test de co-immunoprécipitation, la réticulation, ou la co-localisation par coloration fluorescente, etc. Une SDS-PAGE ou une analyse Western blot peut être également utilisée (par exemple, en montrant que les cinq sous-unités sont présentes dans une électrophorèse sur gel). Dans certains modes de réalisation, le fragment formant un complexe de la gH (i) forme une partie du complexe pentamère gH/gL/pUL128/pUL130/pUL131 ; (ii) comprend au moins un épitope provenant de SEQ ID NO : 1, SEQ ID NO : 2, SEQ ID NO : 3, SEQ ID NO : 4, ou SEQ ID NO : 5 ; et/ou (iii) peut déclencher des anticorps *in vivo* qui présentent une réactivité croisée immunologique avec un virion du CMV.

[0042] D'autres protéines gH appropriées peuvent être des variants de la gH qui présentent divers degrés d'identité avec SEQ ID NO : 1, 2, 3, 4, ou 5, comme au moins identiques à 60 %, 70 %, 80 %, 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % ou 99 % avec la séquence citée dans SEQ ID NO : 1, SEQ ID NO : 2, SEQ ID NO : 3, SEQ ID NO : 4, ou SEQ ID NO : 5. Dans certains modes de réalisation, les protéines variantes de la gH : (i) forment une partie du complexe pentamère gH/gL/pUL128/pUL130/pUL131 ; (ii) comprennent au moins un épitope provenant de SEQ ID NO : 1, SEQ ID NO : 2, SEQ ID NO : 3, SEQ ID NO : 4, ou SEQ ID NO : 5 ; et/ou (iii) peuvent déclencher des anticorps *in vivo* qui présentent une réactivité croisée immunologique avec un virion du CMV.

[0043] La glycoprotéine L (gL) du CMV humain est codée par le gène UL115. On pense que la gL est essentielle à la réplication virale et toutes les propriétés fonctionnelles

connues de la gL sont directement associées à sa dimérisation avec la gH. Le complexe gL/gH est nécessaire pour la fusion des membranes virale et plasmatisque menant à la pénétration du virus dans les cellules hôtes. Il a été rapporté que la gL
5 provenant de la souche du CMV Merlin (GI: 39842115, SEQ ID NO : 6) et de la souche du CMV Towne (GI: 239909463, SEQ ID NO : 7) a une longueur de 278 acides aminés. Il a été rapporté que la gL provenant de la souche du HCMV AD169 (GI: 2506510, SEQ ID NO : 8) a une longueur de 278 acides
10 aminés, avec une séquence signal à son extrémité N-terminale (résidus d'acides aminés 1 à 35), deux sites de N-glycosylation (au niveau des résidus 74 et 114), et à laquelle il manque un domaine TM. La séquence signal N-terminale dans SEQ ID NO : 6 est prédite pour comprendre les résidus
15 d'acides aminés 1 à 30. SEQ ID NO : 7 partage 98 % d'identité de séquence d'acides aminés avec SEQ ID NO : 6. Le séquençage du gène de la gL pleine longueur provenant de 22 à 39 isolats cliniques, ainsi que de souches de laboratoire AD169, Towne et Toledo a révélé moins de 2 % de variation dans les
20 séquences d'acides aminés parmi les isolats.

[0044] Généralement, la séquence signal N-terminale de la protéine gL pleine longueur est clivée par un signal peptidase de la cellule hôte pour produire une protéine gL mature. En tant que telle, à la protéine gL exprimée par la cellule hôte
25 décrite ici, il peut manquer la séquence signal N-terminale (par exemple, la gL est codée par une séquence nucléotidique à laquelle il manque la séquence codant pour la séquence signal N-terminale). Un exemple d'une gL à laquelle il manque la séquence signal est SEQ ID NO : 9, qui comprend les résidus
30 d'acides aminés 31 à 278 de SEQ ID NO : 6, et à laquelle il manque une séquence signal N-terminale de SEQ ID NO : 6. La séquence signal d'autres protéines gL peut être déterminée par alignement des séquences ou des outils d'analyse des séquences tels que décrits ci-dessus.

[0045] Sont également englobés dans l'invention des fragments formant un complexe de la gL. Un fragment formant un complexe de la gL peut être toute partie ou portion de la protéine gL qui conserve la capacité de former un complexe avec une autre protéine du CMV. Dans certains modes de réalisation, un fragment formant un complexe de la gL forme une partie du complexe pentamère gH/gL/pUL128/pUL130/pUL131.

[0046] Un fragment formant un complexe approprié de la gL peut être obtenu ou déterminé par des tests classiques connus dans l'art, comme un test de co-immunoprécipitation, la réticulation, ou la co-localisation par coloration fluorescente, etc. Une SDS-PAGE ou une analyse Western blot peut être également utilisée (par exemple, en montrant que les cinq sous-unités sont présentes dans une électrophorèse sur gel). Dans certains modes de réalisation, le fragment formant un complexe de la gL (i) forme une partie du complexe pentamère gH/gL/pUL128/pUL130/pUL131 ; (ii) comprend au moins un épitope provenant de SEQ ID NO : 6, SEQ ID NO : 7, SEQ ID NO : 8, ou SEQ ID NO : 9 ; et/ou (iii) peut déclencher des anticorps *in vivo* qui présentent une réactivité croisée immunologique avec un virion du CMV.

[0047] D'autres protéines gL appropriées peuvent être des variants de la gL (et des fragments de variants) qui présentent divers degrés d'identité avec SEQ ID NO : 6, 7, 8, ou 9, comme au moins identiques à 60 %, 70 %, 80 %, 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % ou 99 % avec la séquence citée par SEQ ID NO : 6, SEQ ID NO : 7, SEQ ID NO : 8, ou SEQ ID NO : 9. Dans certains modes de réalisation, les protéines variantes de la gL : (i) forment une partie du complexe pentamère gH/gL/pUL128/pUL130/pUL131 ; (ii) comprennent au moins un épitope provenant de SEQ ID NO : 6, SEQ ID NO : 7, SEQ ID NO : 8, ou SEQ ID NO : 9 ; et/ou (iii) peuvent déclencher des anticorps

in vivo présentent une réactivité croisée immunologique avec un virion du CMV.

[0048] Il a été rapporté que la pUL128 provenant de la souche du CMV humain Merlin (GI: 39842124, SEQ ID NO : 10) comporte 130 résidus d'acides aminés, et avec 1 substitution de nucléotide provoquant une terminaison prématurée. Il a été rapporté que la pUL128 provenant des souches du CMV humain Towne (GI: 39841882, SEQ ID NO : 11) et AD169 (GI: 59803078, SEQ ID NO : 12) comporte 171 résidus d'acides aminés. SEQ ID NO : 10 et 12 partagent plus de 99 % d'identité de séquence sur la pleine longueur de SEQ ID NO : 10 ; toutefois, à cause de la terminaison prématurée de la traduction, SEQ ID NO : 10 ne comporte pas les 41 résidus d'acides aminés C-terminaux de SEQ ID NO : 12 (environ 75 % d'identité de séquence sur la pleine longueur de SEQ ID NO : 12).

[0049] Il est prédit que la pUL128 comporte une séquence signal N-terminale, qui est localisée au niveau des résidus 1 à 27 de SEQ ID NO : 10, mais il est prédit qu'il lui manque un domaine TM. La séquence signal N-terminale de la protéine pUL128 pleine longueur peut être clivée par une signal peptidase de la cellule hôte pour produire une protéine pUL128 mature. En tant que telle, à la protéine pUL128 exprimée par la cellule hôte décrite ici, il peut manquer la séquence signal N-terminale (par exemple, la pUL128 est codée par une séquence nucléotidique à laquelle il manque la séquence codant pour la séquence signal N-terminale). Un exemple d'une protéine pUL128 mature est SEQ ID NO : 13, à laquelle il manque une séquence signal N-terminale et qui correspond aux résidus d'acides aminés 28 à 171 de SEQ ID NO : 11. SEQ ID NO : 13 correspond également aux acides aminés 28 à 171 de SEQ ID NO : 12.

[0050] Sont également englobés dans l'invention des fragments formant un complexe de la pUL128. Un fragment

formant un complexe de la pUL128 peut être toute partie ou portion de la protéine pUL128 qui conserve la capacité de former un complexe avec une autre protéine du CMV. Dans certains modes de réalisation, un fragment formant un complexe de la pUL128 forme une partie du complexe pentamère gH/gL/pUL128/pUL130/pUL131.

[0051] Un fragment formant un complexe approprié de la pUL128 peut être obtenu ou déterminé par des tests classiques connus dans l'art, comme un test de co-immunoprécipitation, la réticulation, ou la co-localisation par coloration fluorescente, etc. Une SDS-PAGE ou une analyse Western blot peut être également utilisée (par exemple, en montrant que les cinq sous-unités sont présentes dans une électrophorèse sur gel). Dans certains modes de réalisation, le fragment formant un complexe de la pUL128 (i) forme une partie du complexe pentamère gH/gL/pUL128/pUL130/pUL131 ; (ii) comprend au moins un épitope provenant de SEQ ID NO : 10, SEQ ID NO : 11, SEQ ID NO : 12, ou SEQ ID NO : 13 ; et/ou (iii) peut déclencher des anticorps *in vivo* qui présentent une réactivité croisée immunologique avec un virion du CMV.

[0052] D'autres protéines pUL128 appropriées peuvent être des variants de la pUL128 (et des fragments de variants) qui présentent divers degrés d'identité avec SEQ ID NO : 10, 11, 12, ou 13, comme au moins identiques à 60 %, 70 %, 80 %, 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % ou 99 % avec la séquence citée par SEQ ID NO : 10, SEQ ID NO : 11, SEQ ID NO : 12, ou SEQ ID NO : 13. Dans certains modes de réalisation, les protéines variantes de la pUL128 : (i) forment une partie du complexe pentamère gH/gL/pUL128/pUL130/pUL131 ; (ii) comprennent au moins un épitope provenant de SEQ ID NO : 10, SEQ ID NO : 11, SEQ ID NO : 12, ou SEQ ID NO : 13 ; et/ou (iii) peuvent

déclencher des anticorps *in vivo* qui présentent une réactivité croisée immunologique avec un virion du CMV.

[0053] UL130 est le gène central et le plus grand (214 codons) du locus UL131A-128. La traduction conceptuelle du gène prédit une longue séquence signal N-terminale (25 acides aminés) qui précède une protéine hydrophile, avec deux sites potentiels de N-glycosylation (Asn85 et Asn118) au sein d'un domaine de chimiokine putatif (acides aminés 46 à 120), et un site supplémentaire de N-glycosylation (Asn201) proche de l'extrémité d'une région C-terminale unique. Il est prédit qu'à la pUL130 il manque un domaine TM. Il a été rapporté qu'il s'agit d'une glycoprotéine luminaire qui est sécrétée de façon inefficace à partir de cellules infectées mais qui est incorporée dans l'enveloppe du virion sous une forme mature après passage dans l'appareil de Golgi. Les séquences de la pUL130 provenant des souches du CMV humain Merlin et Towne sont disponibles au public (GI: 39842125, SEQ ID NO : 14, 214 résidus d'acides aminés ; et GI: 239909473, SEQ ID NO : 15, 229 résidus d'acides aminés, respectivement). Il a été rapporté que SEQ ID NO : 15 contient une mutation de décalage du cadre dans la région C-terminale de la pUL130, et elle partage 94 % d'identité avec la SEQ ID NO : 14 du HCMV sur la pleine longueur de SEQ ID NO : 14.

[0054] La séquence signal N-terminale de la protéine pUL130 pleine longueur peut être clivée par une signal peptidase de la cellule hôte pour produire une protéine pUL130 mature. En tant que telle, à la protéine pUL130 exprimée par la cellule hôte décrite ici, il peut manquer la séquence signal N-terminale (par exemple, la pUL130 est codée par une séquence nucléotidique à laquelle il manque la séquence codant pour la séquence signal N-terminale). Un exemple d'une protéine pUL130 mature est SEQ ID NO : 16, à laquelle il

manque une séquence signal N-terminale, et qui correspond aux résidus d'acides aminés 26 à 214 de SEQ ID NO : 14.

[0055] Sont également englobés dans l'invention des fragments formant un complexe de la pUL130. Un fragment formant un complexe de la pUL130 peut être toute partie ou portion de la protéine pUL130 qui conserve la capacité de former un complexe avec une autre protéine du CMV. Dans certains modes de réalisation, un fragment formant un complexe de la pUL130 forme une partie du complexe pentamère gH/gL/pUL128/pUL130/pUL131.

[0056] Un fragment formant un complexe approprié de la pUL130 peut être obtenu ou déterminé par des tests classiques connus dans l'art, comme un test de co-immunoprécipitation, la réticulation, ou la co-localisation par coloration fluorescente, etc. Une SDS-PAGE ou une analyse Western blot peut être également utilisée (par exemple, en montrant que les cinq sous-unités sont présentes dans une électrophorèse sur gel). Dans certains modes de réalisation, le fragment formant un complexe de la pUL130 (i) forme une partie du complexe pentamère gH/gL/pUL128/pUL130/pUL131 ; (ii) comprend au moins un épitope provenant de SEQ ID NO : 14, SEQ ID NO : 15, ou SEQ ID NO : 16 ; et/ou (iii) peut déclencher des anticorps *in vivo* qui présentent une réactivité croisée immunologique avec un virion du CMV.

[0057] D'autres protéines pUL130 appropriées peuvent être des variants de la pUL130 (et des fragments de variants) qui présentent divers degrés d'identité avec SEQ ID NO : 14, 15, ou 16, comme au moins identiques à 60 %, 70 %, 80 %, 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % ou 99 % avec la séquence citée dans SEQ ID NO : 14, SEQ ID NO : 15, ou SEQ ID NO : 16. Dans certains modes de réalisation, les protéines variantes de la pUL130 : (i) forment une partie du complexe pentamère gH/gL/pUL128/pUL130/pUL131 ; (ii) comprennent au moins un épitope provenant de

SEQ ID NO : 14, SEQ ID NO : 15, ou SEQ ID NO : 16 ; et/ou
(iii) peuvent déclencher des anticorps *in vivo* qui présentent
une réactivité croisée immunologique avec un virion du CMV.

[0058] La fonction de la pUL131A est requise pour la
5 réplication du CMV humain non seulement dans les cellules
endothéliales mais également dans les cellules épithéliales.
Il a été rapporté la pUL131A provenant des souches du CMV
humain Merlin (GI: 39842126, SEQ ID NO : 17, 129 acides
aminés) et Towne (GI: 239909474, SEQ ID NO : 18, 129 acides
10 aminés) et AD169 (GI: 219879712, SEQ ID NO : 19, 76 acides
aminés). Il est prédit que la PUL131A contienne une séquence
signal N-terminale, qui est localisée au niveau des résidus
1 à 18 de SEQ ID NO : 18, et à laquelle il manque un domaine
TM. Il a été rapporté que la pUL131A provenant de la souche
15 AD169 contient une insertion de 1 paire de bases, qui provoque
un décalage du cadre. SEQ ID NO : 17 est identique à 96 %
avec SEQ ID NO : 19 sur les 28 acides aminés N-terminaux,
mais elle est seulement identique à 36 % avec SEQ ID NO : 19
sur la pleine longueur de SEQ ID NO : 17, à cause du décalage
20 du cadre dans le gène UL131A de AD169.

[0059] La séquence signal N-terminale de la protéine
pUL131 pleine longueur peut être clivée par une signal
peptidase de la cellule hôte pour produire une protéine pUL131
mature. En tant que telle, à la protéine pUL131 exprimée par
25 la cellule hôte décrite ici il peut manquer la séquence signal
N-terminale (par exemple, la pUL131 est codée par une séquence
nucléotidique à laquelle il manque la séquence codant pour la
séquence signal N-terminale). Un exemple de protéine pUL130
mature est SEQ ID NO : 20, à laquelle il manque une séquence
30 signal N-terminale et qui correspond aux résidus d'acides
aminés 19 à 129 de SEQ ID NO : 17. SEQ ID NO : 35 correspond
également aux résidus d'acides aminés 19 à 129 de
SEQ ID NO : 18.

[0060] Sont également englobés dans l'invention des fragments formant un complexe de la pUL131. Un fragment formant un complexe de la pUL131 peut être toute partie ou portion de la protéine pUL131 qui conserve la capacité de former un complexe avec une autre protéine du CMV. Dans certains modes de réalisation, un fragment formant un complexe de la pUL131 forme une partie du complexe pentamère gH/gL/pUL128/pUL130/pUL131.

[0061] Un fragment formant un complexe approprié de la pUL131 peut être obtenu ou déterminé par des tests classiques connus dans l'art, comme un test de co-immunoprécipitation, la réticulation, ou la co-localisation par coloration fluorescente, etc. Une SDS-PAGE ou une analyse Western blot peut être également utilisée (par exemple, en montrant que les cinq sous-unités sont présentes dans une électrophorèse sur gel). Dans certains modes de réalisation, le fragment formant un complexe de la pUL131 (i) forme une partie du complexe pentamère gH/gL/pUL128/pUL130/pUL131 ; (ii) comprend au moins un épitope provenant de SEQ ID NO : 17, SEQ ID NO : 18, SEQ ID NO : 19, ou SEQ ID NO : 20 ; et/ou (iii) peut déclencher des anticorps *in vivo* qui présentent une réactivité croisée immunologique avec un virion du CMV.

[0062] D'autres protéines pUL131 appropriées peuvent être des variants de la pUL131 (et des fragments de variants) qui présentent divers degrés d'identité avec SEQ ID NO : 17, 18, 19, ou 20, comme au moins identiques à 60 %, 70 %, 80 %, 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % ou 99 % avec la séquence citée dans SEQ ID NO : 17, SEQ ID NO : 18, SEQ ID NO : 19, ou SEQ ID NO : 20. Dans certains modes de réalisation, les protéines variantes de la pUL131 : (i) forment une partie du complexe pentamère gH/gL/pUL128/pUL130/pUL131 ; (ii) comprennent au moins un épitope provenant de SEQ ID NO : 17, SEQ ID NO : 18,

SEQ ID NO : 19, ou SEQ ID NO : 20 ; et/ou (iii) peuvent déclencher des anticorps *in vivo* qui présentent une réactivité croisée immunologique avec un virion du CMV.

[0063] Dans certains modes de réalisation, les gH, gL, pUL128, pUL130, pUL131 (ou l'un de leurs fragments) décrites ici peuvent contenir des résidus d'acides aminés supplémentaires, comme des extensions N-terminales ou C-terminales. De telles extensions peuvent comprendre un ou plusieurs marqueurs, qui peuvent faciliter la détection (par exemple, un marqueur épitopique pour la détection par des anticorps monoclonaux) et/ou la purification (par exemple, un marqueur polyhistidine pour permettre la purification sur une résine chélatrice de nickel) des protéines. Les exemples de marqueurs pour la purification par affinité comprennent, par exemple, le marqueur His (hexahistidine (SEQ ID NO : 36), se lie aux ions métalliques), la protéine de liaison du maltose (MBP) (se lie à l'amylose), la glutathion-S-transférase (GST) (se lie au glutathion), le marqueur FLAG (Asp-Tyr-Lys-Asp-Asp-Asp-Lys (SEQ ID NO : 37), se lie à un anticorps anti-flag), le marqueur Strep (Ala-Trp-Arg-His-Pro-Gln-Phe-Gly-Gly (SEQ ID NO : 38), ou Trp-Ser-His-Pro-Gln-Phe-Glu-Lys (SEQ ID NO : 39), se lie à la streptavidine ou l'un de ses dérivés).

[0064] Dans un certain mode de réalisation, des lieux clivables peuvent être utilisés. Ceci permet de séparer le marqueur du complexe purifié, par exemple par l'addition d'un agent capable de cliver le lieu. Un certain nombre de lieux clivables différents sont connus de l'homme du métier. De tels lieux peuvent être clivés, par exemple, par irradiation d'une liaison photolabile ou hydrolyse catalysée par un acide. Il existe également des lieux polypeptidiques qui incorporent un site de reconnaissance de protéase et qui peuvent être clivés par l'addition d'une enzyme protéase appropriée.

[0065] Dans d'autres modes de réalisation, il peut être souhaitable d'avoir les protéines gH, gL, pUL128, pUL130, pUL131 (ou l'un de leurs fragments) qui ne comprennent pas de séquence marqueur exogène, par exemple, pour des raisons de sécurité clinique ou d'efficacité.

[0066] Bien qu'il soit parfois fait référence aux protéines gH, gL, pUL130 comme à des glycoprotéines, cette nomenclature ne devra pas être prise comme signifiant que ces protéines doivent être glycosylées lorsqu'elles sont utilisées avec l'invention. Alors qu'il a été fait référence ci-dessus à des souches spécifiques, il faudra comprendre que les protéines du CMV gH, gL, pUL128, pUL130, pUL131 (ou leurs fragments) de souches différentes du CMV peuvent être utilisées. A titre d'exemple non limitatif, les autres souches du CMV peuvent comprendre les souches Towne, Toledo, AD169, Merlin, TB20, et VR1814.

B. Acide nucléique codant pour les protéines et les complexes du CMV

[0067] Sont également proposés ici des acides nucléiques codant pour les protéines et les complexes du CMV pour une intégration génomique, et l'expression subséquente du pentamère du CMV.

[0068] Une ou plusieurs constructions d'acide nucléique codant pour les protéines et les complexes du CMV décrits ici peuvent être utilisées pour une intégration génomique. Par exemple, une seule construction d'acide nucléique codant pour les cinq sous-unités, gH, gL, pUL128, pUL130, pUL131 (ou leurs fragments), peut être introduit dans une cellule hôte. En variante, les séquences codant pour les cinq sous-unités (ou leurs fragments) peuvent être portées par deux constructions d'acide nucléique ou plus, qui sont ensuite introduites dans la cellule hôte simultanément ou séquentiellement.

[0069] Par exemple, dans un exemple de mode de réalisation, l'invention propose une seule construction d'acide nucléique codant pour : l'ectodomaine des gH, gL, pUL128, pUL130, et pUL131. Dans un autre exemple de mode de réalisation, l'invention propose deux constructions d'acide nucléique codant pour : l'ectodomaine des gH, gL, pUL128, pUL130, et pUL131. Voir, figure 1. Dans les deux exemples, il a été obtenu une intégration génomique couronnée de succès.

[0070] La construction d'acide nucléique peut comprendre de l'ADN génomique qui comprend un ou plusieurs introns, ou de l'ADNc. Certains gènes sont exprimés plus efficacement lorsque des introns sont présents. La séquence génomique native codant pour la pUL128 comprend deux introns, la séquence génomique native codant pour la pUL131 comprend un exon, tandis que la séquence génomique native codant pour la pUL130 ne comprend aucun intron. La séquence génomique native codant pour la pUL128 comprend trois exons, la séquence génomique native codant pour la pUL131 comprend deux exons, et la séquence génomique native codant pour la pUL130 comprend un exon. En particulier, la séquence d'acide nucléique est appropriée pour l'expression de polypeptides exogènes dans ladite cellule de mammifère.

[0071] Sont également proposés ici, des vecteurs qui comprennent des séquences codant pour les gH, gL, pUL128, pUL130, et/ou pUL131 (ou l'un de leurs fragments). Les exemples de vecteurs comprennent des plasmides qui sont capables de se répliquer de façon autonome ou d'être répliqués dans une cellule de mammifère. Les vecteurs d'expression types contiennent des promoteurs, des amplificateurs, et des terminateurs appropriés qui sont utiles pour la régulation de l'expression de la ou des séquences codantes dans la construction d'expression. Les vecteurs peuvent également comprendre des marqueurs de sélection pour fournir un trait phénotypique pour la sélection des cellules hôtes

transformées (comme la transmission d'une résistance à des antibiotiques tels que l'ampicilline ou la néomycine).

[0072] Les promoteurs appropriés comprennent, par exemple, le promoteur du CMV, un adénovirus, EF1 α , le promoteur de la GAPDH métallothionéine, le promoteur précoce du SV-40, le promoteur tardif du SV-40, le promoteur du virus de la tumeur mammaire de la souris, le promoteur du virus du sarcome de Rous, le promoteur de la polyhédriane, etc. Les promoteurs peuvent être constitutifs ou inductibles. Un ou plusieurs vecteurs peuvent être utilisés (par exemple, un vecteur codant pour les cinq sous-unités ou des fragments de celles-ci, ou deux vecteurs ou plus ensemble codant pour les cinq sous-unités ou des fragments de celles-ci) ; voir, par exemple, la figure 1.

[0073] Lorsque la cellule hôte est une cellule CHO, le promoteur, l'amplificateur, ou le terminateur est actif dans les cellules CHO. Un promoteur fréquemment utilisé est le promoteur pour le gène immédiat précoce (IE) du cytomégalovirus humain (hCMV). Le promoteur de ce gène dirige des taux élevés d'expression de transgène dans une large diversité de types cellulaires. L'activité de ce promoteur dépend d'une série de répétitions imparfaites de 17, 18, 19, et 21 pb, dont certaines lient des facteurs de transcription des familles de la protéine de liaison sensible à l'AMPc (CREB) NF- κ B et du facteur nucléaire 1.

[0074] Un promoteur fort est un promoteur qui provoque l'initiation des ARNm à une fréquence élevée égale ou supérieure à celle du fragment promoteur/amplificateur central du hCMV (décrit dans le brevet US No. 5 168 062) dans une cellule CHO. Un tel promoteur peut être un promoteur fort dépendant du type cellulaire, comme ceux décrits dans le brevet US No. 5 589 392, ou un promoteur fort actif de façon ubiquitaire. Les exemples de promoteurs viraux actifs de façon constitutive comprennent, par exemple, les promoteurs précoce

et tardif du virus SV40, le promoteur immédiat précoce du cytomégalovirus humain (hCMV) ou du cytomégalovirus murin (mCMV), le promoteur de la thymidine kinase (TK) ou le virus Herpes Simplex, ou le promoteur des longues répétitions terminales du virus du sarcome de Rous (RS-LTR). D'autres exemples comprennent, par exemple, le promoteur MIE du hCMV tel que défini par le fragment Pst I de 2,1 kb décrit dans le brevet US No. 5 385 839 et/ou dans le document EP-323 997-A1 ou une partie fonctionnelle de celui-ci possédant une activité de promoteur.

[0075] Des séquences de sites internes d'entrée du ribosome (IRES) et de peptide 2A peuvent être également utilisées. Les IRES et peptide 2A fournissent des variantes d'approches pour la coexpression de séquences multiples. IRES est une séquence nucléotidique qui permet l'initiation de la traduction dans le milieu de la séquence d'un ARN messager (ARNm) faisant partie du plus grand processus de synthèse des protéines. Habituellement, chez les eucaryotes, la traduction peut être initiée uniquement à l'extrémité 5' de la molécule d'ARNm. Les éléments IRES permettent l'expression de gènes multiples dans un transcrit. Des vecteurs polycistroniques à base d'IRES, qui expriment des protéines multiples à partir d'un transcrit, peuvent réduire l'échappement des clones non exprimant de la sélection.

[0076] Le peptide 2A permet la traduction de protéines multiples dans un seul cadre de lecture ouvert dans une polyprotéine qui est ensuite clivée en protéines individuelles par l'intermédiaire d'un mécanisme de saut ribosomique. Le peptide 2A peut fournir une expression plus équilibrée de multiples produits protéiniques.

[0077] Les exemples de séquences d'IRES comprennent, par exemple, les IRES de l'EV71, les IRES de l'EMCV, les IRES DU VHC.

[0078] Pour l'intégration génomique, l'intégration peut être spécifique du site ou aléatoire. Une recombinaison spécifique du site peut être obtenue par introduction d'une ou de plusieurs séquences homologues dans les constructions d'acide nucléique décrites ici. Une telle séquence homologue correspond potentiellement à la séquence endogène au niveau d'un site cible spécifique dans le génome de l'hôte. En variante, une intégration aléatoire peut être utilisée. Parfois, le taux d'expression d'une protéine peut varier selon le site d'intégration. Par conséquent, il peut être souhaitable de sélectionner un certain nombre de clones selon le taux d'expression de la protéine recombinante pour identifier un clone qui atteint le taux d'expression souhaité.

3. Cellules hôtes

[0079] Dans un autre aspect, l'invention propose des cellules hôtes dans lesquelles les séquences codant pour le pentamère sont intégrées de façon stable dans le génome des cellules hôtes, et, lorsqu'elles sont cultivées dans des conditions appropriées, elles expriment le pentamère du CMV tel que divulgué ici. Dans certains modes de réalisation, la cellule hôte est une cellule de mammifère. Dans certains modes de réalisation, la cellule hôte est une cellule de rongeur.

[0080] Les exemples de lignées cellulaires de rongeurs comprennent par exemple, les lignées cellulaires de rein du hamster bébé (BHK) (par exemple, BHK21, BH TK), de Sertoli de souris (TM4), de foie du rat buffalo (BRL 3A), de tumeur mammaire de la souris (MMT), d'hépatome du rat (HTC), de myélome de la souris (NS0), d'hybridome murin (Sp2/0), de thymome de la souris (EL4), d'ovaire du hamster de Chine (CHO) et des dérivés des cellules CHO, embryonnaires murines (NIH/3T3, 3T3 Li), myocardiques du rat (H9c2), de myoblastes de la souris (C2C12), et de rein de la souris (miMCD-3).

[0081] Dans un mode de réalisation, la cellule de rongeur est une cellule CHO. Les cellules CHO appropriées comprennent, par exemple, les lignées DUXB11 et DG44. Ces deux lignées cellulaires sont déficientes en activité dihydrofolate réductase (DHFR), et par conséquent dépendantes d'une source exogène de précurseurs nucléotidiques pour leur croissance. La déficience en DHFR est un phénotype facilement manipulé approprié pour sélectionner l'intégration génomique et l'expression stable d'ADN exogène. L'intégration génomique est accomplie par transfection des cellules avec des cassettes d'expression pour le gène d'intérêt et un gène DHFR. Après la transfection, les cellules sont placées dans un milieu de sélection auquel il manque les précurseurs nucléotidiques.

[0082] L'expression de protéines de recombinaison dans des lignées cellulaires déficientes en DHFR peut être en outre amplifiée par l'addition de méthotrexate (MTX) aux cultures, de telle façon qu'il est possible de sélectionner un nombre élevé de copies du vecteur d'expression introduit. Le MTX est un inhibiteur compétitif de l'enzyme DHFR. L'application de cette pression de sélection supplémentaire en plus de l'absence des précurseurs nucléotidiques permet la sélection et l'isolement de la population mineure de cellules qui ont subi une amplification spontanée du vecteur d'expression intégré contenant le marqueur sélectionnable DHFR et, dans la plupart des cas, le gène d'intérêt. La présence de copies multiples du gène permet d'obtenir un taux élevé d'expression des protéines exogènes. En variante, la sélection par le MTX peut être réalisée indépendamment de la déficience en DHFR (c'est-à-dire, l'utilisation de MTX pour sélectionner une cellule hôte qui est à l'origine compétente en DHFR), comme cela est exemplifié dans les exemples divulgués ici.

[0083] Une autre lignée cellulaire CHO appropriée est la lignée cellulaire CHO-K1 de type sauvage, et son dérivé CHO-K1SV.

[0084] Un procédé de sélection couramment utilisé pour les lignées cellulaires CHO-K1 consiste en la sélection par la glutamine synthétase (GS). En l'absence d'une source exogène de glutamine, la survie cellulaire dépend de l'enzyme GS pour produire la glutamine. Avec des lignées de cellules hôtes telles que les cellules NS/O dérivées de myélome murin et les cellules CHO, qui présentent une activité enzymatique GS endogène relativement basse, le procédé permet un schéma de sélection simple lors de l'utilisation d'un marqueur sélectionnable GS dans le vecteur d'expression et d'un milieu de sélection dépourvu de glutamine. Similaire au système DHFR/MTX, l'inhibiteur compétitif de la GS, le méthionine sulfoximine (MSX) peut être ajouté au milieu pour appliquer une pression supplémentaire et sélectionner les cellules CHO qui dirigent des taux élevés d'expression à partir du vecteur intégré.

[0085] Les cellules CHO-K1, ou toutes autres cellules CHO couramment utilisées, peuvent être également sélectionnées en se basant sur la déficience en DHFR comme il a été décrit ci-dessus. Par exemple, une cellule CHO-K1, ou tout autre type de cellule CHO, peut présenter une déficience en DHFR, comme une délétion dans laquelle au moins une copie de la séquence génomique du gène de la dihydrofolate réductase (DHFR), ou au moins 30 % (par exemple, au moins 40 %, au moins 50 %, au moins 60 %, au moins 70 %, au moins 80 %, au moins 90 %, ou 100 %) de la séquence codante dudit gène DHFR, est délétée. D'autres moyens pour introduire une déficience en DHFR comprennent la création de mutations dans le gène DHFR endogène. Les lignées cellulaires peuvent être en outre amplifiées par l'addition de méthotrexate (MTX) aux cultures comme il a été décrit ci-dessus.

[0086] Les cellules CHO-K1, ou toutes autres cellules CHO couramment utilisées, peuvent être également sélectionnées en se basant sur le MTX, avec ou sans déficience en DHFR. Dans

les exemples fournis ici, les cellules CHO ont été sélectionnées en se basant sur le MTX, sans déficience en DHFR (c'est-à-dire que la cellule CHO d'origine utilisée pour l'intégration génomique est compétente pour la DHFR). Dans un tel système, généralement, le nombre de copies des séquences exogènes (par exemple, les séquences codant pour les protéines du CMV) est généralement bas. Il est estimé que les lignées cellulaires dans les exemples décrits ici comportent environ 1 à 10 copies des séquences exogènes codant pour les protéines du CMV, au niveau d'un nombre très limité de sites d'intégration (par exemple, 1 ou 2 sites d'intégration). En général, lorsqu'une lignée cellulaire déficiente en DHFR est utilisée, le nombre de copies des séquences exogènes est généralement de beaucoup supérieur, parfois aussi élevé que quelques centaines de copies. On s'attend à ce que les deux procédés soient appropriés pour la production des lignées cellulaires CHO divulguées ici, bien que lorsque le nombre de copies est élevé, la cellule hôte puisse perdre une ou plusieurs copies des séquences exogènes durant le repiquage et/ou l'expansion de la lignée cellulaire.

[0087] D'autres souches des cellules CHO appropriées pour l'invention décrite ici comprennent, par exemple, des cellules CHO-ICAM-1, et des cellules CHO-hIFN γ . Ces cellules génétiquement modifiées permettent une insertion stable d'ADN recombinant dans un gène spécifique ou une région d'expression des cellules, l'amplification de l'ADN inséré, et la sélection des cellules présentant un taux élevé d'expression de la protéine recombinante.

[0088] Les exemples de lignées cellulaires CHO disponibles à l'European Collection of Cell Cultures (ECACC) sont énumérés dans le tableau 1. Toutes les cellules CHO énumérées dans le tableau 1 peuvent être utilisées.

Tableau 1

BE2016/5877

Nom de la lignée cellulaire	Mots-clés
CHO	Ovaire de hamster de Chine
CHO (dépourvue de protéines)	Ovaire de hamster de Chine
CHO-CHRM1	Récepteur de l'acétylcholine muscarinique M1 humain, CHRM1, récepteur couplé aux protéines G, GPCR, transfectée, InSCREENeX SCREENflex™, hôte CHO-K1.
CHO-CHRM2	Récepteur de l'acétylcholine muscarinique M2 humain, CHRM2, récepteur couplé aux protéines G, GPCR, transfectée, InSCREENeX SCREENflex™, hôte CHO-K1.
CHO-CHRM5	Récepteur de l'acétylcholine muscarinique M5 humain, CHRM5, récepteur couplé aux protéines G, GPCR, transfectée, InSCREENeX SCREENflex™, hôte CHO-K1.
CHO-CNR1	Récepteur cannabinoïde I humain, CNR1, ID du gène 1268, récepteur couplé aux protéines G, GPCR, transfectée, InSCREENeX SCREENflex™, hôte CHO-K1.
CHO-FFAR2	Récepteur 2 des acides gras libres humain, FFAR2, récepteur couplé aux protéines G, GPCR, transfectée, InSCREENeX SCREENflex™, hôte CHO-K1.
CHO-GPR120	Récepteur GPR120 (orphelin) humain, GPR120, récepteur couplé aux protéines

	G, GPCR, transfectée, InSCREENeX SCREENflex™, hôte CHO-K1.	BE2016/5877
CHO-K1	Ovaire de hamster de Chine	
CHO-K1-AC-free	Ovaire de hamster de Chine, dépourvue de sérum	
CHO-K1/SF	Ovaire de hamster de Chine (adaptée au MEM)	
CHO-NPY1R	Récepteur du neuropeptide Y humain, NPY1R, ID du gène 4886, récepteur couplé aux protéines G, GPCR, transfectée, InSCREENeX SCREENflex™, hôte CHO-K1.	
CHO-OPRL1	Récepteur des opiacés de type 1 humain, OPRL1, récepteur couplé aux protéines G, GPCR, transfectée, InSCREENeX SCREENflex™, hôte CHO-K1.	
CHO-SSTR1	Récepteur 1 de la somatostatine humain, SSTR1, récepteur couplé aux protéines G, GPCR, transfectée, InSCREENeX SCREENflex™, hôte CHO-K1.	
CHO/dhFr-	Ovaire de hamster de Chine	
CHO/dhFr-AC-free	Ovaire de hamster de Chine, dépourvue de sérum	
RR-CHOKI	Ovaire de hamster de Chine	
T02J-10/10 (CHO-GCGR (GCGR))	Récepteur du glucagon humain, GCGR, récepteur couplé aux protéines G, GPCR, transfectée, InSCREENeX SCREENflex™, hôte CHO-K1.	

[0089] Diverses lignées cellulaires CHO sont également disponibles auprès de l'American Type Culture Collection (ATCC), comme les lignées cellulaires CHO hCBEl1 (ATCC® PTA-3357™), E77.4 (ATCC® PTA-3765™), hLT-B : R-hG1 CHO No. 14 (ATCC® CRL-11965™), MOR-CHO- MORAb-003-RCB (ATCC® PTA-7552™),

AQ.C2 clone 11B (ATCC® PTA-3274™), AQ.C2 clone 11B (ATCC® PTA-3274™), hsAQC2 dans CHO-DG44 (ATCC® PTA-3356™), xrs5 (ATCC® CRL-2348™), CHO-K1 (ATCC® CCL-61™), Lec1 [nommée à l'origine 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™), et pgsB-618 (ATCC® CRL-2241™). L'une quelconque de ces lignées cellulaires CHO peut être utilisée.

10 **[0090]** D'autres lignées cellulaires CHO disponibles dans le commerce comprennent, par exemple, les cellules FreeStyle™ CHO-S et la lignée cellulaire Flp-In™-CHO de Life Technologies.

15 **[0091]** Des procédés pour exprimer des protéines recombinantes dans des cellules CHO en général ont été divulgués. Voir, par exemple, dans les brevets US No. 4 816 567 et No. 5 981 214.

20 **[0092]** En plus des cellules CHO, d'autres cellules de mammifères peuvent être également utilisées en tant qu'hôtes. Les exemples de cellules de rongeurs comprennent des cellules BHK21, des cellules NS0, des cellules Sp2/0, des cellules EL4, des cellules NIH/3T3, des cellules 3T3-L1, des cellules ES-D3, des cellules H9c2, des cellules C2C12, YB2/0, des cellules mimcd 3, etc. Les exemples de cellules humaines comprennent : des cellules SH-SY5Y, des cellules IM 32, des cellules LAN, des cellules MCF10A, des cellules 293T, des cellules SK-BR3, des cellules huvec, des cellules huasmc, des cellules HKB-1, des cellules hmsc, des cellules U293, des cellules HE 293, des cellules PERC6®, des cellules Jurkai, 25 des cellules HT-29, des cellules Incap.FGC, des cellules A549, des cellules MDA MB453, des cellules hepg2, des cellules THP-1, des cellules bxpc-3, des cellules Capan-1, des cellules DU145, et des cellules PC-3.

[0093] Par exemple, la lignée cellulaire PERC60, les cellules NS0 de myélome de souris, les cellules de rat de hamster bébé (BHK), et la lignée cellulaire de rein embryonnaire humain (HEK293) ont reçu une approbation réglementaire pour la production de protéines recombinantes.

[0094] Les exemples de lignées cellulaires de primates non humains utiles dans des procédés fournis ici comprennent les lignées cellulaires de rein de singe (CVI-76), de rein de singe vert d'Afrique (VERO-76), de fibroblastes de singe vert (COS-1), et des cellules de rein de singe (CVI) transformées par le SV40 (COS-7). D'autres lignées cellulaires de mammifères sont connues d'une personne ayant des compétences moyennes dans le domaine et sont cataloguées à l'American Type Culture Collection (Manassas, VA).

[0095] Dans certains modes de réalisation, les cellules hôtes sont appropriées pour une croissance dans des cultures en suspension. Les cellules hôtes compétentes en suspension sont généralement monodispersées ou se développent en agrégats lâches sans agrégation substantielle. Les cellules hôtes compétentes en suspension comprennent des cellules qui sont appropriées pour une culture en suspension sans adaptation ni manipulation (par exemple, des cellules hématopoïétiques, des cellules lymphoïdes) et des cellules qui ont été rendues compétentes en suspension par modification ou adaptation de cellules dépendantes de la fixation (par exemple, des cellules épithéliales, des fibroblastes).

[0096] Dans certains modes de réalisation, la cellule hôte est une cellule dépendante de la fixation qui est cultivée et maintenue dans une culture adhérente. Les exemples de lignées cellulaires adhérentes humaines utiles dans des procédés fournis ici comprennent les lignées cellulaires de neuroblastome humain (SH-SY5 Y, IMR32, et LANS), de carcinome du col de l'utérus humain (HeLa), de l'épithélium de sein

humain (MCF10A), de rein embryonnaire humain (293T), et de BE2016/5877
carcinome du sein humain (SK-BR3).

Gènes C12orf35 et protéines C12orf35

5 **[0097]** Dans certains modes de réalisation, la cellule
hôte est une cellule dans laquelle le taux d'expression ou
l'activité de la protéine C12orf35 est réduit,
comparativement à un témoin. Dans un certain mode de
réalisation, une telle cellule est une cellule CHO. La demande
10 de brevet US provisoire No. 61/919 313, déposée le 20 décembre
2013, et incorporée ici en référence, fournit une description
détaillée de cellules de mammifères dans lesquelles le taux
d'expression ou l'activité de la protéine C12orf35 est réduit
comparativement à un témoin.

15 **[0098]** Divers témoins peuvent être utilisés. Le taux
d'expression ou l'activité de la protéine C12orf35 à partir
d'une cellule de type sauvage correspondante peut être utilisé
en tant que témoin. En variante, un témoin peut être un taux
prédéterminé ou un taux seuil qui peut être identifié dans la
20 littérature ou une base de données.

[0099] Le gène C12orf35 humain se rapporte à la séquence
nucléotidique codant pour le cadre de lecture ouvert 35 du
chromosome 12. La protéine C12orf35 codée n'est pas
caractérisée. On pense que l'homologue chez *Cricetulus*
25 *griseus* (hamster de Chine) du gène C12orf35 humain, nommé
Kiaa1551, est localisé sur le chromosome 8. On pense que
l'homologue chez *Mus musculus* du gène C12orf35 humain est
localisé sur le chromosome 6. L'ID de gène pour le gène CHO
C12orf35 est publié comme l'ID de gène GenBank No. 100762086 ;
30 et pour le gène C12orf35 humain, il est publié comme l'ID de
gène GenBank No. 55196. Les informations concernant le gène
C12orf35 chez *Cricetulus griseus*, la séquence codante et la
protéine C12orf35 prédite sont également disponibles à la
GenBank par le numéro d'accèsion NCBI : XM_003512865.

[0100] Le gène C12orf35 est exprimé de façon endogène dans des cellules eucaryotes telles que, par exemple, d'espèces de mammifères comme l'être humain, la souris et le hamster. La protéine prédite codée par le gène C12orf35 est une grosse protéine dépassant 1500 résidus. La liste des séquences présente des exemples de séquences d'acides aminés ou de séquences d'acides aminés putatives de la protéine codée par le gène C12orf35 endogène de différentes espèces de mammifères comme le hamster (SEQ ID NO : 21 et 22), l'être humain (SEQ ID NO : 23 et 24), et la souris (SEQ ID NO : 35). La CDS (séquence d'ADN codante) de C12orf35 du hamster de Chine est représentée par SEQ ID NO : 25. En outre, une section de l'UTR en 5' (voir SEQ ID NO : 26) et de l'UTR en 3' (voir SEQ ID NO : 27) de l'ARNm de C12orf35 du hamster de Chine a été séquencée.

[0101] Chez l'être humain, le gène C12orf35 est également appelé KIAA1551. Le gène C12orf35 est également appelé de type C12orf35 ou homologue de C12orf35 chez le hamster ou 2810474019Rik chez la souris. Différents noms peuvent être alloués dans différentes espèces pour la protéine ou le gène et des variantes non limitantes de noms (alias) sont également énumérées ci-dessus dans le tableau 2. Pour simplifier, dans cette divulgation, les homologues et les orthologues de différentes espèces sont tous appelés « gène C12orf35 » ou « protéine C12orf35 ».

[0102] A l'encontre de ce contexte scientifique, il est surprenant et inattendu que lorsque le taux d'expression ou l'activité de la protéine C12orf35 est réduit comparativement à un témoin (par exemple, par délétion du gène C12orf35, ou par introduction de mutations), le rendement de la protéine recombinante est significativement amélioré. En tant que telles, des cellules hôtes de mammifères (par exemple, des cellules CHO) avec un taux d'expression ou une activité réduit

de la protéine C12orf35 sont particulièrement appropriées pour la production recombinante du complexe pentamère.

[0103] La réduction du taux d'expression ou de l'activité d'une protéine C12orf35 peut être obtenue par divers moyens.

5 Par exemple, le taux d'expression ou l'activité d'une protéine C12orf35 peut être réduit par inactivation de gène, mutation de gène, délétion de gène, silençage de gène, ou une combinaison quelconque des précédents. L'inactivation de gène est une technique génétique par laquelle un gène est rendu
10 inopérant par perturbation de sa fonction. Par exemple, un acide nucléique peut être inséré dans la séquence codante, perturbant de cette façon la fonction du gène. En outre, le gène C12orf35 pleine longueur (ou l'un de ses fragments) peut être délété, grâce à quoi l'expression de la protéine C12orf35 fonctionnelle est substantiellement éliminée. Par exemple, la
15 délétion peut être d'au moins 30 %, d'au moins 40 %, d'au moins 50 %, d'au moins 55 %, d'au moins 60 %, d'au moins 65 %, d'au moins 70 %, d'au moins 75 %, d'au moins 80 %, d'au moins 85 %, d'au moins 90 %, d'au moins 95 %, ou de 100 % de la séquence codante du gène C12orf35. Une autre option est d'introduire une ou plusieurs mutations dans la séquence
20 codante, ce qui rend une protéine C12orf35 non fonctionnelle ou moins fonctionnelle. Par exemple, une ou plusieurs mutations de décalage du cadre peuvent être introduites, produisant une protéine C12orf35 non fonctionnelle ou moins
25 fonctionnelle. En variante ou en outre, un ou plusieurs codons d'arrêt peuvent être introduits dans la séquence codante si bien qu'il est obtenu une protéine tronquée, non fonctionnelle ou moins fonctionnelle. D'autres options comprennent, mais
30 n'y sont pas limitées, une ou plusieurs mutations dans le promoteur, dans l'UTR en 5' et/ou en 3' ou d'autres éléments régulateurs, par exemple par introduction d'une délétion du promoteur ou par introduction d'une construction entre le promoteur et le départ de la transcription. Les procédés pour

la perturbation de gènes afin de supprimer ou d'éliminer l'expression du gène cible sont également bien connus de l'homme du métier.

[0104] Puisque chaque cellule comporte deux copies du gène C12orf35 dans son génome, dans certains modes de réalisation, au moins une copie de la séquence génomique du gène C12orf35, ou au moins 50 % de la séquence codante dudit gène C12orf35, est délétée. Dans certains modes de réalisation, les deux copies des séquences génomiques du gène C12orf35 (ou au moins 50 % de la séquence codante dudit gène C12orf35 à partir de chaque copie) sont délétées.

[0105] Dans certains modes de réalisation, la séquence délétée comprend une portion de la région télomérique du chromosome 8 d'une cellule CHO. Une région télomérique est une région de séquences nucléotidiques répétitives à chaque extrémité d'une chromatide, qui protège l'extrémité du chromosome contre la détérioration ou contre la fusion avec des chromosomes voisins. Dans certains modes de réalisation, au moins 30 %, au moins 40 %, au moins 50 %, au moins 55 %, au moins 60 %, au moins 65 %, au moins 70 %, au moins 75 %, au moins 80 %, au moins 85 %, au moins 90 %, au moins 95 %, ou 100 % de la séquence nucléotidique de la région télomérique du chromosome 8 d'une cellule CHO est délétée.

[0106] Dans certains modes de réalisation, la séquence délétée comprend en outre la délétion d'un gène choisi dans le groupe constitué de : Bicd1, Amn1, protéine 20 de type méthyltransférase, Dennd5b, FAM60A, Caprin2, Ipo8, RPS4Y2, et l'une de leurs combinaisons.

[0107] Dans certains modes de réalisation, la délétion du gène C12orf35 (ou de l'un de ses fragments) est due à une cassure chromosomique. Une cassure chromosomique peut être induite par exemple en traitant les cellules eucaryotes avec un agent toxique qui favorise la cassure chromosomique, comme par exemple le MTX, l'aphidicoline ou l'hygromycine. D'autres

options pour induire des cassures chromosomiques comprennent, mais n'y sont pas limitées, des radiations, des irradiations, des mutagènes, des substances cancérigènes et la bléomycine. Des cassures chromosomiques peuvent également se produire spontanément durant la transfection, par exemple, l'électroporation. Des procédés pour induire une cassure chromosomique sont également connus de l'homme du métier et ainsi, n'ont pas besoin d'une description détaillée quelconque ici. Après l'induction d'une cassure chromosomique, les cellules eucaryotes comportant le point de rupture souhaité (qui produit une délétion du gène C12orf35, ou de l'un de ses fragments) peuvent être identifiées, par exemple, en analysant l'ADN ou en utilisant le procédé selon le cinquième aspect de la présente divulgation. Par exemple, le profil d'expression des cellules traitées peut être analysé pour déterminer si le gène C12orf35 ou les gènes localisés en position centromérique du gène C12orf35 sont exprimés, si l'expression est réduite ou si les gènes ne sont pas exprimés. Par exemple, dans le cas de cellules de souris ou de hamster, il peut être analysé si le gène C12orf35 est exprimé et en variante ou en outre à ceci, il peut être analysé si un ou plusieurs gènes choisis dans le groupe constitué de la protéine 20 de type méthyltransférase, Dennd5b, FAM60A, Caprin2, Ipo8, Tmtc1 ou des gènes qui sont localisés en position télomérique des gènes susmentionnés (où télomérique a cet égard signifie dans la direction de l'extrémité télomérique) sont exprimés par la cellule et/ou si l'expression est réduite ou substantiellement éliminée.

[0108] La réduction du taux d'expression de la protéine C12orf35 peut être obtenue par un silençage du gène après la transcription, par exemple, par des molécules d'acide nucléique antisens, ou des molécules qui médient une interférence par ARN. Des exemples non limitants comprennent

des ARNsi, des ARNsh, des miARN, des oligonucléotides antisens, etc., tous étant bien connus dans l'art.

[0109] Le taux d'expression de la protéine C12orf35 peut être estimé par des procédés connus dans l'art, par exemple, en mesurant le taux d'ARNm codant pour la protéine C12orf35, ou la protéine C12orf35 elle-même. De tels procédés comprennent, par exemple, les analyses Northern blot, FACS, ImageStream, Western blot, qPCR, RT-PCR, qRT-PCR, ELISA, Luminex, Multiplex, etc.

[0110] Dans certains modes de réalisation, le taux d'expression ou l'activité de la protéine C12orf35 est réduit d'au moins 3 fois, d'au moins 5 fois, d'au moins 10 fois, d'au moins 20 fois, d'au moins 30 fois, d'au moins 40 fois, d'au moins 50 fois, d'au moins 60 fois, d'au moins 70 fois, d'au moins 75 fois, d'au moins 80 fois, d'au moins 90 fois, d'au moins 100 fois, comparativement à un témoin.

Gènes FAM60A et protéines FAM60A

[0111] Dans certains modes de réalisation, la cellule hôte est une cellule dans laquelle le taux d'expression ou l'activité de la protéine FAM60A est réduit, comparativement à un témoin. Dans un certain mode de réalisation, une telle cellule est une cellule CHO. La demande de brevet US provisoire No. 61/919 340, déposée le 20 décembre 2013, et incorporée ici en référence, fournit une description détaillée de cellules de mammifères dans lesquelles le taux d'expression ou l'activité de la protéine FAM60A est réduit.

[0112] Divers témoins peuvent être utilisés comme il a été discuté ci-dessus. Le taux d'expression ou l'activité de la protéine FAM60A à partir d'une cellule de type sauvage correspondante peut être utilisé en tant que témoin. En variante, un témoin peut être un taux prédéterminé ou un taux seuil qui peut être identifié dans la littérature ou une base de données.

[0113] La protéine FAM60A est une sous-unité du complexe SIN3-histone désacétylase (HDAC) (complexe SIN3/HDAC) qui fonctionne dans la répression de la transcription (Munoz *et al.*, 2012, THE Journal of Biological Chemistry VOL. 287, NO. 39, pp. 32346-32353 ; Smith *et al.*, 2012, Mol Cell Proteomics 11 (12): 1815-1828). Les histone désacétylases (HDAC) catalysent l'élimination des groupes acétyle à partir des histones. L'acétylation des histones sur les lysines est un mécanisme majeur pour la modulation de la conformation de la chromatine. L'acétylation des histones favorise un état de la chromatine relâché, transcriptionnellement actif tandis que la désacétylation catalysée par les histone désacétylases (HDAC) favorise un état inactif silencieux. L'analyse de bases de données a révélé la présence d'au moins un orthologue de la FAM60A chez la plupart des métazoaires, mais pas chez les nématodes. Le gène FAM60A est conservé chez les métazoaires et peut être trouvé dans les génomes de tous les vertébrés et de la plupart des invertébrés qui ont été complètement séquencés. La recherche de similarité de séquence des homologues de la FAM60A indique que principalement, il n'existe qu'un seul membre représentatif de cette famille dans le génome. Il n'existe que quelques exceptions. Selon Smith *et al.*, 2012, la protéine FAM60A a une séquence unique à laquelle il manque n'importe quel domaine connu de protéine. En outre, il a été décrit par Smith *et al.* 2012, qu'elle ne présente aucune homologie de séquence avec d'autres protéines connues dans le protéome humain. La comparaison des séquences entre les protéines FAM60A provenant de différentes espèces a montré que la protéine FAM60A comprend généralement trois régions : (1) une extrémité N-terminale comprenant des segments hautement conservés chez tous les métazoaires ; (2) une région moyenne qui est hautement conservée parmi les vertébrés tandis que chez les invertébrés elle est constituée d'un espaceur non conservé d'une longueur variable ; (3) une

extrémité C-terminale comprenant des segments hautement conservés chez tous les métazoaires. Ainsi, la conservation la plus élevée a été observée dans les régions N- et C-terminales de la FAM60A.

5 **[0114]** Des études indiquent que la protéine FAM60A s'associe aux complexes SIN3/HDAC dans divers types de cellules eucaryotes comme en particulier des cellules de mammifères. Toutefois, jusqu'à ce jour, les informations fonctionnelles concernant la protéine FAM60A sont limitées.

10 Des études fonctionnelles récentes (voir Smith *et al.*, 2012) indiquent que la protéine FAM60A peut réprimer l'expression génique et régule un sous-ensemble spécifique de gènes. Smith *et al.* 2012 rapportent un rôle de la protéine FAM60A dans la régulation de la voie de signalisation du TGF-bêta, qui joue

15 un rôle pivot dans des processus comme la progression d'un cancer, les métastases, la migration cellulaire et la surveillance immunitaire. Il existe des découvertes indiquant que la protéine FAM60A agit comme un répresseur de la transcription des composants de la voie de signalisation du

20 TGF-bêta tandis que cette fonction de la protéine FAM60A semble être permise par l'intermédiaire de son rôle dans le complexe SIN3-HDAC. L'appauvrissement de la protéine FAM60A dans différentes lignées de cellules cancéreuses en utilisant des ARNsi contre la séquence codante de la FAM60A a produit

25 un changement de la morphologie normale des cellules cancéreuses. En outre, il a été trouvé que les taux de protéine FAM60A changent périodiquement au cours du cycle cellulaire dans des cellules U2OS (Munoz *et al.*, 2012). Des expériences de répression de la FAM60A en utilisant des ARNsi

30 dirigés contre la FAM60A dans des cellules d'ostéosarcome osseux humain U2OS ont révélé que la protéine FAM60A restreint l'expression du gène de la cycline D1.

[0115] A l'encontre de ce contexte scientifique, il a été trouvé de façon surprenante que la réduction de l'expression

ou de l'activité de la protéine FAM60A dans une cellule de mammifère augmente la stabilité de l'expression des protéines recombinantes, sans affecter négativement d'autres caractéristiques de la cellule qui sont importantes pour l'expression recombinante. Cette corrélation entre les effets de la protéine FAM60A et la stabilité de l'expression durant une culture prolongée des cellules était inattendue. En tant que telles, des cellules hôtes de mammifères (par exemple, des cellules CHO) avec un taux d'expression ou une activité réduit de la protéine FAM60A sont particulièrement appropriées pour la production recombinante du complexe pentamère.

[0116] Le gène FAM60A est exprimé de façon endogène chez les métazoaires et par conséquent dans des espèces de mammifères telles que l'être humain, la souris, le rat et le hamster, et la séquence d'acides aminés de la FAM60A est hautement conservée dans les espèces de mammifères ainsi que chez les vertébrés. Différents noms peuvent être alloués dans différentes espèces pour la protéine ou le gène et des variantes de noms non limitantes (alias) sont également énumérées dans le tableau 2 (ci-dessous). Pour simplifier, dans cette divulgation, les homologues et les orthologues de différentes espèces sont tous appelés « gène FAM60A » ou « protéine FAM60A ».

[0117] La liste des séquences présente des exemples de séquences d'acides aminés de protéines FAM60A connues et/ou prédites de différentes espèces de vertébrés, notamment Homo sapiens (SEQ ID NO : 28), Mus musculus (SEQ ID NO : 29), Cricetulus griseus (SEQ ID NO : 30). L'ADNc de FAM60A prédit de Cricetulus griseus est représenté par SEQ ID NO : 31 (séquence codante de 14 à 679 ; voir également la séquence de référence du NCBI : XM_003505482.1). La protéine FAM60A n'a pas été décrite en détail dans la littérature. Ainsi, il était surprenant que la stabilité de l'expression d'une cellule

hôte recombinante puisse être améliorée, si le génome de la cellule hôte est modifié de telle façon que le taux d'expression ou l'activité de la protéine endogène FAM60A est réduit, comparativement à un témoin. Il était inattendu que la protéine FAM60A influence la stabilité de l'expression d'un produit recombinant d'intérêt.

[0118] Les séquences du gène FAM60A codant pour la protéine FAM60A ont été rapportées chez Homo sapiens (ID du gène du NCBI : 58516) ; Rattus norvegicus (ID du gène du NCBI : 686611) ; Mus musculus (ID du gène du NCBI : 56306) ; Bos Taurus (ID du gène du NCBI : 538649) et d'autres. Des variants de transcrits peuvent exister d'une façon dépendante de l'espèce et dans des nombres différents (par exemple, le gène FAM60A humain exprime trois isoformes de transcrits putatifs qui diffèrent dans les UTR mais qui codent pour la même protéine).

[0119] La réduction du taux d'expression ou de l'activité d'une protéine FAM60A peut être obtenue par divers moyens. Par exemple, le taux d'expression ou l'activité d'une protéine FAM60A peut être réduit par inactivation de gène, mutation de gène, délétion de gène, silençage de gène ou une combinaison quelconque des précédents. L'inactivation de gène est une technique génétique qui rend un gène inopérant par perturbation de sa fonction. Par exemple, un acide nucléique peut être inséré dans la séquence codante, perturbant de cette façon la fonction du gène. En outre, le gène FAM60A pleine longueur (ou l'un de ses fragments) peut être délété, grâce à quoi l'expression de la protéine FAM60A fonctionnelle est substantiellement éliminée. Par exemple, la délétion peut être d'au moins 30 %, d'au moins 40 %, d'au moins 50 %, d'au moins 55 %, d'au moins 60 %, d'au moins 65 %, d'au moins 70 %, d'au moins 75 %, d'au moins 80 %, d'au moins 85 %, d'au moins 90 %, d'au moins 95 %, ou de 100 % de la séquence codante du gène FAM60A. Une autre option est d'introduire une ou

ou 100 % de la séquence nucléotidique de la région télomérique du chromosome 8 d'une cellule CHO est délétée.

[0122] Dans certains modes de réalisation, la séquence délétée comprend en outre une délétion d'un gène choisi dans le groupe constitué de : Caprin2 et Ipo8, et l'une de leurs combinaisons.

[0123] Dans certains modes de réalisation, la délétion du gène FAM60A est due à une cassure chromosomique. Une cassure chromosomique peut être induite par des procédés décrits ci-dessus. Après l'induction d'une cassure chromosomique, les cellules comportant le point de rupture souhaité (qui entraîne une délétion du gène FAM60A) peuvent être identifiées, par exemple, en analysant l'ADN ou en utilisant le procédé selon le cinquième aspect de la présente divulgation. Par exemple, le profil d'expression des cellules traitées peut être analysé pour déterminer si le gène FAM60A ou des gènes localisés en position centromérique du gène FAM60A sont exprimés, si l'expression est réduite ou si les gènes ne sont pas exprimés. Par exemple, dans le cas de cellules de souris ou de hamster, il peut être analysé si le gène FAM60A est exprimé et en variante ou en outre à ceci, il peut être analysé si un ou plusieurs gènes choisis dans le groupe constitué de Bicdl, C12orf35, protéine 20 de type méthyltransférase, Dennd5b, Caprin2, Ipo8, Tmtc1 ou des gènes qui sont localisés en position télomérique des gènes susmentionnés (où télomérique a cet égard signifie dans la direction de l'extrémité télomérique) sont exprimés par la cellule et/ou si l'expression est réduite ou substantiellement éliminée.

[0124] La réduction du taux d'expression de la protéine FAM60A peut être obtenue par un silençage de gène après la transcription, par exemple, par des molécules d'acide nucléique antisens, ou des molécules qui médient une interférence par ARN. Des exemples non limitants comprennent

des ARNsi, des ARNsh, des miARN, des oligonucléotides antisens, etc., tous étant bien connus dans l'art.

[0125] Le taux d'expression de la protéine FAM60A peut être estimé par des procédés bien connus, par exemple, en mesurant le taux d'ARNm codant pour la protéine FAM60A, ou la protéine FAM60A elle-même. De tels procédés comprennent, par exemple, les analyses Northern blot, FACS, ImageStream, Western blot, qPCR, RT-PCR, qRT-PCR, ELISA, Luminex, Multiplex, etc.

[0126] Dans certains modes de réalisation, le taux d'expression ou l'activité de la protéine FAM60A est réduit d'au moins 3 fois, d'au moins 5 fois, d'au moins 10 fois, d'au moins 20 fois, d'au moins 30 fois, d'au moins 40 fois, d'au moins 50 fois, d'au moins 60 fois, d'au moins 70 fois, d'au moins 75 fois, d'au moins 80 fois, d'au moins 90 fois, d'au moins 100 fois, comparativement à un témoin.

Gènes de matriptases et matriptases

[0127] Dans certains modes de réalisation, la cellule hôte est une cellule dans laquelle le taux d'expression ou l'activité de la matriptase est réduit, comparativement à un témoin. Dans un certain mode de réalisation, une telle cellule est une cellule CHO. La demande de brevet US provisoire No. 61/985 589, déposée le 29 avril 2014 et incorporée ici en référence, et la demande de brevet US provisoire No. 61/994 310, déposée le 16 mai 2014 et incorporée ici en référence, fournissent une description détaillée de cellules de mammifères dans lesquelles le taux d'expression ou l'activité de la matriptase est réduit.

[0128] Divers témoins peuvent être utilisés comme il a été discuté ci-dessus. Le taux d'expression ou l'activité de la matriptase provenant d'une cellule de type sauvage correspondante peut être utilisé en tant que témoin. En variante, un témoin peut être un taux prédéterminé ou un taux

seuil qui peut être identifié dans la littérature ou une base de données.

[0129] La matriptase a été décrite pour la première fois en 1993 comme une nouvelle activité gélatinolytique dans des cellules cultivées de cancer du sein. La matriptase appartient à la famille des sérine protéases transmembranaires de type II (TTSP). Des orthologues de la matriptase sont présents dans différentes espèces de vertébrés, y compris des espèces de mammifères, et ils ont été identifiés, par exemple, chez l'être humain, le chimpanzé, le chien, la souris, le rat, le poulet, le poisson zèbre, le poisson-globe à taches vertes et le poisson-globe tigre, ce qui suggère une fonction conservée dans l'évolution. La matriptase est énoncée dans la nomenclature des enzymes IUBMB sous la référence EC 3.4.21.109. La matriptase est également connue comme une sérine protéase de type membranaire 1 (MT-SP1) et un suppresseur de la tumorigénicité-14 (ST14) (voir Chen *et al.*, The Transmembrane Serine Protease Matriptase: Implications for Cellular Biology and Human Diseases *J Med Sci* 2012; 32 (3): 097-108). Il s'agit d'une protéine intégrale de la membrane avec un seul domaine transmembranaire proche de l'extrémité N-terminale cytoplasmique. La partie extracellulaire est constituée d'une région de tige (comprenant un seul domaine SEA, 2 domaines CUB et 4 domaines LDLRA) et du domaine sérine protéase C-terminal qui est structurellement hautement similaire aux autres TTSP et comprend une triade catalytique conservée histidine/acide aspartique/sérine (HDS) essentielle à l'activité catalytique (voir, par exemple, List *et al.*, Matriptase: Potent Proteolysis on the cell Surface; *Mol Med* 12 (1-3) 1-7, January-March 2006 et Chen *et al.*, The Transmembrane Serine Protease Matriptase: Implications for Cellular Biology and Human Diseases *J Med Sci* 2012; 32 (3): 097-108). La matriptase est décrite comme étant exprimée dans les épithéliums dans de

nombreux systèmes d'organes tels que la peau, le sein, le poumon, l'épiderme, la cornée, la glande salivaire, les cavités orales et nasales, la thyroïde, le thymus, l'œsophage, la trachée, les bronchioles, les alvéoles, l'estomac, le pancréas, la vésicule biliaire, le duodénum, l'intestin grêle, le côlon, le rectum, le rein, les surrénales, la vessie, l'uretère, les vésicules séminales, l'épididyme, la prostate, les ovaires, l'utérus et le vagin (voir, List et al., 2006 et Chen et al., 2012). La matriptase est synthétisée sous la forme d'un zymogène inactif et elle est convertie en sa forme active par l'intermédiaire d'un processus compliqué. Des détails concernant le processus d'activation qui implique des clivages endoprotéolytiques sont décrits pour la matriptase humaine dans List et al. 2006 et Chen et al. 2012. La matriptase est liée à la membrane en tant que protéine transmembranaire de type II avec le domaine catalytique orienté dans l'espace extracellulaire. En outre, il est décrit dans la littérature qu'une excrétion significative de la matriptase, respectivement sa partie extracellulaire, se produit *in vivo* (voir List et al., 2006 et Chen et al. 2012). Il est décrit dans la littérature que la matriptase est excrétée sous la forme d'un complexe, par exemple, complexée à l'inhibiteur de la sérine protéase de type Kunitz HAI-1. Différentes études suggèrent que dans des cellules humaines, l'inhibiteur spécifique HAI-1 facilite le transport de la matriptase vers la membrane cellulaire car il a été montré que l'élimination voire même des mutations ponctuelles simples dans HAI-1 mènent à une accumulation de la matriptase dans le compartiment de Golgi. Dans la littérature, plusieurs inhibiteurs endogènes différents de la matriptase en dehors de HAI-1 ont été décrits comme HAI-2, l'antithrombine, l'alpha-1-antitrypsine et l'alpha-2-antiplasmine. En outre, également d'autres inhibiteurs de la matriptase ont été décrits (voir, par exemple, Chen et al., 2012). Il est décrit

dans la littérature que la matriptase peut jouer de nombreux rôles dans la physiologie animale comme la fonction de barrière cutanée, l'intégrité épithéliale, le développement des follicules pileux, et l'homéostasie du thymus, et dans des pathologies humaines, comme l'arthrose, l'athérosclérose, et la progression, l'invasion et les métastases tumorales.

[0130] A l'encontre de ce contexte scientifique, qui n'est pas apparenté à la production recombinante d'une protéine, la présente découverte que la matriptase est une protéase impliquée dans la coupure des protéines produites de façon recombinante qui sont sécrétées par les cellules hôtes dans le milieu de culture cellulaire a été hautement surprenante. En prenant en considération le grand nombre et la diversité des protéases exprimées dans les cellules des vertébrés, comme en particulier les cellules de mammifères, il a été encore plus surprenant que la réduction du taux d'expression de l'activité de cette protéine puisse réduire significativement la coupure du polypeptide d'intérêt sécrété dans le milieu de culture cellulaire. Ces effets avantageux ne sont pas observés avec d'autres protéases, même des protéases étroitement apparentées. En tant que telles, des cellules hôtes de mammifères (par exemple, des cellules CHO) avec un taux d'expression ou une activité réduit de la matriptase sont particulièrement appropriées pour la production recombinante du complexe pentamère.

[0131] La liste des séquences présente des exemples de séquences d'acides aminés de la matriptase provenant de différentes espèces de vertébrés comme le hamster (SEQ ID NO : 32 - séquence de référence du NCBI : XP_003495890), l'être humain (SEQ ID NO : 33 - séquence de référence du NCBI : NP_068813), la souris (SEQ ID NO : 34 - séquence de référence du NCBI : NP_035306).

[0132] Les séquences nucléotidiques codant pour la matriptase provenant de différentes espèces de mammifères ont

été également rapportées. Voir, par exemple, le hamster de Chine (ID du gène du NCBI : 100755225) ; Homo sapiens (ID du gène du NCBI : 6768) ; Mus musculus (ID du gène du NCBI : 19143) ; Rattus norvegicus (ID du gène du NCBI : 114093) ; Pan Troglodytes (ID du gène du NCBI : 100188950) et d'autres. Des synonymes pour certains du gène de la matriptase sont énumérés dans le tableau 3, parmi lesquels sont couramment utilisés « ST14 » ou « St14 ».

[0133] Comme cela est montré dans le tableau 3, la matriptase est également appelée « protéine suppresseur de la tumorigénicité 14 » (par exemple, pour l'être humain) et « homologue de la protéine suppresseur de la tumorigénicité 14 » (par exemple, chez la souris et le hamster de Chine). Pour simplifier, dans cette divulgation, les homologues et les orthologues de différentes espèces sont tous appelés « gène de la matriptase » ou « matriptase » (protéine).

[0134] Sont également décrits ici des variants fonctionnels d'une matriptase (des variants qui possèdent substantiellement les mêmes activités catalytiques qu'une matriptase de type sauvage). Par exemple, un variant de la matriptase peut comprendre une séquence qui est identique au moins à 50 %, au moins à 60 %, au moins à 70 %, au moins à 75 %, au moins à 80 %, au moins à 85 %, au moins à 90 %, au moins à 95 %, au moins à 96 %, au moins à 97 %, au moins à 98 % ou au moins à 99 % avec l'une quelconque des séquences de SEQ ID NO : 32 à 34, et qui possède les mêmes ou substantiellement les mêmes activités catalytiques qu'une protéine matriptase de type sauvage. L'activité catalytique d'un variant de la matriptase peut être estimée, par exemple, par la réaction chimique pour cliver divers substrats synthétiques avec Arg ou Lys en position P1 et il est préféré des acides aminés à petite chaîne latérale, tels que Ala et Gly, en position P2 (voir, EC 3.4.21.109).

[0135] Sont également décrits ici des fragments fonctionnels d'une matriptase (des fragments qui possèdent substantiellement les mêmes activités catalytiques qu'une matriptase pleine longueur). Un fragment fonctionnel d'une matriptase peut être un sous-ensemble d'acides aminés consécutifs de la matriptase pleine longueur divulguée ici, et qui possède les mêmes ou substantiellement les mêmes activités catalytiques que la séquence de la protéine pleine longueur. L'activité catalytique d'un fragment de la matriptase peut être estimée, par exemple, par la réaction chimique pour cliver divers substrats synthétiques avec Arg ou Lys en position P1 et il est préféré des acides aminés à petite chaîne latérale, tels que Ala et Gly, en position P2 (voir, EC 3.4.21.109).

[0136] La réduction du taux d'expression ou de l'activité d'une matriptase peut être obtenue par divers moyens. Par exemple, le taux d'expression ou l'activité d'une matriptase peut être réduit par inactivation de gène, mutation de gène, délétion de gène, silençage de gène, ou une combinaison quelconque des précédents. L'inactivation de gène est une technique génétique par laquelle un gène est rendu inopérant par perturbation de sa fonction. Par exemple, un acide nucléique peut être inséré dans la séquence codante, perturbant de cette façon la fonction du gène. En outre, le gène de la matriptase pleine longueur (ou l'un de ses fragments) peut être délété, grâce à quoi l'expression de la matriptase fonctionnelle est substantiellement éliminée. Par exemple, la délétion peut être d'au moins 30 %, d'au moins 40 %, d'au moins 50 %, d'au moins 55 %, d'au moins 60 %, d'au moins 65 %, d'au moins 70 %, d'au moins 75 %, d'au moins 80 %, d'au moins 85 %, d'au moins 90 %, d'au moins 95 %, ou de 100 % de la séquence codante du gène de la matriptase. Une autre option est d'introduire une ou plusieurs mutations dans la séquence codante, ce qui rend une matriptase non fonctionnelle

ou moins fonctionnelle. Par exemple, une ou plusieurs mutations de décalage du cadre peuvent être introduites, produisant une matriptase non fonctionnelle ou moins fonctionnelle. En variante ou en outre, un ou plusieurs codons d'arrêt peuvent être introduits dans la séquence codante de telle façon qu'il est obtenu une protéine tronquée, non fonctionnelle ou moins fonctionnelle. D'autres options comprennent, mais n'y sont pas limitées, une ou plusieurs mutations dans le promoteur, dans les UTR en 5' et/ou en 3' ou d'autres éléments régulateurs, par exemple par introduction d'une délétion du promoteur ou par introduction d'une construction entre le promoteur et le départ de la transcription. Des procédés pour la perturbation de gènes pour supprimer ou éliminer l'expression du gène cible sont bien connus de l'homme du métier.

[0137] Puisque chaque cellule comporte deux copies du gène de la matriptase dans son génome, dans certains modes de réalisation, au moins une copie de la séquence génomique du gène de la matriptase, ou au moins 50 % de la séquence codante dudit gène de la matriptase, ou d'un fragment fonctionnel gène de la matriptase, est déléetée. Dans certains modes de réalisation, les deux copies des séquences génomiques du gène de la matriptase (ou au moins 50 % de la séquence codante dudit gène de la matriptase, ou d'un fragment fonctionnel dudit gène de la matriptase, à partir de chaque copie) sont déléetées.

[0138] Dans certains modes de réalisation, la cellule hôte comprend une mutation dans l'exon 2 du gène de la matriptase. L'exon 2 est particulièrement approprié en tant que cible pour modifier l'activité matriptase parce qu'il existe plusieurs variants d'épissage fonctionnels différents. Ainsi, les exons proches de l'extrémité N-terminale de la matriptase comme par exemple l'exon 1, l'exon 2 et l'exon 3, sont avantageux pour introduire une ou plusieurs mutations,

en particulier une ou plusieurs mutations de décalage du cadre. Une mutation de décalage du cadre dans l'un de ces exons mène le plus vraisemblablement à un codon d'arrêt précoce dans la séquence. Des mutations peuvent être également introduites dans l'un des exons subséquents, par exemple choisis parmi les exons 4 à 19.

[0139] Dans certains modes de réalisation, la matriptase comprend une mutation dans le domaine catalytique. Le domaine catalytique est la région d'une enzyme qui interagit avec son substrat pour provoquer la réaction enzymatique. Une ou plusieurs mutations peuvent être introduites dans ce domaine de telle façon que l'activité catalytique de la protéine est réduite comparativement à un témoin. Le domaine catalytique comprend des séquences codées par les exons 16, 17, 18 et 19. Les mutations peuvent être une délétion, une insertion, une substitution, ou l'une de leurs combinaisons. Les mutations peuvent provoquer une mutation de décalage du cadre, une mutation ponctuelle spécifique, une mutation de codon d'arrêt, ou l'une de leurs combinaisons, dans la séquence codant pour le domaine catalytique. Des mutants inactifs du point de vue catalytique de la matriptase comme par exemple la G827R-matriptase ou la S805A-matriptase ont été également décrits dans la littérature (voir Désilets *et al.*, The Journal of Biological Chemistry Vol. 283, No. 16, pp. 10535-10542, 2008). En outre, la structure cristalline du domaine catalytique d'une matriptase recombinante est connue. D'après cette structure et les données sur la séquence, l'homme du métier peut dériver d'autres cibles spécifiques pour des mutations afin d'altérer la fonction catalytique de la matriptase.

[0140] La réduction du taux d'expression de la matriptase peut être obtenue par un silençage de gène après la transcription, par exemple par des molécules d'acide nucléique antisens, ou des molécules qui médient une

interférence par ARN. Les exemples non limitants comprennent des ARNsi, des ARNsh, des miARN, des oligonucléotides antisens, etc., tous étant bien connus dans l'art.

[0141] Le taux d'expression de la matriptase peut être estimé par des procédés connus dans l'art, par exemple, en mesurant le taux d'ARNm codant pour la matriptase, ou la matriptase elle-même. De tels procédés comprennent, par exemple, les analyses Northern blot, FACS, ImageStream, Western blot, qPCR, RT-PCR, qRT-PCR, ELISA, Luminex, Multiplex, etc. L'activité de la matriptase peut être estimée, par exemple, selon son activité enzymatique.

[0142] Dans certains modes de réalisation, le taux d'expression ou l'activité de la matriptase est réduit d'au moins 3 fois, d'au moins 5 fois, d'au moins 10 fois, d'au moins 20 fois, d'au moins 30 fois, d'au moins 40 fois, d'au moins 50 fois, d'au moins 60 fois, d'au moins 70 fois, d'au moins 75 fois, d'au moins 80 fois, d'au moins 90 fois, d'au moins 100 fois, comparativement à un témoin.

5

10

15

20

25

30

Tableau 2

Abréviations et variantes de noms (alias) de produits codés par des gènes localisés dans le chromosome 8 du hamster de Chine ou le chromosome 6 de la souris.

5

Ab ré vi at io n	Dans l'ann otati on publi que du hamst er de Chine , le produ it géniqu ue est annot é en tant que	Dans l'anno tation publiq ue de la souris , le produi t géniqu e est annoté en tant que	Alias (voir www.genecards.org)
Cc dc 91	Proté ine 91 conte nant un domai ne suren roulé	91 conten ant un domain e surenr oulé	91 contenant un domaine surenroulé P56 Partenaire de liaison de GGA Protéine accessoire P56 DKFZp779L1558 FLJ11088 Protéine 91 contenant un domaine surenroulé Partenaire de liaison de GGA GGABP

Fa r2	Isofo rme 1 de l'acy l gras- CoA- réduc tase 1	Acyl gras- CoA- réduct ase 2	Acyl gras-CoA- réductase 2 MLSTD1 SDR10E2 Protéine 1 contenant le domaine de stérilité mâle EC 1.2.1.1.N2 FLJ10462 1 contenant le domaine de stérilité mâle Chaîne courte de l'acyl gras-CoA- réductase 2 Famille 10E des déshydrogénases/réd uctases, membre 2
Er gi c2	Proté ine 2 du compa rtime nt inter média ire du rétic ulum endop lasmi que- Golgi	ERGIC et Golgi 2	ERGIC et Golgi 2 PTX1 Erv41 Cd002 Protéine CD14 Protéine 2 du compartiment intermédiaire du réticulum endoplasmique-Golgi ERV41 CDA14

<p>RP S4 Y2</p>	<p>Sous-unité 40S de la protéine ribosomique S4, type isoforme X</p>	<p>Protéine ribosomique S4, 2 lié à Y</p>	<p>Protéine ribosomique S4, 2 lié à Y RPS4Y2P Protéine ribosomique S4, pseudogène 2 lié à Y Sous-unité 40S de la protéine ribosomique S4, Y Sous-unité 40S de la protéine ribosomique S4, isoforme 2 liée à Y</p>	BE2016/5877
<p>Tm tc 1</p>	<p>Protéine 1 transmembranaire et contenant des répétitions TPR</p>	<p>1 transmembranaire et contenant des répétitions de tétratricopeptides</p>	<p>1 transmembranaire et contenant des répétitions de tétra-tricopeptides OLF ARG99 FLJ31400 FLJ41625 TMTC1A 1A transmembranaire et contenant des répétitions de tétratricopeptides Protéine 1 transmembranaire et contenant des répétitions TPR</p>	

Zfp1	Protéine 1 à doigt de zinc et contenant un domaine HIT		Protéine à doigt de zinc ZFP1 ZNF475 Protéine à doigt de zinc 475 FLJ34243 Protéine 1 à doigt de zinc Homologue de la protéine 1 à doigt de zinc (souris) Zfp-1 Homologue de la protéine 1 à doigt de zinc	BE2016/5877
IP08	Type importine 8	Importine 8	Importine 8 RANBP8 Protéine 8 de liaison de RAN Protéine 8 de liaison de Ran IMP8 Imp8 Importine 8 RanBP8	

Ca pr in 2	Proté ine de type Capri ne 2	Membre 2 de la famill e des caprin es	Membre 2 de la famille des caprines ClQDC1 EEG1 RNG140 Caprine 2 Protéine 2 cytoplasmique associée à l'activation/ prolifération Protéine de cancer gastrique associée à la multirésistance médicamenteuse Protéine 1 contenant le domaine Clq Protéine à granules d'ARN 140 FLJ11391 FLJ22569 1 contenant le domaine Clq EEG-1 KIAA1873 Protéine EEG-1	BE2016/5877
FA M6 0A	Proté ine de type FAM60 A	Famill e avec simila rité de séquen ce 60, membre A	Famille avec similarité de séquence 60, membre A Cl2orf14 TERA Homologue de la protéine TerA Cadre de lecture ouvert 14 du chromosome 12 Protéine FAM60A	

Denn5b	Protéine de type 5B contenant le domaine Denn	5B contenant le domaine DENN/MADD	5B contenant le domaine DENN/MADD Protéine de type Rab6IP1 MGC24039 Protéine 5B contenant le domaine DENN	BE2016/5877
METTL20	Protéine 20 de type méthyltransférase	483344 2J19Rik	METTL20 C12orf72 DKFZp451L235 MGC50559 Cadre de lecture ouvert 72 du chromosome 12 Protéine 20 de type méthyltransférase EC 2.1.1.	
AMN1	Protéine putative de type protéine AMN1	Antagoniste 1 du réseau de sortie de mitose	Homologue de l'antagoniste 1 du réseau de sortie de mitose (S. Cerevisiae) Homologue de la protéine AMN1	

	Protéine de type récepteur du facteur de croissance des opioïdes	Protéine de type récepteur du facteur de croissance des opioïdes	
C12orf35	Homologue de la protéine non caractérisée C12orf35	Orthologue vraisemblable du cadre de lecture ouverte 35 du chromosome 12 de H. sapiens (C12orf35) ; 2810474019Risk	KIAA1551, C12orf35, FLJ10652, FLJ20696 Cadre de lecture ouvert 35 du chromosome 12 Protéine non caractérisée C12orf35 Protéine non caractérisée KIAA1551

Bicd1	Protéine bicaudale D putative	Homologue 1 de la protéine bicaudale D	Homologue 1 de la protéine bicaudale D (Drosophile) Bic-D 1 Homologue 1 de la protéine bicaudale D (Drosophile) BICD Homologue 1 de la protéine bicaudale D de type cytosquelette Homologue 1 de la protéine bicaudale D	BE2016/5877
-------	-------------------------------	--	---	-------------

Tableau 3

Exemples de variantes de noms du gène de la matriptase et/ou du produit de la protéine codée de la matriptase utilisés dans la littérature (ordre alphabétique)

5

Protéase de 80 kDa du cancer du sein
CAP3
Protéase 3 activatrice des canaux
EC 3.4.21.109
Epithine
HAI
Matriptase
Matriptase-1
Protéine membranaire 1 de type sérine protéase
MT-SP1 ; MTSP1
Prostamine
PRSS14 g.p. (Homo sapiens), PRSS14
Sérine endopeptidase SNC19

Sérine protéase 14
Sérine protéase TADG-15 ; TADG-15 ; TADG15
SNC19
ST14 (nom officiel du gène chez les êtres humains selon le HGNC)
St14
Protéine de suppression de la tumorigénicité-14 ; de suppression de la tumorigénicité 14 (carcinome du côlon, matriptase, épithine)
Protéine suppresseur de la tumorigénicité 14
Homologue de la protéine suppresseur de la tumorigénicité (souris)
TMPRSS14
Protéine du gène 15 exprimé de façon différentielle associé aux tumeurs

[0143] Une cellule hôte appropriée peut comporter des combinaisons quelconques des modifications décrites ici, par exemple, une cellule dans laquelle à la fois le taux d'expression ou l'activité de la protéine C12orf35 est réduit dans ladite cellule, comparativement à un témoin, et le taux d'expression ou l'activité de la protéine FAM60A est réduit dans ladite cellule, comparativement à un témoin. D'autres combinaisons comprennent, par exemple, la réduction de du taux d'expression ou de l'activité de la protéine C12orf35, et la réduction du taux d'activité ou de l'activité de la matriptase ; la réduction du taux d'expression ou de l'activité de la protéine FAM60A, et la réduction du taux

d'expression ou de l'activité de la matriptase ; la réduction du taux d'expression ou de l'activité de la protéine C12orf35, la réduction du taux d'expression ou de l'activité de la protéine FAM60A, et la réduction du taux d'expression ou de l'activité de la matriptase ; la réduction du taux d'expression ou de l'activité de la protéine C12orf35, et l'inclusion de la séquence de la dihydrofolate réductase (DHFR) en tant que marqueur de sélection, etc.

[0144] Les cellules hôtes décrites ici sont appropriées pour une culture à grande échelle. Par exemple, les cultures cellulaires peuvent être 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 ou plus grandes. Dans certains modes de réalisation, la taille de la culture cellulaire peut se situer dans la plage de 10 l à 5000 l, de 10 l à 10,000 l, de 10 l à 20 000 l, de 10 l à 50 000 l, de 40 l à 50 000 l, de 100 l à 50 000 l, de 500 l à 50 000 l, de 1000 l à 50 000 l, de 2000 l à 50 000 l, de 3000 l à 50 000 l, de 4000 l à 50 000 l, de 4500 l à 50 000 l, de 1000 l à 10 000 l, de 1000 l à 20 000 l, de 1000 l à 25 000 l, de 1000 l à 30 000 l, de 15 l à 2000 l, de 40 l à 1000 l, de 100 l à 500 l, de 200 l à 400 l, ou tout nombre entier entre ces valeurs.

[0145] Les composants des milieux pour culture cellulaire sont connus dans l'art, et ils peuvent comprendre, par exemple, un tampon, une teneur en acides aminés, une teneur en vitamines, une teneur en sels, une teneur en minéraux, une teneur en sérum, une teneur en source de carbone, une teneur en lipides, une teneur en acides nucléiques, une teneur en hormones, une teneur en oligo-éléments, une teneur en ammoniac, une teneur en cofacteurs, une teneur en indicateurs, une teneur en petites molécules, une teneur en hydrolysats et une teneur en modulateurs enzymatiques.

4. Purification du complexe pentamère à partir de la culture cellulaire

[0146] Le complexe pentamère produit conformément au procédé décrit ici peut être récolté à partir des cellules hôtes, et purifié en utilisant tout procédé approprié. Les procédés appropriés comprennent la précipitation et divers types de chromatographie, comme l'interaction hydrophobe, l'échange d'ions, l'affinité, la chélation, et l'exclusion par la taille, tous étant connus dans l'art. Des schémas de purification appropriés peuvent être créés en utilisant deux ou plus de ces procédés appropriés ou d'autres. Si c'est souhaité, une ou plusieurs des sous-unités du complexe pentamère peuvent comprendre un « marqueur » qui facilite la purification, comme un marqueur épitopique ou un marqueur HIS, un marqueur Strep. De tels polypeptides marqués peuvent être purifiés de façon pratique, par exemple à partir de milieux conditionnés, par chromatographie chélatrice ou chromatographie d'affinité. Eventuellement, la séquence du marqueur peut être clivée après la purification.

[0147] Par exemple, le document WO 2014/005959 divulgue des exemples de procédés de purification du complexe pentamère par chromatographie d'affinité.

[0148] Dans certains modes de réalisation, une ou plusieurs sous-unités du complexe pentamère comprennent un marqueur pour une purification par affinité. Les exemples de marqueurs pour une purification par affinité comprennent, par exemple, le marqueur His (se lie aux ions métalliques), un anticorps (se lie à la protéine A ou à la protéine G), la protéine de liaison du maltose (MBP) (se lie à l'amylose), la glutathion-S-transférase (GST) (se lie au glutathion), le marqueur FLAG (Asp-Tyr-Lys-Asp-Asp-Asp-Asp-Lys (SEQ ID NO : 37)) (se lie à un anticorps anti-flag), le marqueur Strep (se lie à la streptavidine ou à l'un de ses dérivés).

[0149] La structure du complexe pentamère gH/gL/pUL128/pUL130/pUL131 est inconnue. Toutefois, si le marqueur pour la purification par affinité est fixé à un site qui interfère avec la formation du complexe pentamère, ou à un site qui est enterré au sein du complexe, la purification par affinité ne sera pas couronnée de succès. On pense que les sites suivants sont appropriés pour fixer un marqueur pour la purification par affinité, car le marqueur ne semble pas interférer avec la formation du complexe pentamère, et semble être exposé à la surface d'un pentamère assemblé :

5
10
15

(i) la région C-terminale de la pUL130, (ii) la région N-terminale de la pUL130, (iii) la région C-terminale de la pUL131, (iv) la région N-terminale de la pUL131, (v) la région C-terminale de la pUL128, (vi) la région N-terminale de la pUL128, ou l'une de leurs combinaisons.

[0150] Dans certains modes de réalisation, le complexe pentamère ne comprend pas de marqueur pour la purification.

[0151] Un autre procédé approprié est la chromatographie par échange d'ions. Les exemples de matériaux utiles dans la chromatographie par échange d'ions comprennent la DEAE-cellulose, la QAE-cellulose, la DEAE-céphalose, la QAE-céphalose, DEAE-Toyopearl, QAE-Toyopearl, Mono Q, Mono S, Q sepharose, SP sepharose, etc. Dans un exemple de mode de réalisation, le procédé utilise une colonne de Mono S. Dans un autre exemple de mode de réalisation, le procédé utilise une colonne de Mono Q.

20
25

[0152] Dans certains modes de réalisation, le rendement du complexe pentamère du CMV est d'au moins environ 0,01 g/l, d'au moins environ 0,02 g/l, d'au moins environ 0,03 g/l, d'au moins environ 0,05 g/l, d'au moins environ 0,06 g/l, d'au moins environ 0,07 g/l, d'au moins environ 0,08 g/l, d'au moins environ 0,09 g/l, d'au moins environ 0,1 g/l, d'au moins environ 0,15 g/l, d'au moins environ 0,20 g/l, d'au moins environ 0,25 g/l, d'au moins environ 0,3 g/l, d'au moins

30

environ 0,35 g/l, d'au moins environ 0,4 g/l, d'au moins
environ 0,45 g/l, d'au moins environ 0,5 g/l, d'au moins
environ 0,55 g/l, d'au moins environ 0,6 g/l, d'au moins
environ 0,65 g/l, d'au moins environ 0,7 g/l, d'au moins
5 environ 0,75 g/l, d'au moins environ 0,8 g/l, d'au moins
environ 0,85 g/l, d'au moins environ 0,9 g/l, d'au moins
environ 0,95 g/l, ou d'au moins environ 1,0 g/l.

5. Définitions

10 **[0153]** Le terme « fragment formant un complexe » d'une
protéine du CMV (comme gH, gL, pUL128, pUL130, ou pUL131) se
rapporte à toute partie ou portion de la protéine qui conserve
la capacité de former un complexe avec une autre protéine du
CMV. De tels complexes comprennent, par exemple, le complexe
15 pentamère gH/gL/pUL128/pUL130/pUL131. Les fragments qui
conservent la capacité de former le complexe pentamère sont
également appelés « fragments formant le pentamère ».

[0154] Une « culture à grande échelle » se rapporte à une
culture qui a une taille d'au moins environ 10 litres (par
20 exemple, un volume d'au moins environ 10 l, d'au moins environ
20 l, d'au moins environ 30 l, d'au moins environ 40 l, d'au
moins environ 50 l, d'au moins environ 60 l, d'au moins
environ 70 l, d'au moins environ 80 l, d'au moins environ
90 l, d'au moins environ 100 l, d'au moins environ 150 l,
25 d'au moins environ 200 l, d'au moins environ 250 l, d'au moins
environ 300 l, d'au moins environ 400 l, d'au moins environ
500 l, d'au moins environ 600 l, d'au moins environ 700 l,
d'au moins environ 800 l, d'au moins environ 900 l, d'au moins
environ 1000 l, d'au moins environ 2000 l, d'au moins environ
30 3000 l, d'au moins environ 4000 l, d'au moins environ 5000 l,
30 d'au moins environ 6000 l, d'au moins environ 10 000 l, d'au
moins environ 15 000 l, d'au moins environ 20 000 l, d'au
moins environ 25 000 l, d'au moins environ 30 000 l, d'au
moins environ 35 000 l, d'au moins environ 40 000 l, d'au

moins environ 45 000 l, d'au moins environ 50 000 l, d'au
moins environ 55 000 l, d'au moins environ 60 000 l, d'au
moins environ 65 000 l, d'au moins environ 70 000 l, d'au
moins environ 75 000 l, d'au moins environ 80 000 l, d'au
5 moins environ 85 000 l, d'au moins environ 90 000 l, d'au
moins environ 95 000 l, d'au moins environ 100 000 l, etc.).

[0155] Un complexe pentamère « soluble » se rapporte au
complexe gH/gL/pUL128/pUL130/pUL131 dans lequel la sous-unité
gH ne comprend pas le domaine transmembranaire.

10 **[0156]** Dans tout le mémoire, y compris les
revendications, où le contexte le permet, le terme
« comprenant » et ses variantes comme « comprend » ou
« comprenant » doivent être interprétés comme incluant le
nombre entier ou les nombres entiers indiqués sans exclure
15 nécessairement tout autre nombre entier.

[0157] L'identité de séquence est calculée selon le
pourcentage de correspondances de résidus entre
deux séquences polypeptidiques, ou de correspondances de
nucléotides entre deux séquences polynucléotidiques, alignées
20 en utilisant un algorithme normalisé. Un tel algorithme peut
insérer, d'une façon normalisée et reproductible, des brèches
dans les séquences comparées afin d'optimiser l'alignement
entre deux séquences, et par conséquent, obtenir une
comparaison plus significative des deux séquences. Le
pourcentage d'identité peut être mesuré sur la longueur d'une
25 séquence définie entière, ou il peut être mesuré sur une
longueur plus courte, par exemple, sur la longueur d'un
fragment pris à partir d'une séquence définie plus grande,
par exemple, un fragment d'au moins 45, d'au moins 60, d'au
30 moins 90, d'au moins 120, d'au moins 150, d'au moins 210 ou
d'au moins 450 résidus ou nucléotides consécutifs. Si la
longueur n'est pas spécifiée, l'identité de séquence est
calculée sur la pleine longueur de la plus courte des deux
séquences.

[0158] Cette invention est en outre illustrée par les exemples suivants qui ne devront pas être interprétés comme limitants. BE2016/5877

5 Exemples

[0159] Cet exemple concerne la production du complexe protéinique pentamère du CMV par des lignées cellulaires CHO dans lesquelles les séquences codantes des sous-unités du pentamère ont été intégrées de façon stable dans le chromosome. Les lignées cellulaires sont également qualifiées de lignées cellulaires CHO stables.

10 **[0160]** Comme cela est montré dans cet exemple, ces lignées cellulaires CHO stables ont produit un pentamère fonctionnel du CMV, avec les cinq sous-unités assemblées dans la conformation naturelle. Le rendement est élevé, permettant de cette façon la production du complexe pentamère à une grande échelle commerciale. Ces lignées cellulaires sont particulièrement appropriées pour la fabrication à grande échelle de vaccins contre le CMV utilisant le pentamère.

15 **[0161]** Comme cela est montré dans cet exemple, les séquences nucléotidiques codant pour l'ectodomaine de la gH, la gL, la pUL128, la pUL130 et la pUL131 ont été clonées dans des vecteurs d'expression simples ou doubles, avec des promoteurs subgénomiques et des IRES pour diriger l'expression du composant individuel (figure 1). De façon spécifique, les séquences codant pour l'ectodomaine de la gH, les gL, pUL128, pUL130 et pUL131 pleine longueur ont été optimisées pour les codons pour une expression en culture de cellules de mammifères, synthétisées et clonées dans un vecteur d'expression simple (avec sélection par Neo et DHFR) dirigé par le promoteur du CMV (pour les gH et UL128) et l'IRES de l'EV71 (pour les gL, UL130, UL131) pour l'expression ; ou des vecteurs d'expression doubles, un avec sélection par Neo et DHFR pour la gH dirigé par le promoteur

20

25

30

du CMV, et pour la gL dirigé par le promoteur du CMV ou l'IRES de l'EV71, et un avec sélection par Hyg et Fr α pour la UL128 dirigé par le promoteur du CMV, pour les UL130 et UL131 dirigé par l'IRES de l'EV71 (figure 1). Les marqueurs pour la purification par affinité comme le marqueur His et/ou le marqueur Strep ont été introduits à l'extrémité C-terminale de l'ectodomaine de la gH et/ou de UL130 pour faciliter la purification. Le codon d'arrêt et le domaine transmembranaire de la gH ont été également introduits à l'extrémité C-terminale de l'ectodomaine de la gH pour faciliter le sous-clonage par FACS (figure 1).

[0162] Le plasmide ou les plasmides d'expression ont été transfectés dans un panel de cellules hôtes CHO, comprenant les lignées cellulaires KO pour la matriptase (i) CHOC8TD, (ii) CHOHP3 et (iii) CHOC8TD.

[0163] CHOC8TD est dérivée d'une lignée cellulaire CHO K1 et elle est en outre modifiée par délétion de la région télomérique du chromosome, si bien que cette lignée cellulaire a les caractéristiques suivantes : (i) télomère du chromosome 8 délété ; (ii) productivité élevée et bonne croissance cellulaire ; et (iii) problèmes potentiels de dégradation protéolytique élevée. CHOHP3 est une lignée cellulaire CHO K1 montrant une activité protéolytique réduite, si bien que cette lignée cellulaire a les caractéristiques suivantes : (i) dégradation protéolytique inférieure à CHOC8TD, et (ii) croissance, productivité et efficacité du clonage inférieures à CHOC8TD. CHOC8TD KO pour la matriptase a été dérivée de CHOC8TD, avec la matriptase inactivée, si bien que cette lignée cellulaire a les caractéristiques suivantes : (i) matriptase (sérine protéase) inactivée ; et (ii) dégradation protéolytique moindre.

[0164] Le plasmide ou les plasmides d'expression comprenant les séquences codant pour les gH, gL, UL128, UL130 et UL131 ont été transfectés dans les trois cellules hôtes

avec Nucleofection AMAXA. Pour la stratégie à vecteur simple, les cellules transfectées ont été sélectionnées par G418 et MTX séquentiellement. Pour la stratégie à vecteur double, les cellules transfectées ont été sélectionnées par Hyg puis MTX séquentiellement.

[0165] Les groupes de cellules survivantes ont été cultivés en culture fermée pour produire les protéines du pentamère jusqu'à l'évaluation des protéines pour le rendement avec un test ELISA indirect, la contamination par gH/gL et la dégradation de gL avec les analyses Western et SDS-PAGE. En se basant sur le rendement du pentamère et la croissance des cellules, le groupe des cellules CHO transfectées avec la stratégie de vecteur 2 (figure 1), en utilisant des cellules hôtes CHO8TD, a été sélectionné pour le clonage des cellules simples (plasmide d'expression représenté sur la figure 2). Les clones sélectionnés montrent un rendement élevé et moins de contamination par gH/gL.

[0166] Les clones simples ont été triés avec une analyse FACS en utilisant des Acm spécifiques du pentamère (~ 200). Les 30 clones principaux ont été d'abord sélectionnés en se basant sur le titre du pentamère évalué par un test ELISA indirect et ils ont été en outre sélectionnés pour les ramener à 12 clones principaux. Une culture fermée suivie d'une purification par affinité à l'aide de Strep a montré que les 12 principaux peuvent produire le pentamère avec un rendement > 300 mg/l et une pureté élevée (figure 3). En fait, les clones principaux ont produit le pentamère avec un rendement d'environ 0,3 g/l à environ 0,5 g/l. Le complexe protéinique purifié a été évalué pour sa liaison avec un panel d'Acm dirigés contre le pentamère, et il a été montré qu'ils présentaient tous les épitopes clés.

[0167] Les 12 clones principaux ont été en outre évalués à l'aide d'études de stabilité sur 14 semaines (figure 4). Les lots contrôlés et non contrôlés ont été évalués pour la

croissance cellulaire et le titre du pentamère et le clone VF7 a été sélectionné.

[0168] La production en bioréacteur du pentamère avec le clone principal VF7 a été également estimée. Ce clone a produit > 100 mg/l de pentamère purifié (figure 5).

[0169] Comme cela est exemplifié ici, les cellules CHO stables peuvent être utilisées pour exprimer le pentamère. Ceci est en contraste avec l'expression transitoire par les cellules HEK293. Les lignées cellulaires CHO stables de l'invention peuvent produire les protéines du pentamère de façon régulière avec un rendement 100 fois supérieur.

[0170] La capacité à produire des lignées cellulaires CHO stables dans lesquelles les séquences codant pour les cinq sous-unités du pentamère sont intégrées dans le chromosome, comme cela est exemplifié ici, est plutôt remarquable. Il existe toujours des incertitudes quant à si une lignée cellulaire stable peut être produite, même pour une seule protéine. Par exemple, lorsqu'une séquence codant pour l'IGF-1 humain a été introduite dans des cellules CHO pour produire une lignée stable, il a été découvert que les lignées cellulaires IGF-1/CHO résultantes présentaient une inhibition de la croissance cellulaire et des titres bas. La mesure du titre maximal a été d'environ 8 µg/ml ce qui correspond à 100 mg/l d'un titre d'anticorps (en se basant sur la masse molaire). Comparativement, les mesures de titre moyen d'un anticorps recombinant dans un procédé en bioréacteur sont autour de 3 g/l. Une cause du titre bas d'IGF-1 a été la réduction de la croissance cellulaire et la faible viabilité cellulaire des cellules exprimant l'IGF-1. Durant un procédé d'expression d'anticorps, des cellules dérivées de CHO-K1 se sont développées jusqu'à 2×10^7 cellules/ml et la viabilité cellulaire est supérieure à 97 % durant les 230 à 260 premières heures du temps de culture. Au contraire, des cellules dérivées de CHO-K1 exprimant l'IGF-1 ne se sont

développées que jusqu'à $0,5 \times 10^7$ cellules/ml et la viabilité cellulaire avait déjà chuté au-dessous de 97 % après deux jours. Pour d'autres détails, voir, la demande US provisoire No. 61/738466, déposée le 18 décembre 2012 et la publication de demande PCT No. WO/2014/097113, déposée le 16 décembre 2013.

[0171] Par conséquent, les inventeurs ont fait face à des difficultés supplémentaires significatives parce que les séquences codant pour les cinq sous-unités du pentamère du CMV doivent être intégrées de façon stable dans le génome de la cellule CHO. Les inventeurs ont surmonté ces difficultés, comme cela est mis en évidence par les clones sélectionnés montrant des rendements élevés, avec le pentamère produit de façon recombinante dans sa conformation naturelle et avec tous les épitopes clés.

[0172] Les diverses caractéristiques et les divers modes de réalisation de la présente invention, auxquels il est fait référence dans des sections individuelles ci-dessus s'appliquent, comme il est approprié, aux autres sections, *mutatis mutandis*. Par conséquent, les caractéristiques précisées dans une section peuvent être combinées avec des caractéristiques spécifiées dans d'autres sections, comme il est approprié.

[0173] Le mémoire est au mieux compris à la lumière des enseignements des références citées au sein du mémoire. Les modes de réalisation au sein du mémoire fournissent une illustration de modes de réalisation de l'invention et ne devront pas être interprétés comme limitant l'étendue de l'invention. L'homme du métier comprend facilement que de nombreux autres modes de réalisation sont englobés par l'invention. La totalité des publications, des brevets, et des séquences de GenBank cités dans cette divulgation sont incorporés en référence dans leur intégralité. Dans la mesure où le matériau incorporé en référence contredit ou est

incohérent avec ce mémoire, le mémoire supplantera tout matériau de ce type. La citation de l'une quelconque des références ici n'est pas une admission que de telles références sont de l'art antérieur à la présente invention.

5 **[0174]** L'homme du métier reconnaîtra, ou sera capable de déterminer en utilisant une expérimentation ne dépassant pas la routine, de nombreux équivalents aux modes de réalisation spécifiques de l'invention décrite ici. De tels équivalents sont censés être englobés par les modes de réalisation
10 suivants.

[0175] Les modes de réalisation particuliers de l'invention comprennent :

1. Une cellule recombinante de mammifère, comprenant :

une ou plusieurs séquences polynucléotidiques codant pour le
15 complexe pentamère du cytomégalovirus (CMV), où ledit complexe pentamère comprend : la gH ou l'un de ses fragments formant un complexe, la gL ou l'un de ses fragments formant un complexe, la pUL128 ou l'un de ses fragments formant un complexe, la pUL130 ou l'un de ses
20 fragments formant un complexe, et la pUL131 ou l'un de ses fragments formant un complexe ;

où lesdites une ou plusieurs séquences polynucléotidiques sont intégrées dans l'ADN génomique de ladite cellule de mammifère.

25 2. La cellule de mammifère du mode de réalisation 1, où ladite cellule de mammifère est une cellule d'ovaire de hamster de Chine (CHO).

30 3. La cellule de mammifère du mode de réalisation 2, où ladite cellule CHO est une cellule CHO-K1, CHO-DUXB11, ou CHO-DG44.

4. La cellule de mammifère de l'un quelconque des modes de réalisation 1 à 3, où le taux d'expression ou l'activité de la protéine C12orf35 est réduit dans ladite cellule, comparativement à un témoin.

5. La cellule de mammifère du mode de réalisation 4, où au moins une copie de la séquence génomique du gène C12orf35, ou au moins 50 % de la séquence codante dudit gène C12orf35, est délétée.
- 5 6. La cellule de mammifère du mode de réalisation 5, où les deux copies des séquences génomiques du gène C12orf35, ou au moins 50 % de la séquence codante dudit gène C12orf35 à partir de chaque copie, sont délétées.
7. La cellule de mammifère du mode de réalisation 4 ou 5, où
10 au moins une copie de la région télomérique du chromosome 8 d'une cellule CHO qui comprend le gène C12orf35 est délétée.
8. La cellule de mammifère du mode de réalisation 7, où ladite région télomérique comprend en outre un gène
15 choisi dans le groupe constitué de : Bicd1, Amn1, protéine 20 de type méthyltransférase, Dennd5b, FAM60A, Caprin2, Ipo8, RPS4Y2, et l'une de leurs combinaisons.
9. La cellule de mammifère de l'un quelconque des modes de réalisation 4 à 8, où ladite protéine C12orf35 comprend
20 une séquence qui est au moins identique à 80 % avec l'une quelconque des séquences choisies dans le groupe constitué des SEQ ID NO : 21, 22, 23, 24 et 35.
10. La cellule de mammifère de l'un quelconque des modes de réalisation 4 à 9, où ledit gène C12orf35 comprend une
25 séquence qui est au moins identique à 80 % avec SEQ ID NO : 25.
11. La cellule de mammifère du mode de réalisation 4, où ladite cellule comprend une mutation dans le promoteur, l'UTR en 5', ou l'UTR en 3' dudit gène C12orf35.
- 30 12. La cellule de mammifère du mode de réalisation 4, où ladite protéine C12orf35 comprend une mutation qui réduit son activité, comparativement à un témoin.
13. La cellule de mammifère de l'un quelconque des modes de réalisation 1 à 12, où le taux d'expression ou

l'activité de la protéine FAM60A est réduite dans ladite cellule, comparativement à un témoin.

14. La cellule de mammifère du mode de réalisation 13, où au moins une copie de la séquence génomique du gène FAM60A, ou au moins 50 % de la séquence codante dudit gène FAM60A, est déléetée.
15. La cellule de mammifère du mode de réalisation 14, où les deux copies des séquences génomiques du gène FAM60A, ou au moins 50 % de la séquence codante dudit gène FAM60A à partir de chaque copie, sont déléetées.
16. La cellule de mammifère du mode de réalisation 13 ou 14, où ladite séquence déléetée comprend une portion de la région télomérique du chromosome 8 d'une cellule CHO.
17. La cellule de mammifère du mode de réalisation 16, où ladite séquence déléetée comprend en outre un gène choisi dans le groupe constitué de : Caprin2 et Ipo8, et l'une de leurs combinaisons.
18. La cellule de mammifère de l'un quelconque des modes de réalisation 13 à 17, où ladite protéine FAM60A comprend une séquence qui est au moins identique à 80 % avec l'une quelconque des séquences choisies dans le groupe constitué des SEQ ID NO : 28, 29, et 30.
19. La cellule de mammifère de l'un quelconque des modes de réalisation 13 à 18, où ledit gène FAM60A comprend une séquence qui est au moins identique à 80 % avec SEQ ID NO : 31.
20. La cellule de mammifère du mode de réalisation 13, où ladite cellule comprend une mutation dans le promoteur, l'UTR en 5', ou l'UTR en 3' dudit gène FAM60A.
21. La cellule de mammifère du mode de réalisation 13, où ladite protéine FAM60A comprend une mutation qui réduit son activité, comparativement à un témoin.
22. La cellule de mammifère de l'un quelconque des modes de réalisation 1 à 21, où le taux d'expression ou

l'activité de la matriptase est réduit dans ladite cellule, comparativement à un témoin.

- 5 23. La cellule de mammifère du mode de réalisation 22, où au moins une copie de la séquence génomique du gène de la matriptase, ou au moins 50 % de la séquence codante du gène de la matriptase, est déléetée.
- 10 24. La cellule de mammifère du mode de réalisation 23, où les deux copies des séquences génomiques du gène de la matriptase, ou au moins 50 % de la séquence codante dudit gène de la matriptase à partir de chaque copie, sont déléetées.
- 15 25. La cellule de mammifère de l'un quelconque des modes de réalisation 22 à 25, où ladite matriptase comprend une séquence qui est au moins identique à 80 % avec l'une quelconque des séquences choisies dans le groupe constitué des SEQ ID NO : 32, 33, et 34.
- 20 26. La cellule de mammifère du mode de réalisation 22, où ladite cellule comprend une mutation dans l'exon 2 du gène de la matriptase.
27. La cellule de mammifère du mode de réalisation 22, où ladite cellule comprend une mutation dans le promoteur, l'UTR en 5', ou l'UTR en 3' du gène de la matriptase.
- 25 28. La cellule de mammifère du mode de réalisation 22, où ladite matriptase comprend une mutation qui réduit son activité, comparativement à un témoin.
29. La cellule de mammifère du mode de réalisation 28, où ladite matriptase comprend une mutation dans le domaine catalytique.
- 30 30. La cellule de mammifère de l'un quelconque des modes de réalisation 1 à 29, où la dihydrofolate réductase (DHFR) endogène de ladite cellule de mammifère est déficiente.
31. La cellule de mammifère de l'un quelconque des modes de réalisation 1 à 29, où la dihydrofolate réductase

(DHFR) endogène de ladite cellule de mammifère est compétente. BE2016/5877

- 5 32. La cellule de mammifère de l'un quelconque des modes de réalisation 1 à 31, où ledit fragment formant un complexe de la gH ne comprend pas la séquence signal d'une protéine gH pleine longueur.
- 10 33. La cellule de mammifère de l'un quelconque des modes de réalisation 1 à 32, où ledit fragment formant un complexe de la gH ne comprend pas le domaine transmembranaire d'une protéine gH pleine longueur.
34. La cellule de mammifère de l'un quelconque des modes de réalisation 1 à 33, où ledit fragment formant un complexe de la gH comprend l'ectodomaine d'une protéine gH pleine longueur.
- 15 35. La cellule de mammifère de l'un quelconque des modes de réalisation 1 à 34, où ladite gH comprend une séquence choisie dans le groupe constitué des SEQ ID NO : 1, 2, 3, 4, et 5.
- 20 36. La cellule de mammifère de l'un quelconque des modes de réalisation 1 à 35, où ledit fragment formant un complexe de la gH (i) forme une partie du complexe pentamère gH/gL/pUL128/pUL130/pUL131 ; et (ii) comprend au moins un épitope provenant de SEQ ID NO : 1, SEQ ID NO : 2, SEQ ID NO : 3, SEQ ID NO : 4, ou SEQ ID NO : 5.
- 25 37. La cellule de mammifère de l'un quelconque des modes de réalisation 1 à 36, où ledit fragment formant un complexe de la gL ne comprend pas la séquence signal d'une protéine gL pleine longueur.
- 30 38. La cellule de mammifère de l'un quelconque des modes de réalisation 1 à 37, où ladite gL comprend une séquence choisie dans le groupe constitué des SEQ ID NO : 6, 7, 8, et 9.

39. La cellule de mammifère de l'un quelconque des modes de réalisation 1 à 38, où ledit fragment formant un complexe de la gL (i) forme une partie du complexe pentamère gH/gL/pUL128/pUL130/pUL131 ; et
5 (ii) comprend au moins un épitope provenant de SEQ ID NO : 6, SEQ ID NO : 7, SEQ ID NO : 8, ou SEQ ID NO : 9.
40. La cellule de mammifère de l'un quelconque des modes de réalisation 1 à 39, où ledit fragment formant un
10 complexe de la pUL128 ne comprend pas la séquence signal d'une protéine pUL128 pleine longueur.
41. La cellule de mammifère de l'un quelconque des modes de réalisation 1 à 40, où ladite pUL128 comprend une
15 séquence choisie dans le groupe constitué des SEQ ID NO : 10, 11, 12 et 13.
42. La cellule de mammifère de l'un quelconque des modes de réalisation 1 à 41, où ledit fragment formant un
20 complexe de la pUL128 (i) forme une partie du complexe pentamère gH/gL/pUL128/pUL130/pUL131 ; et
(ii) comprend au moins un épitope provenant de SEQ ID NO : 10, SEQ ID NO : 11, SEQ ID NO : 12, ou SEQ ID NO : 13.
43. La cellule de mammifère de l'un quelconque des modes de réalisation 1 à 42, où ledit fragment formant un
25 complexe de la pUL130 ne comprend pas la séquence signal d'une protéine pUL130 pleine longueur.
44. La cellule de mammifère de l'un quelconque des modes de réalisation 1 à 43, où ladite pUL130 comprend une
30 séquence choisie dans le groupe constitué des SEQ ID NO : 14, 15, et 16.
45. La cellule de mammifère de l'un quelconque des modes de réalisation 1 à 44, où ledit fragment formant un
complexe de la pUL130 (i) forme une partie du complexe pentamère gH/gL/pUL128/pUL130/pUL131 ; et

- (ii) comprend au moins un épitope provenant de SEQ ID NO : 14, SEQ ID NO : 15, ou SEQ ID NO : 16.
46. La cellule de mammifère de l'un quelconque des modes de réalisation 1 à 45, où ledit fragment formant un complexe de la pUL131 ne comprend pas la séquence signal d'une protéine pUL131 pleine longueur.
47. La cellule de mammifère de l'un quelconque des modes de réalisation 1 à 46, où ladite pUL131 comprend une séquence choisie dans le groupe constitué des SEQ ID NO : 17, 18, 19, et 20.
48. La cellule de mammifère de l'un quelconque des modes de réalisation 1 à 47, où ledit fragment formant un complexe de la pUL131 (i) forme une partie du complexe pentamère gH/gL/pUL128/pUL130/pUL131 ; et (ii) comprend au moins un épitope provenant de SEQ ID NO : 17, SEQ ID NO : 18, SEQ ID NO : 19, ou SEQ ID NO : 20.
49. La cellule de mammifère de l'un quelconque des modes de réalisation 1 à 48, où ledit complexe pentamère est soluble.
50. La cellule de mammifère de l'un quelconque des modes de réalisation 1 à 49, où ledit complexe pentamère est sécrété à partir de la cellule hôte.
51. Une culture à grande échelle comprenant la cellule de mammifère de l'un quelconque des modes de réalisation 1 à 50, où ladite culture a une taille d'au moins 20 litres.
52. Une culture à grande échelle comprenant la cellule de mammifère de l'un quelconque des modes de réalisation 1 à 51, où ladite culture a une taille d'au moins 50 litres.
53. La culture à grande échelle du mode de réalisation 51 ou 52, où le rendement du complexe pentamère du CMV est d'au moins 0,05 g/l.

54. La culture à grande échelle du mode de réalisation 51 ou 52, où le rendement du complexe pentamère du CMV est d'au moins 0,1 g/l.
55. Un complexe pentamère du cytomégalovirus (CMV) produit par la cellule de mammifère de l'un quelconque des modes de réalisation 1 à 50, ou la culture à grande échelle de l'un quelconque des modes de réalisation 51 à 54.
56. Une composition comprenant le complexe pentamère du mode de réalisation 55.
57. Un procédé de production du complexe pentamère du cytomégalovirus (CMV), où ledit complexe pentamère comprend : la gH ou l'un de ses fragments formant un complexe, la gL ou l'un de ses fragments formant un complexe, la pUL128 ou l'un de ses fragments formant un complexe, la pUL130 ou l'un de ses fragments formant un complexe, et la pUL131 ou l'un de ses fragments formant un complexe, comprenant :
- (i) la culture de la cellule de mammifère de l'un quelconque des modes de réalisation 1 à 50 dans des conditions appropriées, exprimant de cette façon ledit complexe pentamère ; et
- (ii) la récolte dudit complexe pentamère à partir de la culture.
58. Le procédé du mode de réalisation 57, comprenant en outre la purification dudit complexe pentamère.
59. Un complexe pentamère du cytomégalovirus (CMV) produit par le procédé du mode de réalisation 57 ou 58.
60. Une composition comprenant le complexe pentamère du mode de réalisation 59.
61. La composition du mode de réalisation 56 ou 60, où ledit complexe pentamère a une pureté d'au moins 95 %, en masse.
62. La composition du mode de réalisation 61, comprenant en outre un adjuvant, comme un sel d'aluminium, ou le MF59.

63. La composition du mode de réalisation 61 ou 62, pour une utilisation dans l'induction d'une réponse immunitaire contre le CMV.

5 64. Le complexe pentamère du cytomégalovirus (CMV) du mode de réalisation 55, où ledit complexe pentamère est immunogène.

10 65. Le complexe pentamère du cytomégalovirus (CMV) du mode de réalisation 64, où les anticorps dirigés contre ledit complexe pentamère sont des anticorps neutralisants.

15

20

25

30

Légendes des figures

Figure 1	
Dual Vector	Vecteur double
Multiple IRES	IRES multiples
gH-6His-stuffer	gH-6His-fragment de remplissage
CMV Prom	Promoteur du CMV
Single Vector	Vecteur simple
Multiple Promoters	Promoteurs multiples
C54S Dual vector	Vecteur double C54S
Figure 2	
1st CMV Promoter	1 ^{er} promoteur du CMV
Anchor domain	Domaine d'ancrage
EV71 (7423) IRES	IRES de l'EV71 (7423)
T3 promoter	Promoteur T3
SV40polyA	PolyA du SV40
phage f1 region	Région du phage f1
SV40enhprom	Amplificateur/promoteur du SV40
SV40 minimum origin of replication	Origine de réplication minimale du SV40
synth polyA	PolyA synthétique
SV40prom	Promoteur du SV40
bp	pb
CMVprom/enh	Promoteur/amplificateur du CMV
c-DNA Hyg	ADNc Hyg
Figure 3	
MW	PM
gH/N-terminal gL	gH/gL N-terminale

C-terminal gL/UL128	gL C-terminale/UL128
Non-boiled/Non- reduced	Non porté à ébullition/non réduit
#	No.
Figure 4	
Back-up cells	Cellules de secours
Controlled batches	Lots contrôlés
Uncontrolled batches	Lots non contrôlés
(14-day batch, 36.5C in DM122+MTX)	(lot de 14 jours, 36,5 °C dans du DM122 + MTX)
2w	2 sem
Figure 5	
CMV bioreactor run VF7	Passage du VF7 dans le bioréacteur du CMV
viability	viabilité
day	jour
hours since reactor start	heures depuis le départ du réacteur
titer	titre
mL	ml
harvest	récolte
viable cells [vc/mL]	cellules viables [cv/ml]

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

5 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

10 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

15 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
 20 340 345 350

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

25 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

30 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
 35 420 425 430

His Leu Ile Pro Gln Trp Ala Leu Arg Gln Ile Ala Asp Phe Ala Leu
 435 440 445
 5
 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
 10 465 470 475 480
 Thr Thr Glu Arg Arg Glu Ile Phe Ile Val Glu Thr Gly Leu Cys Ser
 485 490 495
 15 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
 20
 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
 25 545 550 555 560
 Ser Thr Leu Glu Thr Phe Pro Asp Leu Phe Cys Leu Pro Leu Gly Glu
 565 570 575
 30 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
 35

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
 5 625 630 635 640

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

10 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

15 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
 20 705 710 715 720

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

25 Arg Met Leu Lys Thr Cys
 740

<210> 2
 <211> 742

30 <212> PRT
 <213> Herpèsvirus 5 humain

<220>
 <221> AUTRE(S) CARACTERISTIQUE(S)

35 <222> (1)..(742)

<223> /note="Souche Towne"

<400> 2

```

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

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

10 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

15 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
20 85 90 95

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

25 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

30 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
35 165 170 175

```

Trp Thr Glu Ser His Thr Thr Ser Gly Leu His Arg Pro His Phe Asn
 180 185 190
 5
 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
 10 210 215 220
 Val Lys Ile Thr Leu Thr Glu Asp Phe Phe Val Val Thr Val Ser Ile
 225 230 235 240
 15 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
 20 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
 25 290 295 300
 Asn Tyr Leu Asp Leu Ser Ala Leu Leu Arg Asn Ser Phe His Arg Tyr
 305 310 315 320
 30 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
 35

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

5 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

10 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

15 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
 20 450 455 460

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

25 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

30 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
 35 530 535 540

Pro Thr Thr Val Pro Ala Ala Leu Ser Ile Leu Ser Thr Met Gln Pro
 545 550 555 560
 5
 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
 10 580 585 590
 Asn Gln Tyr Leu Ile Lys Gly Ile Ser Tyr Pro Val Ser Thr Thr Val
 595 600 605
 15 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
 20 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
 25 660 665 670
 Asp Asp Val Leu Phe Ala Leu Asp Pro Tyr Asn Glu Val Val Val Ser
 675 680 685
 30 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
 35

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
5 740

<210> 3
<211> 743
<212> PRT
10 <213> Herpèsvirus 5 humain

<220>
<221> AUTRE(S) CARACTERISTIQUE(S)
<222> (1)..(743)
15 <223> /note="Souche AD169"

<400> 3
Met Arg Pro Gly Leu Pro Pro Tyr Leu Thr Val Phe Thr Val Tyr Leu
1 5 10 15
20 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
25 35 40 45
Pro Ile Arg Phe Leu Arg Glu Asn Thr Thr Gln Cys Thr Tyr Asn Ser
50 55 60
30 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
35

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

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

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

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

Ala Arg Gln Glu Glu Ala Gly Thr Glu Ile Ser Ile Pro Arg Ala Leu
 15 355 360 365

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

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

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

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

Gln His Leu Ile Pro Gln Trp Ala Leu Arg Gln Ile Ala Asp Phe Ala
 30 435 440 445

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

35 Arg Gln Glu Leu Tyr Leu Met Gly Ser Leu Val His Ser Met Leu Val

465 470 475 480
 His Thr Thr Glu Arg Arg Glu Ile Phe Ile Val Glu Thr Gly Leu Cys
 485 490 495
 5
 Ser Leu Ala Glu Leu Ser His Phe Thr Gln Leu Leu Ala His Pro His
 500 505 510
 10 515 520 525
 His Glu Tyr Leu Ser Asp Leu Tyr Thr Pro Cys Ser Ser Ser Gly Arg
 Arg Asp His Ser Leu Glu Arg Leu Thr Arg Leu Phe Pro Asp Ala Thr
 530 535 540
 15 545 550 555 560
 Val Pro Ala Thr Val Pro Ala Ala Leu Ser Ile Leu Ser Thr Met Gln
 Pro Ser Thr Leu Glu Thr Phe Pro Asp Leu Phe Cys Leu Pro Leu Gly
 565 570 575
 20
 Glu Ser Phe Ser Ala Leu Thr Val Ser Glu His Val Ser Tyr Val Val
 580 585 590
 25 595 600 605
 Thr Asn Gln Tyr Leu Ile Lys Gly Ile Ser Tyr Pro Val Ser Thr Thr
 Val Val Gly Gln Ser Leu Ile Ile Thr Gln Thr Asp Ser Gln Thr Lys
 610 615 620
 30 625 630 635 640
 Cys Glu Leu Thr Arg Asn Met His Thr Thr His Ser Ile Thr Ala Ala
 Leu Asn Ile Ser Leu Glu Asn Cys Ala Phe Cys Gln Ser Ala Leu Leu
 645 650 655
 35

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
 5 675 680 685

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

10 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

15 Tyr Arg Met Leu Lys Thr Cys
 740

<210> 4
 20 <211> 715
 <212> PRT
 <213> Herpèsvirus 5 humain

<220>
 25 <221> AUTRE(S) CARACTERISTIQUE(S)
 <222> (1)..(715)
 <223> /note="Souche Merlin"

<400> 4
 30 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

35

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

5 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

10 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

15 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

25 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

30 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
5	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
10	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		
15		290	295 300
	Asn Tyr Leu Asp Leu Ser Ala Leu Leu Arg Asn Ser Phe His Arg Tyr		
		305	310 315 320
20	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
25	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		
30		370	375 380
	Ser Gln Thr Pro Pro Arg Thr Thr Leu Leu Leu Tyr Pro Thr Ala Val		
		385	390 395 400
35	Asp Leu Ala Lys Arg Ala Leu Trp Thr Pro Asn Gln Ile Thr Asp Ile		


```

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
5      610                      615                      620

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

10  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

15  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
20      690                      695                      700

Glu Val Thr Asp Val Val Val Asp Ala Thr Asp
705                      710                      715

25  <210> 5
      <211> 692
      <212> PRT
      <213> Herpèsvirus 5 humain

30  <220>
      <221> AUTRE(S) CARACTERISTIQUE(S)
      <222> (1)..(692)
      <223> /note="Souche Merlin"

35  <400> 5

```

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
 5 20 25 30

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

10 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

15 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
 20 100 105 110

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

25 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

30 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
 35 180 185 190

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

5 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

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

Leu Val Lys Lys Asp Gln Leu Asn Arg His Ser Tyr Leu Lys Asp Pro
 15 260 265 270

Asp Phe Leu Asp Ala Ala Leu Asp Phe Asn Tyr Leu Asp Leu Ser Ala
 275 280 285

20 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

25 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
 30 340 345 350

Ile Gln Glu Phe Met Ile Thr Cys Leu Ser Gln Thr Pro Pro Arg Thr
 355 360 365

35 Thr Leu Leu Leu Tyr Pro Thr Ala Val Asp Leu Ala Lys Arg Ala Leu

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
 5 580 585 590

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

10 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

15 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
 20 660 665 670

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

25 Asp Ala Thr Asp
 690

<210> 6
 <211> 278
 30 <212> PRT
 <213> Herpèsvirus 5 humain

<220>
 <221> AUTRE(S) CARACTERISTIQUE(S)
 35 <222> (1)..(278)

<223> /note="Souche Merlin"

<400> 6

```

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

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

10 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

15 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
20 85 90 95

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

25 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

30 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
35 165 170 175

```

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

5 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

10 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
 15 245 250 255

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

20 Gln Ala Val Asp Ala Arg
 275

<210> 7
 <211> 278

25 <212> PRT
 <213> Herpèsvirus 5 humain

<220>
 <221> AUTRE(S) CARACTERISTIQUE(S)

30 <222> (1)..(278)
 <223> /note="Souche Towne"

<400> 7

Met Cys Arg Arg Pro Asp Cys Gly Phe Ser Phe Ser Pro Gly Pro Val
 35 1 5 10 15

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

5 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

10 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
15 85 90 95

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

20 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

25 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
30 165 170 175

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

35 Val Ala Ile Arg Asn Glu Ala Thr Arg Thr Asn Arg Ala Val Arg Leu

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

5 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

10 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

15 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

25 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

30 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

<210> 10
 <211> 130
 <212> PRT
 5 <213> Herpèsvirus 5 humain

 <220>
 <221> AUTRE(S) CARACTERISTIQUE(S)
 <222> (1)..(130)
 10 <223> /note="Souche Merlin"

 <400> 10
 Met Ser Pro Lys Asp Leu Thr Pro Phe Leu Thr Ala Leu Trp Leu Leu
 1 5 10 15
 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
 20 35 40 45
 Cys Asn Arg Phe Thr Val Ala Leu Arg Cys Pro Asp Gly Glu Val Cys
 50 55 60
 25 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
 30 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
 35 115 120 125

Pro Tyr
130

5 <210> 11
<211> 171
<212> PRT
<213> Herpèsvirus 5 humain

10 <220>
<221> AUTRE(S) CARACTERISTIQUE(S)
<222> (1)..(171)
<223> /note="Souche Towne"

15 <400> 11
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 20 25 30

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

25 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
30
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
35 100 105 110

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

5 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

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

<210> 12

15 <211> 171
 <212> PRT
 <213> Herpèsvirus 5 humain

<220>

20 <221> AUTRE(S) CARACTERISTIQUE(S)
 <222> (1)..(171)
 <223> /note="Souche AD169"

<400> 12

25 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

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
 35 50 55 60

His Pro Asn Leu Ile Val
 210

<210> 15
 5 <211> 229
 <212> PRT
 <213> Herpèsvirus 5 humain

<220>
 10 <221> AUTRE(S) CARACTERISTIQUE(S)
 <222> (1)..(229)
 <223> /note="Souche Towne"

<400> 15
 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 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
 25 50 55 60

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

30 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

35

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
 5 130 135 140

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

10 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

15 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
 20 210 215 220

Glu Asn Arg Ala Ser
 225

25 <210> 16
 <211> 189
 <212> PRT
 <213> Herpèsvirus 5 humain

30 <220>
 <221> AUTRE(S) CARACTERISTIQUE(S)
 <222> (1)..(189)
 <223> /note="Souche Merlin"

35 <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
 5 20 25 30

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

10 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

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

Leu Gln Arg Met Pro Arg Thr Ala Ser Lys Pro Ser Asp Gly Asn Val
 20 100 105 110

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

25 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

30 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
 35 180 185

<210> 17
 <211> 129
 <212> PRT
 5 <213> Herpèsvirus 5 humain

 <220>
 <221> AUTRE(S) CARACTERISTIQUE(S)
 <222> (1)..(129)
 10 <223> /note="Souche Merlin"

 <400> 17
 Met Arg Leu Cys Arg Val Trp Leu Ser Val Cys Leu Cys Ala Val Val
 1 5 10 15
 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
 20 35 40 45
 Arg Tyr Lys Tyr Val Glu Gln Leu Val Asp Leu Thr Leu Asn Tyr His
 50 55 60
 25 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
 30
 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
 35 115 120 125

Asn

5 <210> 18
 <211> 129
 <212> PRT
 <213> Herpèsvirus 5 humain

10 <220>
 <221> AUTRE(S) CARACTERISTIQUE(S)
 <222> (1)..(129)
 <223> /note="Souche Towne"

15 <400> 18
 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 20 25 30

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

25 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

30 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
 35 100 105 110

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

5 Asn

<210> 19
 <211> 74

10 <212> PRT
 <213> Herpèsvirus 5 humain

<220>
 <221> AUTRE(S) CARACTERISTIQUE(S)
 15 <222> (1)..(74)
 <223> /note="Souche AD169"

<400> 19
 Met Arg Leu Cys Arg Val Trp Leu Ser Val Cys Leu Cys Ala Val Val
 20 1 5 10 15

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

25 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

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

<210> 20
 35 <211> 111

<212> PRT
 <213> Herpèsvirus 5 humain

 <400> 20
 5 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
 10
 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
 15 50 55 60

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

 20 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
 25
 <210> 21
 <211> 1547
 <212> PRT
 <213> Cricetulus griseus
 30
 <400> 21
 Met Asn Trp Asn Ala Lys Pro Glu Asn Ala Ala Pro Asn Pro Pro Tyr
 1 5 10 15

 35 Ser Lys Ser Gln Ser Ser Leu Leu Gln Gln Phe Leu Met Pro Ser Thr

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

5 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

10 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
 15 275 280 285

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

20 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

25 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
 30 355 360 365

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

35 Asn Gln Ile Leu Asp Thr Asn Val Thr Lys Glu Lys Leu Val Arg Asp

	Ala	Pro	Glu	Val	Val	Ser	Glu	Thr	Val	Tyr	Pro	Val	Val	Lys	Glu	Gly
				580					585					590		
5	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				
10	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
15					645					650					655	
	Pro	Gln	Gly	Ala	Glu	Ala	Leu	Pro	Asn	Pro	Arg	Asp	Ser	Thr	Ile	Val
				660					665					670		
20	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				
25	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
30					725					730					735	
	Leu	Lys	Asp	Lys	Cys	Met	Gln	Cys	Thr	Asp	Val	Pro	His	Glu	Val	Thr
				740						745				750		
35	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

5 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

10 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
 15 835 840 845

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

20 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

25 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
 30 915 920 925

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

35 Met Leu Asp Leu Glu Lys Asp Lys Ile His Cys Cys Ala Leu Gly Trp

	Lys Lys	His Leu Phe Gln Asp	Ser Asp Pro Val Lys	Gly His Val
	1130	1135	1140	
5	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	
10	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
15	1190	1195	1200	
	Gln Glu	Gln Asn Lys Asn Ala	Gly Gly Thr Leu Lys	Leu Cys Ser
	1205	1210	1215	
20	Thr Leu	Thr Glu Pro Asn Glu	Arg Ala Cys Ala Lys	Glu Lys Ile
	1220	1225	1230	
	Val Thr	Asn Ser Glu Pro Ser	Asp Ser Lys Gly Ser	Ser Ser Lys
	1235	1240	1245	
25	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
30	1265	1270	1275	
	Glu Asn	Val Pro Cys Asp Ser	Glu Pro Met Lys Ser	Ser Lys His
	1280	1285	1290	
35	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								
5	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								
10	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							
15	1370			1375			1380								
	Tyr Leu	Asn Arg	Val Ala	Phe Lys	Cys Thr	Glu Gln	Glu Ser	Ile							
	1385			1390			1395								
20	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								
25	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							
30	1445			1450			1455								
	Pro Glu	Lys Glu	Arg Ala	Pro Val	Gln Val	Ser Gly	Ile Lys	Thr							
	1460			1465			1470								
35	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							
5	Lys	Arg	Ser	Phe	Ser	Ala	Asp	Glu	Phe	Glu	Thr	Leu	Gln	Asn	Pro				
	1505						1510					1515							
10	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							
15	<210>	22																	
	<211>	1515																	
	<212>	PRT																	
	<213>	Cricetinae	sp.																
20	<220>																		
	<221>	RESIDU	MODIFIE																
	<222>	(879)..	(879)																
	<223>	Tout	acide	aminé															
25	<400>	22																	
	Met	Asn	Trp	Asn	Ala	Lys	Pro	Glu	Asn	Ala	Ala	Pro	Asn	Pro	Pro	Tyr			
	1			5						10				15					
30	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						
35	Ile	Tyr	Pro	Thr	Asn	Ser	Asn	Ser	Val	Ser	Gln	Pro	Leu	Leu	Asn	Val			

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
 5 260 265 270

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

10 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
 15 305 310 315 320

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
 20 340 345 350

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

25 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 Glu Lys Leu Val Arg Asp
 385 390 395 400

30 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
 35 420 425 430

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

5 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

10 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
 15 500 505 510

Asp Ser Ser Asp Val Ala Val Gly Lys Gly Thr Glu Leu Gln Ile Ala
 515 520 525

20 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
 25 545 550 555 560

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
 30 580 585 590

Thr Val Ser Ala Glu Leu Ser Leu Pro Gly His Lys Glu Lys Gln His
 595 600 605

35 Lys Pro Thr Gln Thr Asp Leu Asp Thr Ala Asp Gly Ser Leu Gly Lys

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

5 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

10 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

15 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
 20 900 905 910

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

25 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

30 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
 35 980 985 990

	Gly His Phe Ser Asp Ile His Gly Glu Lys Ile Lys Thr Ser Glu Thr		
	995	1000	1005
5	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
10	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		
15	1055	1060	1065
	Pro Glu Thr Phe Asn Val Val Thr Ser Asn Ser Asp Lys Asn Met		
	1070	1075	1080
20	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
25	Leu Pro Asn Lys Asp Pro Arg Arg Arg Asn Thr Phe Leu Val Gln		
	1115	1120	1125
	Ser Val Ser Pro Glu Lys Lys Lys Leu Lys Phe Lys Ser Gly Ser		
30	1130	1135	1140
	Ser Lys Leu Lys Tyr Phe Glu Lys Arg Lys Met Asp His Leu Leu		
	1145	1150	1155
35	Ile Ser Asp Val Glu Ile Lys Lys Lys Lys Tyr Glu Lys Gln Glu		

	Lys Ser	Ser Gln Leu Gln Val	Lys Glu Lys Arg Lys	Gln Tyr Leu
	1340	1345	1350	
5	Asn Arg	Val Ala Phe Lys Cys	Thr Glu Gln Glu Ser	Ile Cys Leu
	1355	1360	1365	
	Thr Lys	Leu Asp Ser Ala Ser	Lys Lys Leu Ser Lys	Glu Lys Glu
	1370	1375	1380	
10	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
15	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
20	1430	1435	1440	
	Glu Asp	Trp Leu Lys Cys Ile	Pro Thr Arg Thr Lys	Met Pro Glu
	1445	1450	1455	
25	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
30	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	
35	1505	1510	1515	

<210> 23
 <211> 1747
 <212> PRT
 5 <213> Homo sapiens

<400> 23
 Met Asn Trp Asn Glu Lys Pro Lys Ser Ala Thr Leu Pro Pro Leu Tyr
 1 5 10 15
 10 Pro Lys Ser Gln Pro Pro Phe Leu His Gln Ser Leu Ile Asn Gln Ile
 20 25 30
 15 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
 20 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
 25 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
 30 115 120 125
 His Ser His Ile Gly Ala Thr Val Ser His Gln Thr Asp Phe Gly Ala
 130 135 140
 35 Asn Val Pro Asn Met Pro Ala Leu Gln Ser Gln Leu Ile Thr Ser Asp

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

5 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

10 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

15 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
 20 435 440 445

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

25 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

30 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
 35 515 520 525

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

5 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

10 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
 15 595 600 605

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

20 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

25 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
 30 675 680 685

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

35 Ile Ser Lys Pro Ala Asn Ile His Val Lys Ser Pro Cys Ser Val Val

	Thr Ser	Glu Asp Gln Thr Ala	Asp Lys Thr Ser Ser	Asp Ser Lys
	1085	1090	1095	
5	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	
10	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
15	1145	1150	1155	
	Asp Ser	Val Ile Leu Asp Ser	Glu Lys Asp Asp Ile	His Cys Cys
	1160	1165	1170	
20	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	
25	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
30	1220	1225	1230	
	Leu Val	Asn Asn Pro Lys Thr	Pro Pro Asp Gly Lys	Ser His Phe
	1235	1240	1245	
35	Pro Glu	Leu Gln Asp Asp Ser	Arg Lys Asp Thr Pro	Lys Thr Lys

	Lys Ser	Val Asp Thr Lys Ala	Ser Ser Ser Lys Phe	Ser Arg Ile
	1430	1435	1440	
5	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	
10	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	
15	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
20	1520	1525	1530	
	Arg Asn	Val Lys Glu Lys Val	Gly Gly Lys Gln Pro	Asp Lys Ile
	1535	1540	1545	
25	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
30	1565	1570	1575	
	Lys Leu	Tyr Leu Asn Arg Val	Gly Phe Lys Cys Thr	Glu Arg Glu
	1580	1585	1590	
35	Ser Ile	Ser Leu Thr Lys Leu	Glu Ser Ser Pro Arg	Lys Leu His

<213> Homo sapiens

<400> 24

```

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

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

10 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

15 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
20 85 90 95

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

25 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

30 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
35 165 170 175

```

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

5 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

10 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
 15 245 250 255

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

20 Thr Asp Lys Arg Pro 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

25 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
 30 325 330 335

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

35 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							
5		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			
10					405					410						415				
		Asn	Lys	Asp	Leu	Leu	Met	Ala	Ala	Gly	Cys	Ile	Lys	Met	Thr	Asn	Thr			
					420					425					430					
15		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			
20			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			
25					485						490					495				
		Phe	Asn	Gln	Val	Asp	Ser	Val	Leu	Pro	Asn	Pro	Val	Tyr	Ser	Glu	Lys			
					500						505					510				
30		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			
35			530					535						540						

	Ser	Glu	Asn	Val	Thr	Lys	Val	Glu	Gln	Asn	Ser	Pro	Ala	Val	Cys	Glu
	545					550					555					560
5	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	
10	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
15		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
20					645					650						655
	Thr	Lys	Gly	Met	Pro	Ala	Lys	Ser	Asp	Ser	Ser	Cys	Ser	Met	Glu	Val
				660						665					670	
25	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
30		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
35					725						730					735

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

5 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

10 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
15 805 810 815

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

20 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

25 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
30 885 890 895

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

35 Phe Ser Ser Leu Pro Leu Lys Met Val Glu Pro Gln Lys Pro Ser Leu

	Asp Pro	Ala Asp	Gln Ile	Gln Ile	Thr Ile	Leu Ser	Ser Glu	Gln					
	1100			1105			1110						
5	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						
10	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						
15	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	Ser Glu	Glu Glu	Glu Lys	Gln				
20	1190			1195			1200						
	Lys Glu	Gln Cys	Ser Pro	Leu Asp	Thr Asn	Ser Cys	Lys Gln	Gly					
	1205			1210			1215						
25	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						
30	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					
35	1265			1270			1275						

	Phe Ser	Ser Lys Cys Asp Lys	Leu Asn Pro Leu Gln	Asn His Lys
	1280	1285	1290	
5	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	
10	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
15	1340	1345	1350	
	Gly Gln	Ser Leu Ser Pro Glu	Lys Ile Lys Leu Lys	Leu Lys Ser
	1355	1360	1365	
20	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	
25	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
30	1415	1420	1425	
	Lys Ser	Val Asp Thr Lys Ala	Ser Ser Ser Lys Phe	Ser Arg Ile
	1430	1435	1440	
35	Leu Thr	Pro Lys Glu Tyr Leu	Gln Arg Gln Lys His	Lys Glu Ala

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
 5 1640 1645 1650
 Asp Val Lys Pro His Pro Arg Lys Glu Gln Ala Pro Leu Gln Val
 1655 1660 1665
 10 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
 15 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
 20 1715 1720 1725
 Tyr Lys Gln Met Tyr Leu Glu Lys Arg Ser Arg Ser Leu Gly Ser
 1730 1735 1740
 25 Ser Pro Val Lys
 1745
 <210> 25
 <211> 4644
 30 <212> ADN
 <213> *Cricetulus griseus*
 <400> 25
 atgaattgga atgcaaaacc agagaatgct gcccctaaacc caccatattc taaaagccag 60
 35

	tcgtctcttt tgcagcagtt tttaatgcct tccacaactt ctcaaagttc tttcagctgt	120
	ctcccacata accaagaagc atgcatatat cccactaatt caaattcagt ttcacagcca	180
5	cttctgaaag tcaggagttt cataaatcct ccgatctctg tttctaattgt gcataatagg	240
	acagttgtgg cctcacagac ctcagtagaa agagtcacat atacaaatgt taaaggagcc	300
	caacaaccaa accacaattt gcaaacagtg tcttctggag ttgtgcaaaa tgcoctggatg	360
10	aattcaacaa tgaggaattt tatgccttct cttacagagg caaccatata tcataaacct	420
	gatgggtggc ctagtatgcc atatatgcat gcaccacaga gtcactctgt cacatcagac	480
15	acctactctg tgcaactaca gatgactcct tcaaactctg taagaggccc tgtaaacttac	540
	caaggaaatt atcaaggaaa tccgggactt aaccactoga tggcagggtga gcttggetgg	600
	gtacaatgtg catccagtga acttacttat ccagattaca gaccacotcc aaagcaatat	660
20	ccttatttac cacaaagctt tgtgcaagac acttctgttc agaaacaaaa ctttgtgtca	720
	tctacatcat tacaagtaa aaataatcag cttccacott ctacacagac cttaccatca	780
25	aaggccctg tacctgtgtc gtcatatcag tatgctgcag aaaccagcaa aagactcctt	840
	ccccccctt acagctgtag atatggaagc caacatgtgc aaaattctca gtctgtttct	900
	agacacttgc ctgtggaagt tctcagagt tcagaaatgc actcgtctga aaaaaagaaa	960
30	gatgcttaca aagtctttca acagcagtg cagagcacta gtaaaaatgt cagtacaata	1020
	ggaaaattct gtgagttgaa aattaatata aaacagtctt acaatgactc tgctggctct	1080
35	tctggggatg gtgttcatac tcttgttcaa aataatcaag aagaagaaa gtattcttat	1140

	aatccaagta caaatcaaat actagacaca aatgtcacia aagaaaagct ggtgaggat	1200
	attaaatcac tagtagaaat taaaaagaaa ttttcagaac ttgcaagaaa aattaaatc	1260
5	aacaaaaagc ttttgatggc agctggttgc agtaaaacag ctaatacttc ttatactgaa	1320
	ccaactcggc attctgaatt ttcagcaaaa gaaatgtctg ctaaaagga caatcagtgc	1380
10	tccatggaat tgctagcaac atgcctttct ctttgaaaa accaacctcc aaaaaccaca	1440
	gaagaaaatg tttcaaaacc tttagaagaa aaacaatata atgcatcaag aactagtaca	1500
	acagcgggtg gcccttcaaa tcccatgaat gaagttcatg tgaagaatth ttgttcaggt	1560
15	gttagaatt ctcagaaaat aaccacctcg tcacaaacag tcttqtcaat tctcacacca	1620
	gtttacgatt cttcagatgt agctgttga aaaggaacag agcttcagat tgetgtggtt	1680
20	tcacctttaa ttctttcaga tgtcagtact gtacctggga aagagttagc tctgaagtc	1740
	gtatctgaaa ctgtatatcc agttgtgaag gaaggcagtg tttgtagctt acaaaaccag	1800
	caggcagaaa atgcaacagt aactgctggc ttgccctttg atgttatcag agcagtagca	1860
25	agtgctactg taccagctga gctatcactg cctgggcata aagaaaagca gcacaaacca	1920
	acacagagtg atctagacat cgctgatggc agcctagga aacactctcc ccagggtgct	1980
30	gaagctttgc ctaaccctag ggacagcacc attgtgagtg ggcttatatt acagattgaa	2040
	agtatctggt ctcttcgaga aggtgatgta tcttacaatt cccaaatagc agagatatc	2100
	aactctgtac aaaatgagcc ccagaaacct tcacctgatc agcaagtaat taatagtcaa	2160
35		

caagaagaac aagtagataa ggttgctgaa aataaagact taagttttct gaaagacaag 2220
tgtatgcagt gtacagatgt tcctcatgaa gtcactgaac agccagagcc actgcagcct 2280
5 ttagagacaa catctgatga gtatggtgaa gcaaacggag aaatcctaga ggaaagcagt 2340
aaggagaatc ctggtgaaaa agagatgact aaggacatat tgtgttcacc agctgctggt 2400
cagcaagatc ctcaacctca ggaaattgac acagccagca gtaagtcagg acacagtttt 2460
10 tctacagtaa atgagattaa tgatgaaaaat gaacctgtct catacctaca tgaccagctg 2520
ttagaacttc taaaagagtt tccttatggc attgaaacta ttgccaggcc tgaagtttat 2580
15 gtgggccaac aaaagacaca tgaaatotta gaaaatcaaa ctggtagtaa aactggtaat 2640
gtgtctgggg ataacacaga ccaaataaaa attacagtat taaactcaga acaaatcaaa 2700
gaattatttc ctgaagagga tcagccatgt gatgtagaca aattggcaga acccgagaat 2760
20 acaaaaatca ttgcagaagt aaagagcctg tgtgattcac aggtcccag agaagaaagt 2820
cacaacctg gaatgttga tctggagaaa gataaaatcc attgctgtgc cttgggctgg 2880
25 ctctcaatgg tttatgaagg tgtgccacag tgtcagtgca gttccatgga agagaaagag 2940
aaagaccagt gttctttgga aatctctaata tgcaacaag gagagcaggc ctgcaatagt 3000
ggaatcacta tttttgaaat taatcctatt tctaataact caaaaagtcc tctgatccaa 3060
30 gaatctgaga aaggccattt ttctgacata catggtgaaa agataaaaac atctgaaaca 3120
aaaaacagca gctcaccaag ggtagaacag gaattaactg gtcatttttc aatgaaatgt 3180
35 taccagaaag ataaatctac aacaaaacag gatagctcac tgaaaacaga gcaaaaaata 3240

aaaaatcttt cttctaaatg tgacaaacca aatcccttaa aaagcagtaa aataccaacc 3300

5 cctgaaacat ttaatgtggt aacttccaac tctgataaaa atatgccagc atttttctaaa 3360

caagattctc agggaagcct gcagaagaaa cacctattcc aagactcaga tccagtaaaa 3420

ggacatgtat ggcttttgcc aaataaagat ccacgcagga ggaatacctt tttagtacag 3480

10 tcagtatcac cagaaaagaa aaagttaaaa ttcaaatogg gtagctccaa actgaaatat 3540

tttgaaaaaa gaaaaatgga ccatttgctt atctcagatg tggaaataaa aaagaagaaa 3600

15 tacgaaaaac aagagcagaa caaaaatgct ggaggcacac tcaaattatg tagtactctg 3660

actgaaccaa atgaaagagc ctgtgctaaa gaaaagatag tgacaaatc tgagccctca 3720

gactcaaagg gaagctcctc taagagtact agagttataa ctgtgcagga atattttacag 3780

20 cggaaaaaag acaaacatgt aataggaaat aatgcctcca aaaacatctg ttagagaaaat 3840

gtgccatgtg actctgaacc catgaagtc agtaaacatt ctgcatcacc tagtttggga 3900

25 aaattaattg agggccaggg tgtcagtgca gagactttaa aagaagtaga acataattcc 3960

accagccatg gcaaaaatct caagaccac cgttctgagg agactaggcc atacagtgtg 4020

tcaaatagta aagagaaatt ttataggaca catccagaca aatcttacat tgataaagct 4080

30 aaattagaaa gattgaccag tatgagtagt aagtcagcc agctccaggt aaaggaaaaa 4140

aggaaacagt acctgaatog agttgcattc aatgcacag aacaggaaag catttgctc 4200

35 accaaattg acagtgcac caagaagctt agtaagaga aagaaaagag tacagcatgt 4260

gcacccatga caaaagacta cacacacaag cccatgttgg agtttaaatt atgtccagat 4320
 gtgctattga agaatacaag ctccattgac aaaggggatg atccaaggcc tgggctgag 4380
 5 aaggagcgag cacctgtgca agtttcagga ataaaaacta caaaagaaga ctggttaaaa 4440
 tgtatcccaa caaggacaaa gatgcccga tcaagtgaac aacagatcg ggtgactca 4500
 agactcteta agagaagctt cagtgcagat gaatttgaaa ctetacaaaa ccagtaaaa 4560
 10 gactcaaatg tcatgttccg gactttcaaa aagatgtacc tggagaagag aagcaggagc 4620
 ctggggagca gtccagtga gtag 4644
 15 <210> 26
 <211> 112
 <212> ADN
 <213> *Cricetulus griseus*
 20 <400> 26
 tgctgggatt taaggggaaa gctttaataa aagatcttta tttgtatttc ttgcagattt 60
 gtgacattca aaaccacaga ctatgcaaca ctactactaa accaggtcaa at 112
 25 <210> 27
 <211> 36
 <212> ADN
 <213> *Cricetulus griseus*
 30 <400> 27
 cttcgggata gagtggtttt gcttttacca ccagga 36
 <210> 28
 <211> 221
 35 <212> PRT

<213> Homo sapiens

<400> 28

```

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

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

10 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

15 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
20 85 90 95

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

25 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

30 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
35 165 170 175

```


100 105 110
 Lys Glu Phe Lys Arg His Asn Ser Asp Ala His Ser Thr Thr Ser Ser
 115 120 125
 5
 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
 10 145 150 155 160
 Phe Leu Asp Leu Thr Tyr Trp Lys Arg Gln Lys Ile Cys Cys Gly Ile
 165 170 175
 15 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
 20 Gln Gly Pro Ala Pro Leu Pro Ile Ser Thr Gln Glu Trp
 210 215 220
 <210> 30
 25 <211> 221
 <212> PRT
 <213> *Cricetulus griseus*
 <400> 30
 30 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
 35

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
 5 50 55 60

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

10 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

15 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
 20 130 135 140

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

25 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

30 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
 35 210 215 220

<210> 31
 <211> 1236
 <212> ADN
 5 <213> *Cricetulus griseus*

 <400> 31
 ctcagaagaa aagatgtttg gttttcacia gccaaagatg taccgaagta tagagggctg 60
 10 ctgtatctgc agagccaagt cctccagctc tcggttcacg gacagtaaac gttatgaaaa 120
 ggacttcacg agctgttttg ggttgacaga gactcgtca ggagatatct gcaatgcctg 180
 tgtgctgctt gtgaaaagat ggaagaagtt gccagcagga tcaaaaaaaaa actggaatca 240
 15 tgtgtcacac tcaagggcag gaccagctc aaagacaaca ttgaaacca aaaaagttaa 300
 aactctatct ggaaacagga tgaaaagcaa ccagatcagt aactgcaga aggagttaa 360
 20 acgccacaac tctgatgctc acagtaccac ctcaagtgc tcgccagccc agtctccctg 420
 ctacagtaac cagtcagatg atggtcaga cacagagatg gcttcagct ctaacagaac 480
 tccagttttt tccttcttag atcttaccta ctggaaaaga cagaaaatat gttgtgggat 540
 25 catctataag ggccgttttg ggaagtctt catcgacacg catctctca agccttgctg 600
 cagcagtaag aaggcagctc ctgagaagcc tgaggaacag ggaccagcgc ctctgcccac 660
 30 ctctactcag gagtgtgac tgaggttcat gcagaaggga acaaagagca atttaactt 720
 tgaaaagacc acaaagcaac agactgaccc tcctatTTTT aacttgata cctgctattc 780
 tgccaaaaga cttttctag aatagttttt aatgggttac ccatcccccc atccaacaaa 840
 35

ctcggaagcc agttctagct tactgcaaga agagagtgta cataatattt aatagctga 900
 gtatttcata ggaaggctga atgctgctgt aaagtgctct ttaagtcttt tttttttttt 960
 5 aatcccctct aatgaatgag attagggggg tttcagggga cagagatggg atttgttgtg 1020
 tgataaacca tatgtagttt agtctttotg tggagaggca gtggttgggg cattttaaat 1080
 ggctggctac acttgttttc cctcatggt aatttgcac aactcagtag cacgacctgc 1140
 10 ccctagaagt agttaaagat ttttaaagtc taaggcgttg ccaaggttct gatgattcag 1200
 acctgtacta ctgattatta agcaggacag actgag 1236
 15 <210> 32
 <211> 855
 <212> PRT
 <213> *Cricetulus griseus*
 20 <400> 32
 Met Gly Ser Asn Arg Gly Arg Lys Ala Gly Gly Ser Ser Lys Asp Phe
 1 5 10 15
 Gly Ala Arg Leu Lys Tyr Ser Ser Gly Leu Glu Asn Met Asn Gly Phe
 25 20 25 30
 Glu Glu Gly Val Glu Phe Leu Pro Val Asn Asn Ala Lys Lys Val Glu
 35 40 45
 30 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
 35

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
 5 100 105 110

Phe Lys Asp Leu Ala Asn Gln Val Lys Glu Ala Leu Lys Leu Leu Tyr
 115 120 125

10 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

15 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
 20 180 185 190

Leu Thr Ser Val Val Ala Phe Pro Thr Asp Pro Arg Leu Leu Gly Arg
 195 200 205

25 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

30 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
 35 260 265 270

Asn Asp Leu Val Thr Val Tyr Asp Thr Leu Ser Pro Met Glu Pro His
 275 280 285

5 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

10 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
 15 340 345 350

Ser Pro Tyr Tyr Pro Gly His Tyr Pro Pro Asn Ile Asp Cys Thr Trp
 355 360 365

20 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
 25 385 390 395 400

Tyr Val Glu Ile Asn Gly Glu Arg Tyr Cys Gly Glu Lys Ser Gln Phe
 405 410 415

30 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

35

	Ser	Asn	Asp	Pro	Cys	Pro	Gly	Met	Phe	Met	Cys	Asn	Thr	Gly	Arg	Cys
	450						455					460				
5	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	
10	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
			515					520					525			
15	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
20	545					550					555					560
	Arg	Val	Lys	Val	Val	Ser	Cys	Thr	Lys	Tyr	Thr	Tyr	Arg	Cys	His	Asn
					565					570					575	
25	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			
30	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
35	625					630					635					640

Cys Gly Ala Ser Leu Ile Ser Pro Asn Trp Leu Val Ser Ala Ala His
 645 650 655

5 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

10 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
 15 705 710 715 720

Ala Glu Tyr Ser Thr Val Val Arg Pro Ile Cys Leu Pro Asp Thr Thr
 725 730 735

20 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

25 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
 30 785 790 795 800

Cys Gln Gly Asp Ser Gly Gly Pro Leu Ser Ser Val Glu Thr Glu Gly
 805 810 815

35 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
 5
 Ile Lys Glu Gln Thr Gly Val
 850 855
 <210> 33
 10 <211> 855
 <212> PRT
 <213> Homo sapiens
 <400> 33
 15 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 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
 25 50 55 60
 Leu Leu Leu Val Leu Leu Gly Ile Gly Phe Leu Val Trp His Leu Gln
 65 70 75 80
 30 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
 35

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
 5 130 135 140

Ala Phe Ser Glu Gly Ser Val Ile Ala Tyr Tyr Trp Ser Glu Phe Ser
 145 150 155 160

10 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

15 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
 20 210 215 220

Leu Met Arg Phe Thr Thr Pro Gly Phe Pro Asp Ser Pro Tyr Pro Ala
 225 230 235 240

25 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

30 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
 35 290 295 300

				485					490					495				
	Lys	Asn	Lys	Phe	Cys	Lys	Pro	Leu	Phe	Trp	Val	Cys	Asp	Ser	Val	Asn		
				500					505					510				
5	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		
10		530					535						540					
	Asn	Gly	Lys	Asp	Asp	Cys	Gly	Asp	Gly	Ser	Asp	Glu	Ala	Ser	Cys	Pro		
	545					550					555					560		
15	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		
20				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		
25		610					615						620					
	Glu	Trp	Pro	Trp	Gln	Val	Ser	Leu	His	Ala	Leu	Gly	Gln	Gly	His	Ile		
	625					630					635					640		
30	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		
35				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
 5 690 695 700

Asp Phe Thr Phe Asp Tyr Asp Ile Ala Leu Leu Glu Leu Glu Lys Pro
 705 710 715 720

10 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

15 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
 20 770 775 780

Thr Pro Arg Met Met Cys Val Gly Phe Leu Ser Gly Gly Val Asp Ser
 785 790 795 800

25 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

30 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
 35 850 855

<210> 34
 <211> 855
 <212> PRT
 5 <213> Mus musculus

<400> 34
 Met Gly Ser Asn Arg Gly Arg Lys Ala Gly Gly Gly Ser Gln Asp Phe
 1 5 10 15
 10 Gly Ala Gly Leu Lys Tyr Asn Ser Arg Leu Glu Asn Met Asn Gly Phe
 20 25 30
 15 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
 20 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
 25 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
 30 115 120 125
 Asn Glu Val Pro Val Leu Gly Pro Tyr His Lys Lys Ser Ala Val Thr
 130 135 140
 35 Ala Phe Ser Glu Gly Ser Val Ile Ala Tyr Tyr Trp Ser Glu Phe Ser

Ser Phe Lys Cys Ser Asn Gly Lys Cys Leu Pro Gln Ser Gln Lys Cys
 530 535 540

5 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

10 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
 15 595 600 605

Phe Thr Lys Gln Ala Arg Val Val Gly Gly Thr Asn Ala Asp Glu Gly
 610 615 620

20 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

25 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
 30 675 680 685

Val Gln Glu Leu Lys Leu Lys Arg Ile Ile Thr His Pro Ser Phe Asn
 690 695 700

35 Asp Phe Thr Phe Asp Tyr Asp Ile Ala Leu Leu Glu Leu Glu Lys Ser

	Met	Asn	Trp	Asn	Thr	Lys	Gln	Glu	Asn	Val	Pro	Lys	Pro	Pro	Pro	Tyr
	1				5					10						15
5	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			
10	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
15	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
20				100					105					110		
	Gly	Lys	Gly	Pro	Gln	Gln	Pro	Asn	His	Asn	Leu	Gln	Thr	Val	Ser	Ser
			115					120					125			
25	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
30	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
35				180						185				190		

Gly Leu Asn His Phe Ile Pro Asp Gln Leu Val Asp Trp Thr Gln Tyr
 195 200 205

5 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

10 Gln Leu Pro Thr Tyr Thr Gln Ser Leu Gln Ser Lys His Ser Val Pro
 245 250 255

15 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

20 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

25 Trp Gln Asn Pro Asn Glu Lys Val Ser Ile Gly Gln Phe Ser Glu Val
 325 330 335

30 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

35 Thr Tyr Asn Pro Asn Thr Asn Gln Val Ile Asp Thr Asn Ala Thr Lys

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
 5 580 585 590

Glu Thr Val Ala Leu Pro Phe Asp Val Ile Gly Ala Val Ala Ser Asn
 595 600 605

10 Asn Ile Ser Ala Glu Ile Pro Leu Pro Val Asp Lys Glu Lys Gln His
 610 615 620

Lys Pro Ile Gln Gly Asp Pro Asp Ile Ala Asp Ser Ser Leu Gly Lys
 625 630 635 640

15 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
 20 660 665 670

Glu Gly Asp Val Ser Tyr Asn Ser Gln Ile Ala Glu Ile Phe Asn Ser
 675 680 685

25 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

30 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
 35 740 745 750

	Val	Glu	Ala	Asn	Arg	Glu	Ala	Thr	Glu	Glu	Ser	Cys	Arg	Glu	Tyr	Thr	
			755					760					765				
5	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	
10	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	
15				820					825						830		
	Gly	Ile	Glu	Thr	Phe	Asn	Arg	His	Glu	Val	Ser	Leu	Asp	Gln	Gln	Lys	
			835					840					845				
20	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	
25	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	
30				900					905						910		
	Asp	Ser	Gln	Ile	Pro	Arg	Glu	Glu	Ser	His	Asp	Leu	Gly	Met	Leu	Asp	
			915						920					925			
35	Pro	Glu	Lys	Asp	Lys	Ile	His	Cys	Cys	Ala	Leu	Gly	Trp	Leu	Ser	Met	

	Asp Leu Gly Pro Val Lys Ala Pro Ile Glu Leu Ser Ser Asn Thr
	1115 1120 1125
5	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
10	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
15	1175 1180 1185
	Ala Gly Asp Thr Leu Lys Leu Cys Ser Ile Leu Thr Glu Ser Asn
	1190 1195 1200
20	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
25	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
30	1250 1255 1260
	Ser Gly His Thr Lys Thr Ser Lys His Ser Ala Ala Val Ser Trp
	1265 1270 1275
35	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	Ser His	Gly Lys	Asp Phe	Lys Ile		
	1295			1300			1305			
5	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			
10	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		
15	1355			1360			1365			
	Thr Glu	Arg Glu	Ser Ile	Cys Leu	Thr Lys	Leu Asp	Ser Ala	Ser		
	1370			1375			1380			
20	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			
25	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		
30	1430			1435			1440			
	Ser Gly	Ile Lys	Ser Thr	Lys Glu	Asp Trp	Leu Lys	Cys Ile	Pro		
	1445			1450			1455			
35	Thr Arg	Thr Lys	Met Pro	Glu Ser	Ser Ser	Gln Arg	Asp Ser	Ala Asp		

<223> /note="Description de séquence artificielle : Synthétique peptide"

<400> 37

5 Asp Tyr Lys Asp Asp Asp Asp Lys
1 5

<210> 38
<211> 9

10 <212> PRT
<213> Séquence artificielle

<220>
<221> source

15 <223> /note="Description de séquence artificielle : Synthétique peptide"

<400> 38

Ala Trp Arg His Pro Gln Phe Gly Gly
20 1 5

<210> 39
<211> 8
<212> PRT

25 <213> Séquence artificielle

<220>
<221> source

<223> /note="Description de séquence artificielle : Synthétique peptide"

30

<400> 39

Trp Ser His Pro Gln Phe Glu Lys
1 5

REVENDICATIONS AMENDEES

1. Cellule CHO recombinante, comprenant :
une ou plusieurs séquences polynucléotidiques codant pour le
complexe pentamère du cytomégalo virus (CMV), où ledit
complexe pentamère comprend : la gH ou l'un de ses fragments
5 formant un complexe, la gL ou l'un de ses fragments formant
un complexe, la pUL128 ou l'un de ses fragments formant un
complexe, la pUL130 ou l'un de ses fragments formant un
complexe, et la pUL131 ou l'un de ses fragments formant un
complexe ;

10 où lesdites une ou plusieurs séquences polynucléotidiques
sont intégrées dans l'ADN génomique de ladite cellule CHO et
où l'expression ou l'activité de la protéine FAM60A est
réduite dans ladite cellule, comparativement à un témoin.

15 2. Cellule CHO selon la revendication 1, où ladite
cellule CHO est une cellule CHO-K1, CHO-DUXB11, CHO-DG44, ou
CHO-S.

20 3. Cellule CHO selon la revendication 1 ou 2, où au
moins une copie de la séquence génomique du gène FAM60A, ou
au moins 50 % de la séquence codante dudit gène FAM60A, est
délétée.

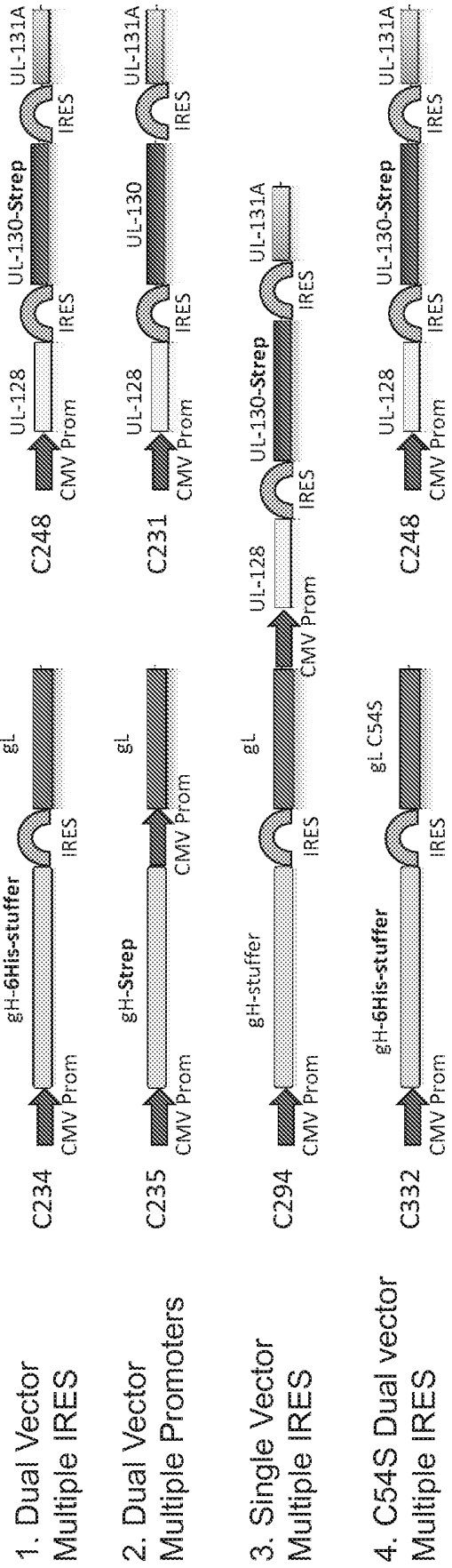
4. Cellule CHO selon l'une quelconque des
revendications 1 à 3, où ledit fragment formant un complexe
de la gH comprend l'ectodomaine d'une protéine gH pleine
longueur.

25 5. Cellule CHO selon l'une quelconque des
revendications 1 à 4, où ledit complexe pentamère est soluble.

6. Cellule CHO selon l'une quelconque des
revendications 1 à 5, où ledit complexe pentamère est sécrété
à partir de la cellule hôte.

7. Culture à grande échelle comprenant la cellule CHO BE2016/5877 selon l'une quelconque des revendications 1 à 6, où ladite culture a une taille d'au moins 20 litres.

5 8. Culture à grande échelle selon la revendication 7, où le rendement du complexe pentamère du CMV est d'au moins 0,05 g/l.



pBWah: Amp/Hyg/Fra

pBWKn: Kan/neo/DHFR*

Figure 1

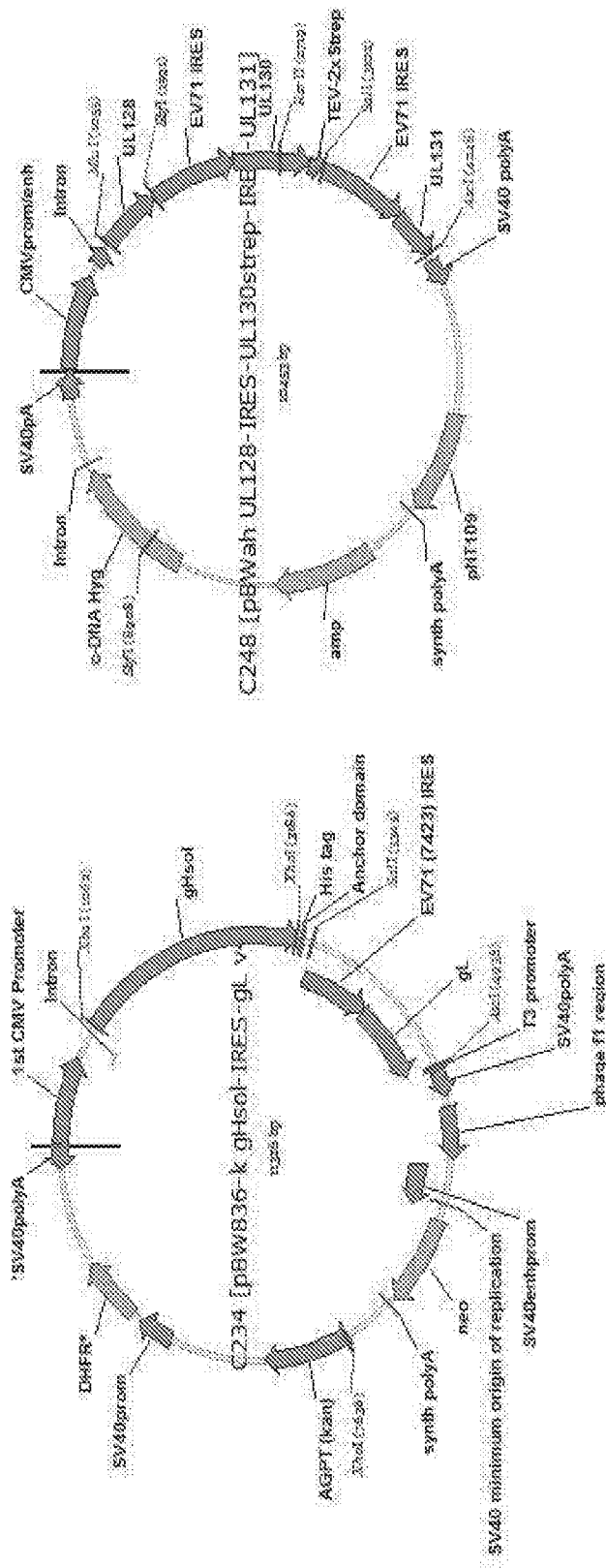


Figure 2

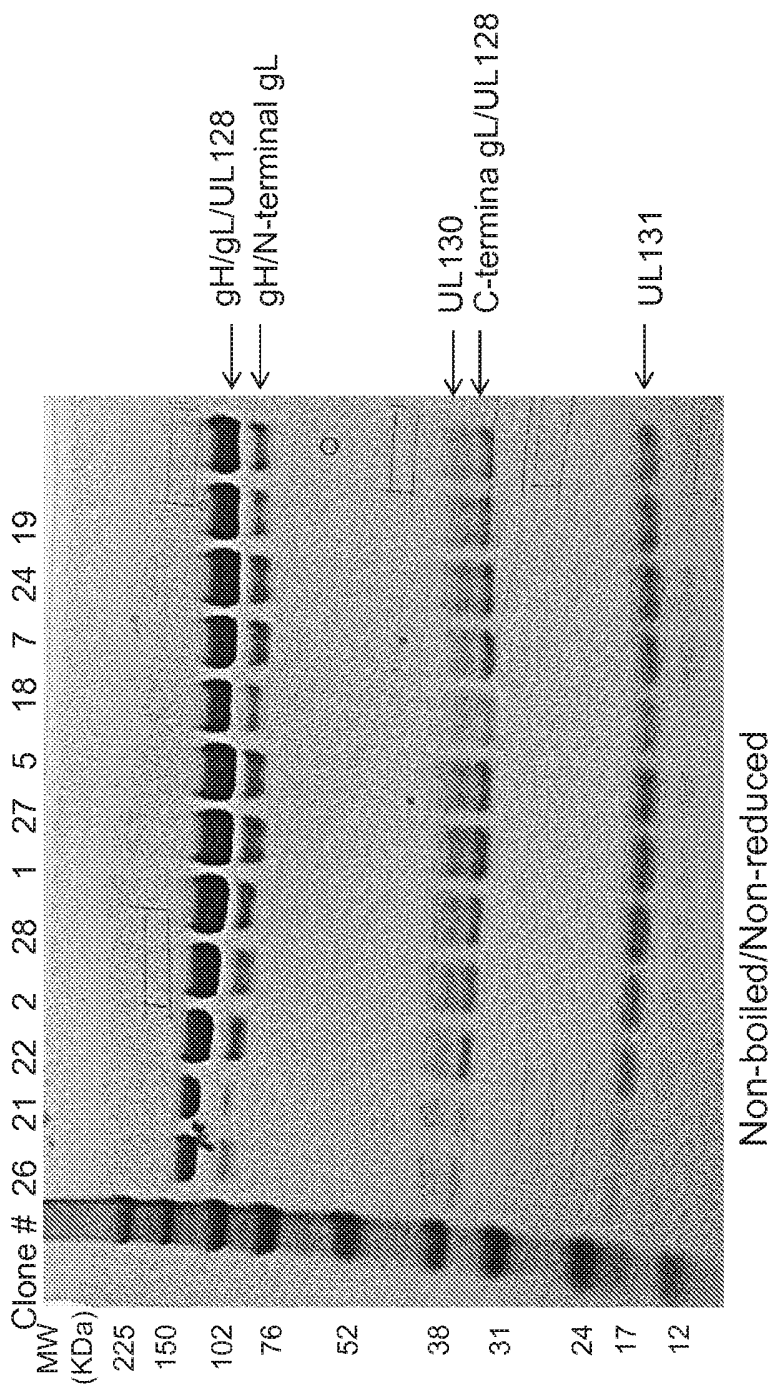


Figure 3

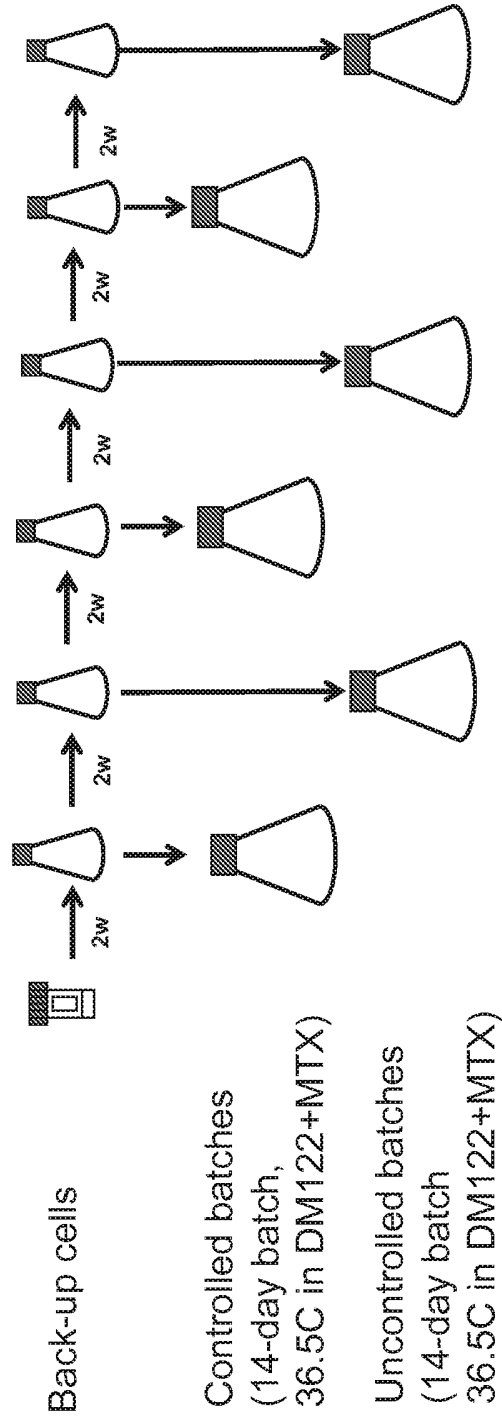


Figure 4

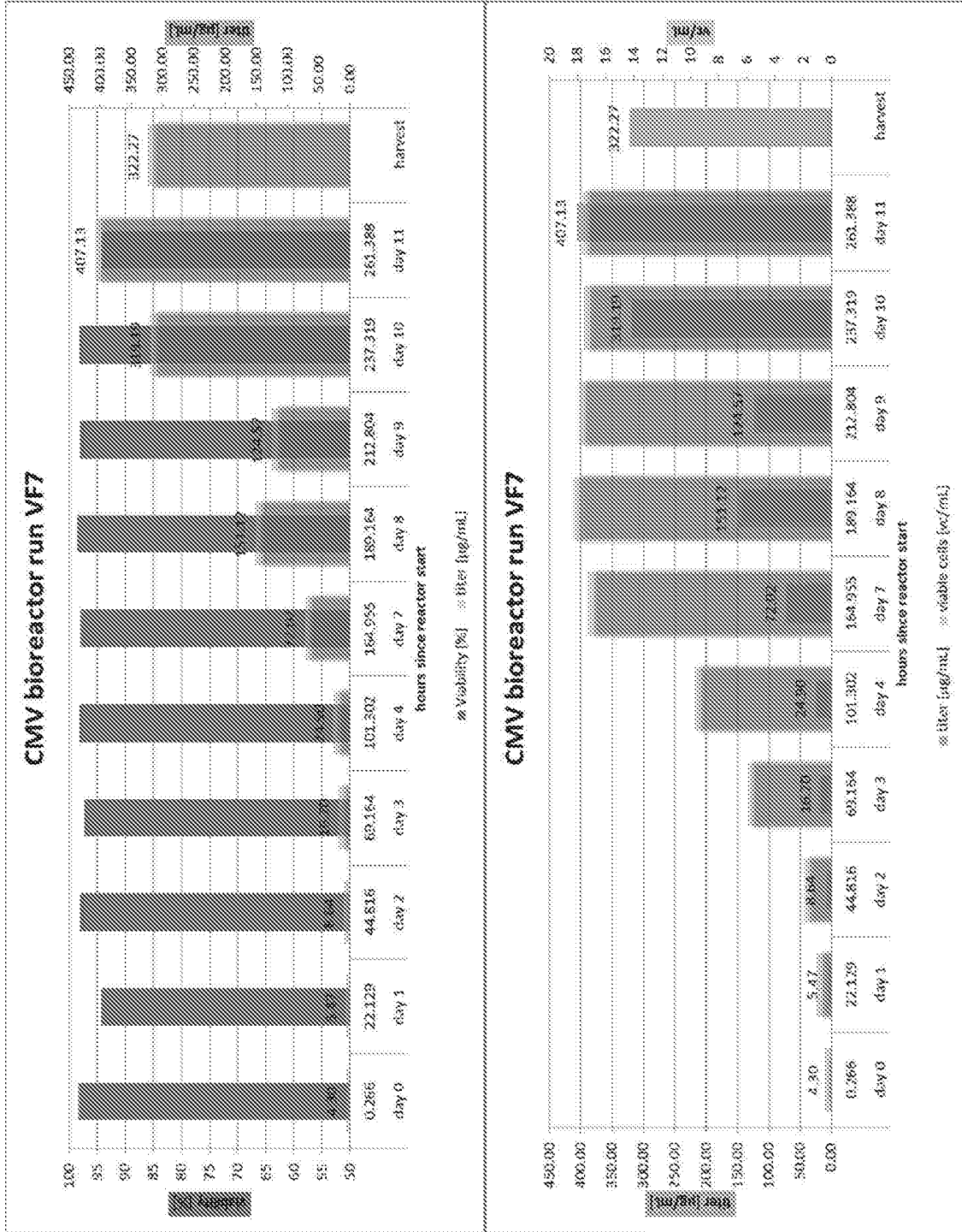


Figure 5

ABREGECELLULES DE MAMMIFERES EXPRIMANT DES ANTIGENES DU
CYTOMEGALOVIRUS

Cette invention concerne des protéines du cytomégalovirus (CMV) appropriées pour des utilisations vaccinales. Sont fournies ici des cellules hôtes de mammifères, en particulier des cellules CHO, dans lesquelles la ou les séquences codant
5 pour les protéines du CMV gH, gL, pUL128, pUL130, pUL131 (ou l'un de leurs fragments formant un complexe) sont intégrées de façon stable dans le génome.



Numero de la demande nationale

RAPPORT DE RECHERCHE
 établi en vertu de l'article 21 § 1 et 2
 de la loi belge sur les brevets d'invention
 du 28 mars 1964

BO 11365
 BE 201605877

DOCUMENTS CONSIDERES COMME PERTINENTS			
Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes	Revendication concernée	CLASSEMENT DE LA DEMANDE (IPC)
X,D	WO 2014/005959 A1 (NOVARTIS AG [CH]) 9 janvier 2014 (2014-01-09) * page 14, ligne 4 - page 17, ligne 35; revendications; exemples * -----	9	INV. C07K14/005 A61K39/245 C12N5/071
X	WEN Y ET AL: "Human cytomegalovirus gH/gL/UL128/UL130/UL131A complex elicits potently neutralizing antibodies in mice", VACCINE, vol. 32, no. 30, 14 mai 2014 (2014-05-14), pages 3796-3804, XP029006326, ISSN: 0264-410X, DOI: 10.1016/J.VACCINE.2014.05.004 * materials and methods; * -----	9	
X,P	HOFMANN I ET AL: "Expression of the human cytomegalovirus pentamer complex for vaccine use in a CHO system", BIOTECHNOLOGY AND BIOENGINEERING., vol. 112, no. 12, 31 juillet 2015 (2015-07-31), pages 2505-2515, XP055241307, US ISSN: 0006-3592, DOI: 10.1002/bit.25670 * abstract; materials and methods; * -----	9	
E	WO 2015/165480 A1 (INST RESEARCH IN BIOMEDICINE [CH]) 5 novembre 2015 (2015-11-05) * revendications; exemples * ----- -/--	9	
		Date d'achèvement de la recherche	Examineur
		22 décembre 2016	Sommer, Birgit
CATEGORIE DES DOCUMENTS CITES		T : théorie ou principe à la base de l'invention E : document de brevet antérieur, mais publié à la date de dépôt ou après cette date D : cité dans la demande L : cité pour d'autres raisons ----- & : membre de la même famille, document correspondant	
X : particulièrement pertinent à lui seul Y : particulièrement pertinent en combinaison avec un autre document de la même catégorie A : arrière-plan technologique Q : divulgation non-écrite P : document interclassaire			

1

EPD FORM 1603 03 02 (15/04/05)



Numero de la demande nationale

RAPPORT DE RECHERCHE
 établi en vertu de l'article 21 § 1 et 2
 de la loi belge sur les brevets d'invention
 du 28 mars 1964

BO 11365
 BE 201605877

DOCUMENTS CONSIDERES COMME PERTINENTS			
Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes	Revendication concernée	CLASSEMENT DE LA DEMANDE (IPC)
A	KIM J Y ET AL: "CHO cells in biotechnology for production of recombinant proteins: current state and further potential", APPLIED MICROBIOLOGY AND BIOTECHNOLOGY, SPRINGER, BERLIN, DE, vol. 93, no. 3, 9 décembre 2011 (2011-12-09), pages 917-930, XP035006081, ISSN: 1432-0614, DOI: 10.1007/S00253-011-3758-5 * le document en entier * -----	1-9	
			DOMAINES TECHNIQUES RECHERCHES (IPC)
		Date d'achèvement de la recherche	Examineur
		22 décembre 2016	Sommer, Birgit
CATEGORIE DES DOCUMENTS CITES			
X : particulièrement pertinent à lui seul Y : particulièrement pertinent en combinaison avec un autre document de la même catégorie A : arrière-plan technologique Q : divulgation non-écrite P : document interclassaire		T : théorie ou principe à la base de l'invention E : document de brevet antérieur, mais publié à la date de dépôt ou après cette date D : cité dans la demande L : cité pour d'autres raisons & : membre de la même famille, document correspondant	

1

EPD FORM 1603 03 02 (P)4C489

**ANNEXE AU RAPPORT DE RECHERCHE
RELATIF A LA DEMANDE DE BREVET BELGE NO.**

BO 11365
BE 201605877

La présente annexe indique les membres de la famille de brevets relatifs aux documents brevets cités dans le rapport de recherche visé ci-dessus.
Lesdits membres sont contenus au fichier informatique de l'Office européen des brevets à la date du
Les renseignements fournis sont donnés à titre indicatif et n'engagent pas la responsabilité de l'Office européen des brevets

22-12-2016

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
WO 2014005959 A1	09-01-2014	AU 2013286093 A1	12-02-2015
		CA 2878344 A1	09-01-2014
		CN 104853771 A	19-08-2015
		EP 2869843 A1	13-05-2015
		GB 2513768 A	05-11-2014
		HK 1210037 A1	15-04-2016
		JP 2015522581 A	06-08-2015
		KR 20150031322 A	23-03-2015
		RU 2015103891 A	27-08-2016
		SG 11201408328V A	27-02-2015
		US 2016159864 A1	09-06-2016
		WO 2014005959 A1	09-01-2014
		-----	-----
WO 2015165480 A1	05-11-2015	AUCUN	
-----	-----	-----	-----



OPINION ÉCRITE

Dossier N° BO11365	Date du dépôt (jour/mois/année) 29.10.2015	Date de priorité (jour/mois/année) 31.10.2014	Demande n° BE201605877
Classification internationale des brevets (CIB) INV. C07K14.005 A61K39.245 C12N5.071			
Déposant GLAXOSMITHKLINE BIOLOGICALS S.A.			

La présente opinion contient des indications et les pages correspondantes relatives aux points suivants :

- Cadre n° I Base de l'opinion
- Cadre n° II Priorité
- Cadre n° III Absence de formulation d'opinion quant à la nouveauté, l'activité inventive et la possibilité d'application industrielle
- Cadre n° IV Absence d'unité de l'invention
- Cadre n° V Déclaration motivée quant à la nouveauté, l'activité inventive et la possibilité d'application industrielle: citations et explications à l'appui de cette déclaration
- Cadre n° VI Certains documents cités
- Cadre n° VII Irrégularités dans la demande
- Cadre n° VIII Observations relatives à la demande

Examineur Sommer, Birgit

OPINION ÉCRITE

Demande n°
BE201605877

Cadre n° I Base de l'opinion

1. Cette opinion a été établie sur la base des revendications déposées avant le commencement de la recherche.
2. En ce qui concerne **la ou les séquences de nucléotides ou d'acides aminés** divulguées dans la demande, le cas échéant, cette opinion a été effectuée sur la base des éléments suivants :
 - a. Nature de l'élément:
 - un listage de la ou des séquences
 - un ou des tableaux relatifs au listage de la ou des séquences
 - b. Type de support:
 - sur papier
 - sous forme électronique
 - c. Moment du dépôt ou de la remise:
 - contenu(s) dans la demande telle que déposée
 - déposé(s) avec la demande, sous forme électronique
 - remis ultérieurement
3. De plus, lorsque plus d'une version ou d'une copie d'un listage des séquences ou d'un ou plusieurs tableaux y relatifs a été déposée, les déclarations requises selon lesquelles les informations fournies ultérieurement ou au titre de copies supplémentaires sont identiques à celles initialement fournies et ne vont pas au-delà de la divulgation faite dans la demande telle que déposée initialement, selon le cas, ont été remises.
4. Commentaires complémentaires :

Cadre n° II Priorité

Cette opinion a été établie en supposant que la date de priorité revendiquée était pertinente sauf indication contraire mentionnée en **feuille séparée**

Cadre n° V Opinion motivée quant à la nouveauté, l'activité inventive et la possibilité d'application industrielle; citations et explications à l'appui de cette déclaration

1. Déclaration

Nouveauté	Oui :	Revendications	1-8
	Non :	Revendications	9
Activité inventive	Oui :	Revendications	1-8
	Non :	Revendications	9
Possibilité d'application industrielle	Oui :	Revendications	1-9
	Non :	Revendications	

2. Citations et explications

voir feuille séparée

Cadre n° VI Certains documents cités

Certains documents publiés

voir le rapport de recherche

Divulgations non écrites

Cadre n° VIII Observations relatives à la demande

voir feuille séparée

Il est fait référence au (aux) document(s) suivant(s):

- D1 WO 2014/005959 A1
- D2 WEN Y et al., VACCINE, vol. 32, 2014, pages 3796-3804
- D3 HOFMANN I et al., BIOTECH. BIOENG., vol. 112, 2015, pages 2505-2515
- D4 WO 2015/165480 A1

La présente demande concerne des cellules CHO, dans lesquelles les séquences codant pour les protéines gH, gL, pUL128, pUL130 et pUL131 (ou pour l'un de leurs fragments formant un complexe) du cytomégalovirus humain (CMV) sont intégrées de façon stable dans le génome. De telles cellules CHO fournissent une source pour la production de complexe pentamère de CMV de façon recombinante. Le taux d'expression des protéines peut être amplifiés dans ces cellules CHO par une délétion au moins de parties de la séquence génomique soit du gène C12orf35 soit du gène FAM60A et/ou par une mutation dans l'exon 2 du gène de la matriptase.

Ad point II:

La date de priorité de la présente demande est valablement revendiquée. Pour cette raison le document D3 mentionné comme document P dans le rapport de recherche n'est pas pris en considération.

Ad point V:

- 1 D1 divulgue l'expression du complexe pentamère du CMV par transfection transitoire de cellules 293 (D1, p.ex. exemples).
D2 divulgue également la production d'un complexe pentamère soluble du CMV par transfection transitoire de cellules 293 (D2, p.ex. matériaux et méthodes).

L'objet de la revendication 15 n'est donc pas nouveau parce que la référence à une méthode de production spécifique n'ajoute pas de caractéristiques techniques, qui définissent la composition revendiquée plus précisément et qui peuvent être utilisés pour différencier la composition revendiquée d'une composition comprenant le complexe pentamère du CMV pouvant être obtenu par un autre processus.

La demande antérieure D4 divulgue la production d'une lignée cellulaire CHO stable produisant le complexe pentamère du CMV (D4, par ex. revendications, exemples). D4 anticipe la nouveauté des revendications 1, 2 et 10-15.

- 2 D1, qui est considéré comme l'état de la technique le plus proche, divulgue l'expression du complexe pentamère du CMV par transfection transitoire de cellules 293 (D1, p.ex. revendications, exemples).

Le problème que la présente invention se propose de résoudre peut être considéré comme la mise à disposition d'autres cellules CHO pour la production d'un complexe pentamère du CMV. La présente demande résout ce problème en mettant à disposition des cellules CHO recombinantes ayant des séquences polynucléotidiques codant pour un complexe pentamère du CMV intégrées dans leur génome et où l'expression ou l'activité de la protéine FAM60A est réduite dans ladite cellule, comparativement à un témoin.

L'objet de la revendication 15 n'est ni nouveau ni inventif (p. ex. D1, D2).

Par contre, l'état de la technique n'enseigne ni suggère l'utilisation de lignées de cellules CHO où l'expression ou l'activité de la protéine FAM60A est réduite pour la production d'un complexe pentamère du CMV. L'activité inventive des revendications 1-8 peut être reconnue.

Ad point VIII:

- 3 Les revendications 1 et 15 manquent de clarté en raison d'expression «... l'un de ses fragments formant un complexe...». Cette expression n'est pas appropriée pour définir clairement la portée des revendications de telle sorte que l'objet n'est pas suffisamment défini par des caractéristiques techniques de l'invention. Un acide nucléique ou aminé en général, doit être considéré comme un composé chimique qui doit être clairement caractérisé par référence aux caractéristiques techniques, par exemple, par sa séquence, mais non seulement par sa fonction ou par le résultat recherché.
- 4 Les revendications 1, 5, 6 et 8 ne satisfont pas aux exigences de clarté car l'objet de la protection demandée n'est pas clairement défini. La revendication tente de définir l'objet par le résultat recherché. Cette formulation n'est pas acceptable en l'espèce, puisqu'il semble possible de définir l'objet en des termes plus concrets, c'est-à-dire en exposant comment l'effet peut être obtenu.
- 5 Le terme «... comparativement à un témoin ...» employé dans la revendication 1 est vague et imprécis, et laisse subsister un doute quant à la signification de la caractéristique technique à laquelle il se rapporte, au point que l'objet de ladite revendication n'est pas clairement défini. Un taux d'expression ou l'activité réduite d'une protéine spécifique ne suffit pas pour caractériser

clairement une cellule CHO et pour la distinguer de l'état de la technique, puisque ledit paramètre dépend fortement du témoin sélectionné pour sa détermination.